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

4 Laura Uelze¹

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

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

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

8 **Abstract**

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

27 INTRODUCTION:

28 Peptidoglycan and teichoic acids are the main structural components of the *Bacillus*
 29 *subtilis* cell wall (Bhavsar and Brown, 2006; Barreteau *et al.*, 2008; Sauvage *et al.*,
 30 2008; Vollmer *et al.*, 2008; Sekiguchi and Yamamoto, 2012). Peptidoglycan is
 31 composed of glycan chains formed by polymerization of a disaccharide of *N*-
 32 acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc). Stem pentapeptides
 33 composed (proximal to distal order) of amino acids L-alanine (position 1), D-glutamate
 34 (position 2), *meso*-diaminopimelate (position 3), D-alanine (position 4) and D-alanine
 35 (position 5) extend from the lactyl group of MurNAc. The glycan strands are linked by
 36 formation of a covalent peptide bond between the carboxyl group of D-alanine (position
 37 4) of one strand [with concomitant loss of D-alanine (position 5)] and the amino group of
 38 *meso*-diaminopimelate (position 3) of a different strand. Glycan strands can be up to 200
 39 disaccharides in length with a stem peptide linkage index of up to ~33%. Thus
 40 peptidoglycan forms a mesh outside the cell to which wall-teichoic acids composed of
 41 poly(ribitol- or glycerol-phosphate) are covalently attached (Weidenmaier and Peschel,
 42 2008; Swoboda *et al.*, 2010; Kawai *et al.*, 2011). This rigid mesh structure confers shape
 43 on the cell and is capable of withstanding turgor pressures of up to 5 atmospheres
 44 (Sekiguchi and Yamamoto, 2012). Interference with peptidoglycan metabolism, either
 45 through inhibition of its synthesis by antibiotics (*e.g.* penicillin and vancomycin) or an
 46 imbalance between synthesis and turnover leads to aberrant cell structure, cell division
 47 and often cell death. Thus maintenance of cell wall integrity is a prerequisite for normal
 48 cell viability.

49 The dramatic changes that occur in cell wall structure during growth and division require
50 exquisite control and coordination of cell wall synthesis and turnover. Autolysins play a
51 key role in sculpting the cell wall during these processes by cleavage of glycosidic
52 linkages in the glycan chains and peptide linkages in the stem peptides (Smith *et al.*,
53 2000; Sekiguchi and Yamamoto, 2012). *B. subtilis* encodes a cohort of 35 enzymes with
54 glucosaminidase, muraminidase, amidase and endopeptidase activities that confer the
55 capability to cleave all covalent linkages of peptidoglycan thereby allowing it to be
56 sculpted, remodeled and turned over during cell growth and division (Smith *et al.*, 2000;
57 Sekiguchi and Yamamoto, 2012). A notable feature of the autolysin cohort of *B. subtilis*
58 is the degree of redundancy of individual enzymatic activities (Smith *et al.*, 2000;
59 Sekiguchi and Yamamoto, 2012). For example, there are seven enzymes (CwlO, CwlS,
60 LytE, LytF, PgdS, YkfC, YqgT) with D,L-endopeptidase activity encoded in the *B. subtilis*
61 genome, each of which cleaves the stem peptide between D-glutamate (position 2) and
62 meso-diaminopimelate (position 3). However these enzymes differ in the type and/or
63 number of domains that interact with cell wall components. Thus the LytE, CwlS and
64 LytF D,L-endopeptidases have 3, 4 and 5 LysM domains respectively located at the
65 amino terminus of the protein, a domain that functions to bind proteins to peptidoglycan
66 (Mesnage *et al.*, 2014). In contrast, the N-terminal domain of CwlO functions to bind the
67 enzyme to the FtsEX ATP-type transporter that is located in the cell membrane
68 (Yamaguchi *et al.*, 2004; Dominguez-Cuevas *et al.*, 2013; Meisner *et al.*, 2013). Thus
69 variance in these proteins regions allows differential localization and binding affinity of
70 these enzymes to elements located within the cell envelope.
71 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 CwO 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 CwO 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 CwO 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 CwO 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 *cwO* 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 CwIS (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); (iii) 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 identity in their catalytic domain, thereby identifying orthologues of CwIS, 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.

Strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Supporting Information Tables S2 and S3 respectively. Oligonucleotides were obtained from IDT, (Belgium) and are listed in the supporting information Table S4. *Escherichia coli* strain TG-1 (Sambrook *et al.*, 1989) was used for propagating plasmids. *E. coli* strain BL21(DE3) was used for protein overexpression and purification. Strains were grown in Luria-Bertani (LB) medium and in salt-free buffered LB (Boylan *et al.*, 1993), to which NaCl was added up to the desired concentration. The high phosphate defined medium (HPDM) and low phosphate defined medium (LPDM) used, are described in Botella *et al.*, (2011). Antibiotics were added at concentrations described by Noone *et al.*, (2014). OD₆₀₀ was measured using a UVmini-1240 UV-VIS spectrophotometer (Shimadzu Scientific Instruments). *B. subtilis* 168 was transformed according to the procedure of Anagnostopoulos and Spizizen (1960). Growth studies were conducted with the BioTek Synergy™2 system (BioTek Instruments) at 37°C with slow shaking. Three technical repeats were performed for each growth curve.

Strain construction

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

by strand overlap PCR (Horton *et al.*, 1989). This linear DNA fragment was then transformed into *B. subtilis* strain 168 with selection for spectinomycin resistance, thereby generating strain LU1 in which the *lytE* gene is replaced by the *spc* gene.

A second *lytE* null mutant strain (LU331) was generated in a similar manner. The strain differentiates from LC01, in that part of the *lytE* signal sequence was preserved in this mutant. The *spc* gene was amplified as described previously. Chromosomal regions located upstream and downstream of *lytE* were amplified by PCR using primer pairs *LytE_Up_Fwd_2* / *Signal_LytE_Rev_Spec* and *LytE_Do_Fwd_Spec* / *LytE_Do_Rev*. The DNA fragments were joined to the fragment encoding the *spc* gene by strand overlap PCR and the linear DNA fragment was transformed into WT168, with selection for spectinomycin resistance, to generate strain LU331.

Strains were constructed that express the *LytE* protein from the *cw/O* promoter (LU17, LU21, LU25). These strains were generated through transcriptional fusion between the *cw/O* promoter and the *lytE* gene. To generate strain LU17, four PCR fragments were amplified. First, the DNA fragments located upstream and downstream of *cw/O* locus were amplified with the primer pairs *Cw/O_UF_Sma* / *Cw/O_UR* and *QE93_Fwd_Ery* / *QE_94_Rev* respectively. Then the *lytE* open reading frame was amplified with the primer pair *LytE_DF* and *LytE_DR_Ery*. Lastly, the erythromycin resistance gene *erm* was amplified with primer pair *Mls_Fwd* / *Mls_Rev* from plasmid pDG6464. All PCR fragments were joined together by strand overlap extension PCR and transformed into strain EB35 with selection for erythromycin resistance. The resulting strain LU17 encodes two copies of the *lytE* gene (one at the *cw/O* locus and one 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 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 *cw/O* locus.

Another similar strain (LU25) was generated with the additional feature that the leader region of *cw/O* 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 *cw/O* 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 *cw/O* mutant strain (LS103) served as proof for the functionality of the HA-tag, given that a failure would lead to a combined *cw/O* / *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 ($\Delta lytF$) was confirmed by PCR.

In order to generate a strain that expresses the LytF protein from the *cw/O* promoter (with *cw/O* leader – LC111, without *cw/O* 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 Cw/O_UF and Cw/O_UR_LytF from WT168 (containing the *cw/O* leader for LU111) or from DN3000 (no *cw/O* 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 *cw/O* promoter (with / without *cw/O* 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 *cw/O* promoter (with / without *cw/O* leader) and are knock-outs for *cw/O* 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 *cw/O* 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 *cw/O* 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 *cw/O* 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

the primer pairs *LytE_Up_Fwd_2* / *Signal_LytE_Rev* and *LytE_Do_Fwd_Ery* / *LytE_Do_Rev* from chromosomal DNA of WT168. The downstream fragment was joined to an erythromycin resistance cassette, which was amplified from pDG646 with the primers *Mls_Fwd* / *Mls_Rev*. The upstream fragment was joined to fragments of the variant *lytE* genes, which were amplified with the primers *Signal_LytE_Fwd* / *Catalytic_LytE_Rev_Ery* from the plasmids pLU40, pLU23, pLU24, pLU25, pLU20, pLU27, pLU37, pLU16, pLU17, pLU14 and pLU29. After joining of the upstream fragment (upstream homology to the *lytE* locus plus variant *lytE* gene sequence) to the downstream fragment (erythromycin resistance cassette plus downstream homology to the *lytE* locus) by strand overlap extension PCR, the linear PCR fragments were then transformed into strain LU331 with selection for erythromycin resistance, to create the strains LU335, LU337, LU339, LU341, LU372, LU374, LU398, LU399, LU401, LU403 and LU405 respectively.

In order to disrupt the *cw/O* gene in a series of strains, expressing a variant *LytE* from the *lytE* promoter, a fragment encompassing the *cw/O* locus was amplified with the primer pair *Cw/O_UF* / *QE94_Rev* from chromosomal DNA of the *cw/O* null mutant strain EB35. The PCR fragment was then transformed by double cross-over into the strains LU335, LU337, LU339, LU341, LU372, LU374, LU399, LU401, LU403 and LU405 with selection for kanamycin resistance to obtain strains LU343, LU350, LU352, LU354, LU383, LU385, LU407, LU409, LU411 and LU413.

In order to investigate variant *LytE* protein expression levels, the variant *LytE* proteins were joined to a HA-tag which can be detected by western analysis. To enable a C-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_Sall* and cloning it into pLU2 via the Sall 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 *cw/O* locus, thereby exchanging the *cw/O* against the *lytE* open reading frame. Plasmid

296 pLU10 was generated in a two-step process: First, upstream homology (containing the
297 *cw/O* leader and promoter and part of the *lytE* signal sequence, with a silent BamHI
298 restriction site) was amplified with Cw/O_UF_LysM and Cw/O_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 Cw/O_DF_LysM and
301 Cw/O_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.

342 Cloning the L2 fragment in pLU18 and pLU14 resulted in the plasmid pLU30 and pLU40.
 343 Cloning the L3 fragment in pLU20 and pLU19 resulted in the plasmids pLU27 and
 344 pLU28.
 345 To create plasmids that would allow the overexpression and purification of variant LytE
 346 proteins in *E. coli*, the respective protein coding sequences (excluding the start codon
 347 and the *lytE* signal sequence) were amplified with the forward primers Exp_L1_Fwd,
 348 Exp_L3_Fwd, Exp_L1_Fwd, Exp_L1_Fwd, Exp_L3_Fwd and Exp_Spac_Fwd from the
 349 plasmids pLU40, pLU20, pLU14, pLU23, pLU25 and pLU37 to generate the plasmids
 350 pLU47, pLU49, pLU48, pLU50, pLU51 and pLU52. The same reverse primer
 351 (Exp_Cat_Rev) was used in all reactions. PCR fragments were cloned into the
 352 NdeI/XhoI sites of the vector pET21b+ thereby re-generating the start codon and
 353 generating a C-terminal fusion between the protein and a 6xHis-tag.

354 **RNA extraction and RT-qPCR**

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

361 **Western analysis**

362 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 β -mercaptoethanol, 1 % TWEEN20) in a 25 μ l 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 μ l assay
383 buffer.

The level of LytE protein that bound to the cell wall, and that which remained in the supernatant 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 similar 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 CwIS, 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 CwIS 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 CwIS 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 CwIS 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 CwIS (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 CwIS orthologs with 3 instead of 4 LysM domains (group X). The relationships between the 185 LysM domains in this cohort of LytE, LytF and CwIS 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, CwIS) and (iv) the LysM domains of LytF and CwIS are more similar to each other than to those of LytE. The three LysM domains of the CwIS 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 CwIS 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 *cw/O* promoter

As discovered by Bisicchia et al. (2007), LytE and Cw/O 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 Cw/O possess similar C-terminal domains, they have different N-terminal domains and are expressed under different circumstances. The regulation of

450 CwIO expression and activity is well described by Noone et al. (2014). In short, the TCS
 451 WalRK is the sole activator for CwIO expression, which is therefore only expressed
 452 during exponential growth, when the WalRK system is active (Salzberg *et al.*, 2013).
 453 LytE expression is also activated by the WalRK system, but LytE expression can also be
 454 triggered by active SigI under stress conditions (Tseng and Shaw, 2008; Tseng *et al.*,
 455 2011; Hashimoto *et al.*, 2012).

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

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

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

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

that overexpressing *LytE* at this high level affects cell morphology with a fraction of cells showing bulges and thickening [Figure 2, b]. In short, both *lytE* null mutants as well as strains overexpressing *LytE*, exhibit difficulties to maintain cell shape and longitudinal growth.

In summary, this results confirm that the *LytE* D,L-endopeptidase can complement *CwlO* when expressed using *cwlO* expression signals - both transcriptional (i.e. promoter) and post-transcriptional (i.e. *cwlO* leader sequence). It is clear however that the two enzymes are not equivalent since the *SigI* regulon (signalling cell wall stress), is induced when *lytE* is expressed using the normal *cwlO* expression signals (i.e. promoter and leader sequence), but to a lesser extent when the *cwlO* leader sequence is removed.

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

The synthetic lethality of the *CwlO* / *LytE* autolysins allows further investigation of whether any other D,L-endopeptidase encoded in the genome can complement their activity. Thus, it can be investigated which D,L-endopeptidase can support growth and viability when expressed under the control of the *cwlO* promoter.

In view of these findings, we sought to establish if the D,L-endopeptidases *LytF* and *CwlS*, - which we previously showed to be closely related to *LytE* (LysM domains), as well as another autolytic enzyme, - *YqgT* (no LysM domain), can fulfil the D,L-endopeptidase requirement when expressed using the *cwlO* promoter.

To establish whether the *CwlS*, *LytF* or *YqgT* D,L-endopeptidases when expressed from the *cwlO* promoter can fulfill the D,L-endopeptidase requirement, we constructed strains 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 *cw/O* 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 *cw/S* or *yqgT*. LytF levels in the strains expressing LytF from the *cw/O* 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 *cw/O* 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 Cw/O 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 Cw/S (4 LysM domains) can support growth of *B. subtilis*, when expressed from the *cw/O* 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

18,246 proteins), all identified orthologs to LytE, CwlS and LytF feature at least 2 LysM domains.

There is little known about how individual LysM domains bind to peptidoglycan and the significance of the fact that the identified D,L-endopeptidase orthologs possess a minimum of 2 LysM domains.

Therefore we sought to investigate the properties of the individual LysM domains of LytE by designing LytE enzymes with a single LysM domain. We then tested the characteristics of these LytE variants proteins in strains which viability was depended on the permutated autolysin. This was achieved by expressing the variant LytE containing either L1, L2 or L3 in single copy at the *cwlO* or the *lytE* locus and attempting to disrupt either the endogenous *lytE* or the *cwlO* gene [Supplementary Figure 2]. For control purposes, we also included a variant LytE without any LysM domain as well as a reconstituted wild type LytE with the three LysM domains in the wild type combination (L1-L2-L3) in the further analysis.

We found that all constructed strains were viable, demonstrating that any single LysM domain of LytE is capable of inferring functionality to the LytE enzyme. A control strain with a LytE protein without a LysM domain could not be constructed, demonstrating that the LytE enzyme requires at least one LysM domain for functionality.

Despite the fact that any individual LysM domain enabled a functional LytE enzyme, we detected a significant change in cell morphology with shorter and thicker cells for the variant strains. Interestingly this effect was greater for strains containing a LytE enzyme with L2 or L3 than L1 [Figure 4, a].

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 (*cwIO* vs *lytE* promoter). Higher expression levels (*cwIO* 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 *cwIO* 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 *cwIO* and *lytE* gene display growth and

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

567 **LysM domain requirement of LytE**

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

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

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

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

LytE protein with a single LysM domain displayed the smallest colonies, which also exhibited strong lysis over the course of the experiment. In general, colonies were bigger and experienced less lysis when they were expressed using the stronger *cw/O* promoter rather than the weaker *lytE* promoter. Interestingly the order of LysM domains also mattered. Those strains that expressed a variant LytE with L1 at the first (N-terminal) position displayed larger colonies, than those in which either L2 or L3 were located adjacent to the signal sequence domain. This correlation was especially strong for LytE variants with a set of 2 LysM domains, as can be seen when comparing L1-L3 to L3-L1.

To address whether the type of variant LytE protein has an effect on the cellular protein level, the level of HA-tagged variant LytE protein was determined by western blot analysis. All variant LytE-HA proteins could be detected and bands migrated according to the predicted MW [Figure 5, b]. The level of variant LytE protein varied significantly in the strains tested. Since the number of LysM domains is proposed to correlate with binding affinity, we expected to find more cell-associated protein in variant proteins with a greater number of LysM domains. Indeed, the level of cell-associated variant LytE protein was higher for proteins with two or more LysM domains, than for proteins with a single LysM domain. If it is assumed that the lower level of cell-associated LytE is caused by lowered binding to the cell wall, and not lowered expression, then we predict the level of variant LytE protein found in the medium to have the reverse profile of that for cell-associated protein. To verify this, we compared the levels of cell-associated variant LytE proteins with those in the culture supernatant. A strong inverse correlation in the level of variant LytE was observed. Variant LytE proteins with one LysM domain

were predominantly found in the culture supernatant, while variant LytE proteins with three LysM domains were predominantly found in association with the cell wall. Moreover the distribution of variant LytE proteins with two LysM domains takes an intermediate between that found for those with three and one LysM domains. We confirmed these results by immunofluorescence microscopy, which demonstrated that HA tagged variant LytE with more LysM domains gave a stronger fluorescence signal than those with fewer LysM domains [data not shown]. Interestingly, the pattern of localization along the cell cylinder was not affected.

We further tested binding affinity of variant LytE protein in vitro in an insoluble peptidoglycan pull down assay. To investigate the effect of salt on the peptidoglycan binding affinity we added increasing amounts of NaCl to the assay buffer (0.05 M - 1 M NaCl) [Figure 5, c]. At a low salt concentration of 0.05 M NaCl, all variant LytE proteins showed an equally strong binding capacity of approximately 100 % (proportion of protein found in the pellet). However, no binding could be detected for the variant LytE protein without LysM domain. Furthermore our results clearly show that the peptidoglycan-binding capacity is negatively correlated to the salt concentration. With an increase in salt concentration (≥ 0.1 M NaCl) binding capacity for the variant LytE proteins is decreased. The rate at which binding capacity is declined varies between proteins and is in negative correlation to the number of LysM domains present. Proteins with more LysM domains exhibited a greater robustness of binding affinity than proteins with a single LysM domain.

In summary, this study demonstrates that the number and order of LysM domains are of great relevance for the functionality of the LytE enzyme. Although a LytE enzyme with

any one of its three LysM domains is functional and can fulfil the D,L-endopeptidase requirement, increasing the number of LysM domains to two or three has a positive effect on normal colony morphology. The order in which the domains are arranged is also of importance in that variant LytE proteins in which the first LysM domain (L1) is localized in its natural position adjacent to the signal peptide fulfil the D,L-endopeptidase requirement to a greater extent than L2 and L3.

Furthermore, the number of LysM domains of a LytE variant protein is positively correlated with peptidoglycan binding affinity and a good determinant of the extent to which the protein is cell-associated or shed into the culture media. We could show that an increase of the number of LysM domains leads to increased binding affinity both *in vivo* and *in vitro*. We further demonstrated that binding affinity is negatively affected by salt *in vitro*.

643 **DISCUSSION:**

644 **The WalRK and SigIRsgl systems work in tandem to ensure adequate expression** 645 **of D,L-endopeptidase activity**

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

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

665 that *LytE* expression was higher when driven by the *cw/O* 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 *cw/O*
 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 *cw/O* promoter with the *cw/O* leader sequence. Based on the
 671 premise that *sigI* is released under stress, removal of the *cw/O* 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 *cw/O* 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 *cw/O*
 678 promoter can fulfil the D,L-endopeptidase requirement for cell viability, it does not
 679 complement Cw/O 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 Cw/O / *LytE* activity when expressed from the *cw/O* promoter. *LytE* naturally fulfils the

D,L-endopeptidase requirement in a *cw/O* mutant, but can also do so when expressed from the *cw/O* locus. LytF cannot complement CwIO / LytE activity when expressed from its natural *lytF* promoter, but gives partial complementation when expressed from the *cw/O* promoter. The fact that LytF is unable to fulfil the D,L-endopeptidase requirement for viability when expressed from the *lytF* promoter, but can do so when expressed from the *cw/O* promoter, indicates that complementation requires either greater than normal LytF protein levels and / or LytF expression at a different time point throughout the growth cycle. However, although a viable strain could be constructed, severe growth and morphology defects were observed. Exacerbation of the aberrant phenotype in strains without *cw/O* leader sequence (causing a further increase in LytF levels) suggests that the level of LytF in this background may be too high for normal growth. We do not know why neither *cw/S* nor *yggT* were able to fulfil the D,L-endopeptidase requirement. However, we ruled out the possibility of a synthetic lethality between these enzymes by constructing viable double mutants.

The fact that the LytE, LytF and CwIS enzymes are highly similar in their C-terminal D,L-endopeptidase catalytic domains, as well as in their N-terminally located LysM domains poses the question as to what causes the inability of CwIS to fulfil the D,L-endopeptidase requirement. We confirmed that there is indeed a high level of CwIS expressed and associated with the cell, when expressed from the *cw/O* promoter [data not shown]. However, it may be of relevance that the CwIS protein is primarily located at the cell septa (Hashimoto *et al.*, 2012), while both LytE and LytF have some presence along the cell cylinder (Hashimoto *et al.*, 2012; Kiriya *et al.*, 2014; Kasahara *et al.*, 2016). This is significant since it has been established that the D,L-endopeptidase

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

The individual LysM domains of LytE are functionally and structurally distinct

We could show that strains expressing a LytE variants with any single LysM domain are viable. However, we observed cell morphology, colony morphology and growth defects in such strains. These phenotypic defects were more pronounced in LytE variants with a single copy of the second (L2) or third (L3) LysM domains as opposed to the N-terminal located L1, indicating that the individual LysM domains are not functionally equivalent. The increased effectiveness of the first LysM domain (L1) is probably due to the fact that it has evolved to be more suitable for interaction with the adjacent signal sequence domain, which perhaps causes the enzyme to be expressed or exported more efficiently.

This finding was confirmed in later experiments when we could show that in strains that express a variant LytE proteins with two or three LysM domains, the order of the

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

LysM domains contribute to peptidoglycan binding in vivo and in vitro

We could show that a variant LytE without LysM domains exhibited no binding affinity to purified cell wall *in vitro* and that it could not sustain cell growth. Thus, LysM domains are essential for cell wall binding and therefore functionality of the LytE enzyme. This observation is supported by several previous studies that have shown that a total loss of LysM domains greatly decreases the peptidoglycan-binding capacity in vitro (Steen *et al.*, 2005; Eckert *et al.*, 2006; Wong *et al.*, 2014). Increasing the number of LysM 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 it 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 binding 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* *cwI*O 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.* 180, 2549–2555. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Ishikawa 1998.pdf.

Kasahara, J., Kiriya, 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.

Kiriya, 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 CsrRS 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 CwIO 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., Furuhashi, K., Fukushima, T., Yamamoto, H., and Sekiguchi, J. (2004). Characterization of a New *Bacillus subtilis* Peptidoglycan Hydrolase Gene, yvcE (Named cwI), 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 2:

a) strain construction

b) microscopy

c) RT-qPCR

d) western blot

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

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 *cw/O* or the *lytE* locus

Figure Legends

Figure 1

Circular phylogenetic tree of the LysM motifs of the identified CwlS, LytE and LytF orthologs.

The amino acid sequence of the LysM motifs of the CwlS, LytE and LytF orthologs were aligned and a phylogenetic tree estimated with Maximum Likelihood method. Bootstrap values are shown (100 replicates). Inner colours correspond to the position of the individual LysM motif within the respective protein starting at the N-terminus (red= first position, blue = second position, green = third position, purple = fourth position, orange = fifth position). Outer colours correspond to the type of the respective catalytic domain (blue=CwlS, pink=LytE, green=LytF). The three LysM motifs of the CwlS-type protein M4KSM2 are highlighted by arrows. The simplified structure of a LytF-type enzyme is shown.

Figure 2

2a: Schematic illustration of strains that express LytE from the *cwlO* promoter.

In both strains, the *cwlO* gene is replaced by the *lytE* open reading frame and the endogenous *lytE* gene is disrupted by a spectinomycin resistance cassette. In LU21 ($P_{cwlO\text{leader}}lytE \Delta cwlO \Delta lytE$) the LytE D,L-endopeptidase is expressed from the *cwlO* promoter with *cwlO* leader sequence. In LU25 ($P_{cwlO}lytE \Delta cwlO \Delta lytE$) the LytE enzyme is expressed from the *cwlO* promoter without *cwlO* leader sequence. Open reading frames for *lytE* and *cwlO* are not contiguous.

2b: Typical cell morphology of strains in which LytE is expressed from the *cwlO* promoter

Comparison of cell morphology of strains in which *lytE* is expressed from the *cwlO* promoter either with (LU17, LU21) or without the *cwlO* leader sequence (LU25) to the wild type (WT168). The presence or absence of the endogenous *lytE* gene is indicated. Cells were grown in LB to an OD600 of 0.1, heat-fixed and examined with an Olympus BX61 microscope. Magnification is 1000x. A scale bar is shown.

2c: Transcript levels of *cwlO*, *lytE*, *sigI* and *mreBH* in strains that express LytE from the *cwlO* promoter.

Transcript levels were determined by RT-qPCR for the following strains: WT168, LU21 ($P_{cwlO\text{leader}}lytE \Delta cwlO \Delta lytE$) and LU25 ($P_{cwlO}lytE \Delta cwlO \Delta lytE$). Cells were grown in HPDM and samples harvested at OD600=0.15, OD600=0.3 and OD600=3.00 respectively. RNA was prepared from the samples and transcribed into cDNA, which was analysed by RT-qPCR. Transcript levels were normalised relative to the level of transcript for the WT168 sample, which was assigned a value of 1. Error bars are shown.

2d: Cell-associated and secreted LytE protein levels in strains in which LytE is expressed from the *cwlO* promoter.

The level of LytE protein associated with the cell (left) and secreted in the medium (right) was established by western blot analysis in the following strains: LU45 (*lytE*-HA),

LU147 (*lytF*-HA Δ *cw/O*), LU161 (*P_{cw/O}lytF*-HA Δ *cw/O* Δ *lytE*) and LU163 (*P_{cw/O}lytF*-HA Δ *cw/O* Δ *lytE*). Cells were grown in HPDM and harvested at OD600=0.2 (t1) and OD600=1.0 (t2) respectively. Total cell protein was prepared from the samples and the HA-tagged *LytE* protein was detected by western analysis. Twenty μ g of total cell protein extract was loaded onto each lane. The amount of protein loaded for medium samples was equivalent to 1 ml medium at an OD600 = 0.3. Molecular weight markers are shown (M, kDa) and the film was exposed for 1 minute. The arrowhead indicates the position of the full-length *LytE* protein (*LytE*, MW = 37 kDa).

Figure 3

3a: Typical cell morphology of strains that express *LytF* from the *cw/O* promoter.

Comparison of cell morphology of strains in which *lytF* is expressed from the *cw/O* promoter either with (LU111, LU116) or without the *cw/O* leader sequence (LU125, LU134). The presence or absence of the endogenous *lytE* gene is indicated. The endogenous *cw/O* and *lytF* genes are disrupted in all strains. Cells were grown in LB to an OD600 of 0.1, heat-fixed and examined with an Olympus BX61 microscope. Magnification is 1000x. A scale bar is shown.

3b: Transcript levels of *lytF*.

LytF transcript levels were determined by RT-qPCR, for the following strains: WT168, LS103 (Δ *cw/O*), LS089 (Δ *lytE*), LU74 (Δ *lytF*), LU111 (*P_{cw/O}leader* *lytF* Δ *lytF* Δ *cw/O*), LU125 (*P_{cw/O}lytF* Δ *lytF* Δ *cw/O*) and LU134 (*P_{cw/O}lytF* Δ *lytF* Δ *cw/O* Δ *lytE*). Cells were grown in HPDM and samples harvested at OD600=0.15. RNA was prepared from the samples and transcribed into cDNA, which was analysed by RT-qPCR. Transcript levels were normalised relative to the level of transcript for the WT168 sample, which was assigned a value of 1. Error bars are shown.

3c: Cell-associated protein levels of *LytF* in WT168 and when expressed from the *cw/O* promoter with and without *cw/O* leader sequence.

Cell-associated protein samples of the following strains are shown: LU147 (*lytF*-HA), LU153 (Δ *lytF* *P_{cw/O}leader* *lytF*-HA Δ *cw/O* Δ *lytE*) and LU157 (Δ *lytF* *P_{cw/O}lytF*-HA Δ *cw/O* Δ *lytE*). A negative control (WT168 without HA-tag) is also displayed. Cells were grown in HPDM and harvested at OD600=0.2 (t1) and OD600=1.0 (t2). Total cell protein was prepared from the samples and the HA-tagged proteins were detected by western analysis. Ten μ g total cell protein extract was deployed per lane. Molecular weight markers are shown (M, kDa). The film was exposed for one minute. The arrowhead indicates the degraded *LytF* protein (full length *LytF*, MW = 51 kDa).

3d: Localization of the *LytF* protein in a strain in which the protein is expressed from the *cw/O* promoter.

Strain LU151 (Δ *lytF* *P_{cw/O}leader* *lytF*-HA Δ *cw/O*) was grown in buffered LB to an OD600 of 0.2 at which point cells were harvested, fixed in 2.6 % paraformaldehyde, immunostained and examined with an Olympus BX61 microscope. Immunostaining was achieved by endowing the *LytF* protein with a HA-tag and treating cells first with a primary anti-HA antibody followed by binding of the primary antibody with a fluorescein isothiocyanate (FITC)-

Figure 4

4a: Typical cell morphology of strains that express a variant LytE protein with one LysM motif from the *lytE* promoter either with or without the endogenous *cw/O* gene.

Comparison of cell morphology of strains in which a variant LytE with one LysM motifs is expressed from the *lytE* promoter in combination with an intact *cw/O* gene (LU337 ($P_{lytE}lytE[L1cat] \Delta lytE$), LU339 ($P_{lytE}lytE[L2cat] \Delta lytE$), LU341 ($P_{lytE}lytE[L3cat] \Delta lytE$)), or in combination with a *cw/O* null mutant (LU350 ($P_{lytE}lytE[L1cat] \Delta lytE \Delta cw/O$), LU352 ($P_{lytE}lytE[L2cat] \Delta lytE \Delta cw/O$), LU354 ($P_{lytE}lytE[L3cat] \Delta lytE \Delta cw/O$)). A control with a reconstituted wild type *lytE* is shown (LU335 ($P_{lytE}lytE[L123cat]R \Delta lytE$), LU343 ($P_{lytE}lytE[L123cat]R \Delta lytE \Delta cw/O$)). Cells were grown in LB to an OD600 of 0.1, heat-fixed and examined with an Olympus BX61 microscope. Magnification is 1000x. A scale bar is shown.

4b: *sigI* and *mreBH* transcript levels for strains that express a variant LytE protein from the *cw/O* promoter.

Transcript levels were established by RT-qPCR for the following strains: WT168, EB35 ($\Delta cw/O$), LS089 ($\Delta lytE$), LU320 ($P_{cw/Oleader}lytE[L123cat]R \Delta cw/O \Delta lytE$), LU262 ($P_{cw/Oleader}lytE[L1cat] \Delta cw/O \Delta lytE$), LU266 ($P_{cw/Oleader}lytE[L2cat] \Delta cw/O \Delta lytE$) and LU270 ($P_{cw/Oleader}lytE[L3cat] \Delta cw/O \Delta lytE$). Cells were grown in HPDM and samples harvested at OD600=0.15. RNA was prepared from the samples and transcribed into cDNA, which was analysed by RT-qPCR with specific primer sets. Transcript levels were normalised relative to the level of transcript for the WT168 sample, which was assigned a value of 1. Error bars are shown.

4c: Growth curves of strains that express a variant LytE protein with one LysM motif from the *cw/O* and the *lytE* promoter at different salt concentrations.

Growth curves were obtained for the following strains: WT168, EB35 ($\Delta cw/O$), LS089 ($\Delta lytE$), LU21 ($P_{cw/Oleader}lytE \Delta cw/O \Delta lytE$), LU320 ($P_{cw/Oleader}lytE[L123cat]R \Delta cw/O \Delta lytE$), LU343 ($P_{lytE}lytE[L123cat]R \Delta cw/O \Delta lytE$), LU262 ($P_{cw/Oleader}lytE[L1cat] \Delta cw/O \Delta lytE$), LU266 ($P_{cw/Oleader}lytE[L2cat] \Delta cw/O \Delta lytE$), LU270 ($P_{cw/Oleader}lytE[L3cat] \Delta cw/O \Delta lytE$), LU350 ($P_{lytE}lytE[L1cat] \Delta cw/O \Delta lytE$), LU352 ($P_{lytE}lytE[L2cat] \Delta cw/O \Delta lytE$), LU354 ($P_{lytE}lytE[L3cat] \Delta cw/O \Delta lytE$). Growth curves were obtained in the following way: Overnight cultures of the strains were grown in salt-free buffered LB and used to inoculate 100 μ l media in a CELLSTAR® 96 well plate to an OD600 of 0.03. NaCl was added to the final concentrations of 0.1 M, 0.2 M, 0.3 M and 0.4 M. The 96 well plate was incubated at 37°C with slow shaking and the OD600 was measured every 10 minutes over a time period of 10 hours. Three technical repeats for each strain were conducted. One representative example is shown.

Figure 5

5a: Colony morphology of strains that express a variant LytE protein from either the *cw/O* or the *lytE* promoter.

LB agar plates (no antibiotic) were inoculated with a toothpick from an overnight LB culture and incubated overnight at 37°C followed by incubation for 14 days at room temperature. Growth of each colony was documented after 1 day, 4 days, 7 days and 14 days. The following strains were characterized: LU262 ($P_{cw/Oleader}lytE[L1cat] \Delta cw/O \Delta lytE$), LU266 ($P_{cw/Oleader}lytE[L2cat] \Delta cw/O \Delta lytE$), LU270 ($P_{cw/Oleader}lytE[L3cat] \Delta cw/O \Delta lytE$), LU205 ($P_{cw/Oleader}lytE[L11cat] \Delta cw/O \Delta lytE$), LU217 ($P_{cw/Oleader}lytE[L21cat] \Delta cw/O \Delta lytE$), LU229 ($P_{cw/Oleader}lytE[L31cat] \Delta cw/O \Delta lytE$), LU213 ($P_{cw/Oleader}lytE[L13cat] \Delta cw/O \Delta lytE$), LU277 ($P_{cw/Oleader}lytE[L331cat] \Delta cw/O \Delta lytE$), LU313 ($P_{cw/Oleader}lytE[L213cat] \Delta cw/O \Delta lytE$), LU320 ($P_{cw/Oleader}lytE[L123cat]R \Delta cw/O \Delta lytE$), LU350 ($P_{lytE}lytE[L1cat] \Delta lytE \Delta cw/O$), LU352 ($P_{lytE}lytE[L2cat] \Delta lytE \Delta cw/O$), LU354 ($P_{lytE}lytE[L3cat] \Delta lytE \Delta cw/O$), LU407 ($P_{lytE}lytE[L11cat] \Delta lytE \Delta cw/O$), LU409 ($P_{lytE}lytE[L21cat] \Delta lytE \Delta cw/O$), LU383 ($P_{lytE}lytE[L31cat] \Delta lytE \Delta cw/O$), LU411 ($P_{lytE}lytE[L13cat] \Delta lytE \Delta cw/O$), LU385 ($P_{lytE}lytE[L331cat] \Delta lytE \Delta cw/O$), LU413 ($P_{lytE}lytE[L213cat] \Delta lytE \Delta cw/O$), LU343 ($P_{lytE}lytE[L123cat]R \Delta lytE \Delta cw/O$), WT168, LS089 ($\Delta lytE$), EB35 ($\Delta cw/O$).

5b: The levels of LytE variant proteins expressed from the *cw/O* promoter, for cell-associated proteins and secreted proteins.

Protein samples of the following strains are shown: LU61 ($\Delta cw/O P_{cw/Oleader}lytE-HA \Delta lytE$), LU421 ($P_{cw/Oleader}lytE-HA [L123cat]R \Delta cw/O \Delta lytE$), LU432 ($P_{cw/Oleader}lytE-HA [L213cat] \Delta cw/O \Delta lytE$), LU418 ($P_{cw/Oleader}lytE-HA [L1cat] \Delta cw/O \Delta lytE$), LU419 ($P_{cw/Oleader}lytE-HA [L2cat] \Delta cw/O \Delta lytE$), LU440 ($P_{cw/Oleader}lytE-HA [L3cat] \Delta cw/O \Delta lytE$), LU414 ($P_{cw/Oleader}lytE-HA [L11cat] \Delta cw/O \Delta lytE$), LU415 ($P_{cw/Oleader}lytE-HA [L13cat] \Delta cw/O \Delta lytE$) and LU417 ($P_{cw/Oleader}lytE-HA [L31cat] \Delta cw/O \Delta lytE$). A negative control nc (WT168 without HA-tag) is also displayed. Cells were grown in salt-free buffered LB and harvested at an OD600 of 0.25. Total cell protein was prepared from the samples and the HA-tagged proteins were detected by western analysis. Five μ g total cell protein extract was deployed per lane (left) and 0.15 OD equivalent media proteins were deployed per lane (right). Molecular weight markers are shown (M, kDa). Both films were exposed for seven minutes. The black arrowhead indicates the full length LytE protein (LytE, MW = 37 kDa), the grey arrowhead indicates a full length dimer protein (variant LytE with two LysM motifs, MW = 30 kDa) and the open arrowhead indicates a full length monomer protein (variant LytE with one LysM motif, MW = 23 kDa).

5c: *In vitro* binding affinity of purified variant LytE proteins to purified *B. subtilis* cell walls at different salt concentrations.

The diagram shows the results of peptidoglycan pull down assays, conducted for purified variant LytE proteins by proportion of protein found in the pellet. The following proteins were analysed: LytE[L123cat], LytE[cat], LytE[L13cat], LytE[L31cat], LytE[L1cat], LytE[L3cat]. Forty-four pmole of variant LytE protein were mixed with purified cell wall of *B. subtilis* to a final absorbance of OD600 = 10, in a 25 μ l reaction volume. A negative

Figure 6

6a: A putative model of the evolutionary development of the *cwlS*, *lytE* and *lytF* D,L-endopeptidases and their LysM motifs.

The proposed model assumes a stable inheritance of genes and LysM motifs. The common ancestor is a gene that encodes for a D,L-endopeptidase type protein with one LysM motif. Through a combination of gene duplication (blue star), LysM motif duplication (red star) as well as LysM motif loss events (orange star), the model arrives at the set of genes observed in this study, with *LytE* orthologs (having either two or three LysM motifs), *CwlS* orthologs (four or three LysM motifs) and *LytF* orthologs (with four or five LysM motifs).

6b: Proposed model for the evolutionary development of groups with different sets of paralogs.

The model assumes a stable inheritance of genes and LysM motifs and describes how through a combination of LysM motif (red star) and gene (blue star) duplication events as well as loss of LysM motifs (orange star) different bacterial species gain different sets of paralogs.

Supplementary Figure 1

Supplementary 1: Basic phylogeny of the *Bacillus* species.

Phylogeny of the investigated 19 *Bacillus* species was inferred through alignment-free genome comparison with feature frequency profiles (FFP). Genome sequences (excluding plasmid sequences) of all species were retrieved from GenBank. FFP was performed with length $l = 24$. A neighbour Joining (NJ) tree was constructed based on the Jensen-Shannon divergence matrix. Bootstrapping was performed to verify the robustness of the tree and bootstrapping values of 100 out of 100 were achieved for all branches. Letters (A, B, C, X) refer to the different groups of paralogs defined in Figure 2b.

Supplementary Figure 2:

Supplementary 2a: schematic structure of the wild type *lytE*

Schematic structure of the WT168 *lytE*, which features a signal sequence followed by three LysM motifs (L1, L2 and L3), separated by linker sequences and a catalytic domain (cat) which is separated from L3 by a spacer sequence.

Supplementary 2b: schematic structure of a variant *lytE* fragment with no LysM motifs

A linear fragment encoding a variant *lytE* without a LysM motif was generated by amplifying the catalytic domain with (forward) primers flanked by the signal sequence.

Supplementary 2c: construction of variant *lytE* fragments with one LysM motifs

Linear fragments encoding variant *lytE* with one LysM motif (III), are constructed by amplifying individual LysM motifs (I), with primers flanked by the signal and the spacer sequence, and joining them to the catalytic domain (II) by long-flanking homology PCR, thereby utilizing the spacer sequence for strand overlap.

Supplementary 2d: construction of variant *lytE* fragments with two LysM motifs

PCR fragments encoding variant *lytE* with two LysM motifs (IV), were constructed by amplifying individual LysM motifs with two set of primers (I and II) (one primer set is flanked by the signal and the linker sequence and the other by the linker and spacer sequence) and joining them alternating to the catalytic domain (III) by long-flanking homology PCR, thereby utilizing the linker and spacer sequence for strand overlap to generate all nine possible permutations.

Supplementary 2e: integration of a variant *lytE* fragment on either the *cw/O* or the *lytE* locus

Schematic representation of strains LU213 and LU411, in which a variant *LytE* open reading frame with 2 LysM motifs in the combination L1-L3 is integrated on the chromosome. In LU213 the variant *LytE* is expressed from the *cw/O* promoter with the *cw/O* leader sequence (replacing the endogenous *cw/O* gene) and in LU411 the variant *lytE* is expressed from from the *lytE* promoter (replacing the endogenous *lytE* gene). The endogenous *lytE* gene in LU213 is disrupted with a chloramphenicol resistance cassette, while the endogenous *cw/O* gene in LU411 is disrupted with kanamycin resistance cassette. Open reading frames for *lytE* and *cw/O* are not contiguous.

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
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> 168	224308	P54421	O07532	O31852
<i>Bacillus subtilis</i> subsp. <i>inaquosorum</i> KCTC 13429	1236548	L8PYS4	L8PUD5	L8PTT2
<i>Bacillus subtilis</i> subsp. <i>natto</i> BEST195	645657	D4G6X8	D4G6X3	D4FXM8
<i>Bacillus subtilis</i> subsp. <i>spizizenii</i> W23	655816	E0TX21	E0TX16	E0TZ07
<i>Bacillus subtilis</i> subsp. <i>spizizenii</i> TU-B-10	1052585	G4NUN9	G4NUN5	G4NQ36
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> SC-8	1089443	G4EYF8	G4EYG3	G4EUZ6
<i>Bacillus subtilis</i> XF-1	1233100	M4KT97	M4KPW4	M4KSM2
<i>Bacillus subtilis</i> subsp. <i>niger</i> PCI 246	1495315	A0A080UQ42	A0A080UT85	A0A080UNC1
<i>Bacillus amyloliquefaciens</i> DSM 7	692420	E1UQ68	E1UQ63	E1UVA8
<i>Bacillus amyloliquefaciens</i> Y2	1126211	I2C2Z6	I2C2Z0	I2C6G3
<i>Bacillus velezensis</i> FZB42 / <i>Bacillus amyloliquefaciens</i> subsp. <i>Plantarum</i> FZB42	326423	A7Z2V6	A7Z2V1	A7Z5K3
<i>Bacillus velezensis</i> UCMB5033 / <i>Bacillus amyloliquefaciens</i> subsp. <i>Plantarum</i> UCMB5033	1338518	S6FIE6	S6FIE3	S6FL69
<i>Bacillus atrophaeus</i> 1942	720555	A0A0H3DZF9	A0A0H3DZF6	A0A0H3E2R9
<i>Bacillus licheniformis</i> DSM 13	279010	Q65LZ8	Q65IH0	none
<i>Bacillus licheniformis</i> CG-B52	1368424	T5HVV1	T5HPU6	none
<i>Bacillus licheniformis</i> S 16	1348578	W7RE25	W7RRW1	none
<i>Bacillus licheniformis</i> WX-02	1126218	A0A0H4X825	A0A0H4XYT8	none
<i>Bacillus paralicheniformis</i> ATCC 9945a	766760	R9TSA0	R9TUX7	none
<i>Bacillus sonorensis</i> L12	1274524	M5PGX3	M5PFH3	none

Table S2: List of strains used in this study

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

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

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

[LFH] long flanking homology fragment; [cm^r] chloramphenicol resistance gene; [erm^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 <i>kan^r</i> cassette	Guerout-Fleury et al., 1995
pDG646	plasmid with <i>erm^r</i> cassette	Guerout-Fleury et al., 1995
pDG1727	plasmid with <i>spc^r</i> cassette	Guerout-Fleury et al., 1995
pDG1661	plasmid with <i>cm^r</i> 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 <i>lytE</i> over single cross-over, <i>kan^r</i> cassette	This work
pLU6	pDG780 derivative enabling C-terminal HA-tagging of <i>lytF</i> over single cross-over, <i>kan^r</i> cassette	This work
pLU9	pDG780 derivative with upstream homology to the <i>cwI</i> O locus (including <i>cwI</i> O promoter and <i>cwI</i> O leader sequence)	This work
pLU10	pLU9 derivative with up- and downstream homology to the <i>cwI</i> O locus (including <i>cwI</i> O promoter and <i>cwI</i> O leader sequence)	This work
pLU14	pLU10 derivative enabling integration of variant <i>lytE</i> [L13cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU15	pLU10 derivative enabling integration of variant <i>lytE</i> [L12cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU16	pLU10 derivative enabling integration of variant <i>lytE</i> [L11cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU17	pLU10 derivative enabling integration of variant <i>lytE</i> [L21cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU18	pLU10 derivative enabling integration of variant <i>lytE</i> [L22cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU19	pLU10 derivative enabling integration of variant <i>lytE</i> [L23cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU20	pLU10 derivative enabling integration of variant <i>lytE</i> [L31cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU21	pLU10 derivative enabling integration of variant <i>lytE</i> [L32cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU22	pLU10 derivative enabling integration of variant <i>lytE</i> [L33cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU23	pLU10 derivative enabling integration of variant <i>lytE</i> [L1cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU24	pLU10 derivative enabling integration of variant <i>lytE</i> [L2cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU25	pLU10 derivative enabling integration of variant <i>lytE</i> [L3cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU26	pLU20 derivative enabling integration of variant <i>lytE</i> [L311cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU27	pLU20 derivative enabling integration of variant <i>lytE</i> [L331cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU28	pLU19 derivative enabling integration of variant <i>lytE</i> [L233cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU29	pLU19 derivative enabling integration of variant <i>lytE</i> [L213cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU30	pLU18 derivative enabling integration of variant <i>lytE</i> [L222cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU31	pLU18 derivative enabling integration of variant <i>lytE</i> [L212cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU37	pLU10 derivative enabling integration of variant <i>lytE</i> [cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU38	pLU14 derivative enabling integration of variant <i>lytE</i> [L113cat] at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU40	pLU14 derivative enabling integration of variant <i>lytE</i> [L123cat] ^R at the <i>cwI</i> O locus by double cross-over, <i>kan^r</i> cassette	This work
pLU46	pLU2 derivative enabling C-terminal HA-tagging of <i>lytE</i> over single cross-over, <i>spc^r</i> 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

16S_Bsu_F	GTCAGCTCGTGTCTGAGATG
16S_Bsu_R	CAACTGAATGCTGGCAACTAAGATC
Aaa_LysM_Fwd	cgcgatccacgttatttgcaggagcggcatctgcacaaGTAAAACCGGTGAATCAG
Aaa_LysM_Rev	atgtagatgacgttttgccttgaagacgatgaagatgaagcaggttttgagctgcttgctgtgtcc TACTTTTAGTACTTGGTTTG
Catalytic_LytE_Fwd_Spac	ttcaagcagcaaaacgtcatctacatcactaatgtgagcaagctggtttctgatgcaaaagcgtagtc GGAACGCCATATAAATGGGG
Catalytic_LytE_Rev_Ery	attatgtcttttgcgcagtcggcGAAAAAACCCTTCCAATGAACGG
Catalytic_LytE_Rev_Xmal	tcccccgggGAAAAAACCCTTCCAATGAACGG
CwIO_DF_LysM	ccgctcgagATTAACATATACCCTTTTTGAATA
CwIO_DR_LysM	cggggtaccCCTTCTTCATTCGTAATGCCCAGGT
CwIO_UF	CCTTCATCTGTGGAGCAG
CwIO_UF_LysM	aaggaaaaagcggccgcATCATTATATAAAACCGGCTT
CwIO_UF_Smal	gcatcccggtCCTTCATCTGTGGAGCAG
CwIO_UR	cgtagctgaatgatttgccttttcatTTTTATATCCTCCCTTTTACTGTGAACC
CwIO_UR_CwIS	aagccggtacatacttcttttcatTTTTATATCCTCCCTTTTACTGTGAA
CwIO_UR_LysM	taacgtggatcctaaacaactgctgctgtagctgaatgatttgccttttcat TTTTATATCCTCCCTTTTAC
CwIO_UR_LytF	gccctgctgctaatttcttttcatTTTTATATCCTCCCTTTTACTGTGAA
CwIO_UR_Yqgt	gtttcacgtaataaacgccatcatTTTTATATCCTCCCTTTTACTGTGAA
CwIS_Do_Fwd	tatattatgtaaactaatgcaagAAAATGAGTATGACAAAAAGCC
CwIS_Do_Rev	TCTGGGGAAATAGGCTGCTCT
CwIS_Fwd	ATGAAAAGAAGATTGTAGCCGGC
CwIS_LysM_Fwd	cgcgatccacgttatttgcaggagcggcatctgcacaaATAAAAGTAAAAAGCGGTGA
CwIS_LysM_Rev	atgtagatgacgttttgccttgaagacgatgaagatgaagcaggttttgagctgcttgctgtgtcc AGGAACCTTAAGACTTTGTC
CwIS_Rev_Ery	tcgattatgtcttttgcgcagtcggcTAAACAAAAAAGCCTATAAAAGGC
CwIS_Tag_F_BamHI	cgggatccGTCGGAGTGCCTTACCGATGGGGC
CwIS_Tag_R_HA_EcoRI	ggaattctaacaacaaagcctataaaaggcttttgcatactattttattatgcataatccg gaacatcatacggataAAAATAACTTCTTGCGCCCAAGTA
CwIS_Up_Fwd	CGGCAAGCATATTAACAAGTGAG
CwIS_Up_Rev	tatattctaggaagtaggtattcacTTCATTCAACCTCCTAAACTG
Exp_Cat_Rev	ccgctcgagGAATCTTTTCGCACCGAGGTAACG
Exp_L1_Fwd	ggaattccatagAGCATTAAGGTGAAAAAAGGCGACA
Exp_L2_Fwd	ggaattccatagACATACAAAGTAAAGAGCGGGACA
Exp_L3_Fwd	ggaattccatagACTTATAAAGTGAAGAGCGGAGACA
Exp_LytF_Fwd	ggaattccatagATTAAGGTCAAAAGCGGAGATTCTTT
Exp_Spac_Fwd	ggaattccatagGGAACAAGCACAAGCAGCTCAAAAC
Exp_XlyB_Fwd	ggaattccatagTATCATGTAAAAAAGGTGACACACTG
L1_Fwd_Link	ggcaaatctactagttcaaaagcagcagttcttctcttcttct AGCATTAAGGTGAAAAAAGGC
L1_Fwd_Sig_BamHI	cgcgatccacgttatttgcaggagcggcatctgcacaa AGCATTAAGGTGAAAAAAGGC
L1_Rev_Link	agaagaagaggaagaagaactgctgcttttgaactagtagatttgcc GTTAATCGATAAAGTTTGTCCC
L1_Rev_Spac	atgtagatgacgttttgccttgaagacgatgaagatgaagcaggttttgagctgcttgctgtgtcc GTTAATCGATAAAGTTTGTCT
L2_Fwd_Link	ggcaaatctactagttcaaaagcagcagttcttctcttcttct ACATACAAAGTAAAGAGCGGG
L2_Fwd_Sig_BamHI	cgcgatccacgttatttgcaggagcggcatctgcacaa ACATACAAAGTAAAGAGCGGG
L2_Rev_Link	agaagaagaggaagaagaactgctgcttttgaactagtagatttgcc TTTCAGTTTCAGGACTTGTC
L2_Rev_Spac	atgtagatgacgttttgccttgaagacgatgaagatgaagcaggttttgagctgcttgctgtgtcc TTTCAGTTTCAGGACTTGTC
L3_Fwd_Link	ggcaaatctactagttcaaaagcagcagttcttctcttcttct ACTTATAAAGTGAAGAGCGGA
L3_Fwd_Sig_BamHI	cgcgatccacgttatttgcaggagcggcatctgcacaa ACTTATAAAGTGAAGAGCGGA
L3_Rev_Link	agaagaagaggaagaagaactgctgcttttgaactagtagatttgcc TTTCACCTTCAATACITGGTTG

L3_Rev_Spac	atgtatgacgttttgcgttgaagacgatgaagatgaagcaggttttgagctgcttgcctgttcc TTTCACCTTCAATACTTGGT
LytE_DF	ATGAAAAAGCAAATCATTACAGCTACG
LytE_Do_Fwd_Ery	gggtgccagagttaaaggatcATTAGACGGAGCACAGGCGTTTT
LytE_Do_Fwd_Spec	caataaaccttgcctcgtctacTACTGGAAGCCTCGTTACCTCGG
LytE_Do_Fwd_Spec	caataaaccttgcctcgtctacATTAGACGGAGCACAGGCGTTTT
LytE_Do_Rev	GTATTTGCCGTATGCGCTCAGGC
LytE_DR_Ery	attatgtcttttgcgcagtcggcGAAAAAACCCTTCCAATGAA
LytE_Probe_F	TATCGATTAAACGGCAAATCTACAAGTT
LytE_Probe_T7_R	taatacgactcactatagggCAGTGCTTGTCTGCCACACTT
LytE_qPCR_F2_	GGAGGAAAATATGAAAAAGCAAATCA
LytE_qPCR_R2	TCGCTCTCATACGTATCAACACTCT
LytE_Spacer_Fwd_Sig	cgcggatccagcttatttgcaggagcggcatctgcacaaGGAACAAGCACAGCAGC
LytE_Tag_F_BamHI	cgggatccTTAGTCGGAACGCCATATAATGG
LytE_Tag_R_HA_EcoRI	ggaattcgaaaaaacccgttccaatgaacgggttttctctaaaaattattatgcata atccggaacatcatcagcgataGAATCTTTTCGCACCGAGGTAACG
LytE_Up_Fwd	AACGGTAGGCTCTTATGTAAGGC
LytE_Up_Fwd_2	AACGGTAGGCTCTTATGTAAGGC
LytE_Up_Rev_Spec	cggtacgttattagcgagccagtcCCTAAAACAACTGCTGTCTGTAGC
LytF_Do_Fwd	tatatattgtaaactaatgcaagTATTAGAGCCTAAAGACGTTGATGC
LytF_Do_Rev	CCGGGATTGTGTTATGTGAGAA
LytF_Fwd	ATGAAAAAGAAATTAGCAGCAGGG
LytF_LysM_Fwd	cgcggatccagcttatttgcaggagcggcatctgcacaaATTAAGGTCAAAGCGGAG
LytF_LysM_Rev	atgtatgacgttttgcgttgaagacgatgaagatgaagcaggttttgagctgcttgcctgttcc TGGAATAGTAAGCGTCTGG
LytF_qPCR_Fwd	TGTCGGAGATTTCGTTTTCTTTTC
LytF_qPCR_fwd	TGTCGGAGATTTCGTTTTCTTTTC
LytF_qPCR_Rev	CCCGTTGCCTAGGTAAATTCC
LytF_qPCR_rev	CCCGTTGCCTAGGTAAATTCC
LytF_Rev_Ery	tcgattatgtcttttgcgcagtcggcGCATCAACGTCTTTAGGCTCTAATA
LytF_Tag_F_BamHI	cgggatccCTTGGGGTTCCGTATCGCTGGGGA
LytF_Tag_R_HA_EcoRI	ggaattcgcatcaacgtcttttaggtcttaataaaaaactgtacagccggttaaggccgcacagtttctg ttttattatgcataatccggaacatcatcagcgataGAAATATCGTTTTCACCGAGATA
LytF_Up_Fwd	CTTTTGGTTTATCCTTCGGCCTT
LytF_Up_Rev	tatatctaggaagtaggtattcacCTTAGCTCTCCTTTTCCCTCC
MreBH_qPCR_F	GCAAAACCCGCTTGAATCTG
MreBH_qPCR_R	TGCCTTTTGCAGCTTATCAATC
OE93_Fwd_Ery	gggtgccagagttaaaggatcGAGCAATCCTTACTGGAAAGCAGC
OE94_Rev	ATCGACAAGCGTTACGTTACCGAC
SigI_Chip_Fwd	AAATTCTCTGCTCGCATTTGC
SigI_Chip_Rev	GGAAAGCTCGGCTTCAATCAG
Signal_LytE_Fwd	ATGAAAAAGCAAATCATTACAGCTACG
Signal_LytE_Rev	TTGTGCAGATGCCGCTCCTGCAA
Signal_LytE_Rev_Spec	cggtacgttattagcgagccagtcTTGTGCAGATGCCGCTCCTGCAA
Spacer_Check_Rev	GATGTAGATGACGTTTTGC
Spec_Fwd_EcorI	ccggaattccggCTAATAACGTAAACGTGACTGGCA
Spec_Rev_Sall	ttccgcgccgctatggccgacgtcgacgcgtAGGGTTTATTGTTTTCTAAAATCTG
YocH_Northern_F	GCAGCAAACACAGCTACTGAGAAT
YocH_Northern_R	GCTCTTTTTGCACAGCCTCTT
YqgT_Do_Fwd	tatatattgtaaactaatgcaagACTGAACTTTTTCTTTTTCTGCTG
YqgT_Do_Rev	TTGGAGCGGCCATTTTCGCTGT
YqgT_Fwd	ATGATGGCGTTTATTACGGTG

YqgT_Rev_Ery	tcgattatgtcttttgcgcagtcggcTCTCATATTTGCTTCGGCAGAA
YqgT_Tag_F_BamHI	cgggatccACGGAGAATCCGCCTGACAGACTG
YqgT_Tag_R_HA_EcoRI	ggaattctctcatatttgccttcggcagaaaaagaaaaagtttcagttattatgcataatccgg aacatcatcacggataCTCATGAAAAAATAACGCCCTCCA
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