

Table 1 Predicted ORFs of S-ESS1 genome and their presumed functions (identified by online BLASTP)

| ORF | Strand | Start | Stop | size(aa) | Predicted protein | Related phage(s) or microbes | E-value (<0.001) |
|-----|--------|-------|-------|----------|-----------------------------------|---|------------------|
| 3 | + | 4476 | 5417 | 313 | hypothetical protein | | 2.00E-36 |
| 8 | + | 13824 | 15515 | 563 | phage portal protein | Bradyrhizobium sp,Ruegeria phage DSS3-P1,Labrenzia,Xylella phage Sano,Enterobacter phage Enc34,Proteus phage pPM_01,Salmonella phage Chi],Salmonella phage iEPS5] | 0 |
| 11 | + | 20010 | 20597 | 195 | minor tail protein | Burkholderia phage AH2 | 2.00E-22 |
| 12 | + | 22686 | 24635 | 649 | phage tape measure protein | Sinorhizobium meliloti,Bradyrhizobium sp. WSM3983,Loktanela phage pCB2051-A,Burkholderia phage AH2,Asticcacaulis sp,Salmonella phage FSL,Salmonella phage | 6.00E-74 |
| 13 | + | 27024 | 28217 | 397 | hypothetical protein DSS3P1_47 | | 4.00E-85 |
| 16 | + | 30444 | 31400 | 318 | tail assembly protein | Caldimonas manganoxidans,Providencia phage Redjac,Gulbenkiania mobilis,Pseudomonas aeruginosa,Hydrogenophaga sp | 6.00E-32 |
| 34 | + | 5605 | 6294 | 226 | DUF2815 domain-containing protein | Klebsiella variicola,Enterobacter aerogenes,Serratia marcescens | 7.00E-16 |
| 35 | + | 6637 | 8607 | 656 | DNA polymerase I | Silicibacter phage DSS3-P1, Ruegeria phage DSS3-P1,Burkholderia phage,Enterobacter phage,Providencia phage,Xylella phage,Salmonella phage, | 1.00E-165 |
| 36 | + | 9274 | 9990 | 238 | DEAD/DEAH box helicase | Lactobacillus fermentum,Megamonas funiformis,Listeria monocytogenes,Anaerovibrio lipolyticus | 3.00E-25 |
| 37 | + | 10558 | 11463 | 301 | putative terminase small subunit | Ruegeria phage DSS3-P1, Burkholderia phage BcepNazgul,Xylella phage Salvo,Enterobacter phage Enc34,Burkholderia phage AH2,Proteus phage,Achromobacter phage | 4.00E-26 |
| 38 | + | 13591 | 13824 | 77 | head-tail joining protein | Labrenzia sp,Ruegeria phage,Achromobacter phage,Agrobacterium,Loktanela phage,Enterobacter phage | 1.00E-14 |
| 40 | + | 17695 | 18429 | 244 | minor capsid protein E | Mesorhizobium plurifarum,Bradyrhizobium sp,Labrenzia,Burkholderia phage,Agrobacterium | 2.00E-79 |
| 42 | + | 18913 | 19305 | 130 | hypothetical protein | | 1.00E-11 |
| 45 | + | 24577 | 25962 | 461 | tail tape measure protein | Ruegeria phage DSS3, Sinorhizobium ,Burkholderia phage,Roseivivax isoporaе,Rhizobium sp | 2.00E-67 |
| 46 | + | 26731 | 27024 | 97 | tail length tape measure protein | Loktanela phage | 2.00E-20 |
| 50 | + | 29323 | 29523 | 66 | hypothetical protein DSS3P1_44 | | 2.00E-15 |
| 51 | + | 31489 | 31857 | 122 | tail assembly structural protein | Pseudomonas phage,Vibrio phage | 7.00E-13 |
| 73 | + | 8429 | 8752 | 107 | DNA polymerase | Ruegeria phage,Silicibacter phage,Burkholderia phage,Xylella phage,Selenomonas sp | 2.00E-25 |
| 74 | + | 8825 | 9277 | 150 | hypothetical protein | | 2.00E-13 |
| 75 | + | 9791 | 10768 | 325 | SNF2-related protein | | 2.00E-95 |

Continued Table 1

| ORF | Strand | Start | Stop | size(aa) | Predicted protein | Related phage(s) or microbes | E-value (<0.001) |
|-----|--------|-------|-------|----------|---|---|------------------|
| 76 | + | 11432 | 13582 | 716 | terminase large subunit | Ruegeria phage DSS3-P1, Loktanella phage pCB2051-A, Salmonella phage FSLSP088, Pseudomonas aeruginosa | 0.00E+00 |
| 77 | + | 15512 | 17323 | 603 | peptidase S49 | Sinorhizobium meliloti, Variovorax sp | 1.00E-102 |
| 79 | + | 19307 | 19924 | 205 | hypothetical protein | | 1.00E-43 |
| 80 | + | 19925 | 20230 | 101 | minor tail protein | Salmonella phage | 1.00E-04 |
| 81 | + | 20483 | 21262 | 259 | tail length tape measure protein | Loktanella phage pCB2051-A | 3.00E-49 |
| 86 | + | 25676 | 26578 | 300 | tail length tape measure protein | Loktanella phage pCB2051-A | 1.00E-26 |
| 87 | + | 28217 | 29065 | 281 | hypothetical protein DSS3P1_46 | | 1.00E-90 |
| 88 | + | 29075 | 29317 | 80 | tail assembly structural protein | Pseudomonas phage MP1412 | 2.00E-09 |
| 89 | + | 29507 | 30208 | 233 | putative tail protein | Ruegeria phage DSS3-P1, Loktanella phage pCB2051-A, Achromobacter phage phiAxp-2, Pseudomonas phage SM1 | 2.00E-09 |
| 91 | + | 32159 | 32863 | 234 | putative tail protein | Ruegeria phage DSS3-P1, Loktanella phage pCB2051-A, Achromobacter phage phiAxp-2, Agrobacterium | 7.00E-59 |
| 92 | + | 32879 | 33484 | 201 | hypothetical protein DSS3P1 | | 1.00E-32 |
| 94 | + | 36113 | 37294 | 393 | Peptidoglycan-binding domain 1 protein | Sinorhizobium meliloti AK83 | 7.00E-92 |
| 106 | + | 48173 | 49624 | 483 | hypothetical protein ruthe_00792 | | 3.00E-14 |
| 107 | + | 49625 | 49939 | 70 | hypothetical protein | | 5.00E-12 |
| 108 | - | 48149 | 49999 | 616 | ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent | Sulfitobacter donghicola, Sulfitobacter, | 0.00E+00 |
| 111 | - | 43406 | 44356 | 316 | FAD-dependent thymidylate synthase | Sulfitobacter donghicola, Sulfitobacter, Orientia tsutsugamushi | 8.00E-108 |
| 114 | - | 39899 | 40630 | 243 | DNA methyltransferase | Tetraselmis viridis virus, Geminicoccus roseus, Fodinicurvata fenggangensis | 2.00E-57 |
| 145 | - | 46012 | 47130 | 372 | DUF932 domain-containing protein | Rhizobium sp. YK2, Pelagibacterium sp, Bradyrhizobium elkanii | 2.00E-99 |
| 148 | - | 42988 | 43416 | 142 | dihydrofolate reductase | Rhizobium phage vB_RleM_P10VF, Methylobacterium, Firmicutes bacterium, Escherichia coli | 1.00E-11 |
| 150 | - | 40975 | 41283 | 102 | hypothetical protein | | 4.00E-36 |
| 160 | - | 29131 | 29382 | 83 | Uncharacterized conserved protein | Janthinobacterium sp | 2.00E-04 |
| 179 | - | 3382 | 3633 | 83 | Cro/C1 family transcriptional regulator | Acinetobacter, Moraxella bovoculi, Proteus mirabilis | 6.00E-04 |

Continued Table 1

| ORF | Strand | Start | Stop | size(aa) | Predicted protein | Related phage(s) or microbes | E-value (<0.001) |
|-----|--------|-------|-------|----------|---|---|------------------|
| 180 | — | 2341 | 3324 | 327 | Phage associated DNA primase | Sulfitobacter geojensis, | 1.00E-109 |
| 190 | — | 41871 | 42311 | 146 | DUF3310 domain-containing protein | Pseudovibrio sp | 1.00E-10 |
| 194 | — | 39678 | 39980 | 100 | hypothetical protein P10VF_021 | | 7.00E-05 |
| 195 | — | 38661 | 39167 | 168 | collagen triple helix repeat family protein | Haemophilus parasuis | 1.00E-04 |
| 218 | — | 483 | 2492 | 669 | Phage associated DNA primase | Sulfitobacter geojensis] | 0.00E+00 |
| 256 | — | 56643 | 56963 | 106 | hypothetical protein | Actinobacteria,Bradyrhizobium retamae,uncultured Mediterranean phage uvMED,Synechococcus phage P60,uncultured Mediterranean phage uvMED | 3.00E-11 |
| 260 | — | 53370 | 53597 | 75 | hypothetical protein | Pseudomonas stutzeri,Yersinia aldovae,Alistipes putredinis,Pseudomonas aeruginosa,Bacteroides oleiciplenus | 8.00E-10 |
| 261 | — | 52485 | 53030 | 181 | hypothetical protein | Pseudoruegeria sabulilitoris,Methylosinus,Leptospirillum ferriphilum | 6.00E-14 |
| 262 | — | 52011 | 52295 | 94 | hypothetical protein | [Sinorhizobium meliloti,Enterobacter cloacae,Yersinia pseudotuberculosis | 7.00E-06 |
| 263 | — | 51285 | 51554 | 89 | hypothetical protein CAPSK01_001762 | Candidatus Accumulibacter sp. SK-01 | 2.00E-08 |
| 265 | — | 50001 | 50636 | 212 | ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent | Rhodobacteraceae bacterium CY02, Nereida ignava,Paenirhodobacter sp. MME-103,Tateyamaria sp. ANG-S1,Donghicola sp. JL3646 | 2.00E-108 |
| 274 | — | 57556 | 59061 | 501 | hypothetical protein | Caulobacter virus Karma,Caulobacter virus Magneto,Rhizobium phage RHEph01,Alistipes putredinis | 6.00E-20 |
| 275 | — | 56893 | 57282 | 129 | hypothetical protein | Bradyrhizobium retamae,Bradyrhizobium elkanii | 2.00E-19 |
| 276 | — | 55660 | 56646 | 328 | hypothetical protein | Agrobacterium,Caulobacter virus Magneto,Caulobacter virus Karma | 2.00E-16 |