

Supplementary Figure S1. Multiple alignment of type I toxin families. (A) LdrD/Fst; (B) EHEC; (C) YhzE; (D) TxpA; (E) Hok/Gef; (F) Ibs; (G) TisB; (H) ShoB. Multiple alignments of protein sequences for each family were constructed by using MUSCLE program (1). Consensus was constructed using CONSENSUS program available on-line (<http://coot.embl.de/Alignment//consensus.html>). Proteins for type I toxins analyzed experimentally in this work are highlighted by red. Additional YhzE homologs present in *B. subtilis subsp. subtilis str.* 168 are underlined, and the putative phage YhzE homolog is in blue lettering. The color coding of amino acids is the following: hydrophobic amino acids (AVLIM) are blue, small amino acids (G) – green; aromatic amino acids (FYW) – magenta; polar amino acids (DEKR) – red; Unaligned N-terminal regions are generally non colored.

Supplementary Figure S1

A

LdrD/Fst (all)

```
SE_SL483_448875_Salmonella enterica SL483 -----MTLTELSITVWHDLAAPTLVGIATGLFLGWHRRK-----
*49176087_Escherichia coli str. K-12 substr. MG1655 -----MTLAQFAMIFWHDLAAPILAGIITAAIVSWWRNRK-----
49176088_Escherichia coli str. K-12 substr. MG1655 -----MTLAQFAMTFWHDLAAPILAGIITAAIVGWWRNRK-----
49176089_Escherichia coli str. K-12 substr. MG1655 -----MTLAQFAMIFWHDLAAPILAGIITAAIVSWWRNRK-----
49176368_Escherichia coli str. K-12 substr. MG1655 -----MTFAELGMAFWHDLAAPVIAGILASMIWNLNKRR-----
EC_CFT073_1509044_Escherichia coli CFT073 -----MTLAQFAVIFWHDLAAPILAGIITAAIVGWWRNRK-----
EC_UMN026_1524026_Escherichia coli UMN026 -----MTLAQFAMIFWHDLAAPILAGIITAAIVSWWRNRK-----
SD_Sd197_1171258_Shigella dysenteriae Sd197 -----MTLAQFAMIFWHDLAAPILAGIITVAIVGWWRNRK-----
EC_IAI39_1621258_Escherichia coli IAI39 -----MTLAQFAMIFWHDLAAPILAGIITAVIVSWWRNRK-----
SF_2457T_1265871_Shigella flexneri 2a str. 2457T -----MTLAQFAMTFWHDLAAPILAGIITAAIVSWWRNRK-----
EC_IAI39_1620723_Escherichia coli IAI39 -----MTLTQFAMTFWHDLAAPILAGIITAAIVSWWRNRK-----
EC_55989_1365538_Escherichia coli 55989 -----MTLAQFAMTFWHDLAAPILAGIITAAIVSWWRNRK-----
EC_ED1a_1351568_Escherichia coli ED1a -----MTLAQFAMTFWHDLAAPILAGIITAAIVGWWRNRK-----
*SB_308394_1259350_Shigella boydii CDC 3083-94 -----MTLAKFAMIFWHDLAAPILAGIITAAIVGWWRNRK-----
EC_HS_1317712_Escherichia coli HS -----MTLAQFAMIFWDDLAAPILAGIITAAIVGWWRNRK-----
EC_APEC01_1282999_Escherichia coli APEC 01 -----MTLAQFAMIFWHDLAAPILAGIITAAIVGWWRNRK-----
SE_CT_02021853_3848063_Salmonella enterica CT_02021853 -----MTLTQLGVVFWHDLAAPIIAGIITASVIVNWLDRK-----
*SE_z4z23_3940185_Salmonella enterica arizonae -----MTLTQLGVAFWHDLAAPIIAGIITASLIVNWLDRK-----
SE_SC-B67_3758987_Salmonella enterica SC-B67 -----MP-LQQGGYMTLTQLGVAFWHDLAAPIIAGIITASVIVNWLDRK-----
SE_AKU12601_3597279_Salmonella enterica AKU_12601 -----MP-LQQGGYMTLTQLGVAFWHDLAAPIIAGIITASVIVNWLDRK-----
EC_IAI39_4212812_Escherichia coli IAI39 -----MTLAEELGMAFWHDLAAPVIAGILASLIVNSLNKRR-----
EC_SMS35_3925001_Escherichia coli SMS-3-5 -----MTLAEELGMAFWHDLAAPVIAGILASLIVNWLNRK-----
EC_APEC01_4000357_Escherichia coli APEC 01 -----MTLAEELGMAFWHDLAAPVIAGILASMIWVNLNRK-----
*EC_E234869_3946592_Escherichia coli O127:H6 str. E2348/69 -----MTLAEELGMAFWHDLAVPVIAGILASMIWVNLNRK-----
SF_8401_3620153_Shigella flexneri 5 str. 8401 -----MTFAELGMAFWHDLAAPVIAGILASMIWVNLNRK-----
EC_UMN026_4161915_Escherichia coli UMN026 -----MTLAEELGMAFWHDLAAPVIAGILASMIWVNLKQK-----
EC_IAI1_3764984_Escherichia coli IAI1 -----MTLAEELGMAFWHDLAAPVIAGILASMIWVNLNRK-----
SS_Ss046_4048784_Shigella sonnei Ss046 -----MTLAEELGMTFWHDLAAPVIAGILASMIWVNLNRK-----
SS_Ss046_4049750_Shigella sonnei Ss046 -----MNNFLCYTQIGDLPKREVHMHYEIVTKILVPIFVGIIVLKLVTIWLEKQNEE-----
SS_Ss046_4049750_Shigella sonnei Ss046 -----MNNFLCYTQIGDLPKREVHMHYEIVTKILVPIFVGIIVLKLVTIWLEKQNEE-----
224545267_Bryantella formatexigens DSM 14469 MFIEILFSICYNSTVTASDTRYGREVLCEMFALSFIIAVTAGVVCHYVIKWLGDGK-----
*Streptococcus pneumoniae R6 1965404..1965499 -----MSLIIEALTTIADVIAIGIILYFVCKRNLDSKE-----
*Streptococcus pneumoniae R6 1965747..1965842 -----MSLIPELALTTIADVIAIGIILYFVCKRNLDSKE-----
GC_1317888_Streptococcus equi subsp. zooepidemicus -----MDAFTSFLVAVSSGIVILFEYVNLNRD-----
GC_1418301_Staphylococcus aureus RF122 -----LLIFVHIIAPVIVSGCAVAYTYWLSKRN-----
GC_1491108_Lactobacillus helveticus DPC 4571 -----SISLIVAPIIVGIVLALFNHNLDRD-----
GC_1516065_Macrocooccus caseolyticus JCSC5402 -----MLTIFVQVLCISIVSGCIVALFESHWLSLNK-----
GC_1570450_Staphylococcus epidermidis ATCC 12228 -----MILDIFVHIITTVISGCIVALFTHWLQRN-----
GC_1702840_Staphylococcus saprophyticus subsp. ATCC 15305 -----MSEIFVNIIMTTAASGCLVALFAHWLRRN-----
GC_1832812_Staphylococcus aureus subsp. aureus NCTC 8325 -----MLIFVHIIAPVIVSGCAIAFFSYWLSRRN-----
GC_1856538_Staphylococcus aureus subsp. aureus N315 -----MLIFVHIIAPVIVSGCAIAFFSYWLSRRN-----
GC_1873359_Staphylococcus aureus subsp. aureus MSSA476 -----MLIFVHIIAPVIVSGCAIAFFSYWLSRRN-----
```

GC_1879142_Staphylococcus aureus subsp. aureus ED98 -----MLIFVHIIALVIGSCAIAFFSYWLSRRN-----
GC_1889650_Staphylococcus aureus subsp. aureus str. Newman -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1894024_Staphylococcus aureus subsp. aureus MW2 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1906941_Staphylococcus aureus subsp. aureus COL -----MLIFVHIIAPVIGSCAIAFFFLAK-----
GC_1929120_Staphylococcus aureus subsp. aureus USA300_FPR3757 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1929870_Staphylococcus aureus subsp. aureus USA300_TCH1516 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1934331_Staphylococcus aureus subsp. aureus Mu50 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1935730_Staphylococcus aureus subsp. aureus Mu3 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1965404_Streptococcus pneumoniae R6 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_1965747_Streptococcus pneumoniae R6 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_1972693_Staphylococcus aureus subsp. aureus MRSA252 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1972906_Streptococcus pneumoniae D39 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_1973249_Streptococcus pneumoniae D39 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_1979535_Staphylococcus aureus subsp. aureus JH1 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_1979661_Staphylococcus aureus subsp. aureus JH9 -----MLIFVHIIAPVIGSCAIAFFSYWLSRRN-----
GC_2003059_Streptococcus pneumoniae G54 -----MSLIXELALTIADVIAGIILYFVCKWLDGKK-----
GC_2036109_Streptococcus pneumoniae Taiwan19F-14 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_2036452_Streptococcus pneumoniae Taiwan19F-14 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2039022_Streptococcus pneumoniae P1031 -----MSLIEELALTIADVIAGIILYFVCRWLDGKK-----
GC_2039365_Streptococcus pneumoniae P1031 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2047370_Streptococcus pneumoniae JJA -----MSLIEELALTIADVIAGIILYFVCRWLDGKK-----
GC_2047713_Streptococcus pneumoniae JJA -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2086132_Streptococcus pneumoniae TIGR4 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_2086475_Streptococcus pneumoniae TIGR4 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2111780_Streptococcus pneumoniae 70585 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_2112123_Streptococcus pneumoniae 70585 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_211586_Streptococcus mutans UA159 -----MWHNFFIYIVAPIFVGVVLALESEWLDKDD-----
GC_2135898_Streptococcus pneumoniae CGSP14 -----MSLIEELALTIADVIAGIILYFVCRWLDGKK-----
GC_2136241_Streptococcus pneumoniae CGSP14 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2147931_Streptococcus pneumoniae ATCC 700669 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_2148274_Streptococcus pneumoniae ATCC 700669 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2171141_Streptococcus pneumoniae Hungary19A-6 -----MSLIEELALTIADVIAGIILYFVCRWLDGKE-----
GC_2171484_Streptococcus pneumoniae Hungary19A-6 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
GC_2251172_Staphylococcus epidermidis RP62A -----MLEILVHITTTVINGCIVALFTHWLNKRN-----
GC_2256364_Staphylococcus epidermidis ATCC 12228 -----MLIFVHIIAPVIGSCAIAFFTHWLNKRN-----
GC_2258937_Staphylococcus epidermidis RP62A -----LLIIFVHIIAPVIGSCAVAYTYWLSKRN-----
GC_227302_Streptococcus pneumoniae ATCC 700669 -----MMELVLKTIIGPIVVGVLRIVDKWLNKDK-----
GC_228970_Streptococcus pneumoniae G54 -----MMELVLKTIIGPIVVGVLRIVDKWLNKDK-----
GC_231905_Streptococcus pneumoniae R6 -----MMELVLKTIIGPIVVGVLRIVDKWLNKDK-----
GC_231906_Streptococcus pneumoniae D39 -----MMELVLKTIIGPIVVGVLRIVDKWLNKDK-----
GC_237220_Streptococcus pneumoniae JJA -----MMELVLKTIIGPIVVGVLRIVDKWLNKDK-----
GC_2444003_Staphylococcus aureus RF122 -----MFNLLINIMTSAISGCLVAFFAHWRTRN-----
GC_2484026_Staphylococcus aureus subsp. aureus MSSA476 -----NLLIDIMTSALSGCLVAFFAHWRTRN-----
GC_2491397_Staphylococcus aureus subsp. aureus N315 -----NLLINIMTSAISGCLVAFFAHWRTRN-----
GC_2491960_Staphylococcus aureus subsp. aureus COL -----NLLINIMTSAISGCLVAFFAHWRTRN-----
GC_2502625_Staphylococcus aureus subsp. aureus NCTC 8325 -----NLLINIMTSAISGCLVAFFAHWRTRN-----
GC_2504858_Staphylococcus aureus subsp. aureus MW2 -----NLLIDIMTSALSGCLVAFFAHWRTRN-----
GC_2508137_Staphylococcus aureus subsp. aureus ED98 -----NLLINIMTSAISGCLVAFFAHWRTRN-----
GC_253506_Streptococcus pneumoniae CGSP14 -----MMELVLKTIIGPIVVGVLRIVDKWLNKDK-----
GC_2555522_Staphylococcus aureus subsp. aureus USA300_FPR3757 -----NLLINIMTSAISGCLVAFFAHWRTRN-----
GC_2560457_Staphylococcus aureus subsp. aureus str. Newman -----NLLINIMTSAISGCLVAFFAHWRTRN-----

GC_2562577_Staphylococcus aureus subsp. aureus Mu50 -----NLLINIMTSALSGCLVAFFAHWRTRN-----
GC_2564215_Staphylococcus aureus subsp. aureus Mu3 -----NLLINIMTSALSGCLVAFFAHWRTRN-----
GC_258421_Streptococcus pneumoniae P1031 -----MMELVLKTIIGPIVVGVLRLVDRKLNKDK-----
GC_2592712_Staphylococcus aureus subsp. aureus JH1 -----NLLINIMTSALSGCLVAFFAHWRTRN-----
GC_2592836_Staphylococcus aureus subsp. aureus JH9 -----NLLINIMTSALSGCLVAFFAHWRTRN-----
GC_2596556_Staphylococcus aureus subsp. aureus MRSA252 -----NLLIDIMTSALSGCLVAFFAHWRTRN-----
GC_264530_Streptococcus pneumoniae Taiwan19F-14 -----MMELVLKTIIGPIVVGVLRLVDRKLNKDK-----
GC_271752_Streptococcus pneumoniae 70585 -----MMELVLKTIIGPIVVGVLRLVDRKLNKDK-----
GC_321727_Streptococcus pneumoniae Hungary19A-6 -----MMELVLKTIIGPIVVGVLRLVDRKLNKDK-----
GC_586686_Staphylococcus epidermidis RP62A -----LLIIFVHITTTVLSCCIIALFTHWICNCN-----
GC_694453_Staphylococcus epidermidis ATCC 12228 -----LLIIFVHITTTVLSCCIIALFTHWICNCN-----
GC_805790_Streptococcus equi subsp. zooepidemicus MGCS10565 -----MDAFTSLVAVSSGIVILFEYLNLRD-----
GC_917116_Streptococcus thermophilus LMG 18311 -----MMDLLFKTIIGPIVVSIIILRLVDRWLNDR-----
GC_943484_Streptococcus thermophilus LMD-9 -----MMDLLFKTIIGPIVVSIIILRLVDRWLNDR-----
SP_SP18BS74_149006854_Streptococcus pneumoniae SP18-BS74 -----MSLIEELALTIADVIAGIILYFVCKWLDGKK-----
SP_SP14BS69_149003024_Streptococcus pneumoniae SP14-BS69 -----MSLIEELALTIITDVIAGIILYFVCKWLDGKK-----
228003276_Anaerococcus prevotii DSM 20548 -----MDNLLYSLTVSIVAGVLSHYICKWLDGNGKDSN-----
228003950_Anaerococcus prevotii DSM 20548 -----MIDPCQGLLYSLIVAIMANVISHYICKWLDSEDE-----
167747809_Anaerostipes caccae DSM 14662 -----MSRNKIYNEIAVDSFQLKGGTVEYILRFLMSVSISSIVNYYIHKWLDHRRKDS-----
229827833_Abiotrophia defectiva ATCC 49176 -----MNLLDLMVSVTAGVATFIICKWLDNRNSNDN-----
225386962_Clostridium asparagiforme DSM 15981 -----MQNLLNFLISIMASVAGNYISKWLNDRKDKS-----
*160937005_Clostridium bolteae ATCC BAA-613 -----MSDVLFVAVMASVAGYYICKWLDNRKDKS-----
224389231_Clostridiales bacterium 1_7_47_FAA -----MQGIIISFIISVMVNVIAYYICRWLDRKNDDN-----
228003046_Anaerococcus prevotii DSM 20548 -----MRFLLDIIFAIIVQVADDCLYKWLKASPSSGY-----
227486129_Anaerococcus lactolyticus ATCC 51172 -----MRLLLDIIFAVVIQIIGNYLYKWLKLNKDDN-----
228003033_Anaerococcus prevotii DSM 20548 -----MREVFSEILDVIVQVLGNLYKLLDNKLNKDDN-----
GC_1683661_Listeria welshimeri serovar 6b str. SLCC5334 -----LIVIKEILPKNLGFIISLVILSIVLYFVIGKWLNGKE-----
GC_1752814_Listeria monocytogenes HCC23 -----LIVMKEILPKNIGFIISLVILSIVLYFVIGKRLNGKH-----
GC_896581_Listeria innocua Clp11262 -----FIVIKEILPKNIGFIISLVILSIVLYFVIGKRLNGKH-----
GC_907654_Listeria monocytogenes EGD-e -----FIVIKEILPKNIGFIISLVILSIVLYFVIGKRLNGKH-----
GC_908009_Listeria monocytogenes Clp81459 -----FIVIKEILPKNIGFIISLVILSIVLYFVIGKRLNGKH-----
116873087_Listeria welshimeri serovar 6b str. SLCC5334 -----LIVIKEILPKNLGFIISLVILSIVLYFVIGKWLNGKEK-----
224499404_Listeria monocytogenes Finland 1988 -----FIVIKEILPKNIGFIISLVILSIVLYFVIGKWLNGKHST-----
46907099_Listeria monocytogenes str. 4b F2365 -----FIVIKEILPKNIGFIISLVILSIVLYFVIGKRLNGKHSR-----
SP_SP19BS75_149011128_Streptococcus pneumoniae SP19-BS75 -----MMDTILKTIIGPIVVGVLRLVDRKLNKDK-----
SP_SP9BS68_148993415_Streptococcus pneumoniae SP9-BS68 -----MMELILKTIIGPIVVGVLRLADKWLKDK-----
SP_CGSP14_182683238_Streptococcus pneumoniae CGSP14 -----MMELVLKTIIGPIVVGVLRLVDRKLNKDK-----
ST_CNZR1066_55822971_Streptococcus thermophilus CNZR1066 -----MMDLLFKTIIGPIVVSIIILRLVDRWLNDRDYC-----
VP_2008_227371732_Veillonella parvula DSM 2008 -----MQELFYALWGTLLAPIIVALVVTYSYWLNNKNNK-----
*SA_Mu50_14141832_Staphylococcus aureus subsp. aureus Mu50 -----MFSIIFVSVIAPIIIVGVIIITLFSHWLNRRDKQ-----
SH_15305_73662179_Staphylococcus saprophyticus ATCC 15305 -----MCITLFTVFIAPIVVGIVVTLFSYWLNNRD-----
SA_COL_57650976_Staphylococcus aureus subsp. aureus COL -----MFNLLINIMTSAISGCLVAFFAHWRTRNNKGGDK-----
SA_RP62A_57865642_Staphylococcus epidermidis RP62A -----MMLIFVHIIAPVISGCAIAFTHWLSKCNK-----
SH_15305_73663746_Staphylococcus saprophyticus ATCC 15305 -----MLMIFVHIIAPVISGCAVAYTYWLSKRNK-----
SAp_17227202_Staphylococcus aureus -----MLIIFVHIIAPVISGCAVAYTYWLSKRNK-----
SAp_90969155_Staphylococcus aureus -----MFTDGIIESPISQASIDEERRCITLLIIFVHIIPVISGCAVAYTYWLSKRNK-----
SA_N315_16119207_Staphylococcus aureus subsp. aureus N315 -----MFTDGIIESPISQASIDEERRCITLLIIFVHIIAPVISGCAVAYTYWLSKRNK-----
SA_MRSA252_49482916_Staphylococcus aureus MRSA252 -----MRFMNEILVHIMTTAISGCLVTLFGYWLHKRDKK-----
SA_MN8_227557251_Staphylococcus aureus subsp. aureus MN8 -----MNEILVHIMTTAISGCLVTLFGYWLHKRDKK-----
SAp_161598529_Staphylococcus aureus -----MLESSISQSTWRREVRFMNEILVHIMTTVISGCLVTLFGYWLHKRDKK-----
SA_TCH959_161508264_Staphylococcus aureus USA300_TCH959 -----MNEILVHIMTTVISGCLVTLFGYWLHKRDKK-----

```

SC_SK14_223043577_Staphylococcus capitis SK14 -----MTREGGAYMLFDIFVHIMATATSGCIVALFAHWRTRNDKRRK---
SH_JCSC1435_70726410_Staphylococcus haemolyticus JCSC1435 -----MADILVHIATVTSGCIVTLFSYWLNRNRNNK-----
SA_12228_27468440_Staphylococcus epidermidis ATCC 12228 -----MILDFVHIITTVISGCIVALFTHWLRQRNDKHR---
SH_JCSC1435_70727615_Staphylococcus haemolyticus JCSC1435 -----MVEALVHITTTVISGCIVALFTYWLTRNRNNKKK---
SH_JCSC1435_70727294_Staphylococcus haemolyticus JCSC1435 -----MLEILVHITTTVISGCIIALFTHWLRNRNRNNK-----
SAP_33416263_Staphylococcus aureus -----MLEILVHITTTVISGCIIALFTHWLRNRNRNNK-----
SH_JCSC1435_70727301_Staphylococcus haemolyticus JCSC1435 -----MLEIFVHITTTVISGCIIALFTHWLRNRKKK-----
SH_JCSC1435_70727603_Staphylococcus haemolyticus JCSC1435 -----MLEILVHITTTVISGCIIALFTHWLRNRKDK-----
LC_334_116494134_Lactobacillus casei ATCC 334 -----MLKDLFTVVVGPVVVGVIIELIKRWLDLRQLL-----
LC_334_116496109_Lactobacillus casei ATCC 334 -----MLRDEFFQIIAPLLVGVMTLVDHNLDDGGQ-----
LC_BL23_191639722_Lactobacillus casei BL23 -----MIDFLNQIIAPLIVGVVLLLEHRLNDRR-----
LC_18071178_Lactobacillus curvatus -----MLKSIPTLLIAPVLAGIATSLFDHNLDDQGRK-----
LG_JVV03_227521224_Lactobacillus gasseri JV-V03 -----MKEFLTLLVVAPILVGMVKSFLFDHNLDDRRNNKKHK-
*LG_MV22_226819732_Lactobacillus gasseri MV-22 -----MKDFLTLVVAPIVVEIVKSLFDHNLDDRHNNKKH--
*6114824|Bacil|Enterococcus faecalis plasmid pAD1 -----MKDLSLVIAPIFVGLVLEMSRVLDEEDDSRK---
227556034|Bacil|Enterococcus faecalis HH22 -----MRQLASIGGVLFVKDLSLVIAPIFVGLVLEMSRVLDEEDDSRK---
29375045|Bacil|Enterococcus faecalis V583 -----MNNFLCYTQIGDLPRKEVHHMYEIVTKILVPIFVGLVLEMSRVLDEEDDSRK---
consensus/90% .....h..hhhplhsslhshhhh.hh..Whpppp..

```

B

EHEC (all)

MDSL LTQ KLTVLIAVLELLVALLRLIDLLK	2954068	2953982	NC_002655	Escherichia coli O157:H7 EDL933
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2953459	2953373	NC_002655	Escherichia coli O157:H7 EDL933
MDSL LTQ KLTVLIAVLELLVALLRLIDLLK	2883837	2883751	NC_002695	Escherichia coli O157:H7 str. Sakai
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2883228	2883142	NC_002695	Escherichia coli O157:H7 str. Sakai
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2213607	2213521	NC_004337	Shigella flexneri 2a str. 301
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2495866	2495780	NC_004431	Escherichia coli CFT073
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2194355	2194269	NC_004741	Shigella flexneri 2a str. 2457T
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	971748	971662	NC_007613	Shigella boydii Sb227
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2345389	2345303	NC_007946	Escherichia coli UTI89
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2245050	2244964	NC_008253	Escherichia coli 536
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2318165	2318079	NC_008563	Escherichia coli APEC O1
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2386542	2386456	NC_009801	Escherichia coli E24377A
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2385934	2385848	NC_009801	Escherichia coli E24377A
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	936559	936645	NC_010498	Escherichia coli SMS-3-5
MDSL LTQ KLTVLIAVLELLVALLRLIDLLK	2880648	2880562	NC_011353	Escherichia coli O157:H7 str. EC4115
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2880039	2879953	NC_011353	Escherichia coli O157:H7 str. EC4115
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2459136	2459050	NC_011415	Escherichia coli SE11
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2458528	2458442	NC_011415	Escherichia coli SE11
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2334978	2334892	NC_011601	Escherichia coli O127:H6 str. E2348/69
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2258771	2258685	NC_011740	Escherichia fergusonii ATCC 35469
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2258163	2258077	NC_011740	Escherichia fergusonii ATCC 35469
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2264582	2264496	NC_011741	Escherichia coli IAI1
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2263974	2263888	NC_011741	Escherichia coli IAI1
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2241213	2241127	NC_011742	Escherichia coli S88
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2423648	2423562	NC_011745	Escherichia coli ED1a
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	947153	947239	NC_011750	Escherichia coli IAI39
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2506708	2506622	NC_011751	Escherichia coli UMN026
MD TLTQ KLTVLIAVLELLVALLRLIDLLK	2506100	2506014	NC_011751	Escherichia coli UMN026

C

YhzE

```
229101755|Bacil|Bacillus_cereus_Rock3-28-----MVVVVAARGEAGKSGFPFIILLFIVLIIIVGVS YIGLGF-----
229171811|Bacil|Bacillus_cereus_MM3VKATYNNKEDIPMGLGKGGSGGTGGGEAGQDKSVSSGYIVLIVLFIVLIVVGVYNGINF-----
229095655|Bacil|Bacillus_cereus_Rock3-29VKAIYNNKEDVHMLGKGGSGGTG---GGEAGQAGKSGFPFIIVLFIVLIIIVGASYSGLGF-----
229068710|Bacil|Bacillus_cereus_F65185-----MGLGGNGSGSGSGNGGGE-----AGQAGKSGFSLIIVLFIVLIIIVGVS YIGLGF-----
89100821|Bacil|Bacillus_sp-_NRR_L_B-14911---MYYPYPYPPPPVPYRAGG-----AWYALIVVLLILLIVFGGWYYGNC-----
206976537|Bacil|Bacillus_cereus_H3081-97MSDFEGVPYWAYVAYTEKRRGWGMN-----SQLILLLVLFILLVIVGIICL-----
196037222|Bacil|Bacillus_cereus_NVH0597-99-----MGICRIYRKEEGMGMN-----SQLILLLVLFILLVIVGIICL-----
229010457|Bacil|Bacillus_mycooides_DSM_2048-KTMYNDKEDFHMGS GKGGS--GGG-EDGKDGKSGSSGLLLIILLFIVFIIIGASYSVFSF-----
229165990|Bacil|Bacillus_cereus_AH621---MYNDKEDFHMGS GKGSGSGGG-EDGKDGKSGSSGLLLLILLFIVFIIIGASYSVFSF-----
192812183|Bacil|Geobacillus_sp-_Y412MC10-----MSADA-----GYGGRSATFVLVLYILLVILAA GFWI-----
167630710|Clost|Heliobacterium_modesticaldum_Ice1-----MGGSS-----YGFSGSGFVVAGLVIALIIIGA FYPHIVGFGYSGD
192812289|Bacil|Geobacillus_sp-_Y412MC10-----MSGAYGGAWG-----NSVGTILVLFILLVILSAWV-----
167462588|Bacil|Paenibacillus_larvae_BRL-230010-----MGYTAGSYGGGNG-----YGLITSTATILVLFILLVIVRSV LGY-----
167461796|Bacil|Paenibacillus_larvaeBRL-230010-----MSGVGGNGYCG-----FTSTGAILVLFILLVIVSR TFWLY-----
192813911|Bacil|Geobacillus_sp-_Y412MC10-----MSCEGGGY-----GYGNVGIIVLVL FILLVILSTFYI-----
192811952|Bacil|Geobacillus_sp-_Y412MC10-----MSYDKCCYGG-----GYGTSVGIIVLVL FILLVIVLSTFWI-----
169190051|Bacil|Paenibacillus_sp-_JDR-2-----MSEIHVSPVNDHR-----GGGFNHMGSILVLFILLVIVCTICGGGLY-----
169185223|Bacil|Paenibacillus_sp-_JDR-2-----MSGY-----VGGVNSTGFILVLFILLV IICGCGFGKFPY-----
169189370|Bacil|Paenibacillus_sp-_JDR-2-----MSYGVGPVAGAGCG-----YGGGSAFGFILVLFILLV IIRSVWL-----
169188819|Bacil|Paenibacillus_sp-_JDR-2-----MSGVY-----GGGVNTVGFILVLFILLV IIRTVCH-----
192808045|Bacil|Geobacillus_sp-_Y412MC10-----MGEVGY-----GGYNSVGA IIVL FILLVITASFWV-----
192812237|Bacil|Geobacillus_sp-_Y412MC10-----MSGVGEYGGGY-----GYGSSVGA IIVL FILLVITRAFI-----
205372673|Bacil|Bacillus_coahuilensis_m4-4-----MVLGT-----GYGSYAILIIVL FILLIILGLAFYPCVGPAAE--
169189592|Bacil|Paenibacillus_sp-_JDR-2-----MSGL-----GHGFNNVGSILVLFILLV IIVCGCGFVGGWGG--
226311789|Bacil|Brevibacillus_brevis_NBRC_100599-----MSTSF CGN-----GFFGDCFAIVLVL FILLVIIGCSCDEC-----
226311290|Bacil|Brevibacillus_brevis_NBRC_100599-----MSA-----GFFDD-FSLILVLFVLLVIVAC SCE-----
226312909|Bacil|Brevibacillus_brevis_NBRC_100599-----MRFFLILFKGVISM-----GFFDDNFALV LVL FVLLTIVACSCDEC-----
226311821|Bacil|Brevibacillus_brevis_NBRC_100599-----MLAKEGKRMSLF-----NGGFDDFAL IIVL FVLLVIVACSCD-----
169826817|Bacil|Lysinibacillus_sphaericus_C3-41QPYGNVGGWNNSYCGNNYSYCGG-----NNSGYGSTFVL IIVL FILLIIVGATYI-----
229543403|Bacil|Bacillus_coagulans_36D1-----MTDGGY-----SMGSSFAL IIVL FILLIIVGAFYMGSSGY-----
229544318|Bacil|Bacillus_sp-_B14905MYYGHEGAYPAYDPCCGYGYGSG---YGWGFAGHRNGFAL IIVL FILLIIGACWIGPN-----
126650401|Bacil|Bacillus_sp-_B14905---MGYGYGGVGNMNNGGCYGG---DNNGSYNGSAFVL IIVL FILLIIVGASFMSGGY-----
56420122|Bacil|Geobacillus_kaustophilus_HTA426---MPYGFYGGWGWGGYGFYPGY-----GGYGYGFVL IIVL FILLIIVGATWC-----
126650221|Bacil|Bacillus_sp-_B14905VGNAGYGSVGNYGCCYSGYGNEYSMGYCMDHGKGNSS FVL IIVL FILLIIVGATFMQKEDC-----
169828252|Bacil|Lysinibacillus_sphaericus_C3-41-----MGYCMDHGKGNSS FVL IIVL FILLIIVGATFMQKEDC-----
239827237|Bacil|Geobacillus_sp-_WCH70---MSFGYGF-GWGGCGCGFHGF---YGGFGHGRS FVL IIVL FILLIIVGSAWF-----
239826324|Bacil|Geobacillus_sp-_WCH70---MGFGCCGFGGYGGFYGFGY-----GGFGHSS FVL IIVL FILLIIVGAA FVY-----
221314202|Bacil|Bacillus_subtilis_NCIB_3610SYIVIRKDENSMYGSYGYGFYGC-----TNTFVL IIVL FILLIIVGAA FIC-----
```

154685616|Bacil|Bacillus_amyloliquefaciens_FZB42 -MGFGYGGFGGGYGGYGGCGGYGGG-----YVGGGYGSTFVLVVVLFILLIIVGASFFG-----
255767264|Bacil|Bacillus subtilis subsp 168 -MGFGYG-FGGGYGGG-CYGGYAGG-----YGGGYGSTFVLLVLFILLIIVGASFF-----
9630273|Caudo|Bacillus phage SPBc2 -----MGFYSGYSGGYSGG-----GYGSSFVLIVVLFILLIIVGATFLY-----
228990077|Bacil|Bacillus_pseudomycoides_DSM_12442 -----MVPLTQGGNVMGFGGSCG-----GFGGFFALLIVLFILLIIGCSCWGGGFGC---
15614747|Bacil|Bacillus_halodurans_C-125 ---MYGYGYGFPVAGAGYDCGYGYG--PVAGVGHCPKGFALIVVLFILLIIGASWGKC-----
15613960|Bacil|Bacillus_halodurans_C-125 -----MSAPTG-----GYGYSGFALLVLFILLVIIIGASWSGWY-----
228991022|Bacil|Bacillus_pseudomycoides_DSM_12442 ---YGYGGGYGYSCGGGYGYGGYGG-----GCGYG-GFALLIVLFILLIIGASCCGGFIGC---
229085235|Bacil|Bacillus_cereus_Rock3-44 ---MEQVMWGYYPSSCCGYGGYGG-----GCCGTGSGFALIVVLFILLIIGASCVR-----
89097803|Bacil|Bacillus_sp-NRRL_B-14911 --MNYGYSGYGCDPCYSGGYGYPAY--PVAGAQQGYA-GFALIVVLFILLIIGASCYGKW-----
52080320|Bacil|Bacillus_licheniformis_ATCC_14580 -----MSGGY-----GHGSGFALIVVLFILLIIGASWFGGGYGGY---
152975136|Bacil|Bacillus_cereus_cytotoxis_NVH39198 PLYLKCTLGKE--EKIMGYEASGF-----GFSGGFALLVLFILLIIGCSCWGGGFGY---
229010357|Bacil|Bacillus_mycoides_DSM_2048 ---MDFMEGGVNMGYGGSCGYGGSCG-----GCCGFGGFFALLVLFILLIIVGASWGGGFGC---
30261724|Bacil|Bacillus_anthraxis_str-Ames IVVHELRLTYSKGGTHMSYGGSCA-----GFGGFFALLIVLFILLIIGCSCWGGGYYG---
89099782|Bacil|Bacillus_sp-NRRL_B-14911 -----MSDGGC-----YGGGFALLVLFILLIIGASWYGGGGYGY--
229541453|Bacil|Bacillus_coagulans_36D1 ---MNNILMKQEEVDGMSGGAAYG-----GGFALLVLFILLIIGASWAW-----
149182279|Bacil|Bacillus_sp-SG-1 APNVNMAPTAVSPAYTAPGYGYCAP-----ARGYGDVFLIVVLFILLIIVGASFVPKC-----
15616531|Bacil|Bacillus_halodurans_C-125 MVAGVMEEPMPYAAAPAFGYGMPAP-----RKHGHGFVLIIVLFILLIIVGSSFVKC-----
239827240|Bacil|Geobacillus_sp-WCH70 -----MMWYYDGYAPHYYGYGYG-----GRFGCSFSLIVVLFILLIIVLA AVR-----
15615254|Bacil|Bacillus_halodurans_C-125 ---MGYYYGYPPFPFPPYRGYCPPY-----GGGFALIVVLFILLIILCGWFK-----
126650315|Bacil|Bacillus_sp-B14905 -----MGYCGNVGGIENGGYGGG-----SNNGSTFALIVVLFILLIIVGATFLKREKDC---
23098568|Bacil|Oceanobacillus_ihayensis_HTE831 -----MGKDHGGE-----NFGSAGFALLVLFILLVIIGTAYSGGDKCGGYCG
56964187|Bacil|Bacillus_clausii_KSM-K16 GYGYGGYGYGPEAYGGYGGYGYGCG-----GYGGGRGFALIVVLFILLIIVGASWFGYGGEEQCGP
255767463|Bacil|Bacillus subtilis 168 -----MYGYSYGYGFGCG-----TNTFVLIVVLFILLIIVGAAAFIC-----
255767336|Bacil|Bacillus subtilis 168 ---MYGYGYGGCCSYGGYGYGGC-----GYGYRTFALIVVLFILLIIVGAAFLGGGCC---
16079078|Bacil|Bacillus subtilis 168 -----MGFYSGYSGGYSGG-----GYGSSFVLIVVLFILLIIVGATFLY-----
255767745|Bacil|Bacillus subtilis 168 -----MGFYNSGGYSGNSG-----YSNGFGSSFALIVVLFILLIIVGAAIFNY-----
255767118|Bacil|Bacillus subtilis 168 (yczM) -----MSG-----GYGTSFALIVVLFILLIIVGTA FVGGY-----
221308193|Bacil|Bacillus subtilis subsp 168 -----MSGYSNCGG-----YGGISSFALIVVLFILLIIVGTA FVGGF-----
2|Bacil|Bacillus subtilis 168 -----MG-----EVFAGGFALLVLFILLIIGASWLY-----
255767231|Bacil|Bacillus subtilis 168 (yhzE) -----MS-----GGYSNGFALLVLFILLIIVGAAFIY-----
161376709|Bacil|Bacillus_megaterium VLSIYGGENMGFGCCGYGGYGYGGG-----GSGFGSGFALIIVLFILLIIGASC FGGFGGGC---
206977001|Bacil|Bacillus_cereus_H3081-97 ---MSEKCEHRHDDCNRRHGCG-----GGFALLIVLFILLIIGASC FGGGGGGC---
228918610|Bacil|Bacillus_thuringiensis_BGSC_4CC1 HGISPSFTLLSKGGP IVMGYGGTCG-----EGCGFAGGFALLVLFILLIIGCTFC-----
196037116|Bacil|Bacillus_cereus_W YMRTPLFSLLIKGS IVMGYGGSCG-----EGCGFTGEFALLVLFILLIIGCSCFC-----
15613742|Bacil|Bacillus_halodurans_C-125 -----MS-----VPHHGGFALIVVLFILLVIIGAAWVC-----
149180386|Bacil|Bacillus_sp-SG-1 -----MS-----ERGYGAGFALIVVLFILLIIVGAA YVGGGY-----
239827238|Bacil|Geobacillus_sp-WCH70 -----MMG-----AGYGGGFALIVVLFILLIIVGCACIGGWGY-----
56419190|Bacil|Geobacillus_kaustophilus_HTA426 -----MS-----APYSGGFALIVVLFILLIIVGCACVIY-----
15614748|Bacil|Bacillus_halodurans_C-125 -----MS-----AGFHGGGFALIVVLFILLVIVGSAFCC-----
15613741|Bacil|Bacillus_halodurans_C-125 -----MS-----GAYHGGFALIVVLFILLVIVGAAWYY-----
23099284|Bacil|Oceanobacillus_ihayensis_HTE831 -----MSG-----GYGYAGGFALIVVLFILLVIVGAAWF-----
154684872|Bacil|Bacillus_amyloliquefaciens_FZB42 -----MSG-----YSGISSFALIVVLFILLVIVGTAYVGGF-----
15615220|Bacil|Bacillus_halodurans_C-125 ---GVGYGAPVGVGYCPPVAGVG-----YGPCKGFALIVVLFILLIIVGASKFGVY-----
196249221|Bacil|Geobacillus_sp-G11MC16 -----MS-----AWHGNSFALIVVLFILLIIVGTA FGRYAY-----
157692134|Bacil|Bacillus_pumilus_SAFR-032 GGYGYGYGGGYGGCCYPSYGGGYG-----YGGGRTFALIVVLFILLIIVGAAFLGGGGGCC---

D

TxpA (all)

```
224547544|Clost|Bryantella_formatexigens_DSM_14469      -----MLTLLSIAIALILVAVTFIIGSLGGVAILIVFGDLI---VAVIVIVFIIRLLTKKKQK-----
*227554839|Bacil|Enterococcus_faecalis_HH22            -----MSVEAALGLMIGFATLV---VTIIFVILALVLDNKNRNS-----
29375319|Bacil|Enterococcus_faecalis_V583              -----MFLSVEAALGLMIGFATLV---VTIIFVILALVLDNKNRNS-----
29377545|Bacil|Enterococcus_faecalis_V583              -----MYDGGSELERQVMPMCIRCTLTKGAFFLSIEATLELMISFATLV---ALLIFGILEATKNDKK-----
229548059|Bacil|Enterococcus_faecalis_ATCC_29200       -----MHFNERSIFLSTIEAALELMISFAAFV---ALLIFGILEATKNNKK-----
238857151|Bacil|Enterococcus_faecalis_TUSoD_Ef11        -----MYPMHFNERSIFLSTIEAALELMISFAAFV---ALLIFGILEATKNDKK-----
29377544|Bacil|Enterococcus_faecalis_V583              -----MYPMHFNERSIFLSTIEAALELMISFAAFV---ALLIFGILEATKNNKK-----
15614759|Bacil|Bacillus_halodurans_C-125                -----MTVYETLSIVVQVNSIL---LATLTFVLSLVVFNKKK-----
224544345|Clost|Bryantella_formatexigens_DSM_14469     -----MESTYDGVVASRLILYIGREVAPMTAYEIVSTILAVASLM---VAVGMFVIALLAFLDKRRKKR-----
166031350|Clost|Dorea_formicigenerans_ATCC_27755        -----MLCYDTRDRETKARRLPFFYGGIKHPPNERKEGDAMVTYSDMIQFCIFI---VALVGLCYEIFKGRR-----
49484189|Bacil|Staphylococcus_aureus_aureus_MRSA252    -----MSKLDNRWNLSEYKEVVPVALLKSLERRRRLMITISTMLQFGLFI---IALIGLVIKLIELSNKK-----
88195852|Bacil|Staphylococcus_aureus_aureus_NCTC_8325  -----MDRWLSEYKEVVPVALLKSLERRRRLMITISTMLQFGLFI---IALIGLVIKLIELSNKK-----
66395573|Caudo|Staphylococcus_phage_42E                 -----MLALLKSLERRRRLMITISTMLQFGLFI---IALIGLVIKLIELSNKK-----
221138600|Bacil|Staphylococcus_aureus_subsp-JKD6008     -----MVALLKSLERRRRLMITISTMLQFGLFI---IALIGLVIKLIELSNKK-----
15924937|Bacil|Staphylococcus_aureus_subsp-aureus_Mu50  -----MVALLKSLERRRRLMITISTMLQFGLFI---IALIGLVIKLIELSNKK-----
221310065|Bacil|Bacillus_subtilis_subsp-subtilis_str-168 -----MTVYESLMIMINFGGLILNTVLLIFNIMMIVTSSQKKK-----
146319745|Bacil|Streptococcus_suis_05ZYH33              -----MVMGKSSKSNRKEKLSLSTFEVLTLMFVAGNFV---IALVMLVLELVKTKKITYVNFGEFTVIF-----
22538135|Bacil|Streptococcus_agalactiae_2603V-R         -----MGNSSKSNRKEELDLSLSTFEVLTLMFVAGNFV---IALVMLVLELVKTKK-----
199599144|Bacil|Lactobacillus_rhamnosus_HN001           -----MSVFEALSLVNQVVDNF---MTLIVLVTVLTAENKRRK-----
227517267|Bacil|Enterococcus_faecalis_TX0104            -----MSISTEYERKPYLSALDVIQLIFDFGLYT---IALISLVVTLKDKKK-----
229824502|Clost|Catonella_morbi_ATCC_51271              -----MCIHQNLGRRTLNSAEAVRRLILEFGTFV---IALIGLCHKLFNKHKDKK-----
223934112|Bacil|Streptococcus_suis_89-1591              -----MGNSSKSDGKEEHLTAFEVVTILGFGSFT---IALIGLCYKIFKDNKDKK-----
77411239|Bacil|Streptococcus_agalactiae_CJB111          -----MTAFEVVQIIIGFGSFT---VALIGLCYKIFKDDDKKK-----
225869355|Bacil|Streptococcus_equi_subsp-zoepidemicus   -----MTAFEVVQITLLGSSFT---IALISLCYQIFKDDNKK-----
225871364|Bacil|Streptococcus_equi_subsp-equi_4047     -----MTAFEVVQILLGSSFT---IALISLCYQIFKDDNKK-----
29377818|Bacil|Enterococcus_faecalis_V583              -----MNVLPKSIERRSLLSVSESLQMLLAFGGFT---LTLITIVAILNYKDKKK-----
227555552|Bacil|Enterococcus_faecalis_HH22              -----MCVLPKSQLERGLLSALETIQILILFGMFT---IALIKLIVDLLKNDKKK-----
227550225|Bacil|Enterococcus_faecium_TX1330             -----MKGVSLSAYETIQITILGFGMFT---IALISLVNMLKNDKKK-----
69245599|Bacil|Enterococcus_faecium_DO                 -----MLKSMKGVSLLSAYETIQITILGFGMFT---IALISLVNMLKNDKKK-----
238858600|Bacil|Enterococcus_faecalis_TUSoD_Ef11        -----MHVFPKFRERRGLLSAYETIQITILGFGMFT---IALIALIVKLLKNDKKK-----
29377690|Bacil|Enterococcus_faecalis_V583              -----MHVFPKFTERRGLLSAYETIQITILGFGMFT---IALIALIVKLLKNDKKK-----
194272120|Bacil|Enterococcus_faecalis_OG1RF             -----MSVLPKFTERRGLLSAYETIQITILGFGMFT---IALIALIVKLLKNDKKK-----
238858724|Bacil|Enterococcus_faecalis_TUSoD_Ef11        -----MSVLPKFTERRGLLSAYETIQITILGFGMFT---IALIALIVKLLKNDKKK-----
16079658|Bacil|Bacillus_subtilis_subsp-subtilis_str-168 -----MSTYESLMVMIGFANLIGGIMTWFVISLLTLFMLRKKDTHPIYITVKEK-----
192812442|Bacil|Geobacillus_sp-_Y412MC10                -----MEVKAALTIMFLFGSFT---LALLTYINNNKKIKTHPKVTRRRVWFH-----
192812492|Bacil|Geobacillus_sp-_Y412MC10                -----MEVKDALTIMFLFGTFI---IALLTYINNNKRR-----
157693513|Bacil|Bacillus_pumilus_SAFR-032               -----MSTFQALMLLAFSGSFT---ITLLTYIDKK-----
226309994|Bacil|Brevibacillus_brevis_NBRC_100599        -----MSKNEGVMLVEVYQALTLFMFGMFI---LALLTYLKKK-----
167462097|Bacil|Paenibacillus_larvae_larvae_BRL-230010  -----MKQAIYAYVGGRITSKGGDAMEFQALQLMFLFGMFI---LAL-TFTRKMK-----
167461049|Bacil|Paenibacillus_larvae_larvae_BRL-230010  -----METFQALQLMFLFGMFI---LALLTFIQMK-----
167461297|Bacil|Paenibacillus_larvae_larvae_BRL-230010  -----METFQALQLMFLFGMFI---LALLTFIQK-----
167461307|Bacil|Paenibacillus_larvae_larvae_BRL-230010  -----METFQALQLMFLFGMFI---LALLTFIQK-----
73662916|Bacil|Staphylococcus_saprophyticus_subsp-      -----MVSIIEVPYMLGFGTFI---VTLGIAITIVKSSHKK-----
149182144|Bacil|Bacillus_sp-_SG-1                       -----MGQRLILPRKGGDTLMVAIEDVLQIMIAFGTLV---AAIIAAATEKKK-----
205375278|Bacil|Bacillus_coahuilensis_m4-4              -----MLTISDVLQILGFGSLL---LAGVSLVVAIIAVTLKK-----
21282640|Bacil|Staphylococcus_aureus_subsp-aureus_MW2   -----MISIANALHMLLSFGMFI---VTFIGIVVAIINLSNKK-----
```

227432872 Bacil Leuconostoc_mesenteroides_subsp-_cremoris	-----MSISDALLVMFAFGGFI---ISILTFVLLIEKISKNNRILIFG----
50915151 Bacil Streptococcus_pyogenes_MGAS10394	-----MSGGGAYVCKSQPKERRRQGLSVYETLTLMIAFGTLI---VAIMNNKKNK-----
199599727 Bacil Lactobacillus_rhamnosus_HN001	-----MSVADALMLMLVFGDFL---LSLIALIITIMLIIQDNQKDRR-----
199599589 Bacil Lactobacillus_rhamnosus_HN001	-----MSVADALMLMPVFGDFV---LSLIALIITIMLIIQDNQKDRR-----
229551349 Bacil Lactobacillus_rhamnosus_LMS2-1	-----MSVADALMLMLVFGDFV---LSLIALIITIMLIIQDNQKDRR-----
116496125 Bacil Lactobacillus_casei_ATCC_334	-----MSVADALMLMLVFGGFI---LSLIALIVTIVVAILDSSKDRL-----
191639567 Bacil Lactobacillus_casei_BL23	-----MSVADALMLMLVFGGFI---LSWIALIITIVVAILDNQKDRH-----
224454503 Bacil Lactobacillus_paracasei_8700:2	-----MSVTDALMLMLVFGGFI---LSLIALIITIVVAILDNQKDRH-----
224454566 Bacil Lactobacillus_paracasei_8700:2	-----MSVADALMLMLVFGGFI---LSLIALIVTIVVAILDNQKDRH-----
191639590 Bacil Lactobacillus_casei_BL23	-----MSVADAITIMLAFGFSFV---LLLVGTVVILVRAILDNQKDRH-----
116496086 Bacil Lactobacillus_casei_ATCC_334	-----MSHAQNPVKGGVWLSVADAITIMLAFGFSFV---LLLVGTVVILVRAILDNQTD-----
224454531 Bacil Lactobacillus_paracasei_8700:2	-----MSHAQNPVKGGVWLSVADAITIMLAFGFSFV---LLLVGTVVILVRAILDNQKDRH-----
224456139 Bacil Lactobacillus_paracasei_8700:2	-----MSVAEAITLMLAFGFSFV---LSLVNTFVTLVKAMQDNQTD-----
224456168 Bacil Lactobacillus_paracasei_8700:2	-----MSVAEAITLMLAFGFSFV---LSLVNTFVTLVKAMQDNQKDRH-----
227522667 Bacil Lactobacillus_hilgardii_ATCC_8290	-----MSVSDAPQLAIGVGLLI---INLIGLVVELIKSQKNNRSLAELAIT-----
227362872 Bacil Lactobacillus_sakei_carnosus_DSM_15831	-----MLSTDALQLMLLFGTLI---INLIALVVELIKNHQQKNNRLLSSGKLAVI-----
116618051 Bacil Leuconostoc_mesenteroides_mesenteroides	-----MSVSDAVQLMLGFGTFV---LMLLGLVVELIKNSNKK-----
116618195 Bacil Leuconostoc_mesenteroides_mesenteroides	-----MSVSDTLQLMIFGTFV---LMLVGLVVELIKNSNKK-----
116494107 Bacil Lactobacillus_casei_ATCC_334	-----MSISDALQLMLGFGMFV---VALLALVVELIKNQKK-----
191637424 Bacil Lactobacillus_casei_BL23	-----MSISDALQLMLGFGFV---VLLALVVELIKNQKK-----
226820159 Bacil Lactobacillus_gasseri_MV-22	-----MSVADALQLMLAFGTFI---VALIALVVELIKSQKK-----
170016344 Bacil Leuconostoc_citreum_KM20	-----MSVSDALQLMFMFGTFI---VALLALVVELIKSQKK-----
170016281 Bacil Leuconostoc_citreum_KM20	-----MSVSDALQLMLMFGTFI---VALLTLVVELIKNQKK-----
160940345 Clost Clostridium_bolteae_ATCC_BAA-613	-MPYAVTLLFCIISTGSSYCWVTSISERRWYNMSTYEALSLMIAFGVLI---VMIIGTKK-----
160940340 Clost Clostridium_bolteae_ATCC_BAA-613	-----MTSISERGWYDMSTYETLSLMIAFGVLI---VMIIGTKK-----
160940344 Clost Clostridium_bolteae_ATCC_BAA-613	-MPYAVTKVFSIINTGSSYCWVTSISERGWYNMSTYEALSLMIAFGVLI---VMIIGTKK-----
199598668 Bacil Lactobacillus_rhamnosus_HN001	---MLNVNQERKYIEPLLVSKRHSSTFYIAAEMSDFEALSLMIMAALLL---IAVIELVLKLIIRD-----
116493973 Bacil Lactobacillus_casei_ATCC_334	-----MSIYEALSLMIMFGLFV---LGLITLVKLNDR-----
29377692 Bacil Enterococcus_faecalis_V583	-----MNVSTKIYERRGLLSIAEALALMISFGSFI---ATLIFGILKVVVKEDKKK-----
194272132 Bacil Enterococcus_faecalis_OG1RF	-----MSIAEALALMISFGSFI---ATLIFGILEAVKENNKK-----
229547340 Bacil Enterococcus_faecalis_TX1332	-----MNVSTKIYERRGLLSIAEALALMISFGSFI---ATLIFGILEAVKENNKK-----
218289529 Bacil Alicyclobacillus_acidocaldarius_LAA1	-----MSDIGILLQMIQTGSLL---ITLLQFVLALVLMKRRK-----
196249249 Bacil Geobacillus_sp-_G11MC16	-----MVMTIADALTLMIAFASLI---VAVIAVAKDKK-----
221088255 Bacil Geobacillus_sp-_Y412MC61	-----MMTVAEALSMLMISFAPLV---VAVIAVSKKHPHTLLGLGEGGLLVPVSGP-----
228006493 Bacil Alicyclobacillus_acidocaldarius_subsp-	-----MSVYDALSLVQFASLM---VAVLALVVTLVLALRRK-----
228005982 Bacil Alicyclobacillus_acidocaldarius_subsp-	-----MSVAEALSMLIQFGSFI---VTLTLVVALVTLKRRK-----
218289522 Bacil Alicyclobacillus_acidocaldarius_LAA1	-----MTVADALSMLIQFGSFI---VALTSLVALVALRRK-----
218289238 Bacil Alicyclobacillus_acidocaldarius_LAA1	-----MTVADAVSLAIQFGSFI---IALLSLVVALVALTRK-----
228004480 Bacil Alicyclobacillus_acidocaldarius_subsp-	-----MTVADALSLLIQFGSFI---MTCLSLVVALVALTRK-----
consensus/95%hph..sl.hhh.hu.hh...h.hl.hh.....
consensus/90%hph.ts1.hhh.Fu.hh...h.ll.hhh.h.....
consensus/80%hoh.ps1tlhl.FG.hl...lslshlhth...p.....
15612907 Bacil Bacillus_halodurans_C-125	-----MTVFEALMFAVAFATLI---IAVLSFHEKK-----
47092607 Bacil Listeria_monocytogenes_str._4b_H7858	-----MYSFTPGKELDLSVEALTLMIAFATLV---VLLIGFEQKK-----
255016714 Bacil Listeria_monocytogenes_FSL_F2-515	----MVVASLFSKKEVMPMFFLFRLSWKGYAMTVEALSFAIAFATLV---LLISSQNEQKKNHT-----
29377706 Bacil Enterococcus_faecalis_V583_1	-----MTVFEALMLAIAFATLI---VKISNKNDKK-----

E

Hok (selected)

```
90580669|Gamma|Vibrio_angustum_S14      --MRVLLAVRQGHKTVSITLLVALPRVRDDLCEIQYQVGTQPFKAVLAFEVV-
27367087|Gamma|Vibrio_vulnificus_CMCP6  -QVMPKKTALAALVVVCITAVIITALYTGSLCDIRYKQDQNDLSIKLAYEVR-
15800674|Gamma|Escherichia_coli_O157-H7_EDL933  ---MPQKTIIVGML--CLTMLLTWVVLHASPCEFRVSMWSEIAAFLQCKP--
212705715|Gamma|Providencia_rustigianii_DSM_4541  ----MTKYAMIAVVSICVTILGCSFLIKDRLSYLHIDNDNTVVQVKLDYEIN-
197283937|Gamma|Proteus_mirabilis_HI4320  ----MTKITLFGITVVVCTVLMFALIVGERLCSLNISNGNMVVQAILMCHQ--
223992212|Gamma|Providencia_rettgeri_DSM_1131  ----MAKIALLLGLVTVCLTVLCFSLLMQERLCSFSSISSGNTLVQATLSCDN--
227358566|Gamma|Yersinia_mirabilis_ATCC_29906  ----MTKLALIGLITVCI TVLCFLLRHERLCSFNISSGNTLVQANLSCDK--
172056019|Gamma|Proteus_mirabilis_HI4320  ----MTKYALIGLIAVCLTVLCFSLLMRDRLCSI SVNSGNTVVQATLSYEER-
157149415|Gamma|Escherichia_coli_E24377A  -----MFAVCATVLCFSLIFRERLCELNIHRGNTVVQVTLAYEARQ
9507717|Plasm|Plasmid_F                  -SKFMTKYALIGLLAVCATVLCFSLIFRERLCELNIHRGNTVVQVTLAYEARK
187736817|Gamma|Escherichia_coli_1520      -SKFMTKYALIGLLAVCATVLCFSLIFRERLCELNIHRGNTVVQVTLAYEARK
9507718|Plasm|Plasmid_F                  ----MTKYALIGLLAVCATVLCFSLIFRERLCELNIHRGNTVVQVTLAYEARK
217975709|Gamma|Shewanella_baltica_OS223  -DMPKRKTTTSLAIVCFLLAALALIRDDLCKVEYRNGAMLLNVVLAYEVRG
217980295|Gamma|Shewanella_baltica_OS223  -MPKRKTTTSLAIVCFLLAALALIRDDLCKVEYRNGAMLLNVVLAYEVRG
238764212|Gamma|Yersinia_kristensenii_ATCC_33638  ---MSQKLLKTAMI-CITAVLIWMI RSHLCEVHIKLGGAETAAFLQCKQQA
170769155|Gamma|Escherichia_albertii_TW07627  ----MSQKSLIAITTLCTVAILIWMMLRGSLSCEIRMSFWGAEFAAFLQCRQ--
168748214|Gamma|Escherichia_coli_O157-H7_str-_EC4113  ----MSQKSLITVTI-CMTVIFTIWMMLHGSLSCEFRRLNLWGAEFAAFLQCKQ--
156932349|Gamma|Cronobacter_sakazakii_ATCC_BAA-894  -GMKPLHYLLACLFMVCITILIFALMNQGTLCBELTIRSGSQEVAAKLACTGK-
209906156|Gamma|Enterobacter_cancerogenus_ATCC_35316  --MTPLKAALGTVSIIICLTIVIFAFINRCKLCELTIKSEHQEVAAKLACISG-
156935098|Gamma|Cronobacter_sakazakii_ATCC_BAA-894  -MAMKDNFTFWCIIVCVTVIMFTTSLRETLCLELRLRGAGMEIVASLACKPRE
157371511|Gamma|Serratia_proteamaculans_568  ---MRLKHVHSHLTVVCI TILIFWMI RDSLCELTIVYQENITILIRLACDVKR
227333763|Gamma|Citrobacter_youngae_ATCC_29220  -AMKQQRFRFITLIIICITVVLVVLITRKDLCEVRLKAGQTEVAVFTAYESE-
37526772|Gamma|Phototrhobdus_luminescens_TTO1  -TMKQQKAI FIAIVI-CIAALA AVLTRKDLCEVRI RSGQTEVAVFMDYEP-
15799696|Gamma|Escherichia_coli_O157-H7_EDL933  -AMKQHKVMIVALIVXCITAVVAALVTRKDLCEVHIRTGQTEVAVFTAYESE-
49175991|Gamma|Escherichia_coli_str-_K-12_MG1655  --MKQHKAMIVALIVICITAVVAALVTRKDLCEVHIRTGQTEVAVFTAYESE-
226955804|Gamma|Escherichia_sp-_1_1_43  -AMKQHKAMIVALIVICITAVVAALVTRKDLCEVHIRTGQTEVAVFTAYESE-
9632542|Caudo|Enterobacteria_phage_933W  -AMKQQKAMLIALIVICLTVIVTALVTRKDLCEVRI RTGQTEVAVFVDYESEK
9632541|Caudo|Enterobacteria_phage_933W  --MKQQKAMLIALIVICLTVIVTALVTRKDLCEVRI RTGQTEVAVFVDYESEK
16129521|Gamma|Escherichia_coli_str-_K-12_MG1655  --MKQQKAMLIALIVICLTVIVTALVTRKDLCEVRI RTGQTEVAVFTAYEPEE
15801499|Gamma|Escherichia_coli_O157-H7_EDL933  -AMKQQKAMLIALIVICLTVIVTALVTRKDLCEVRI RTGQTEVAVFTAYEPEE
9507499|Plasm|Plasmid_ColIb-P9           --MMPQRTFLMMLIVICVTILCFVMV RDSLCLVRLRQQGNTVLVATLAYEVKR
209906319|Gamma|Enterobacter_cancerogenus_ATCC_35316  -GM-PKRALLGLFLICTLLIFWVMV RDSLCELHFRQEKTELA AVLAYEAKR
152972428|Gamma|Klebsiella_pneumoniae_MGH_78578  -RGMPPKYLLFGLVVICFTILLLTWVMV RDSLCELQRRQGNIELVAF LACDIKQ
206580878|Gamma|Klebsiella_pneumoniae_342  ---MPQKYLLFGLVVICFTILLLTWVMV RDSLCELQLRQGDIELVAF LACDIKQ
226957115|Gamma|Escherichia_sp-_1_1_43  -LGMPPKYRLLSLIVICFTLLFFTWMI RDSLCELHIKQGSYELAAFLACNLKE
49176371|Gamma|Escherichia_coli_str-_K-12_sMG1655  ---MPQKYRLLSLIVICFTLLFFTWMI RDSLCELHIKQGSYELAAFLACNLKE
15804103|Gamma|Escherichia_coli_O157-H7_EDL933  -LGMPPKYRLLSLIVICFTLLFFTWMI RDSLCELHIKQGSYELAAFLACNLKE
15800295|Gamma|Escherichia_coli_O157-H7_EDL933  ---MLTKYALVAVIVLCLTVPGFTLLVGD SLCEFTVKERNIEFRAVLAYEPKK
15800296|Gamma|Escherichia_coli_O157-H7_EDL933  -INMLTKYALVAVIVLCLTVLGF TLLVGD SLCEFTVKERNIEFKAVLAYEPKK
110640812|Gamma|Escherichia_coli_536      -INMLTKYALVAVIVLCLTVLGF TLLVGD SLCEFTVKERNIEFKAVLAYEPKK
```

```

152971617|Gamma|Klebsiella_pneumoniae_MGH_78578      -QDMLTKYALVAIIVLCITVLGFTLLVHSSLCELSIKERNIEFKAVLAYESKK
206578783|Gamma|Klebsiella_pneumoniae_342          ---MLTKYALVAIIVLCITVLGFTLLVHSSLCELSIKERNIEFKAVLAYESKK
15803279|Gamma|Escherichia_coli_O157-H7_EDL933     ---MLTKYALVAIIVLCCTVLGFTLMVGDSDLCELSIRERGMIEFKAVLAYESKK
170765967|Gamma|Escherichia_albertii_TW07627      -----MLCFTVLGF TLMVGDSDLCELSIRERGMIEFKAVLAYESKK
145301477|Gamma|Aeromonas_salmonicida_A449        ---MFGKTAVVSLIVCITALGVISLVRDSDLCELEVHQGETEIRLNLAYEAKR
145301527|Gamma|Aeromonas_salmonicida_A449        ---MPGKTAVMALLIVCVTVLAF TALVRDSDLCELSVRQGGTEIRAYLAYEPRE
38703974|Gamma|Escherichia_coli_O157-H7_str-_Sakai  ---MKRNPVVCLLIICITILTFTLLTRQTLYELRFRDSDKEVAALMACTSR-
91210664|Gamma|Escherichia_coli_UTI89             ----LATTANPGPCE-SNTILFT-LTRQSLYELRFRDSDKEVAALMACTSR-
157369192|Gamma|Serratia_proteamaculans_568       ---MQQERVVLRMLIICITLIALMWITRGSLSCELRITLGDTEVAAILAYETEG
157372209|Gamma|Serratia_proteamaculans_568       -----MIICMTLIALMWITRGSLSCELRITLGDTEVAAILAYESKG
238760001|Gamma|Yersinia_aldovae_ATCC_35236       ---MPQKT FVLC LLIVCITVLAFTLITHKSLCELRRLKDNKEVAAILAYESER
206575599|Gamma|Klebsiella_pneumoniae_342        -MKLPSPNLLWCVLIVCLTLLAF TYLTRKSLCEIRYPVFIDEEMALFNADQDA
10955445|Gamma|Escherichia_coli                  -MKQPKNALTWCLLIVCCTLLIF TYLTRNRLCEVRLKGDREVTASLAYESNG
83404855|Gamma|Escherichia_coli                  -MKLPGNALIWCVLIVCCTLLIF TLLTRNRLCEVRLKGDYREVATMAYESGG
9507608|Plasm|Plasmid_R100                        -MKLPRSSLVWCVLIVCLTLLIF TYLTRKSLCEIRYRDGHREVAAFMAYESGK
222104868|Gamma|Escherichia_coli                  -----MCLTLLIF TYLTRKSLCEIRYRDGHREVAAFMAYESGK
157412188|Gamma|Escherichia_coli_APEC_O1          AMKMPNQPIVLCILIVCLTLLIF TWLTRNSLCELRMKDGTREVF AILAYESGK
18466663|Gamma|Salmonella_enterica_Typhi_str-_CT18 -MKLPNQPVVLCVLIVCLTLLIF TWLTRNSLCELRRLKDG TREVSAVMDYESGK
10957290|Gamma|Salmonella_enterica__Typhi         AMKLPNQPVVLCVLIVCLTLLIF TWLTRNSLCELRRLKDG TREVSAVMDYESGK
Consensus      .....p.hhh.h..lChThlhh.hhhttpLCphp...h.h .shhttp..

```

F

Ibs (selected)

--FKYIILVILLALSSPAY	605479	605535	NC_011852	Haemophilus parasuis SH0165
--FKYIILVILLVLSLPAY	1031977	1032033	NC_010519	Haemophilus somnus 2336
--FKYIILVVFLLVSSPAY	33079	33135	NC_008201	Mannheimia phage phiMHaA1
--FKIAIILVILLVVSSNAY	1142163	1142219	NC_011852	Haemophilus parasuis SH0165
--IKIILVLLLLVSSQPAY	152027	152083	NC_000907	Haemophilus influenzae Rd KW20
--IKIILVLLLLVSSQPAY	214235	214291	NC_007146	Haemophilus influenzae 86-028NP
--IKLILVLLLLVSSQPAY	608786	608842	NC_009567	Haemophilus influenzae PittGG
--IKLILVLLLLVSSQPAY	513497	513553	NC_009566	Haemophilus influenzae PitTEE
KLCYITITFVLLMLLSSPAY	1186462	1186530	NC_009566	Haemophilus influenzae PitTEE
KTLQIILILVLLLLVSSPAY	1009618	1009680	NC_010519	Haemophilus somnus 2336
KPLTIIIIILALLVVSSPAY	259006	259062	NC_010519	Haemophilus somnus 2336
--LFKALILILMLVMPAF	1008195	1008251	NC_010519	Haemophilus somnus 2336
--LMKIFLIVLLLLSPAY	260948	261004	NC_010519	Haemophilus somnus 2336
--LVKVLILILMLVSPAY	261404	261460	NC_010519	Haemophilus somnus 2336
--LVRIILVILLVSSPAY	1143595	1143651	NC_011852	Haemophilus parasuis SH0165
--MKCVIILVILLVSSPAY	687324	687380	NC_009792	Citrobacter koseri ATCC BAA-895
--MKFVILVILLVSSPAY	3064331	3064387	NC_011740	Escherichia fergusonii ATCC 35469
--MKILII-VLLVSSPAY	1001209	1001262	NC_010498	Escherichia coli SMS-3-5
--MKILII-VLLVSSPAY	2196637	2196690	NC_011740	Escherichia fergusonii ATCC 35469
--MKLVILVILLVSSPAY	3276253	3276309	NC_011741	Escherichia coli IA11
--MKLVILVILLVSSPAY	3063955	3064011	NC_011740	Escherichia fergusonii ATCC 35469
--MKLVILVILLVSSPAY	3649628	3649684	NC_011751	Escherichia coli UMN026
--MKRLII-VLLVSSPAY	2190475	2190528	NC_008253	Escherichia coli 536
--MKSVIILVILLVSSPAY	2915661	2915717	NC_011740	Escherichia fergusonii ATCC 35469
--MKWIII-VLLVSSPAY	2196964	2197017	NC_011740	Escherichia fergusonii ATCC 35469
--MQQVILVILLVSSPAY	3285265	3285321	NC_011149	Salmonella enterica enterica SL483
--MQQVILVILLVSSPAY	4348582	4348638	NC_010067	Salmonella enterica arizonae
--MRFVILVILLVSSPAY	855795	855851	NC_010468	Escherichia coli ATCC 8739
--MRLIII-VLLVSSPAY	1057068	1057121	NC_010658	Shigella boydii CDC 3083-94
--MTSVIILVILLVSSPAY	786854	786910	NC_010067	Salmonella enterica arizonae
--MKLVILVILLVSSPAY	3420382	3420438	NC_010498	Escherichia coli SMS-3-5
--MKVVIILVILLVSSPAY	2171908	2171964	NC_003198	Salmonella enterica Typhi str. CT18
--MRLVILVILLVSSPAY	3861330	3861386	NC_002655	Escherichia coli O157:H7 EDL933
--MKVVIILVILLVSSPAY	2345568	2345624	NC_011205	Salmonella enterica CT_02021853
--MRLVILVILLVSSPAY	3084404	3084460	NC_009800	Escherichia coli HS
--MKLIII-VLLVSSPAY	2383714	2383767	NC_011748	Escherichia coli 55989
--MQQVILVILLVSSPAY	3324415	3324471	NC_006905	Salmonella entericaSC-B67
--MKLIII-VLLVSSPAY	2240242	2240295	NC_008563	Escherichia coli APEC O1
--MKLVILVILLVSSPAY	4001984	4002040	NC_002655	Escherichia coli O157:H7 EDL933
--MKWIII-VLLVSSPAY	1739368	1739421	NC_010468	Escherichia coli ATCC 8739
--MHQVILVILLVSSPAY	3409390	3409446	NC_011205	Salmonella entericaCT_02021853
--MKILII-VLLVSSPAY	2190803	2190856	NC_008253	Escherichia coli 536
--MKVVIILVILLVSSPAY	2171812	2171868	NC_011149	Salmonella entericaSL483
--MRLAIIIVLLVSSPAY	3282007	3282063	NC_011748	Escherichia coli 55989
--MRLVILVILLVSSPAY	3074755	3074811	NC_008253	Escherichia coli 536
--MKLIII-VLLVSSPAY	2383386	2383439	NC_011748	Escherichia coli 55989
--MRLVILVILLVSSPAY	3282399	3282455	NC_011748	Escherichia coli 55989
--MKHVILVILLVSSPAY	2383054	2383110	NC_011748	Escherichia coli 55989
--MKFVILVILLVSSPAY	3306927	3306983	NC_008253	Escherichia coli 536
--MKLVILVILLVSSPAY	3543516	3543572	NC_011748	Escherichia coli 55989
--MKLLVIVLMLVLTTPLY	946840	946893	CP001072	Helicobacter pylori Shi470
--MRL--VIVLMLVLTTPLY	432397	432350	CP001072	Helicobacter pylori Shi470
--MKLLVIVLMLVLTTPLY	931110	931163	CP000241	Helicobacter pylori HPAG1
--MKLLVIVLMLVLTTPLY	931049	931102	AE001439	Helicobacter pylori J99
--MKLLVIVLMLVLTTPLY	930742	930795	AE001439	Helicobacter pylori J99
--MKLPLVIVLMLVLTTPLY	442707	442654	AE001439	Helicobacter pylori J99
--MKLLVIVLMLVLTTPLY	887193	887246	FM991728	Helicobacter pylori B38
--MKLLAIVLMLVLTTPLY	403303	403356	FM991728	Helicobacter pylori B38
--MKLLVIVLMLVLTTPLY	930625	930678	CP001173	Helicobacter pylori G27
--MKLLAIVLMLVLTTPLY	415957	416007	CP001173	Helicobacter pylori G27
--FKLIMIIVLMLVLTTPLY	1648187	1648240	CP001173	Helicobacter pylori G27
--MRL--VIVLMLVLTTPLY	448662	448615	CP001173	Helicobacter pylori G27
--MKLLVIVLMLVLTTPLY	962263	962316	CP001217	Helicobacter pylori P12
--MRL--VIVLMLVLTTPLY	447826	447779	CP001217	Helicobacter pylori P12
--MKLSLIVLMLVLTTPLY	1284766	1284819	AM260522	Helicobacter acinonychis str. Sheeba
--MRL--VIVLMLVLTTPLY	479869	479822	AE000511	Helicobacter pylori 26695
--MRL--VIVLMLVLTTPLY	1071947	1071994	AE000511	Helicobacter pylori 26695

G

TisB (all)

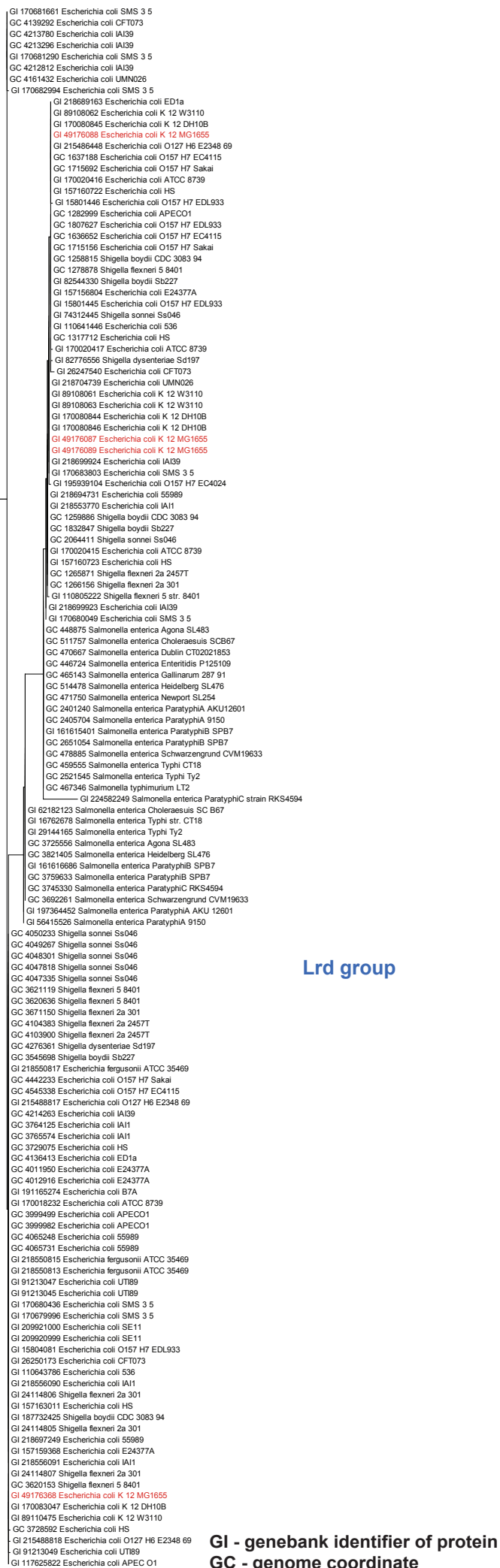
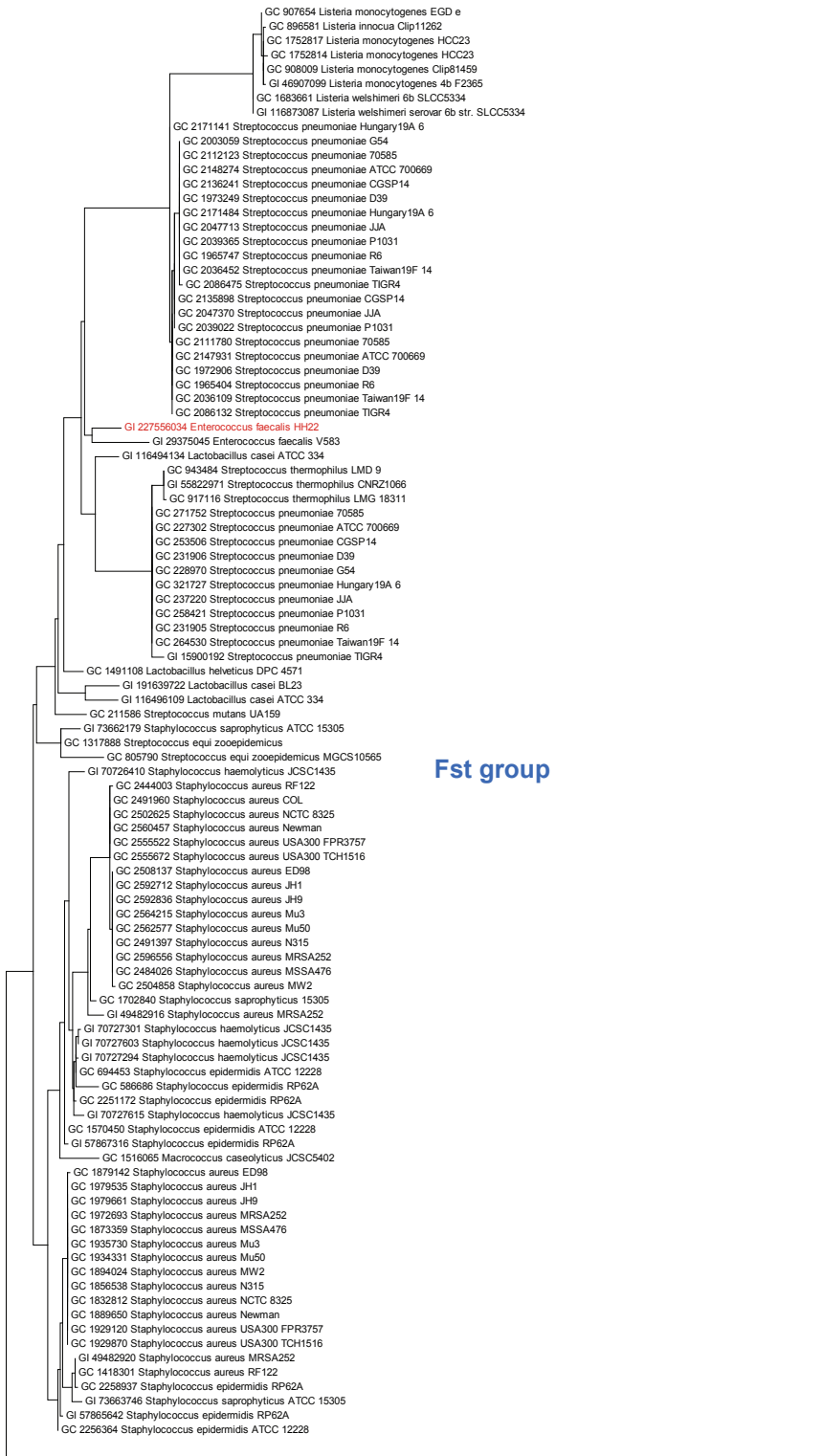
MSLVDIAILILKLIVAALQLLDAVLKFLR	12857	12943	NC_009792	Citrobacter koseri ATCC BAA-895
MGAVDMVILILRLIVAILQLLDAVLKFLR	26581	26495	NC_009436	Enterobacter sp. 638
MSGVDLFIILILKLIVAVLQLLDAVLKFLR	3899304	3899390	NC_009778	Enterobacter sakazakii ATCC BAA-894
MSLVDIAILILKLIVAALQLLDAVLKFLR	27814	27728	NC_010468	Escherichia coli ATCC 8739
MSLVDIAILILKLIVAALQLLDAVLKFLR	4170173	4170259	NC_009801	Escherichia coli E24377A
MSLVDIAILILKLIVAALQLLDAVLKFLR	3883134	3883220	NC_009800	Escherichia coli HS
MNLVDIAILILKLIVAALQLLDAVLKFLR	3949160	3949246	NC_010473	Escherichia coli str. K-12 substr. DH10B
MNLVDIAILILKLIVAALQLLDAVLKFLR	3851576	3851662	NC_000913	Escherichia coli str. K-12 substr. MG1655
MNLVDIAILILKLIVAALQLLDAVLKFLR	3786862	3786776	AC_000091	Escherichia coli str. K-12 substr. W3110
MSLVDIAILILKLIVAALQLLDAVLKFLR	4644028	4644114	NC_002695	Escherichia coli O157:H7 str. Sakai
MSLVDIAILILKLIVAALQLLDAVLKFLR	4713011	4713097	NC_002655	Escherichia coli O157:H7 EDL933
MSLVDIAILILKLIVAALQLLDAVLKFLR	4130354	4130440	NC_010498	Escherichia coli SMS-3-5
MSLVDIAILILKLIVAVLQLLDAVLKFLR	4073697	4073783	NC_011740	Escherichia fergusonii ATCC 35469
MGGMDIILILKLMVAVLQLLDAVLKQFR	26378	26301	NC_011283	Klebsiella pneumoniae 342
MGGMDIILILKLMVAVLQLLDAVLKQFR	4464305	4464382	NC_009648	Klebsiella pneumoniae subsp. pneumoniae MGH 78578
MSVVDITILILKLIVAALQLLDAVLKFLR	3783539	3783453	NC_010067	Salmonella enterica subsp. arizonae serovar 62:z4,z23:--
MRVVDITILILKLIVAALQLLDAVLKFLR	3938099	3938185	NC_006905	Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
MSVVDITILILKLIVAALQLLDAVLKFLR	3786643	3786729	NC_006511	Salmonella enterica serovar Paratyphi A str. ATCC 9150
MSVVDITILILKLIVAALQLLDAVLKFLR	3922945	3923031	NC_011149	Salmonella enterica subsp. enterica serovar Agona str. SL483
MSVVDITILILKLIVAALQLLDAVLKFLR	4031471	4031557	NC_011205	Salmonella enterica subsp. enterica serovar Dublin str. CT_02021853
MSVVDITILILKLIVAALQLLDAVLKFLR	3825787	3825701	NC_011274	Salmonella enterica subsp. enterica serovar Gallinarum str. 287/91
MSVVDITILILKLIVAALQLLDAVLKFLR	4012653	4012739	NC_011083	Salmonella enterica subsp. enterica serovar Heidelberg str. SL476
MSVVDITILILKLIVAALQLLDAVLKFLR	3969028	3969114	NC_011080	Salmonella enterica subsp. enterica serovar Newport str. SL254
MSVVDITILILKLIVAALQLLDAVLKFLR	3782322	3782408	NC_011147	Salmonella enterica enterica serovar Paratyphi A str. AKU_12601
MSVVDITILILKLIVAALQLLDAVLKFLR	3946617	3946703	NC_010102	Salmonella enterica subsp. enterica serovar Paratyphi B str. SPB7
MSVVDITILILKLIVAALQLLDAVLKFLR	3880235	3880321	NC_011094	Salmonella enterica enterica serovar Schwarzengrund str. CVM19633
MSVVDITILILKLIVAALQLLDAVLKFLR	3835774	3835688	NC_004631	Salmonella enterica subsp. enterica serovar Typhi Ty2
MSVVDITILILKLIVAALQLLDAVLKFLR	3850284	3850198	NC_003198	Salmonella enterica subsp. enterica serovar Typhi str. CT18
MSVVDITILILKLIVAALQLLDAVLKFLR	3998444	3998530	NC_003197	Salmonella typhimurium LT2
MSLVDIAILILKLIVAALQLLDAVLKFLR	3968862	3968776	NC_010658	Shigella boydii CDC 3083-94
MSLVDIAILILKLIVAALQLLDAVLKFLR	3712508	3712422	NC_007613	Shigella boydii Sb227
MSLVDIAILILKLIVAALQLLDAVLKFLR	3891354	3891440	NC_007606	Shigella dysenteriae Sd197
MSLVDIAILILKLIVAALQLLDAVLKFLR	3875416	3875502	NC_004741	Shigella flexneri 2a str. 2457T
MSLVDIAILILKLIVAALQLLDAVLKFLR	3933281	3933195	NC_008258	Shigella flexneri 5 str. 8401
MSLVDIAILILKLIVAALQLLDAVLKFLR	3796569	3796655	NC_007384	Shigella sonnei Ss046
MSVVDITILILKLIVAALQLLDAVLKFLR	3932891	3932977	NC_012125	Salmonella entericaenterica serovar Paratyphi C strain RKS4594
MSVVDITILILKLIVAALQLLDAVLKFLR	3873101	3873187	NC_011294	Salmonella enterica subsp. enterica serovar Enteritidis str. P125109
MSLVDIAILILKLIVAALQLLDAVLKFLR	4336773	4336859	NC_011751	Escherichia coli UMN026
MSLVDIAILILKLIVAALQLLDAVLKFLR	4452494	4452580	NC_011750	Escherichia coli IAI39
MSLVDIAILILKLIVAALQLLDAVLKFLR	4228573	4228659	NC_011748	Escherichia coli 55989
MSLVDIAILILKLIVAALQLLDAVLKFLR	3924698	3924784	NC_011741	Escherichia coli IAI1
MSLVDIAILILKLIVAALQLLDAVLKFLR	4102427	4102513	NC_011415	Escherichia coli SE11
MSLVDIAILILKLIVAALQLLDAVLKFLR	4747132	4747218	NC_011353	Escherichia coli O157:H7 str. EC4115
MNLVDIAILILKLIVAALQLLDAVLKFLR	3739909	3739995	NC_012759	Escherichia coli BW2952
MGGMDIILILKLMVAVLQLLDAVLKQFR	5181738	5181815	NC_012731	Klebsiella pneumoniae NTUH-K2044

H

ShoB (all)

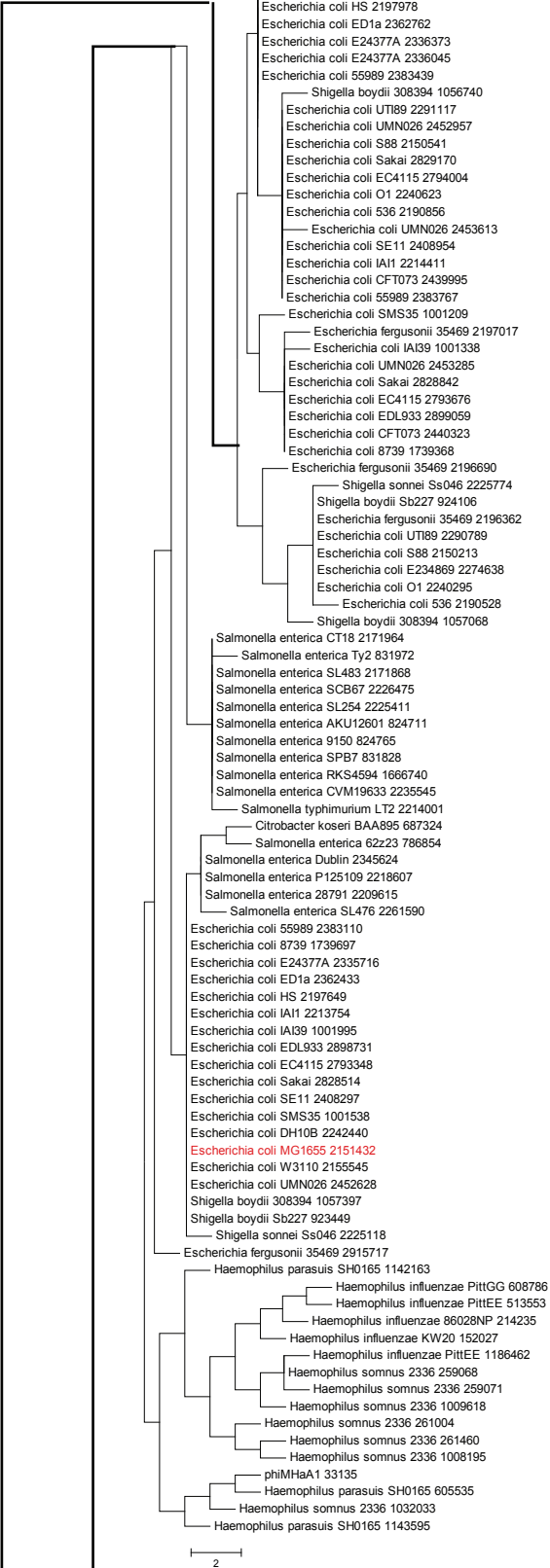
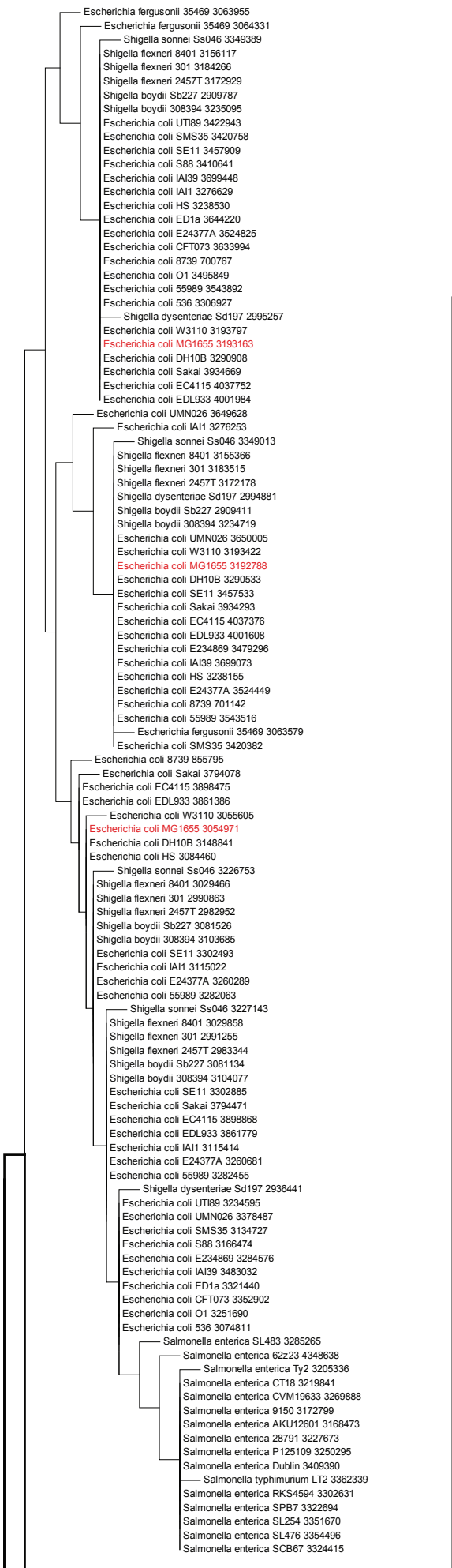
MTDCRYLIKRVIKIIIAVLQLLLFL	2705612	2705535	NC_008253	Escherichia coli 536
MTDCRYLIKRVIKIIIAVLQLLLFL	2854277	2854200	NC_008563	Escherichia coli APEC O1
MTDCRYLIKRVIKIIIAVLQLLLFL	1212726	1212803	NC_010468	Escherichia coli ATCC 8739
MTDCRYLIKRVIKIIIAVLQLLLFL	2963064	2962987	NC_004431	Escherichia coli CFT073
MTDCRYLIKRVIKIIIAVLQLLLFL	2850000	2849923	NC_009801	Escherichia coli E24377A
MTDCRYLIKRAIKIIIAVLQLLLFL	2738533	2738456	NC_009800	Escherichia coli HS
MTDCRYLIKRVIKIIIAVLQLLLFL	2789984	2789907	NC_010473	Escherichia coli str. K-12 substr. DH10B
MTDCRYLIKRVIKIIIAVLQLLLFL	2698219	2698142	NC_000913	Escherichia coli str. K-12 substr. MG1655
MTDCRYLIKRVIKIIIAVLQLLLFL	2698853	2698776	AC_000091	Escherichia coli str. K-12 substr. W3110
MTDCRYLIKRVIKIIIAVLQLLLFL	3420580	3420503	NC_002695	Escherichia coli O157:H7 str. Sakai
MTDCRYLIKRVIKIIIAVLQLLLFL	3490419	3490342	NC_002655	Escherichia coli O157:H7 EDL933
MTDCRYLIKRVIKIIIAVLQLLLFL	2795418	2795341	NC_010498	Escherichia coli SMS-3-5
MTDCRYLIKRVIKIIIAVLQLLLFL	2841944	2841867	NC_007946	Escherichia coli UTI89
MTDCRYLIKRVIKVVIIVLQLLLFL	511367	511444	NC_011740	Escherichia fergusonii ATCC 35469
MTDCRYLIKRVIKIIIAVLQLLLFL	2735311	2735234	NC_010658	Shigella boydii CDC 3083-94
MTDCRYLIKRVIKIIIAVLQLLLFL	2597276	2597199	NC_007613	Shigella boydii Sb227
MTDCRYLIKRAIKIIIAVLQLLLFL	2594263	2594186	NC_007606	Shigella dysenteriae Sd197
MTDCRYLIKRVIKIIIAVLQLLLFL	2697413	2697336	NC_004337	Shigella flexneri 2a str. 301
MTDCRYLIKRVIKIIIAVLQLLLFL	2690857	2690780	NC_004741	Shigella flexneri 2a str. 2457T
MTDCRYLIKRVIKIIIAVLQLLLFL	2693893	2693816	NC_008258	Shigella flexneri 5 str. 8401
MTDCRYLIKRVIKIIIAVLQLLLFL	2845461	2845384	NC_007384	Shigella sonnei Ss046
MTDCRYLIKRVIKIIIAVLQLLLFL	2584024	2583947	NC_012759	Escherichia coli BW2952
MTDCRYLIKRVIKIIIAVLQLLLFL	2991718	2991641	NC_011751	Escherichia coli UMN026
MTDCRYLIKRVIKIIIAVLQLLLFL	2874716	2874639	NC_011750	Escherichia coli IAI39
MTDCRYLIKRVIKIIIAVLQLLLFL	2923434	2923357	NC_011748	Escherichia coli 55989
MTDCRYLIKRVIKIIIAVLQLLLFL	2765974	2765897	NC_011741	Escherichia coli IAI1
MTDCRYLIKRVIKIIIAVLQLLLFL	2958999	2958922	NC_011415	Escherichia coli SE11
MTDCRYLIKRVIKIIIAVLQLLLFL	3524697	3524620	NC_011353	Escherichia coli O157:H7 str. EC4115
MTDCRYLIKRVIKIIIAVLQLLLFL	2957223	2957146	NC_011745	Escherichia coli ED1a
MTDCRYLIKRVIKIIIAVLQLLLFL	2737604	2737527	NC_011742	Escherichia coli S88
MTDCRYLIKRVIKIIIAVLQLLLFL	2929861	2929784	NC_011601	Escherichia coli O127:H6 str. E2348/69

Supplementary Figure S2. Phylogenetic tree of the LdrD/Fst toxin family. Maximum likelihood (ML) phylogenetic tree was constructed from an alignment of non-identical by using the MOLPHY program (2) with the JTT substitution matrix to perform local rearrangement of an original Fitch tree (3). The sequences that are identical to those used for tree reconstruction were added to the tree manually and have zero branch length. Known representatives from the Ldr family of *E. coli* K12 and the Fst family of *E. faecalis* are highlighted.



GI - genbank identifier of protein
 GC - genome coordinate

Supplementary Figure S3. Phylogenetic tree of the Ibs toxin family. Maximum likelihood (ML) phylogenetic tree was constructed from an alignment of non-identical by using the MOLPHY program (2) with the JTT substitution matrix to perform local rearrangement of an original Fitch tree (3). The sequences that are identical to those used for tree reconstruction were added to the tree manually and have zero branch length. Known representatives from Ibs family of *E.coli* K12 are highlighted.



Supplementary Figure S4. Mapping of *z3289*-sRNA-1 and *z3290*-sRNA-2 toxin-antitoxin genes of *E. coli* O157:H7 EDL933. (A) Sequence of *yehI* and *yehL* intergenic region. Nucleotide coordinates are given on the left. The coding sequences of the toxins are colored in red and the sRNA sequences are colored in blue. The regions of complementarity between the toxins and antitoxin RNAs are shaded (light gray for *z3289*-sRNA-1 and dark gray for *z3290*-sRNA-2). The transcription start sites are indicated by the +1 and the corresponding -10 and -35 sequences are boxed. (B) Northern analysis detecting the expression of the *z3289* and *z3290* mRNAs. Total RNA (10 µg) isolated from *E. coli* O157:H7 EDL933 cells grown to OD₆₀₀ ≈ 0.4 (E) and OD₆₀₀ ≈ 5.0 (overnight, S) in LB medium and from cells grown to OD₆₀₀ ≈ 0.4 (E) and OD₆₀₀ ≈ 2.5 (overnight, S) in M9 media supplemented with 0.2% glucose were loaded in each lane. The labeled oligonucleotide used in the northern analysis cannot distinguish between the two highly homologous mRNAs. (C) Primer extension analysis of *z3289* and sRNA-1/sRNA-2. Primer extension assays were performed using total RNA isolated from *E. coli* O157:H7 cells grown in either LB media to OD₆₀₀ ≈ 0.4 (E) and OD₆₀₀ ≈ 5 (overnight, S) or in M9 media supplemented with 0.2% glucose to OD₆₀₀ ≈ 0.4 (E) and OD₆₀₀ ≈ 2.5 (overnight, S). Given the high degree of homology between the two sRNAs, we are unable to distinguish them by this assay.

Supplementary Figure S4

A

yehK stop

2953244 aaccaagaaatagaataaataactctagataattaaagtcgcagcacatgcaacttgaag
ttggttctttatcttattatatgagatctattaatttcagcgtcgtgtacgttgaacttc

2953304 tatgacgagtatagctttcttaaaggggaaacgggtattcccccttagatgcgtttcaat
atactgctcatatcgaaagaatttcccccttgccataagggggaatctacgcaaagtta

2953364 ttgccggttacttcaacaaatcaatcaaccgtaacagagccactaataactccagtacgg
aacgggc

2953424 caatgagcacggttaacttttgtgtcagcgtgtccatccgctctccttattaaggagcgc
gttactcgtgccaatgaaaacacagtcgcacaggtaggcgagaggaataattcctcgcg
z3289 start

2953484 aacaccagcaccttacccttactgcctgtcgatactggtttacgtatgttggtacgaa
ttgtggtcgtggaatgggaatgtgacggacagctatgaccaaatgcatacaaccatgctt

2953544 acgttgctctccggatcagggcaccgcctgcggcaccaccggtccggcagccaggactta
tgcaacgagaggcctagtcctcgtggcggacgccgtggtgggcaggccggtcggctcctgaat

2953604 aaattggtccgatccggcacaacctccaacgggcaggattataaatctcactcactgaa
tttaaccaggctaggccgtggtggagggttgcccgctcctaaatattagagtgagtgactt
+1 z3289 -10

2953664 ⁻³⁵ ttgcaatggagtgatgatggttataat⁻¹⁰ttattc^{+1 sRNA-1}gttggtacgaaatggtgcacaatca
aacgttacctcacacctataccaatattaataagcaaccatgctttacaacgtgtagt
₋₃₅

2953724 gcaagttagcgtcactgtagtaataaaagaacctgcctcaccagcaggtttttttat
cgttcaatcgcagtgacatcattttttttcttggacggagtggtcgtccaaaaataaa

2953784 tctgcatctactttccagatatttttttgccaatagcccaccataaatattcacatccc
agacgctagatgaaaggctataaaaaaacggttatcgggtggtatttataagttaggg

2953844 ggtcaaattggtgcggtgaatataaaaatcgatgaatttataataaaaatgagttatatga
ccagtttaccacgcaacttatatttttagctacttaaatatttatttttactcaatatact

2953904 tgaattaataaaaacagtaatcattttattaaagggggaagcggttttcccccttagatg
acttaattatttttgtcattagtaaaataatttcccccttcgccaaaagggggaatctac

2953964 cgtttcaatttgcccgttacttcaacaaatcaatcaaccgtaacagagctactaatagct
gcaaagttaaacgggcaatgaagttggttagttggtgattgtctcgatgattatcga

2954024 ccagtacagcaatgagcacggttaacttttgtgtcagcgagtccatccgctctccttatt
ggatcatgtcgttactcgtgccaattgaaaacacagtcgctcaggtaggcgagaggaataa
z3290 start

2954084 aaggagcgcacaccagcaccttaccttacactgcctgtcgatactggtttacgtatgt
ttcctcgcggtgtggctcgtggaatgggaatgtgacggacagctatgaccaaatagcataca

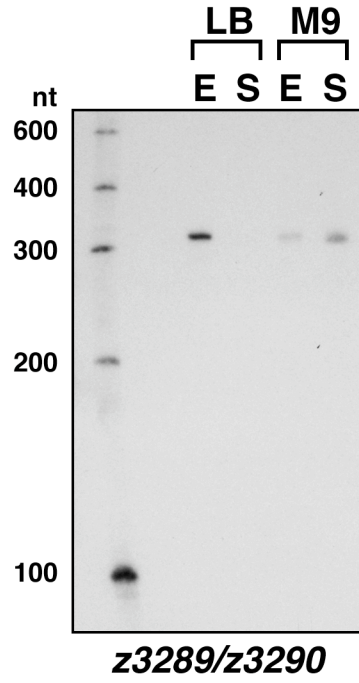
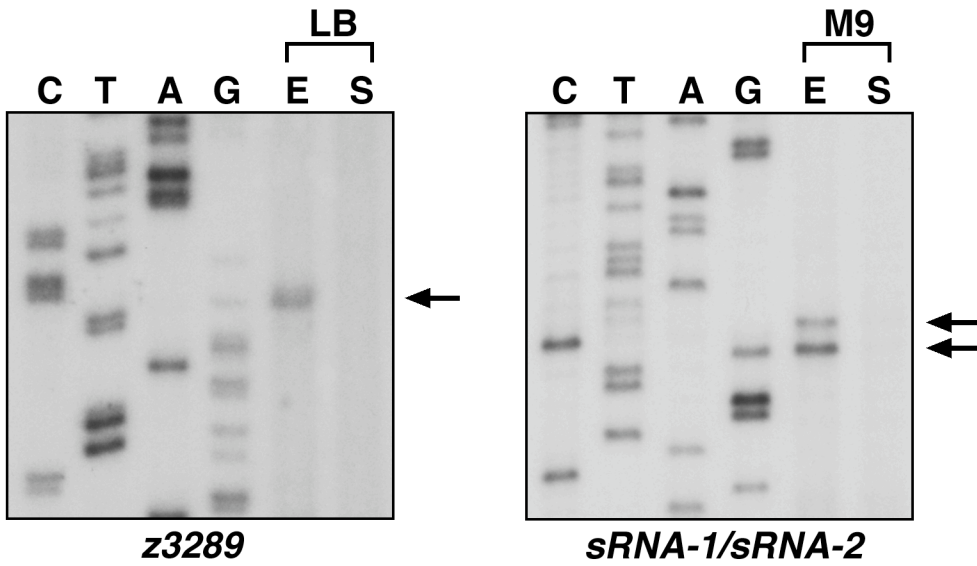
2954144 tggaccacgatggtgcactccggatcagggcaccacctgcagcaccctccggttagc
acctggtgctacaacgtgaggcctagtcctcgtggtagctcgtggggggcaggccatcg

2954204 caggacttgaaataggttcgatccggcgcacacgtcccaacagtcaggattataaatctca
gtcctgaactttatccaagctaggccgcggtgcagggttgtcagtcctaatatttagagt
+1 z3290 -10

2954264 ttatcttaattgcaacgggggtccagccgtggtataatcgttcgttgggaacacgatggt
aatagaattaacgttgccccaggctcggcaccaatattaagcaagcaaccttgtgctacaa
-35 -10 +1 SRNA-2
-35

2954324 gcacaggctgtggtgtaggcctgaaaatagtaataaaagaacctgcctcaccagcagg
cgtgtccgacaccacaatccggacttttatcattttttcttgacggagtggtcgtcc

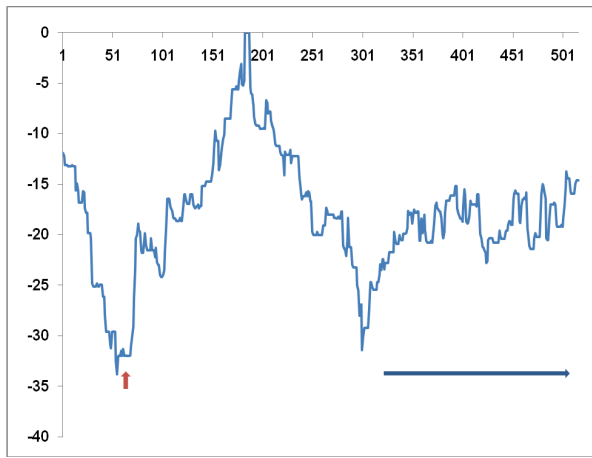
2954384 tttttttatctactgtgatctgcttcccagatatttttcgctcaacaactaa
aaaaaataaatgacactagacgaagggtctataaaaagcgagtttggtgatt

B**C**

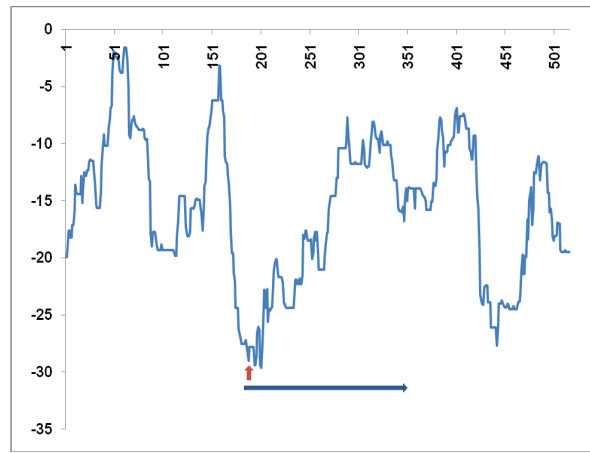
Supplementary Figure S5. Free energy plot analysis for experimentally characterized type I toxins of different families. Free energy profiles for RNA local secondary structures along nucleotide sequences in previously tested RNA antitoxin systems in *E.coli*, including the *E. coli* R plasmid (A-E) and *B.subtilis* (F). X axis: nucleotide positions; Y axis: free energy of RNA folding. Window lengths for free energy estimations corresponded to the lengths of RNA antitoxins in the system. Red arrows show the location of sRNA antitoxins.

Supplementary Figure S5

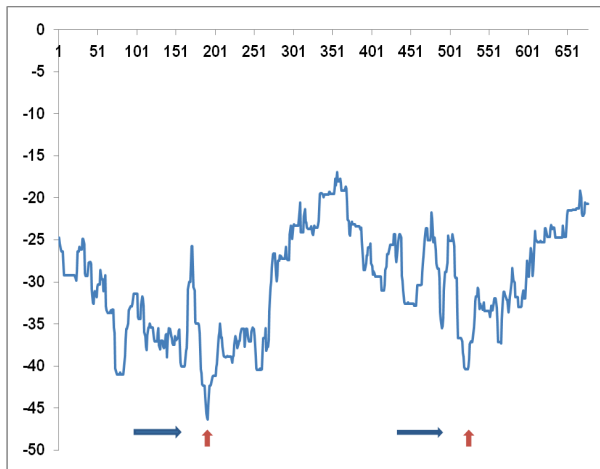
A *E. coli* TisB



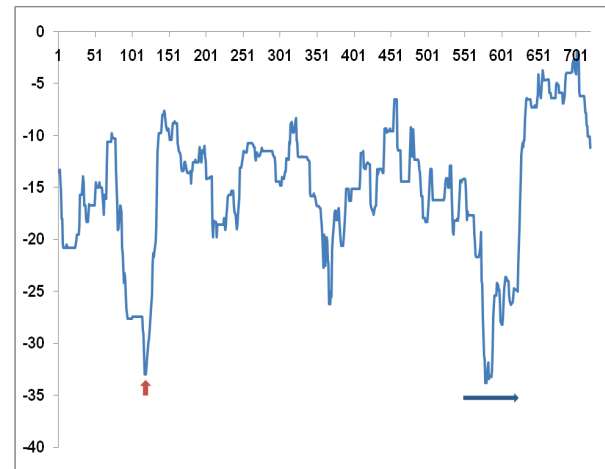
B R plasmid Hok



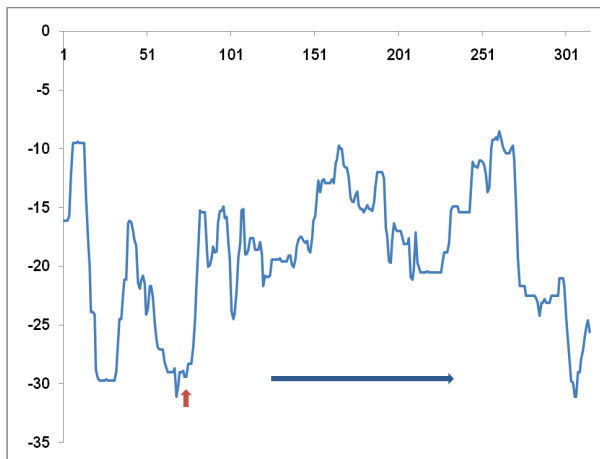
C *E. coli* lbsA, lbsB



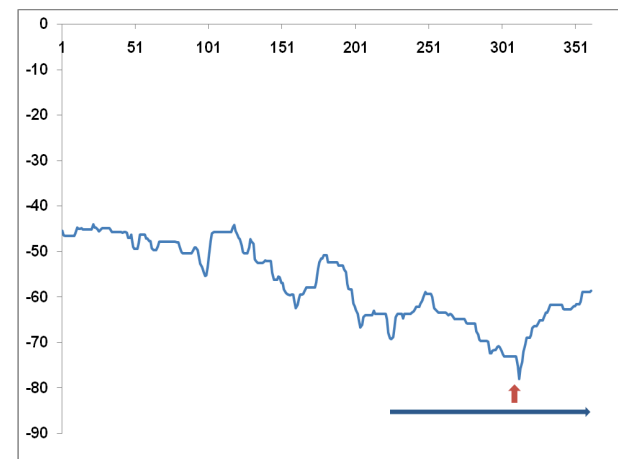
D *E. coli* ShoB



E *E. coli* LdrD

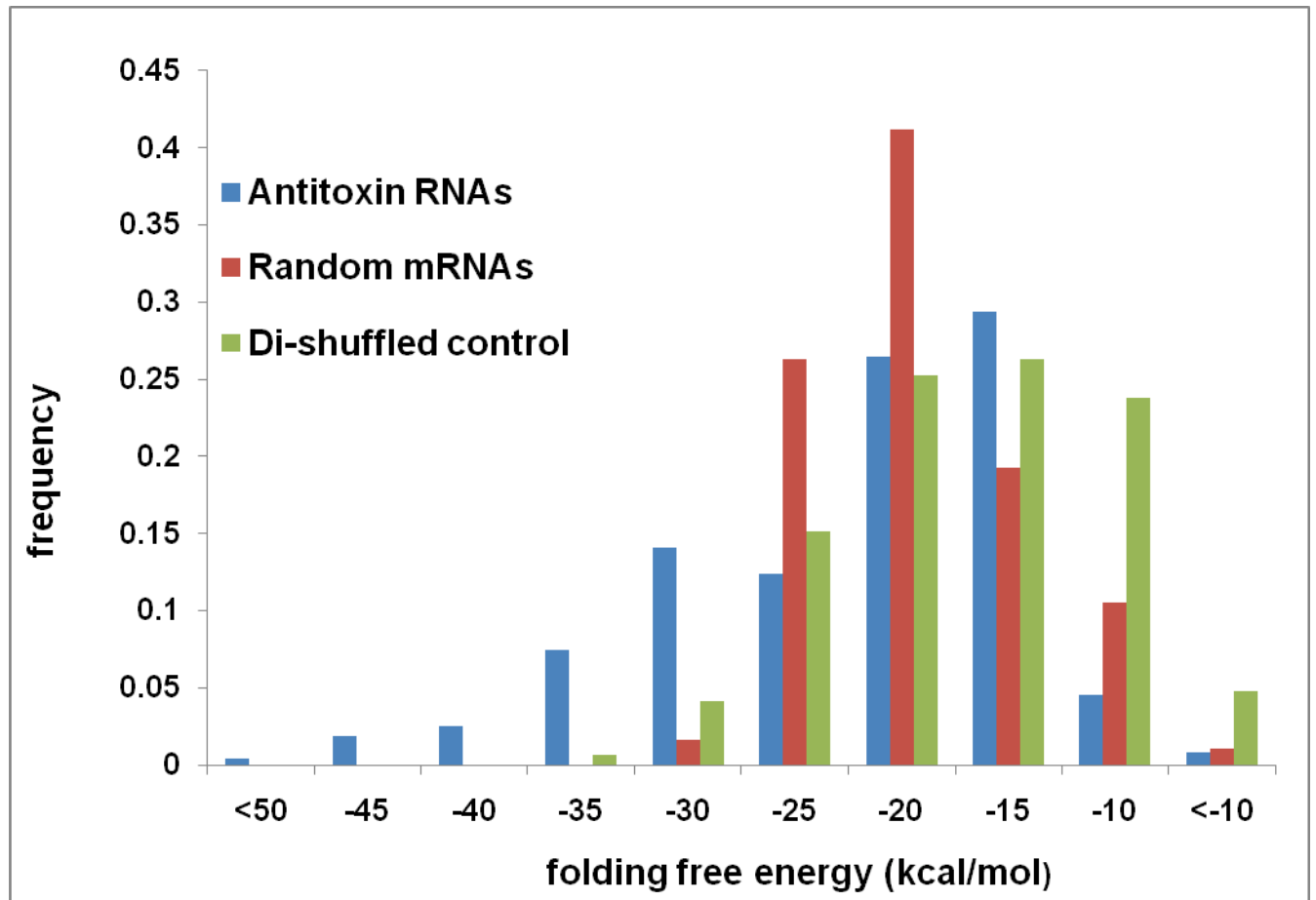


F *B. subtilis* TxpA



Supplementary Figure S6. Comparison of the free energy distributions for the antisense RNAs of type I toxins loci predicted in this study with random controls. Folding free energy estimations for the putative antisense RNAs, control sequences from the same genomes and randomly shuffled dinucleotide sequences were calculated using the Afold program (4). Di-shuffled controls are sequences with randomly shuffled dinucleotides, retaining nucleotide composition of native RNAs.

Supplementary Figure S6



REFERENCES

1. Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.
2. Adachi, J. and Hasegawa, M. (1992), *Computer Science Monographs No. 27*, Institute of Statistical Mathematics, Tokyo.
3. Felsenstein, J. (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol.*, **266**, 418-427.
4. Ogurtsov, A.Y., Shabalina, S.A., Kondrashov, A.S. and Roytberg, M.A. (2006) Analysis of internal loops within the RNA secondary structure in almost quadratic time. *Bioinformatics*, **22**, 1317-1324.