Supplemental data for:

Unleashing natural competence in *Lactococcus lactis* by induction of the competence regulator ComX.

Joyce Mulder, Michiel Wels, Oscar P. Kuipers, Michiel Kleerebezem, Peter A. Bron

Table S1: Comparison of competence protein identities and length of *L. lactis* KF147 to the competent strain *S. thermophilus* LMD-9.

	comC	*comEA*	*comEC*	*comFA*	*comFC*	*comGA*	*comGB*	*comGC*	*comGD*	*comGE*	*comGF*	*comGG*	*comX*	*coiA*	*ssbA*	*ssbB*	*recA*	*DprA*
Protein identity (%)	40	45	44	47	44	51	45	57	36	33	47	24	28	37	60	73	78	59
Positve AA score (%)	58	64	60	63	62	69	67	77	52	64	65	48	51	61	77	81	87	78
Alignment length (%)	96	67	97	87	100	99	97	97	96	98	97	55	90	89	98	100	98	99
Total length protein KF147 (AA)	221	215	736	440	216	312	291	127	143	98	141	94	163	329	129	166	387	282
Total length protein LMD-9(AA)	215	231	746	439	220	313	296	108	120	93	145	105	165	347	130	172	379	279

Fig. S1: Genetic organization of the canonical competence genes in *L. lactis* KF147 (A). The *cin*-box sequences indicated in Wydau *et al.,* 2006 were used to extract *cin*-box sequences in *L. lactis* KF147 (20). These sequences are depicted underneath the black arrow which reflects the promotor region. Numbers surrounding the gene clusters correspond with start and stop codons of the first and last gene of the competence gene/operon. Colors of the gene-representing arrows correspond with the protein-icons in the graphical representation of the core competence machinery (analogous to Johnston *et al.,* 2014, Muschiol *et al.,* 2015 and Mann *et al.,* 2013 (1-3)). Black gene-representing arrows indicate genes present within the *com*-associated operons that are predicted not to have a direct structural role in the core competence machinery.

Upon competence development, the pilus comprising ComGA, ComGB and ComGC entraps the exogenous dsDNA which subsequently connects with ComEA. One strand of the dsDNA is degraded by an unknown domain or protein in *L. lactis*, and the other strand translocates into the cell by transfer through the pore protein ComEC and is further processed and internalized by ComFA and ComFC. In the cytosol, the ssDNA is protected by single stranded binding proteins (SsbA and SsbB) and DprA, which interacts with RecA in order to facilitate DNA integration into the genome. CoiA is thought to fulfill a role in DNA processing and ComC and the minor pilins ComGD-GG are proposed to fulfill a role in priming of pilus assembly (8-10).

Figure S1:

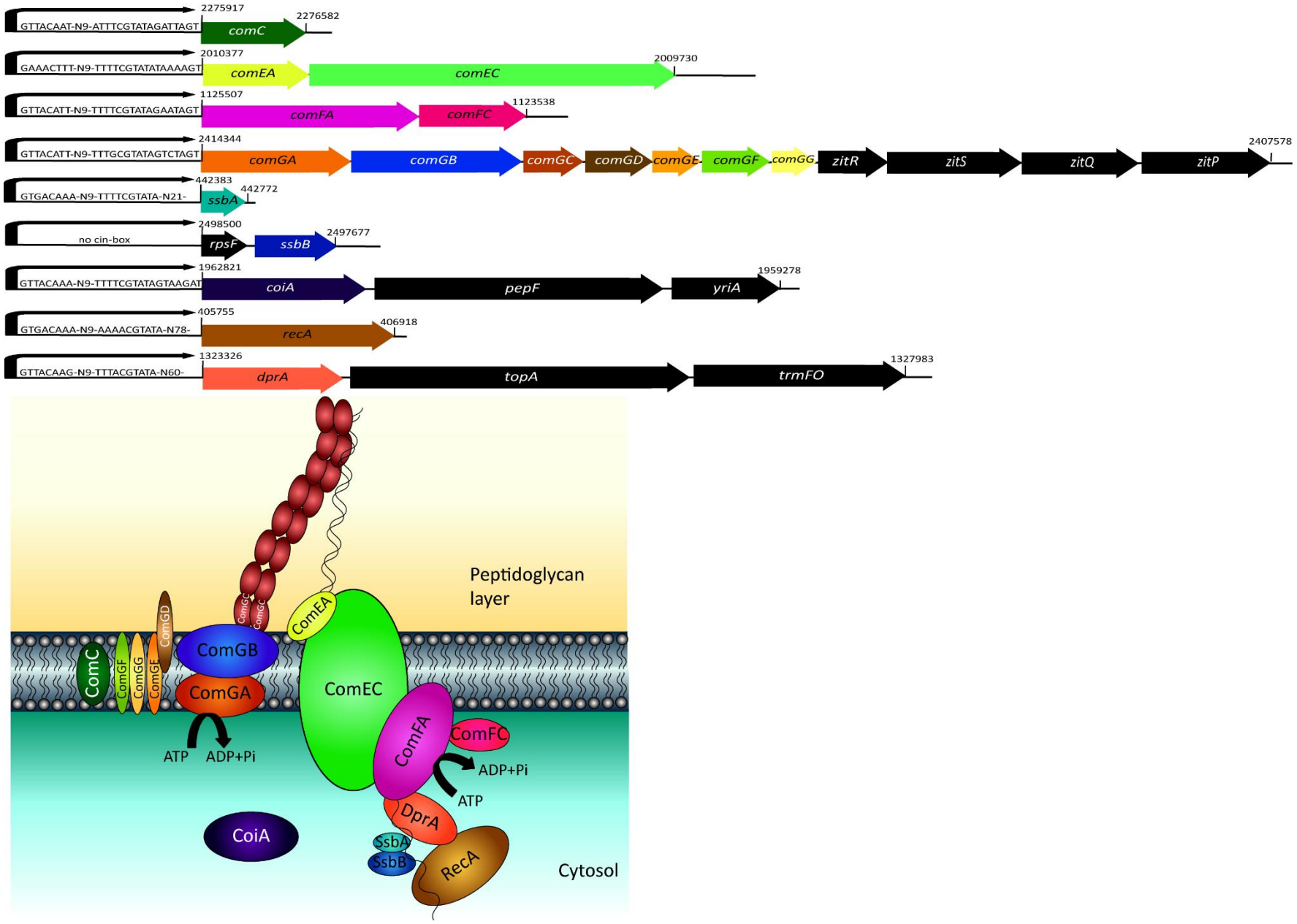

2275917
GTTACAAT-N9-ATTTCGTATAGATTAGT
comC
2276582
2010377
GAAACTTT-N9-TTTTCGTATATAAAAGT
comEA
comEC
2009730
1125507
GTTACATT-N9-TTTTCGTATAGAATAGT
comFA
comFC
1123538
2414344
GTTACATT-N9-TTTGCGTATAGTCTAGT
comGA
comGB
comGC
comGD
comGE
comGF
comGG
zitR
zitS
zitQ
zitP
2407578
442383
GTGACAAA-N9-TTTTCGTATA-N21-
ssbA
442772
2498500
no cin-box
rpsF
ssbB
2497677
1962821
GTTACAAA-N9-TTTTCGTATAGTAAGAT
coiA
pepF
yriA
1959278
405755
GTGACAAA-N9-AAAACGTATA-N78-
recA
406918
1323326
GTTACAAG-N9-TTTACGTATA-N60-
dprA
topA
trmFO
1327983
Peptidoglycan layer
ComGC
ComGC
ComEA
ComGD
ComC
ComGF
ComGG
ComGE
ComGB
ComGA
ComEC
ComFA
ComFC
ATP
ADP+Pi
ADP+Pi
ATP
DprA
CoiA
SsbA
SsbB
RecA
Cytosol

Fig. S2: Genomic analysis of 43 *L. lactis* strains to assess genetic capacity to develop natural competence. Analogous to Figure 1 of the main paper, except that full-length protein identity-scores (%) for the selected subset of late competence associated proteins are displayed in comparison to their homologues in strain *L. lactis* KW2 that was used as a reference. All other analyses, all labels used and references for strains are analogous to those presented in Figure 1. Genetic events leading to competence gene decay (black cells in the figure) are specified as premature stopcodon within the first 90% of the gene (a), transposon insertion (b), prophage insertion (c) absence of gene, mutated/alternative start or lengthened/fused protein; at least more than 25% of its total length (d) , followed by the position within the protein sequence where the event is detected relative to its N-terminus. The right-hand column represents the source of isolation of the strains in which P= plant, D= dairy, S= soil, W= water, H= human body and F= fruit (4).

Figure S2:

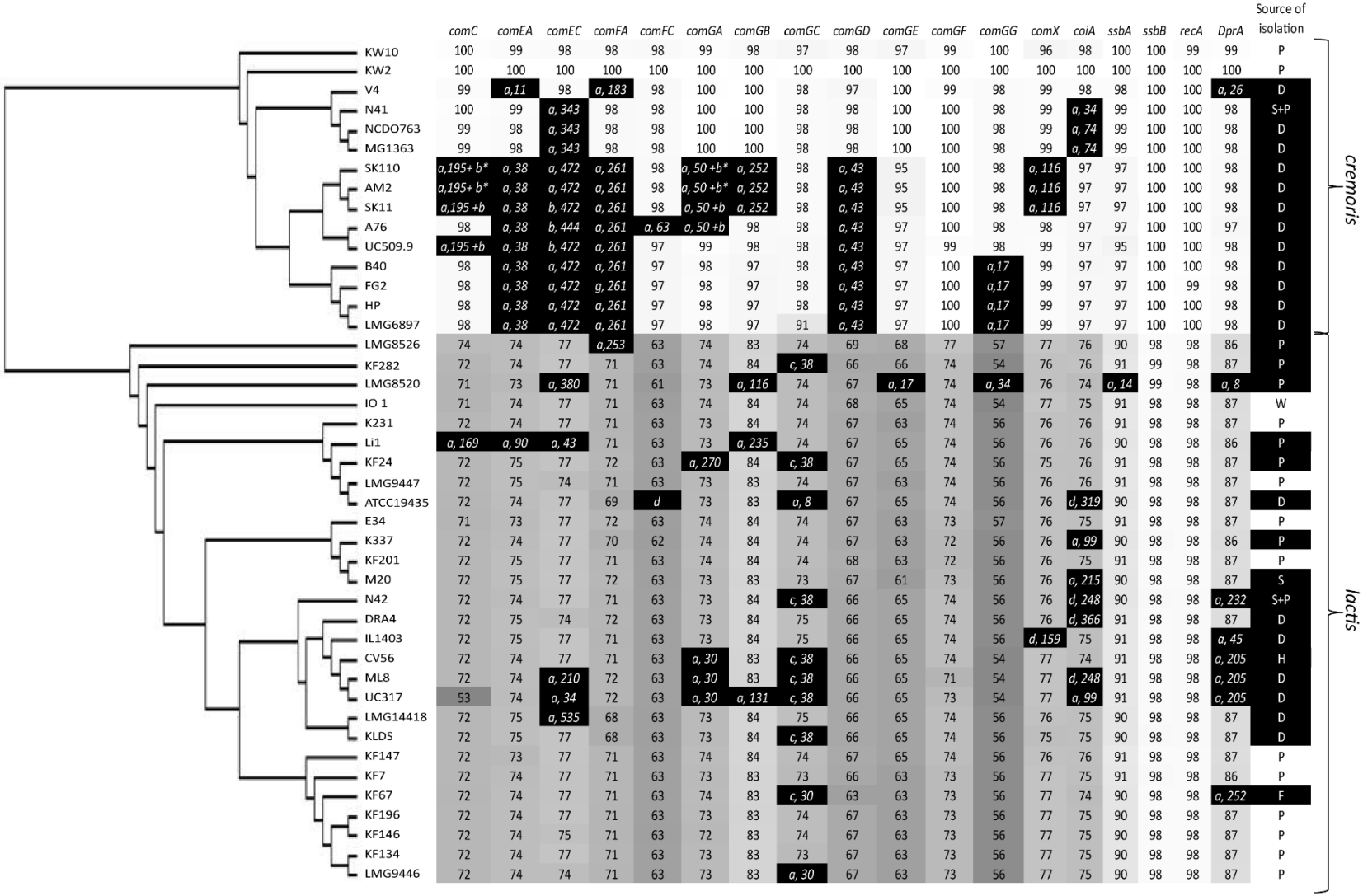

comC comEA comEC comFA comFC comGA comGB comGC comGD comGE comGF comGG comX coiA ssbA ssbB recA DprA Source of isolation
KW10 100 99 98 98 98 99 98 97 98 97 99 100 96 98 100 100 99 99 P
KW2 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 P
V4 99 a,11 98 a, 183 98 100 100 98 97 100 99 98 99 98 98 100 100 a, 26 D
N41 100 99 a, 343 98 98 100 100 98 98 100 100 98 99 a, 34 99 100 100 98 S+P
NCDO763 99 98 a, 343 98 98 100 100 98 98 100 100 98 99 a, 74 99 100 100 98 D
MG1363 99 98 a, 343 98 98 100 100 98 98 100 100 98 99 a, 74 99 100 100 98 D
SK110 a,195+ b* a, 38 a, 472 a, 261 98 a, 50 +b* a, 252 98 a, 43 95 100 98 a, 116 97 97 100 100 98 D
AM2 a,195+ b* a, 38 a, 472 a, 261 98 a, 50 +b* a, 252 98 a, 43 95 100 98 a, 116 97 97 100 100 98 D
SK11 a,195 +b a, 38 b, 472 a, 261 98 a, 50 +b a, 252 98 a, 43 95 100 98 a, 116 97 97 100 100 98 D
A76 98 a, 38 b, 444 a, 261 a, 63 a, 50 +b 98 98 a, 43 97 100 98 98 97 97 100 100 97 D
UC509.9 a,195 +b a, 38 b, 472 a, 261 97 99 98 98 a, 43 97 99 98 99 97 95 100 100 98 D
B40 98 a, 38 a, 472 a, 261 97 98 97 98 a, 43 97 100 a,17 99 97 97 100 100 98 D
FG2 98 a, 38 a, 472 g, 261 97 98 97 98 a, 43 97 100 a,17 99 97 97 100 99 98 D
HP 98 a, 38 a, 472 a, 261 97 98 97 98 a, 43 97 100 a,17 99 97 97 100 100 98 D
LMG6897 98 a, 38 a, 472 a, 261 97 98 97 91 a, 43 97 100 a,17 99 97 97 100 100 98 D
LMG8526 74 74 77 a,253 63 74 83 74 69 68 77 57 77 76 90 98 98 86 P
KF282 72 74 77 71 63 74 84 c, 38 66 66 74 54 76 76 91 99 98 87 P
LMG8520 71 73 a, 380 71 61 73 a, 116 74 67 a, 17 74 a, 34 76 74 a, 14 99 98 a, 8 P
IO 1 71 74 77 71 63 74 84 74 68 65 74 54 77 75 91 98 98 87 W
K231 72 74 77 71 63 73 84 74 67 63 74 56 76 76 91 98 98 87 P
Li1 a, 169 a, 90 a, 43 71 63 73 a, 235 74 67 65 74 56 76 76 90 98 98 86 P
KF24 72 75 77 72 63 a, 270 84 c, 38 67 65 74 56 75 76 91 98 98 87 P
LMG9447 72 75 74 71 63 73 83 74 67 63 74 56 76 76 91 98 98 87 P
ATCC19435 72 74 77 69 d 73 83 a, 8 67 65 74 56 76 d, 319 90 98 98 87 D
E34 71 73 77 72 63 74 84 74 67 63 73 57 76 75 91 98 98 87 P
K337 72 74 77 70 62 74 84 74 67 63 72 56 76 a, 99 90 98 98 86 P
KF201 72 75 77 71 63 74 84 74 68 63 72 56 76 75 91 98 98 87 P
M20 72 75 77 72 63 73 83 73 67 61 73 56 76 a, 215 90 98 98 87 S
N42 72 74 77 71 63 73 84 c, 38 66 65 74 56 76 d, 248 90 98 98 a, 232 S+P
DRA4 72 75 74 72 63 73 84 75 66 65 74 56 76 d, 366 91 98 98 87 D
IL1403 72 75 77 71 63 73 84 75 66 65 74 56 d, 159 75 91 98 98 a, 45 D
CV56 72 74 77 71 63 a, 30 83 c, 38 66 65 74 54 77 74 91 98 98 a, 205 H
ML8 72 74 a, 210 72 63 a, 30 83 c, 38 66 65 71 54 77 d, 248 91 98 98 a, 205 D
UC317 53 74 a, 34 72 63 a, 30 a, 131 c, 38 66 65 73 54 77 a, 99 91 98 98 a, 205 D
LMG14418 72 75 a, 535 68 63 73 84 75 66 65 74 56 76 75 90 98 98 87 D
KLDS 72 75 77 68 63 73 84 c, 38 66 65 74 56 75 75 90 98 98 87 D
KF147 72 73 77 71 63 74 84 74 67 65 74 56 76 76 91 98 98 87 P
KF7 72 74 77 71 63 73 83 73 66 63 73 56 77 75 91 98 98 86 P
KF67 72 74 77 71 63 74 83 c, 30 63 63 73 56 77 74 90 98 98 a, 252 F
KF196 72 74 77 71 63 73 83 74 67 63 73 56 77 75 90 98 98 87 P
KF146 72 74 75 71 63 72 83 74 67 63 73 56 77 75 90 98 98 87 P
KF134 72 74 77 71 63 73 83 73 67 63 73 56 77 75 90 98 98 87 P
LMG9446 72 74 74 71 63 73 83 a, 30 67 63 73 56 77 75 90 98 98 87 P
cremoris
lactis

Fig S3: Premature stop-codons and transposon interruption within *comEC* among *L. lactis* subsp. *cremoris* strains SK11, A76 and UC509.9.

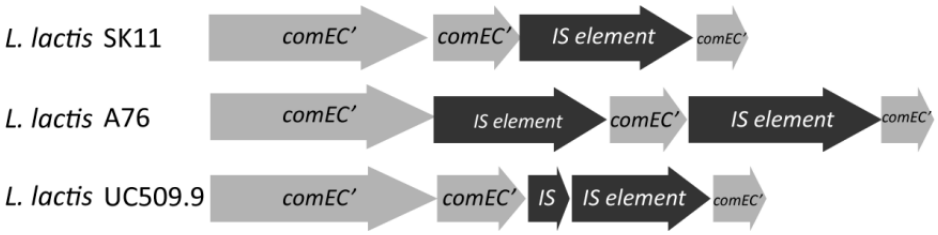


Fig. S4: Mutational events occurring in *L. lactis* subsp. *cremoris* strains that are focused on the late competence gene-sets used in this study. A mutational sequence of events was tentatively reconstructed for the late competence genes in different subspecies *cremoris* strains where the dot resembles the *cremoris* strain ancestor. The core-genome based SNP phylogenetic relatedness displays considerable topological congruency with the late-competence related evolutionary relatedness.

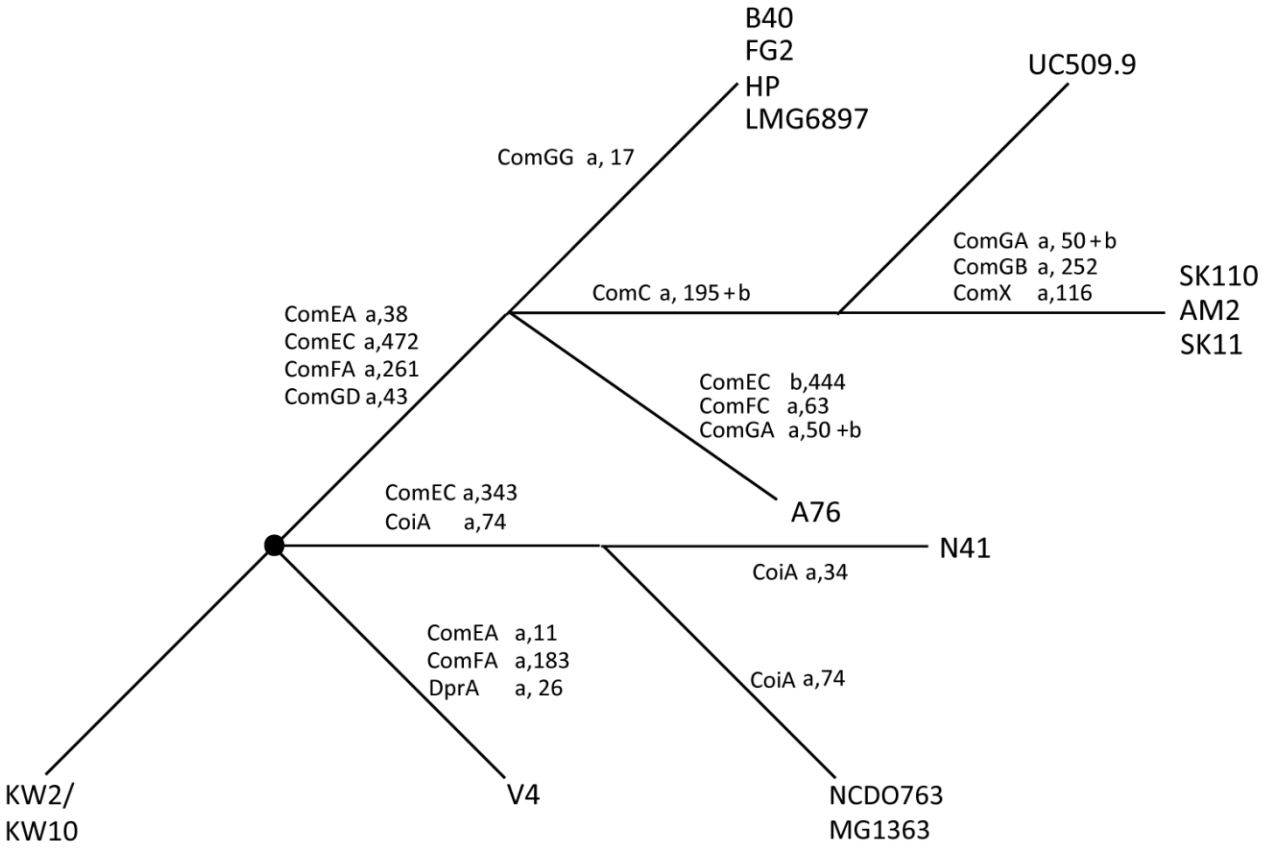


Fig. S5: Representative image of phage sequence insertions within *comGC* of *L. lactis* subsp. *lactis* strains KF282, KF24, N42, CV56, ML8, UC37, KLDS and KF67. Phage sequence arrows are scaled to the KF282 phage sequence arrow, whereas the *com* gene arrows are scaled to the KF282 *comGB* gene. On the right, the conserved insertion sequence within *comGC* in *L. lactis* subsp. *lactis* strains is depicted.

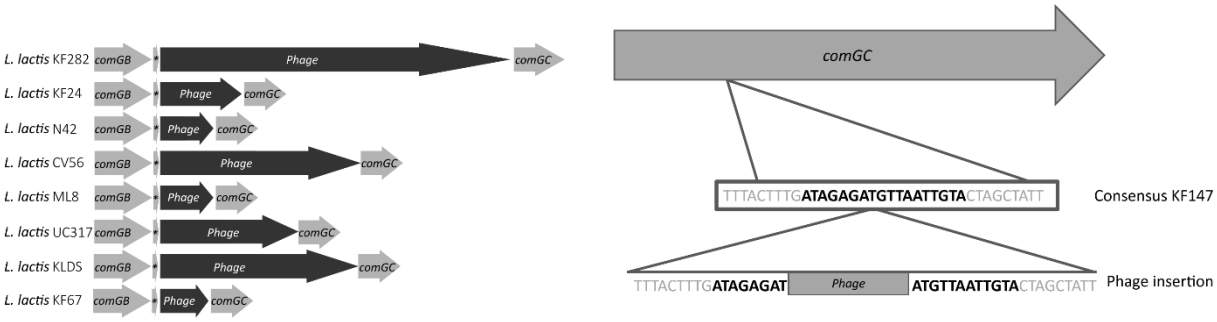


Fig. S6: Growth inhibition upon full induction with nisin is only observed in *L. lactis* KF147 harboring pNZ6200. Full induction with nisin in *L. lactis* KF147 harboring pNZ8150 (empty vector control) and *L. lactis* KF147 harboring pNZ8040 (*pepN* under nisin control) shows the same growth curve as uninduced conditions showing that nisin induction itself and overexpression of a protein does not lead to growth inhibition. In contrast, growth inhibition of *L. lactis* KF147 harboring pNZ6200 can be observed relative to uninduced conditions in this strain.

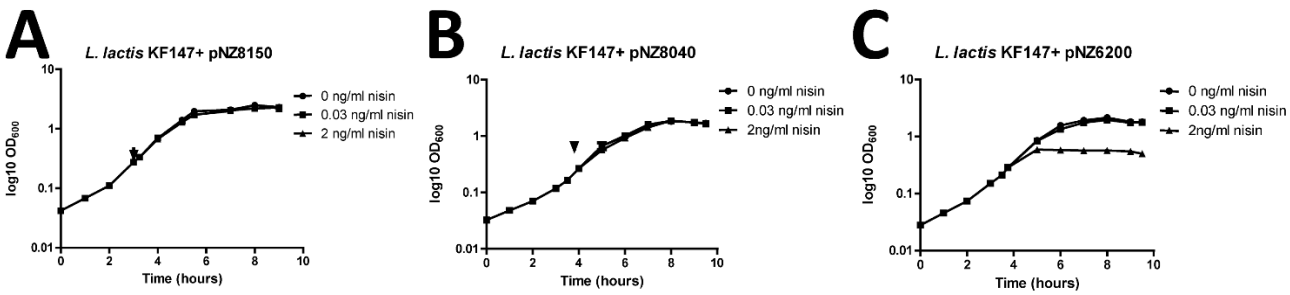
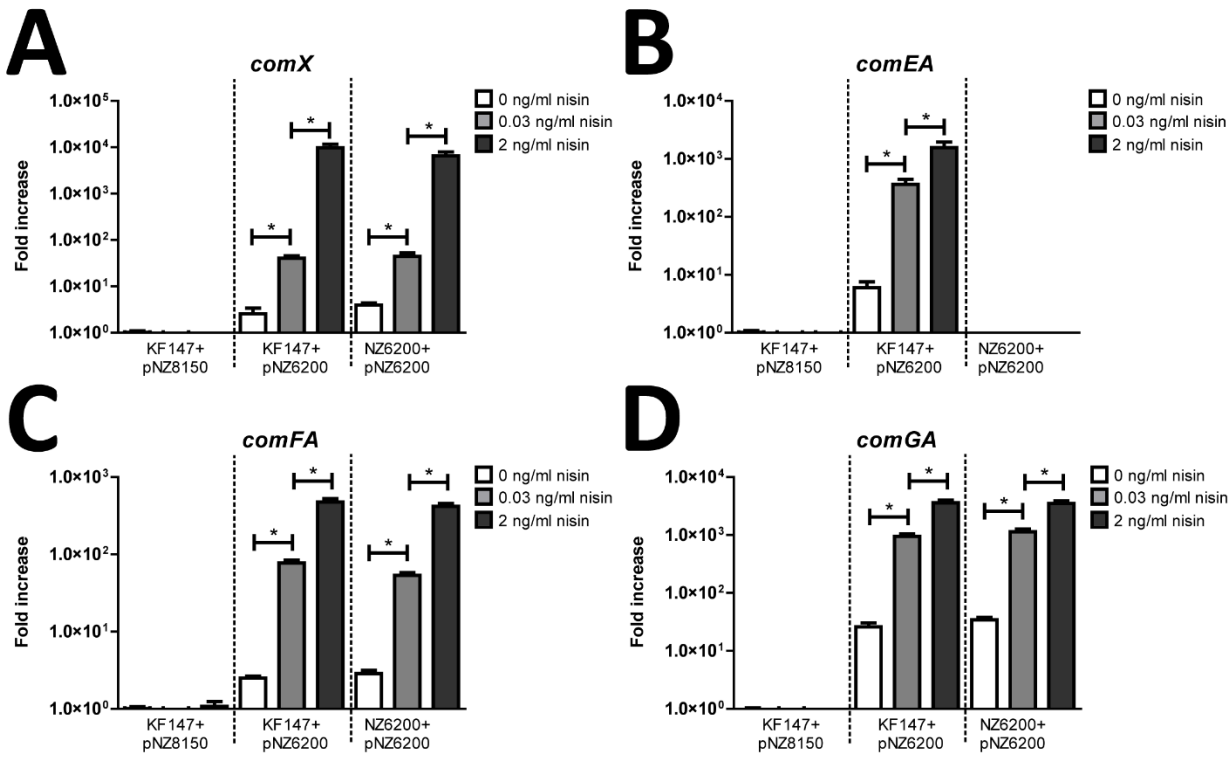


Fig. S7: Gene expression analysis of competence genes in KF147 and NZ6200. Transcript levels of *comX* (A), *comEA* (B), *comFA* (C) and *comGA* (D) in *L. lactis* KF147 harboring pNZ8150 or pNZ6200 and *L. lactis* NZ6200 harboring pNZ6200 in uninduced (0 ng/ml nisin), moderately induced (0.03 ng/ml nisin) and fully induced (2 ng/ml nisin) conditions. Values were obtained from triplicate measurements and statistical significance (P-value<0,05; indicated by *) is based on the non-parametric Mann-Whitney U-test (one-tailed).



References

1. **Mann JM, Carabetta VJ, Cristea IM, Dubnau D.** 2013. Complex formation and processing of the minor transformation pilins of *Bacillus subtilis*. Mol Microbiol **90:**1201-1215.
2. **Johnston C, Martin B, Fichant G, Polard P, Claverys JP.** 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol **12:**181-196.
3. **Muschiol S, Balaban M, Normark S, Henriques-Normark B.** 2015. Uptake of extracellular DNA: competence induced pili in natural transformation of *Streptococcus pneumoniae*. Bioessays **37:**426-435.
4. **Dijkstra AR, Setyawati MC, Bayjanov JR, Alkema W, van Hijum SA, Bron PA, Hugenholtz J.** 2014. Diversity in robustness of *Lactococcus lactis* strains during heat stress, oxidative stress, and spray drying stress. Appl Environ Microbiol **80:**603-611.