

Table S3: Top results generated with AnDom webserver and HH-suite for chosen candidates of *S. aureus*. SACOL0051, SACOL0052, SACOL0238, SACOL0239 and SACOL0694 show distant homology to known proteins with regulatory domains.

S. aureus candidates			AnDom 2.0 webserver				HH-suite				Additional information
Identifier	Description	Sequence	1. result	2. result	3. result	4. result	1. result	2. result	3. result	4. result	Regulatory domains
SA1642	Extracellular enterotoxin type G precursor	5 10 15 20 25 30 35 40 45 50 55 60 MKKLSSTVIIILLIIVFHMNMVUNAPDPKKDELNKVSDYKNGKVTGMGVNMLTSPFVE 65 70 75 80 85 90 95 100 105 110 115 120 GRGVINSRQFLSHDLIFPIEYKSNEVTELETELANNVKKKVDIFGVFPYFTTCIIPK 125 130 135 140 145 150 155 160 165 170 175 180 SEPDINQFPGGCCMYGGLFPMSSNERDKLITVQVITDNKQSLGFTTITNKMNVTIQELD 185 190 195 200 205 210 215 220 225 230 235 240 YKARHMLTEKKLVYEPDGSAPESGYIKTEKNNTSFWFLDPKKELVPPVPVKPLIYIGD 245 250 255 NRVVDSSKIKMEVFLMTH	SCOP: b.40.2.2 Query: 25 - 139 / 258 Template: 2 - 119 / 121 E-value = 1e-26	SCOP: d.15.6.1 Query: 146 - 256 / 258 Template: 2 - 109 / 109 E-value = 1e-24	SCOP: b.40.2.2 Query: 141 - 257 / 258 Template: 1 - 115 / 117 E-value = 2e-23	SCOP: b.40.2.2 Query: 147 - 257 / 258 Template: 1 - 112 / 113 E-value = 6e-22	Query: 1 - 255 / 258 Template: 10 - 247 / 261 E-value = 1.9e-75	Query: 47 - 256 / 258 Template: 30 - 237 / 247 E-value = 1.3e-51	Query: 87 - 256 / 258 Template: 1 - 165 / 167 E-value = 2e-50	Query: 24 - 235 / 258 Template: 21 - 220 / 220 E-value = 9e-50	No homologous regulatory domains found
SA1643	Enterotoxin SeN	5 10 15 20 25 30 35 40 45 50 55 60 MKNKIKMLRFLYIAAIIITLCLLNNNVNVDKDLKKKSDLSDSKFLNFTSYTDTDI 65 70 75 80 85 90 95 100 105 110 115 120 WQLDESNIKISTDQLLNTIILKNIDISVLKTSLSKVFNSDLANQFKKNIDIVGLYFG 125 130 135 140 145 150 155 160 165 170 175 180 NKCVGLTEKTSICLYGGVVIDNQGLDEEKVIGVNVFDDVGQSGFVIKTKAKVTVQEL 185 190 195 200 205 210 215 220 225 230 235 240 DTKVRFLKLENLYIKNDGNIQGCCIFFPHSHHQDQSFYDILNVKSGVGAEPFQYDSD 245 250 255 NRTVSSNVHIDVFLYKD	SCOP: d.15.6.1 Query: 147 - 255 / 258 Template: 2 - 108 / 109 E-value = 6e-24	SCOP: d.15.6.1 Query: 142 - 257 / 258 Template: 1 - 115 / 117 E-value = 8e-23	SCOP: d.15.6.1 Query: 148 - 257 / 258 Template: 1 - 112 / 113 E-value = 5e-22	SCOP: b.40.2.2 Query: 50 - 146 / 258 Template: 2 - 119 / 120 E-value = 5e-22	Query: 7 - 258 / 258 Template: 9 - 252 / 261 E-value = 2.5e-74	Query: 7 - 255 / 258 Template: 30 - 237 / 247 E-value = 1.1e-56	Query: 94 - 258 / 258 Template: 1 - 167 / 167 E-value = 7.9e-53	Query: 14 - 257 / 258 Template: 15 - 262 / 262 E-value = 9.2e-53	No homologous regulatory domains found
SA1644	Enterotoxin YENT2	5 10 15 20 25 30 35 40 45 50 55 60 MKTKTCMYGGVTEHGDQIKDNKSTDNSNLIILKVVYERNRSLSFDIPTTKKKNTAQEYDI 65 70 75 80 85 90 95 100 105 110 115 120 KVRNLLKKNLYEPMSSPYETGYIKTEGSGHSFWDLMPESGKKFYPTKLLIYDNKN 125 130 135 TVESKSNVEVHLTKK	SCOP: d.15.6.1 Query: 21 - 136 / 136 Template: 2 - 116 / 117 E-value = 2e-25	SCOP: d.15.6.1 Query: 25 - 134 / 136 Template: 2 - 109 / 109 E-value = 8e-23	SCOP: d.15.6.1 Query: 27 - 135 / 136 Template: 2 - 112 / 113 E-value = 5e-22	SCOP: d.15.6.1 Query: 25 - 134 / 136 Template: 4 - 111 / 112 E-value = 5e-22	Query: 2 - 134 / 136 Template: 125-248 / 261 E-value = 4.6e-46	Query: 2 - 136 / 136 Template: 37-167 / 167 E-value = 9.1e-45	Query: 3 - 134 / 136 Template: 99-237 / 247 E-value = 6.7e-40	Query: 1 - 136 / 136 Template: 1-136 / 136 E-value = 1e-39	No homologous regulatory domains found
SA1645	Enterotoxin Yent1	5 10 15 20 25 30 35 40 45 50 55 60 MKLFAFIPICVKSCLFLPMANGPKPEQIHKASPTGLMDNMRYLDDKRVSETNIKSPQE 65 70 75 80 85 90 95 100 105 110 115 120 KFLQHDLLFKINGSKILKTEPFNNKLSLSQYKXKNVLDLFGNTYNGCYFSLDNMELNDGRL 125 130 IEKNVYVVRCDRA	SCOP: b.40.2.2 Query: 23 - 125 / 133 Template: 5 - 112 / 121 E-value = 4e-26	SCOP: b.40.2.2 Query: 29 - 113 / 133 Template: 2 - 88 / 100 E-value = 5e-17	SCOP: b.40.2.2 Query: 21 - 106 / 133 Template: 4 - 96 / 120 E-value = 8e-17	SCOP: b.40.2.2 Query: 39 - 106 / 133 Template: 3 - 70 / 83 E-value = 5e-15	Query: 1 - 112 / 133 Template: 10 - 142 / 261 E-value = 1.1e-42	Query: 1 - 133 / 133 Template: 10 - 142 / 156 E-value = 8.4e-40	Query: 19 - 133 / 133 Template: 1 - 115 / 115 E-value = 7.3e-38	Query: 6 - 113 / 133 Template: 17 - 129 / 214 E-value = 2.4e-36	No homologous regulatory domains found
SA1648	Enterotoxin SeO	5 10 15 20 25 30 35 40 45 50 55 60 MKNSKVMNLVLLILMLIAICSVNNAYANEEDPKIESLCKSSVDPJALHNINDDYINNKR 65 70 75 80 85 90 95 100 105 110 115 120 FTTVKSIVSTTEKFLDPLLFSKINMLDQISAEFKDLKVEFSSSAISKEFLGKTVDIYVG 125 130 135 140 145 150 155 160 165 170 175 180 YYKACSBGHRGVDTACTYGGVTFPEHNNKLSPEKNIGVAVYKDNVNVVITIVTTDKRVKVA 185 190 195 200 205 210 215 220 225 230 235 240 QELDLKIVTKLNNAYKLVDMNTSDVQKGYIKFHSSEHKESFYDLYTIKGNLPQGYLQI 245 250 255 YNDNKTIDSSDHYHIDVLYFT	SCOP: d.15.6.1 Query: 145 - 260 / 260 Template: 1 - 115 / 117 E-value = 8e-26	SCOP: d.15.6.1 Query: 151 - 258 / 260 Template: 3 - 108 / 109 E-value = 8e-26	SCOP: b.40.2.2 Query: 29 - 150 / 260 Template: 3 - 120 / 120 E-value = 3e-25	SCOP: d.15.6.1 Query: 151 - 260 / 260 Template: 1 - 112 / 113 E-value = 2e-24	Query: 25 - 260 / 260 Template: 25 - 252 / 261 E-value = 1.2e-71	Query: 1 - 260 / 260 Template: 3 - 262 / 262 E-value = 7.2e-63	Query: 97 - 260 / 260 Template: 1 - 166 / 167 E-value = 1.4e-52	Query: 47 - 258 / 260 Template: 31 - 262 / 247 E-value = 5.4e-52	No homologous regulatory domains found
SACOL0051	Hypothetical protein	5 10 15 20 25 30 35 40 45 50 55 60 MEFEHKEKLISEVNNKTEINNVVFFSLGKSSVKAQVKLLKNTYLIKODISKALKFKKK 65 70 75 80 85 90 95 100 105 110 115 120 SGEFPPEIKLIDVSTSEKILFKELKTLINTRNRYVDGFIAPDSQHNFAVLPEEINANAF 125 130 135 140 145 150 155 160 165 170 175 180 VRPDNTTEKLPFLSKKNINMYLKRVTNNKASFPSEYNEKEVIFKPYTGQFPGDEBEVHELY 185 190 195 200 205 210 215 220 225 230 235 240 SEGYKGLKRVLDNWLNEIDQLIESSTFWLQNMLLDNGKVIYGYFPFDFNEIGFVYVLRHS 245 250 255 260 265 270 275 280 285 290 295 300 SSTYALIEGLSYGKSLQPVKEADIVYILMQLFEIGDKAYIFDDTEEANEIKLGQNASFI 305 310 315 320 325 330 335 340 345 350 355 360 FAVCEYLKHEDNPKFLSESAQKVAKGILSMIDETTYETHLLNYPDLVSKEKFRIIYDGE 365 370 375 380 385 390 395 400 405 410 415 420 ALALALLKQDBEMLKMTYVEMLDORSEKKTWYBDRHSCYCNELWYVHQDYKTE 425 430 435 440 445 450 455 460 465 470 475 480 GKKNVNYLDYIKNRETTFFPLEMNAIYRLVQKADTGREELVNNLIDEQYLIDVNI 485 490 495 500 505 510 515 520 525 530 535 RADYQRVGFPEIAMIYFNKPSRLISGFFIKHHGVRVRIDIEHYLSGYVQYQLAFNR	No structural similarity found				Query: 1 - 538 / 538 Template: 8 - 541 / 552 E-value = 9.1e-130	Query: 1 - 537 / 538 Template: 11 - 546 / 585 E-value = 5.3e-117	Query: 22 - 537 / 538 Template: 60 - 581 / 1067 E-value = 5.3e-117	Query: 20 - 537 / 538 Template: 27 - 549 / 1044 E-value = 1.3e-99	26. result Query: 223 - 404 / 538 Template: 431 - 618 / 716 E-value = 1.3e-23
SACOL0052	Glycosyl transferase, group 1 family protein; identified by match to protein family HMM PF00534	5 10 15 20 25 30 35 40 45 50 55 60 MIYSVGRNLGNKLIGIEKAMINRHNRLFKNSITSKLIASFASVSPRLHFNASLQIEDDDLL 65 70 75 80 85 90 95 100 105 110 115 120 SLYDVLGDSINIKSIHKDWITHTWTECHNYTLKFIENTNDIKIYKSNRYMVFHNDNSYQ 125 130 135 140 145 150 155 160 165 170 175 180 HLDIINHFVNRKIRKIDILNSRQSLSCSLTLTHQKVVCSEIYTFPGGIRKQKFNPEL 185 190 195 200 205 210 215 220 225 230 235 240 ENSGQILIIYNTDSQKYFNNDQELLAPAEIKLYKNIDGIFLSDKNINIAPIFNNTSETIP 245 250 255 260 265 270 275 280 285 290 295 300 VLAVLHSTHVKNIDMVESDIKNTYKRVFNMLNRYSGIIVSTQGGQLDISARINNEIPVH 305 310 315 320 325 330 335 340 345 350 355 360 TIPVYIDEHFNPKNNHNSNNMKIISVARYSPKQLNMQELVSKLIKFEPPHIGLRLY 365 370 375 380 385 390 395 400 405 410 415 420 GPGKEEKYKQLIYEVNLNENVFLRPGKRNLSAEIQDAYNSLITSNMBGFWLGLLETITE 425 430 435 440 445 450 455 460 465 470 475 480 GIPPVGVNSKYGPSLELNNNGYLINKNDKELNVRNLLLDKTLRDTFSQECIKHKS 485 490 495 500 AFSSKIVMKLWEDRFSSTINSSYE	SCOP: c.87.1.8 Query: 329 - 483 / 503 Template: 201 - 355 / 370 E-value = 9e-07	Lipopolysaccharide core biosynthesis protein RfaG { <i>Escherichia coli</i> [Taxid: 562]}	No further structural similarity found		Query: 1 - 499 / 503 Template: 3 - 505 / 550 E-value = 6.8e-91	Poly(Glycerol-phosphate) alpha-glycosyltransferase; glycosyl transferase; lipopolysaccharide biosynthesis protein...	Poly(Glycerol-phosphate) alpha-glycosyltransferase; glycosyl transferase; lipopolysaccharide biosynthesis protein...	Mur ligase middle domain-containing protein; mur ligase middle domain protein.	Six-hairpin glycosidase; transcriptional regulator; glycoside hydrolase family 76; transmembrane protein; glt214; thymidylate kinase.
SACOL0238	Teichoic acid biosynthesis protein, putative; identified by match to protein family HMM PF04464	5 10 15 20 25 30 35 40 45 50 55 60 MKTKQAIHIDNIYERVQLYIEGHSEGDVLTSGQFVLRNLTEKTLLEANMKIDGNTFI 65 70 75 80 85 90 95 100 105 110 115 120 CRFNVALDDGYLLPMKDYLFVYHDQLEYIQGLNPNIIDQAYALNEEQIEEYNELTQNI 125 130 135 140 145 150 155 160 165 170 175 180 KGVNVLAYDAKVRKGGVSGHVITITPEIASDVNEVFVDIEITLPQKRSQVIAISAHN 185 190 195 200 205 210 215 220 225 230 235 240 LKQGHKASFEPSRPLFKAIPIKTLKLIHKSRTILFTSDSRPKLSGNPKYVVDLRLQK 245 250 255 260 265 270 275 280 285 290 295 300 VDFDVIDKTVFKENIDRRKWRDKFLPYLLGKADYIFVDDFHLIYTVFRPQSQEIYQV 305 310 315 320 325 330 335 340 345 350 355 360 WHAVGAFKTVGFSRTGKKGGPFIIDSLNHRSTYKAVSSETDIPFYAEAPGIREENVVPTG 365 370 375 380 385 390 395 400 405 410 415 420 VPRTDVLFDEATATQIKQEMEDLPIIKGKVLIPAPFTFGNGHGTANPPFKLIDPELA 425 430 435 440 445 450 455 460 465 470 475 480 RYCEKHNAVVLFKMHPFVKNRLMSREHRQYFIDVSDHREVNDILFVTDLLISDYSSLIY 485 490 495 500 505 510 515 520 525 530 535 540 EYAVFKPMIFYADLEDYIITRDFYEPFESFVPGKIVQSFALMDALDNEDEYEVKVPV 545 550 555 560 FLDKHFKYQDGSRERLVKDLFRR	No structural similarity found				Query: 1 - 564 / 564 Template: 1 - 562 / 565 E-value = 2.2e-125	Glycosyl/ glycerophosphate transferase; CDP-glycerol: poly(glycerophosphate) glycerophospho-transferase; teichoic acid biosynthesis protein B...	Teichoic acid biosynthesis protein B; teichoic acid biosynthesis family protein.	CDP-glycerol:poly-glycerophosphate transferase; glycosyl transferase, group 2 family.	Teichoic acid biosynthesis protein, putative.
SACOL0239	TagF domain-containing protein; identified by match to protein family HMM PF04464	5 10 15 20 25 30 35 40 45 50 55 60 MIKNTIKKLIHSIYITFKLLSKLPKNKLIYFESPHGQYSDNPKALYEYLTEHSDAGLI 65 70 75 80 85 90 95 100 105 110 115 120 NGVKVGYEHIPOQNVYPTVKFSKMFILAMPRAKAMINTRTPDWLTKSPKTTLTQWNG 125 130 135 140 145 150 155 160 165 170 175 180 ETPLKIGLDISNUVLGTQNTQYQDQPKESQRMDLYVSPNPYSTSIQNAFHYVSRDKIL 185 190 195 200 205 210 215 220 225 230 235 240 TGYPRNDKLSHKNDTEYINGIKTRLNIPLDKKVIMYAPTWRDEAIRREGSYQFNVNF 245 250 255 260 265 270 275 280 285 290 295 300 IEALQALDDYVILLRMHLYLVTRIDEDHDPVKDVSDEYIDSLYILISDALVTDYSSVM 305 310 315 320 325 330 335 340 345 350 355 360 PFDGVAERKQIFIALDKYDELRGFDMDIKKELPGQIVENQALIDALKQIDETANEY 365 370 375 380 385 IEARTVYFQKCSLEDQASRIQCTIFK	No structural similarity found				Query: 1 - 389 / 389 Template: 53 - 433 / 441 E-value = 1.9e-57	Teichoic acid biosynthesis protein F; TagF protein; TagF domain-containing protein; glycosyl transferase...	teichoic acid biosynthesis protein F; CDP-glycerol:Poly(glycerophosphate) glycerophospho-transferase family...	TagF domain protein; teichoic acid biosynthesis protein; CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase.	Polyglycerol phosphate polymerase; Major teichoic acid biosynthesis protein F; CGPTase; CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase.
SACOL0694	Teichoic acids export protein ATP-binding subunit; with TagG is involved in the export of teichoic acids	5 10 15 20 25 30 35 40 45 50 55 60 MNVSVNKKVTKETRYNNKREMKDALIPHKNNKTFPALDDISLKAYEGDVGILGVGIG 65 70 75 80 85 90 95 100 105 110 115 120 SKGKSTLNIIGGSLSPYQKVQVNRNGESVIAISAGLSGQLTGIEINIEFHLKMGCFKRKEI									