**SUPPLEMENTAL MATERIAL**

**Whole genome sequencing and comparative genomics of two nematicidal *Bacillus* strains reveals a wide range of possible virulence factors**

Nik Susič, Sandra Janežič, Maja Rupnik, Barbara Gerič Stare

G3: Genes|Genomes|Genetics, 2020

**SUPPLEMENTARY RESULTS AND DISCUSSION**

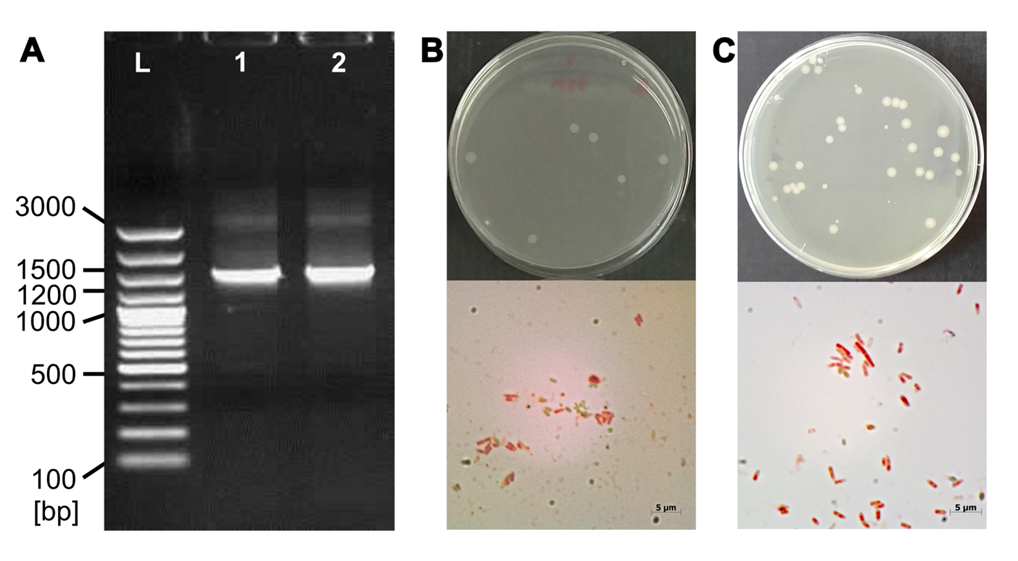
**Species identification**

Fragments of the 16S rRNA gene sequences were amplified in both strains (Supplementary Figure 1A) and the resulting PCR amplicons were sequenced. EzBioCloud 16S rRNA sequence analysis initially supported the *B. firmus* denomination of both strains (Supplementary Table 1). Tetra analysis was in accordance with 16S rRNA results, and a TCS search for related genomes returned *B. firmus* DS1 as the closest genome in the database to I-1582 and ZZV12-4809, with Z-scores of 0.99866 and 0.99709, respectively (Supplementary Table 2B,C). However, it is known that the method performs unsatisfactorily for whole genome phylogenetic tree reconstructions, due to the weakness of the phylogenetic signal within tetranucleotide usage patterns (Teeling *et al*. 2004). ANI values (%) calculated for *B. firmus* I-1582, *Bacillus* sp. ZZV12-4809 and 6 other *Bacillus* genome assemblies currently publicly available in GenBank (see Materials and Methods) showed that I-1582 and ZZV12-4809 shared 89.7/90.8% (ANIb/ANIm) identity with each other, as well as 97.4/98.1% and 89.9%/90.8% identity respectively with the type *B. firmus* strain NBRC 15306 (Supplementary Table 2). According to the strict 95% cut-off value to delineate species boundary (Richter and Rosselló-Móra 2009; Chun *et al*. 2018), of the following complete genomes in the public domain designated as *B. firmus*, only strains NBRC 15306, NCTC 10335 and I-1582 could be considered as *B. firmus*, while strains DS1 and LK28 showed transitional characteristics. Following this logic, ZZV12-4809 and 14\_TX do not belong to *B. firmus* or the same species, as they do not share high ANI (86.4/88%). On the other hand, strain NCTC 10335 shares virtually identical 99.9/100% ANI with the type *B. firmus* NBRC 15306, with >99% of aligned assembly nucleotides. Both assemblies are from different collections of type material, but due to their high level of identity the two might share a common origin (information on origin is lacking). Phylogenetic analysis of the *B. firmus* group supported ANI calculations. *B. firmus* 14\_TX was the most distant from the other *B. firmus* accessions (slightly more distant than *B. oceanisediminis* 2691) and formed an outgroup with *B. oceanisediminis* 2691, due to high identity value between the two (Figure 1). Since the identity of ZZV12-4809 could not be confirmed as *B. firmus*, we performed phylogenetic analysis within the genus *Bacillus* to discern the phylogenetic position of this strain. The results, accompanied by the ANI values, showed that ZZV12-4089 was still most closely related to the *B. firmus* group (Supplementary Figure 2).ANI calculations and phylogenetic analysis indicated previously unreported variability within the strains of *B. firmus* that points to incorrect species circumscription in some strains. Therefore, description of a novel species could be warranted. Based on these results, strain ZZV12-4809 could not be considered as *B. firmus,* so we assigned it a provisional designation as *Bacillus* sp. ZZV12-4809.

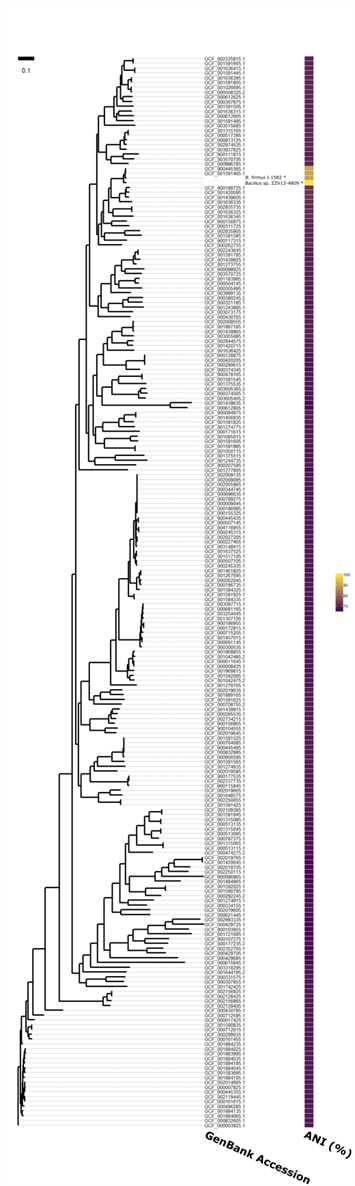
**Comparative genomics**

Comparisons of orthologous clusters can provide information relevant for taxonomic classification and phylogeny (Wang *et al*. 2015). The strain-specific clusters that were found in *B. firmus* I-1582 and *Bacillus* sp. ZZV12-4809 could potentially be involved in important biological processes, so affiliations across GO terms were checked for differences between the strains. In I-1582, more than 50% of these genes were involved in ion binding (GO:0043167; 14.3%), nucleic acid binding (GO:0003676; 14.3%), nucleotide binding (GO:0000166; 12.2%), binding (GO:0005488; 10.2%) and nucleoside binding (GO:0001882; 10.2%) processes, while in ZZV12-4809 more than 50% of strain-specific clusters were involved in transferase activity (GO:0016740; 13.8%), ion binding (GO:0043167; 12.6%), oxidoreductase activity (GO:0016491; 10.3%), transporter activity (GO:0005215; 9.2%) and hydrolase activity (GO:0016787; 9.2%). Additionally, 4.6% of ZZV12-4809-specific clusters were not assigned a GO term, so the function of these remained unclear.

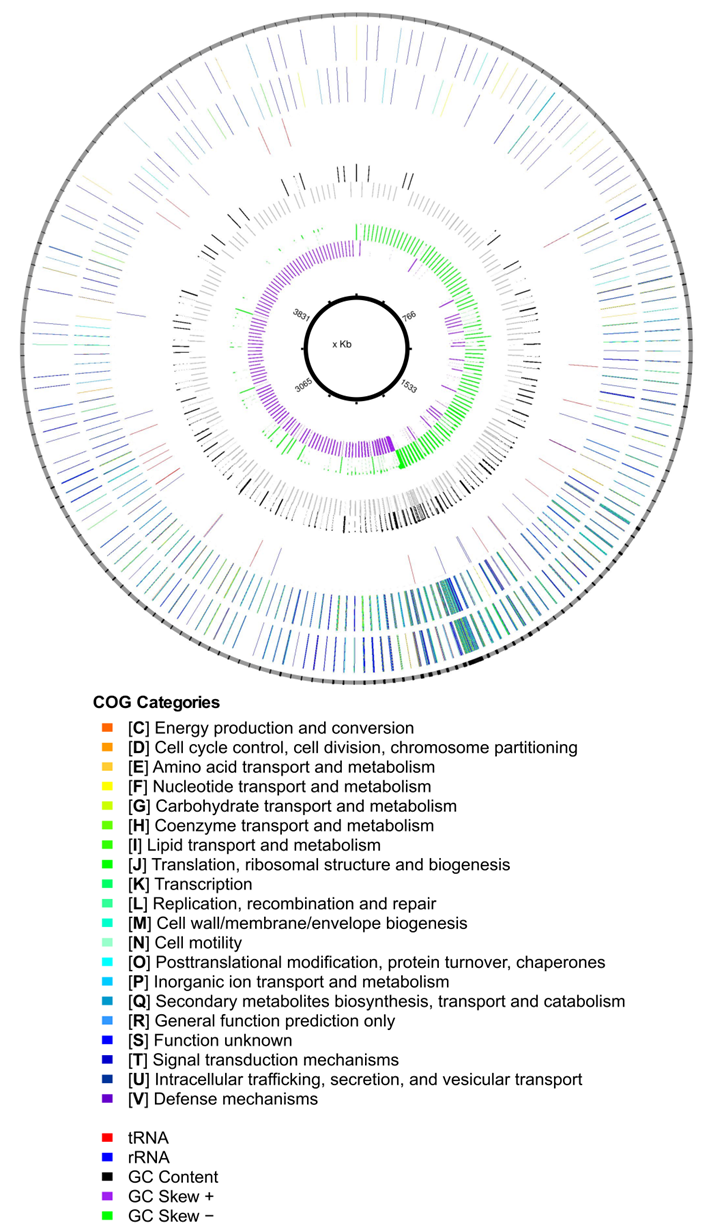
**SUPPLEMENTARY FIGURES**



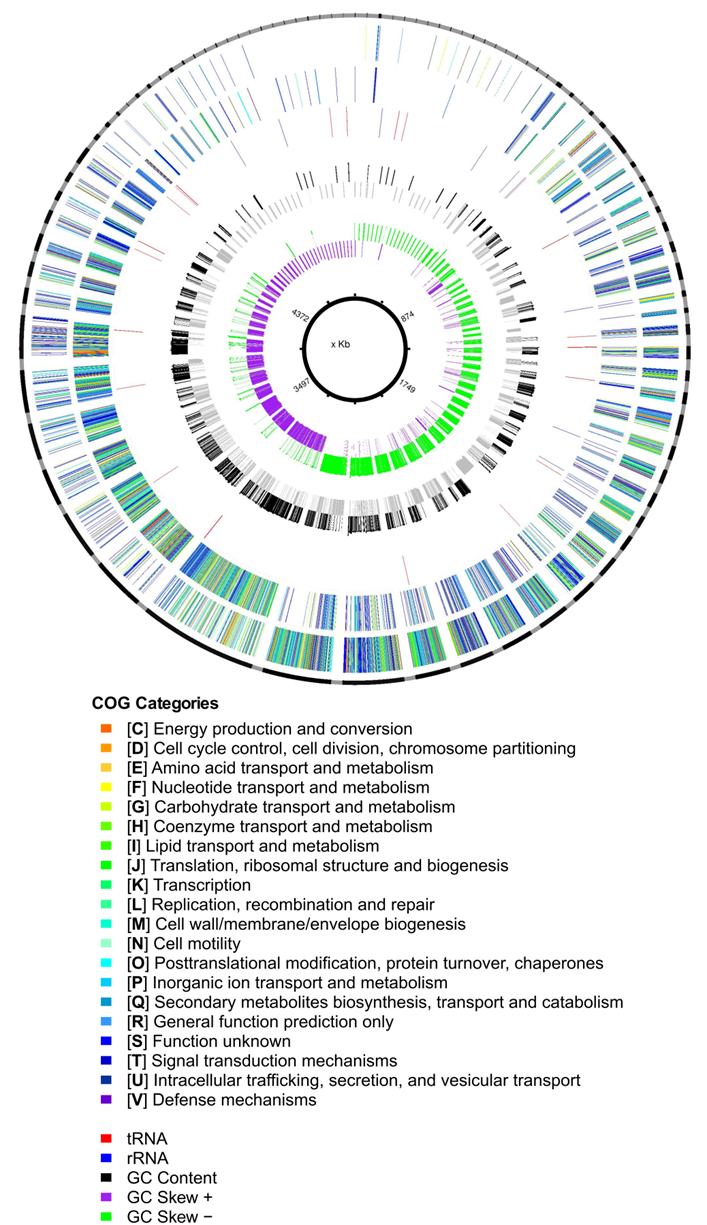
**Supplementary Figure 1.** Characteristics of the two *Bacillus* strains studied. A) PCR amplicons of the 16S rRNA gene sequence from the two *Bacillus* strains: L – DNA ladder GeneRuler™ 100 bp Plus (Fermentas); 1 – *Bacillus firmus* I-1582; and 2 – *Bacillus* sp. ZZV12-4809. Colony morphology and differential spore stain in B) strain I-1582 and C) strain ZZV12-4809, with spores stained green and vegetative cells in red.

****

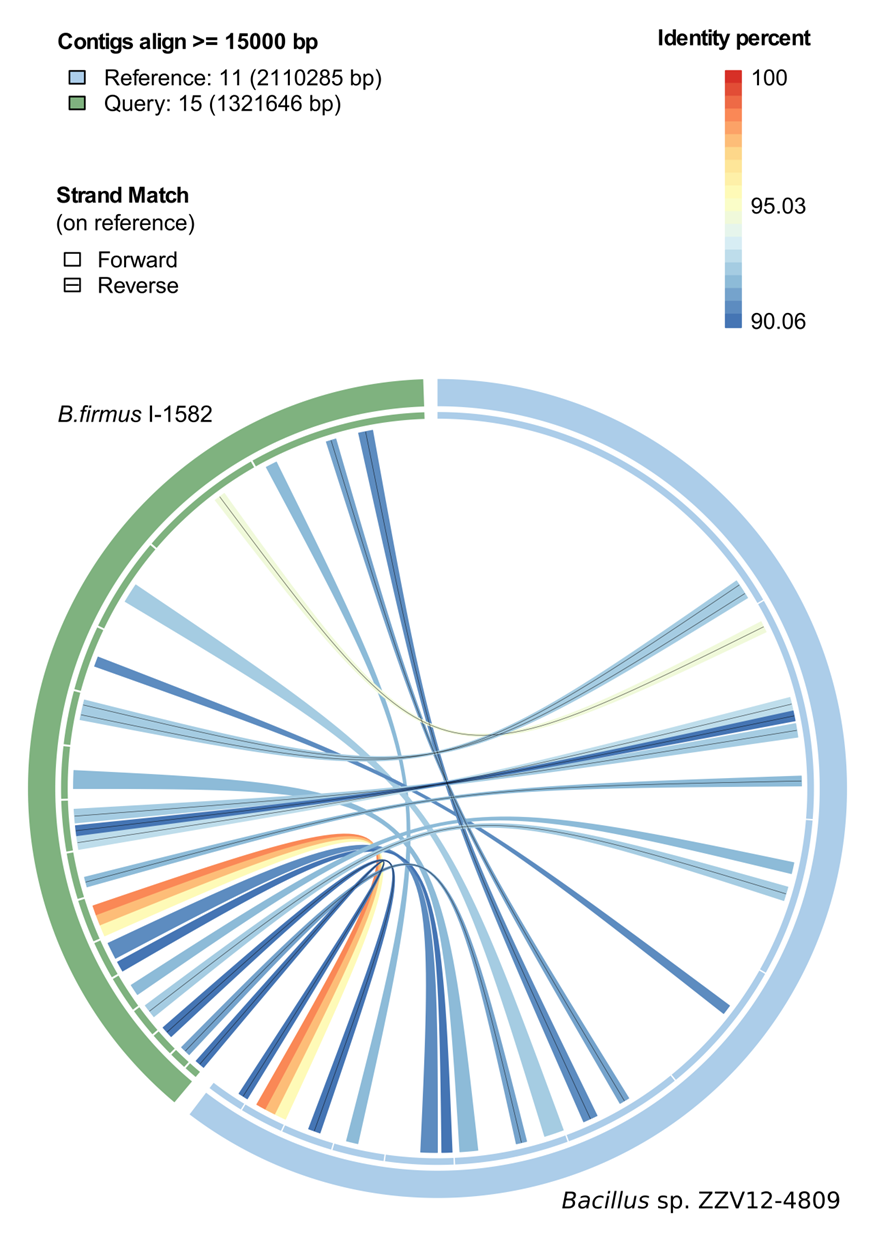
**Supplementary Figure 2.** Maximum likelihood phylogenetic tree of the genus *Bacillus*, constructed from 222 available genome assemblies in GenBank from *Bacillus* type strains. The tree was annotated with ANI (%) values for all assemblies calculated against *Bacillus* sp. ZZV12-4809, which showed the highest identity only with members of the *B. firmus* group. The *Bacillus* strains presented in this study are highlighted with asterisks (\*). For a list of all the assembly accessions see Supplementary Table 3.

****

**Supplementary Figure 3.** Circular map of the *Bacillus firmus* I-1582 genome assembly. Circle tracks from the center to the outside of both maps represent the following features: assembly size (kbp), GC-skew (violet/green), GC-content (black/grey), rRNA sequences (blue), tRNA sequences (red), reverse- and forward-strand genes colored according to COG category, and assembled contigs (black).



**Supplementary Figure 4.** Circular map of the *Bacillus* sp. ZZV12-4809 genome assembly. Circle tracks from the center to the outside of both maps represent the following features: assembly size (kbp), GC-skew (violet/green), GC-content (black/grey), rRNA sequences (blue), tRNA sequences (red), reverse- and forward-strand genes colored according to COG category, and assembled contigs (black).



**Supplementary Figure 5.** Selected syntenic regions between the *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809 genome assemblies. Only high identity (> 90%) blocks longer than 15 kbp are shown.

**SUPPLEMENTARY TABLES**

**Supplementary Table 1.** Results of 16S rRNA sequence analysis based on a similarity search against quality-controlled databases of 16S rRNA sequences in the EzBioCloud Identify service (Yoon *et al*. 2017; available at: https://www.ezbiocloud.net/). Using ContEst16S (Lee *et al*. 2017), 16S rRNA gene sequences were extracted from the I-1582 and ZZV12-4809 genome assemblies and checked for possible contaminants. Extracted 16S rRNA sequences from strains I-1582 and ZZV12-4809 returned the best hit to the type strain *Bacillus firmus* NBRC 15306(T). Based solely on 16S rRNA, both strains could be denominated as *B. firmus*. Only the best hits to type strain sequences are shown. The sequence similarity cut-off for species delineation was 98.7%.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hit taxon name | Hit strain name | Accession (GenBank) | Similarity (%) | Diff./ total nt | Hit taxonomy | Comple-teness (%) |
| **I-1582** | | | | | | |
| *Bacillus firmus* | NBRC 15306(T) | BCUY01000205 | 100.00 | 0/1215 | Bacteria; Firmicutes; Bacilli;  Bacillales; Bacillaceae; Bacillus | 100.0 |
| *Bacillus oceanisediminis* | H2(T) | GQ292772 | 99.58 | 5/1183 | Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus | 94.7 |
| *Bacillus infantis* | NRRL B-14911 | CP006643 | 98.85 | 14/1215 | Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus | 100.0 |
| *Bacillus dakarensis* | Marseille-P3515(T) | LT707409 | 98.60 | 17/1215 | Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus | 100.0 |
| **ZZV12-4809** | | | | | | |
| *B. firmus* | NBRC 15306(T) | BCUY01000205 | 99.23 | 11/1427 | Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus | 100.0 |
| *B. oceanisediminis* | H2(T) | GQ292772 | 99.07 | 13/1393 | Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus | 94.7 |
| *B. infantis* | NRRL B-14911 | CP006643 | 97.97 | 29/1427 | Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus | 100.0 |

**Supplementary Table 2.** Pairwise genome comparisons of *Bacillus* strains based on a) average nucleotide identity (ANI%) and b) Tetra calculations for the two *B. firmus* I-1582 and *Bacillus* sp. ZZV12-4809 genomes and six other *Bacillus* sp. genomes obtained from GenBank (see Methods). Additionally, c) the TCS method was used to query genomes similar to I-1582 and ZZV12-4809 assemblies. Analysis was carried out using the JSpecies Web Server (Richter *et al*. 2017).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A) ANIb (ANIm) % | *B. firmus* 14\_TX | *B. firmus* DS1 | *B. firmus* LK28 | *B. firmus* NBRC 15306 | *B. firmus* NCTC 10335 | *B. oceanisediminis* 2691 | *B. firmus* I-1582 | *Bacillus* sp. ZZV12-4809 |
| *B. firmus* 14\_TX | \* | 88.14 (89.30) | 88.02 (89.03) | 88.26 (89.40) | 88.37 (89.41) | 93.72 (94.64) | 88.67 (89.49) | 86.35 (87.95) |
| *B. firmus* DS1 | 87.70  (89.31) | \* | 92.57 (93.65) | 93.58 (94.47) | 93.69 (94.46) | 88.30 (89.65) | 93.60 (94.36) | 88.88 (90.29) |
| *B. firmus* LK28 | 87.58 (89.03) | 92.53  (93.65) | \* | 93.85 (94.64) | 93.82 (94.65) | 87.88 (89.17) | 93.87 (94.57) | 88.98 (90.10) |
| *B. firmus* NBRC 15306 | 87.74 (89.40) | 93.55 (94.47) | 93.65 (94.64) | \* | **99.87 (99.99)** | 88.40 (89.75) | **97.45 (98.09)** | 89.55 (90.80) |
| *B. firmus* NCTC 10335 | 87.74 (89.41) | 93.54 (94.46) | 93.66 (94.64) | **100.00 (99.99)** | \* | 88.44 (89.84) | **97.48 (98.09)** | 89.54 (90.82) |
| *B. oceanisediminis* 2691 | 93.60 (94.64) | 88.54 (89.64) | 88.18 (89.17) | 88.64 (89.74) | 88.80 (89.84) | \* | 89.23 (90.00) | 86.45 (88.13) |
| *B. firmus* I-1582 | 87.96 (89.49) | 93.38 (94.37) | 93.66 (94.56) | **97.42 (98.09)** | **97.35 (98.09)** | 88.61 (90.01) | \* | 89.32 (90.76) |
| *Bacillus* sp. ZZV12-4809 | 86.20 (87.96) | 89.09 (90.29) | 89.26 (90.10) | 89.88 (90.81) | 89.93 (90.82) | 86.39 (88.13) | 89.73 (90.76) | \* |
| B) Tetra Z-score | *B. firmus* 14\_TX | *B. firmus* DS1 | *B. firmus* LK28 | *B. firmus* NBRC 15306 | *B. firmus* NCTC 10335 | *B. oceanisediminis* 2691 | *B. firmus* I-1582 | *B. firmus* ZZV12-4809 |
| *B. firmus* 14\_TX | 1.0 | 0.995 | 0.994 | 0.996 | 0.995 | **0.999** | 0.996 | 0.997 |
| *B. firmus* DS1 | 0.995 | 1.0 | 0.998 | 0.999 | 0.998 | 0.996 | 0.999 | 0.997 |
| *B. firmus* LK28 | 0.994 | 0.998 | 1.0 | 0.998 | 0.996 | 0.994 | 0.998 | 0.996 |
| *B. firmus* NBRC 15306 | 0.996 | 0.999 | 0.998 | 1.0 | 0.998 | 0.995 | **0.999** | 0.997 |
| *B. firmus* NCTC 10335 | 0.995 | 0.998 | 0.996 | 0.998 | 1.0 | 0.996 | 0.998 | 0.997 |
| *B. oceanisediminis* 2691 | **0.999** | 0.996 | 0.994 | 0.995 | 0.997 | 1.0 | 0.996 | 0.996 |
| *B. firmus* I-1582 | 0.996 | 0.999 | 0.998 | **0.999** | 0.998 | 0.996 | 1.0 | 0.998 |
| *Bacillus* sp. ZZV12-4809 | 0.997 | 0.997 | 0.996 | 0.997 | 0.997 | 0.996 | 0.998 | 1.0 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| C) TCS Z-score | Query genomes | | | |
| Results | ***B. firmus* I-1582** | Z-Score | ***Bacillus* sp. ZZV12-4809** | Z-Score |
| 1 | *Bacillus firmus* DS1 | 0.99866 | *Bacillus firmus* DS1 | 0.99709 |
| 2 | *Bacillus firmus* LK28 | 0.99813 | *Bacillus* sp. 2\_A\_57\_CT2 | 0.99681 |
| 3 | *Bacillus* sp. 2\_A\_57\_CT2 | 0.99599 | *Bacillus firmus* LK28 | 0.99643 |
| 4 | *Bacillus oceanisediminis* 2691 | 0.99563 | *Bacillus oceanisediminis* 2691 | 0.99635 |
| 5 | *Sporosarcina globispora* DSM 4 | 0.99054 | *Sporosarcina globispora* DSM 4 | 0.99104 |
| 6 | *Bacillus* sp. NRRL B-14911 | 0.94923 | *Bacillus* sp. SG-1 | 0.94413 |
| 7 | *Bacillus infantis* NRRL B-14911 | 0.94764 | *Bacillus* sp. NRRL B-14911 | 0.94244 |
| 8 | *Bacillus* sp. SG-1 | 0.94746 | *Bacillus infantis* NRRL B-14911 | 0.94055 |
| 9 | *Fictibacillus enclensis* NIO-1003 | 0.93927 | *Bacillus* sp. FJAT-27986 | 0.93723 |
| 10 | *Bacillus* sp. FJAT-27986 | 0.93403 | *Fictibacillus enclensis* NIO-1003 | 0.93189 |

**Supplementary Table 3.** List of 224 GenBank assembly accessions used in phylogenetic positioning of *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809 within the genus *Bacillus*, with corresponding ANI (%) values computed between each accession and *Bacillus* sp. ZZV12-4809 assembly.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Assembly accession | ANI (%)a | Assembly accession | ANI (%)a | Assembly accession | ANI (%)a |
| GCF\_000429725.1 | 65.9373 | GCF\_001457015.1 | 68.2396 | GCF\_001042475.2 | 68.5777 |
| GCF\_000429685.1 | 66.0743 | GCF\_003097715.1 | 68.2453 | GCF\_002009135.1 | 68.5865 |
| GCF\_001721685.1 | 66.2334 | GCF\_001461825.1 | 68.2475 | GCF\_001591985.1 | 68.5976 |
| GCF\_900096965.1 | 66.3842 | GCF\_000245315.1 | 68.2649 | GCF\_000307855.1 | 68.6013 |
| GCF\_001484965.1 | 66.4377 | GCF\_900186955.1 | 68.2786 | GCF\_001274775.1 | 68.6084 |
| GCF\_002019765.1 | 66.4879 | GCF\_000420205.1 | 68.2806 | GCF\_004116955.1 | 68.6123 |
| GCF\_001420645.1 | 66.5208 | GCF\_002019605.1 | 68.2931 | GCF\_900445355.1 | 68.6124 |
| GCF\_900103955.1 | 66.5782 | GCF\_003605405.2 | 68.3062 | GCF\_000003925.1 | 68.6146 |
| GCF\_000615945.1 | 66.5808 | GCF\_002014665.1 | 68.326 | GCF\_001274935.1 | 68.6148 |
| GCF\_002250115.1 | 66.5859 | GCF\_001742425.1 | 68.3415 | GCF\_000299035.1 | 68.6199 |
| GCF\_002019705.1 | 66.598 | GCF\_001583695.1 | 68.3521 | GCF\_002009095.1 | 68.6329 |
| GCF\_002993335.1 | 66.6687 | GCF\_001592005.1 | 68.3529 | GCF\_001884105.1 | 68.6342 |
| GCF\_000429705.1 | 66.8781 | GCF\_000300535.1 | 68.3533 | GCF\_001884185.1 | 68.6391 |
| GCF\_001439635.1 | 66.9752 | GCF\_002128405.1 | 68.3591 | GCF\_002128425.1 | 68.6401 |
| GCF\_001274915.1 | 67.0966 | GCF\_000011645.1 | 68.36 | GCF\_001884045.1 | 68.6412 |
| GCF\_900107275.1 | 67.1452 | GCF\_001267695.1 | 68.3651 | GCF\_000007825.1 | 68.6457 |
| GCF\_001277805.1 | 67.1523 | GCF\_900445435.1 | 68.3692 | GCF\_001591545.1 | 68.6567 |
| GCF\_002352765.1 | 67.1729 | GCF\_001042485.2 | 68.3695 | GCF\_003148415.1 | 68.6641 |
| GCF\_001592025.1 | 67.2819 | GCF\_000161615.1 | 68.3946 | GCF\_001243895.1 | 68.6674 |
| GCF\_001590785.1 | 67.3613 | GCF\_000374345.1 | 68.4013 | GCF\_002027305.1 | 68.668 |
| GCF\_000612805.1 | 67.3936 | GCF\_000696635.1 | 68.4162 | GCF\_000196735.1 | 68.6736 |
| GCF\_000621445.1 | 67.4559 | GCF\_000227465.1 | 68.4193 | GCF\_000161455.1 | 68.677 |
| GCF\_000177235.2 | 67.6088 | GCF\_001375535.1 | 68.4293 | GCF\_900115845.1 | 68.6864 |
| GCF\_001315065.1 | 67.6828 | GCF\_000344745.1 | 68.486 | GCF\_001517105.1 | 68.6924 |
| GCF\_000513115.1 | 67.725 | GCF\_000009045.1 | 68.4888 | GCF\_001884235.1 | 68.7018 |
| GCF\_001315045.1 | 67.7323 | GCF\_001244735.1 | 68.4907 | GCF\_001884035.1 | 68.7054 |
| GCF\_000513095.1 | 67.7334 | GCF\_000715205.1 | 68.5001 | GCF\_900478165.1 | 68.7098 |
| GCF\_000334155.1 | 67.7804 | GCF\_003605365.2 | 68.5132 | GCF\_001590835.1 | 68.7101 |
| GCF\_000474275.2 | 67.8755 | GCF\_001584325.1 | 68.5138 | GCF\_002156925.1 | 68.7205 |
| GCF\_001278705.1 | 67.8761 | GCF\_000832605.1 | 68.5188 | GCF\_002156865.1 | 68.7334 |
| GCF\_000513135.1 | 67.9596 | GCF\_000245335.1 | 68.5195 | GCF\_001637525.1 | 68.7438 |
| GCF\_001315085.1 | 67.9726 | GCF\_000172815.1 | 68.522 | GCF\_000171615.1 | 68.7569 |
| GCF\_000292245.2 | 67.9961 | GCF\_000186085.1 | 68.524 | GCF\_000832985.1 | 68.7638 |
| GCF\_000787375.1 | 68.0052 | GCF\_000789275.1 | 68.5286 | GCF\_001591525.1 | 68.8197 |
| GCF\_001591945.1 | 68.0939 | GCF\_000017425.1 | 68.5316 | GCF\_000507145.1 | 68.8198 |
| GCF\_900128875.1 | 68.1392 | GCF\_003318295.1 | 68.5334 | GCF\_000430785.1 | 68.8219 |
| GCF\_000008425.1 | 68.1594 | GCF\_000712595.1 | 68.5443 | GCF\_000331575.1 | 68.823 |
| GCF\_003254445.1 | 68.1667 | GCF\_000712615.1 | 68.5464 | GCF\_001884025.1 | 68.8313 |
| GCF\_002119445.1 | 68.1685 | GCF\_000155325.1 | 68.5483 | GCF\_002237735.1 | 68.8668 |
| GCF\_002109385.1 | 68.1786 | GCF\_000496285.1 | 68.5514 | GCF\_900177535.1 | 68.8709 |
| GCF\_001307105.1 | 68.1787 | GCF\_002055965.1 | 68.5529 | GCF\_000956595.1 | 68.8829 |
| GCF\_001644195.2 | 68.1918 | GCF\_001584335.1 | 68.5531 | GCF\_000764085.1 | 68.8861 |
| GCF\_001969855.1 | 68.2001 | GCF\_001884135.1 | 68.5571 | GCF\_001648575.1 | 68.8873 |
| GCF\_000290615.1 | 68.205 | GCF\_001375515.1 | 68.5613 | GCF\_001050115.1 | 68.9459 |
| GCF\_001969815.1 | 68.2138 | GCF\_000507105.1 | 68.5615 | GCF\_900445485.1 | 68.9532 |
| GCF\_000262045.1 | 68.214 | GCF\_000374565.1 | 68.5664 | GCF\_000305495.1 | 69.0203 |
| GCF\_000691165.1 | 68.2218 | GCF\_001884065.1 | 68.5691 | GCF\_002019665.1 | 69.0258 |
| GCF\_000691145.1 | 68.2339 | GCF\_001591925.1 | 68.5698 | GCF\_001883995.1 | 69.0663 |
| GCF\_900207585.1 | 69.0979 | GCF\_003570725.1 | 69.4973 | GCF\_000508325.2 | 70.5606 |
| GCF\_001591605.1 | 69.1012 | GCF\_001439965.1 | 69.5117 | GCF\_001636395.1 | 70.5819 |
| GCF\_000708755.2 | 69.1361 | GCF\_900098925.1 | 69.5168 | GCF\_002335815.1 | 70.6535 |
| GCF\_001439925.1 | 69.1541 | GCF\_002019595.1 | 69.5379 | GCF\_000311725.1 | 70.6605 |
| GCF\_000321185.1 | 69.1887 | GCF\_001420715.1 | 69.578 | GCF\_000612625.1 | 70.6708 |
| GCF\_002243645.1 | 69.1941 | GCF\_002009555.1 | 69.6088 | GCF\_900117315.1 | 70.699 |
| GCF\_900156865.1 | 69.2241 | GCF\_001183985.1 | 69.6329 | GCF\_000813125.1 | 70.7061 |
| GCF\_000380245.2 | 69.2492 | GCF\_002019635.1 | 69.6759 | GCF\_001591505.1 | 70.7085 |
| GCF\_001439915.1 | 69.2741 | GCF\_003515685.1 | 69.7062 | GCF\_001591805.1 | 70.7181 |
| GCF\_001591625.1 | 69.2741 | GCF\_003989135.1 | 69.7537 | GCF\_003937825.1 | 70.9069 |
| GCF\_002844575.1 | 69.2924 | GCF\_900111815.1 | 69.7588 | GCF\_900156875.1 | 70.939 |
| GCF\_001591785.1 | 69.2955 | GCF\_001636425.1 | 69.818 | GCF\_002874535.1 | 70.952 |
| GCF\_001685015.1 | 69.3121 | GCF\_000504145.1 | 69.842 | GCF\_001315165.1 | 70.9769 |
| GCF\_900094975.1 | 69.3315 | GCF\_001591585.1 | 69.913 | GCF\_000517385.1 | 71.0042 |
| GCF\_001456935.1 | 69.3381 | GCF\_000612665.1 | 69.9311 | GCF\_002835735.1 | 71.2512 |
| GCF\_002250055.1 | 69.3626 | GCF\_003073175.1 | 69.9578 | GCF\_001636325.1 | 71.4711 |
| GCF\_001889165.1 | 69.3666 | GCF\_001591485.1 | 70.0585 | GCF\_000262755.1 | 71.635 |
| GCF\_002734215.1 | 69.3697 | GCF\_003055085.1 | 70.0785 | GCF\_001636345.1 | 71.8104 |
| GCF\_000430765.1 | 69.3731 | GCF\_001636415.1 | 70.2145 | GCF\_001636335.1 | 72.3227 |
| GCF\_000285535.1 | 69.4004 | GCF\_003570705.1 | 70.2187 | GCF\_001439605.1 | 72.5692 |
| GCF\_002019645.1 | 69.4092 | GCF\_001591665.1 | 70.2765 | GCF\_001420595.1 | 72.6571 |
| GCF\_001591425.1 | 69.4105 | GCF\_001591445.1 | 70.303 | GCF\_900199725.1 | 72.7647 |
| GCF\_001591825.1 | 69.4405 | GCF\_000986785.1 | 70.3448 | GCF\_001591465.1 | 90.145 |
| GCF\_001591565.1 | 69.4434 | GCF\_001636315.1 | 70.3502 | *B. firmus* I-1582 | 90.147 |
| GCF\_900104555.1 | 69.4442 | GCF\_000307875.1 | 70.3927 | GCF\_900445365.1 | 90.2647 |
| GCF\_001273755.1 | 69.4531 | GCF\_002835805.1 | 70.4581 | *Bacillus* sp. ZZV12-4809 | 100 |
| GCF\_001887185.1 | 69.489 | GCF\_001026695.1 | 70.4865 |  |  |

a computed against *Bacillus* sp. ZZV12-4809 assembly

**Supplementary Table 4.** Strain-specific homologous clusters and their GO-term affiliation as determined by OrthoVenn analysis of *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***B. firmus* I-1582** | | | | ***Bacillus* sp. ZZV12-4809** | | | | |
| Slimmed GO | Name | | Unique-input acc. | Slimmed GO | | Name | | Unique-input acc. |
| **Slimmed biological process GO** | | | | | **Slimmed biological process GO** | | | |
| GO:0005975 | carbohydrate metabolic process | | 3 | GO:0005975 | | carbohydrate metabolic process | | 5 |
| GO:0005976 | polysaccharide metabolic process | | 2 | GO:0005976 | | polysaccharide metabolic process | | 2 |
| GO:0006082 | organic acid metabolic process | | 4 | GO:0006066 | | alcohol metabolic process | | 2 |
| GO:0006139 | nucleobase-containing compound metabolic process | | 3 | GO:0006082 | | organic acid metabolic process | | 9 |
| GO:0006259 | DNA metabolic process | | 6 | GO:0006112 | | energy reserve metabolic process | | 1 |
| GO:0006629 | lipid metabolic process | | 2 | GO:0006139 | | nucleobase-containing compound metabolic process | | 5 |
| GO:0006725 | cellular aromatic compound metabolic process | | 4 | GO:0006259 | | DNA metabolic process | | 3 |
| GO:0006793 | phosphorus metabolic process | | 3 | GO:0006396 | | RNA processing | | 1 |
| GO:0006807 | nitrogen compound metabolic process | | 6 | GO:0006629 | | lipid metabolic process | | 2 |
| GO:0006811 | ion transport | | 1 | GO:0006725 | | cellular aromatic compound metabolic process | | 5 |
| GO:0007049 | cell cycle | | 1 | GO:0006807 | | nitrogen compound metabolic process | | 9 |
| GO:0007059 | chromosome segregation | | 1 | GO:0006810 | | transport | | 1 |
| GO:0008150 | biological process | | 12 | GO:0006811 | | ion transport | | 1 |
| GO:0008152 | metabolic process | | 11 | GO:0006865 | | amino acid transport | | 1 |
| GO:0008643 | carbohydrate transport | | 1 | GO:0007154 | | cell communication | | 3 |
| GO:0009117 | nucleotide metabolic process | | 2 | GO:0007155 | | cell adhesion | | 1 |
| GO:0009225 | nucleotide-sugar metabolic process | | 2 | GO:0008150 | | biological process | | 16 |
| GO:0009987 | cellular process | | 7 | GO:0008152 | | metabolic process | | 17 |
| GO:0015074 | DNA integration | | 4 | GO:0008643 | | carbohydrate transport | | 5 |
| GO:0016043 | cellular component organization | | 1 | GO:0009292 | | genetic transfer | | 2 |
| GO:0016070 | RNA metabolic process | | 1 | GO:0009308 | | amine metabolic process | | 2 |
| GO:0032196 | transposition | | 5 | GO:0009987 | | cellular process | | 12 |
| GO:0032502 | developmental process | | 2 | GO:0015031 | | protein transport | | 2 |
| GO:0032989 | cellular component morphogenesis | | 1 | GO:0015074 | | DNA integration | | 2 |
| GO:0042180 | cellular ketone metabolic process | | 1 | GO:0015833 | | peptide transport | | 2 |
| GO:0043170 | macromolecule metabolic process | | 5 | GO:0016070 | | RNA metabolic process | | 6 |
| GO:0044237 | cellular metabolic process | | 8 | GO:0017144 | | drug metabolic process | | 1 |
| GO:0044238 | primary metabolic process | | 2 | GO:0032502 | | developmental process | | 5 |
| GO:0044255 | cellular lipid metabolic process | | 2 | GO:0043170 | | macromolecule metabolic process | | 9 |
| GO:0046483 | heterocycle metabolic process | | 3 | GO:0044237 | | cellular metabolic process | | 15 |
| GO:0050896 | response to stimulus | | 1 | GO:0044238 | | primary metabolic process | | 7 |
| GO:0051186 | cofactor metabolic process | | 1 | GO:0044255 | | cellular lipid metabolic process | | 4 |
| GO:0051301 | cell division | | 1 | GO:0044419 | | interspecies interaction between organisms | | 1 |
| GO:0065007 | biological regulation | | 2 | GO:0046483 | | heterocycle metabolic process | | 6 |
| GO:0071555 | cell wall organization | | 3 | GO:0050896 | | response to stimulus | | 5 |
| **Slimmed molecular function GO** | | | | GO:0051179 | | localization | | 1 |
| GO:0000166 | | nucleotide binding | 6 | GO:0051704 | | multi-organism process | | 5 |
| GO:0001882 | | nucleoside binding | 5 | GO:0065007 | | biological regulation | | 3 |
| GO:0003674 | | molecular function | 1 | GO:0071555 | | cell wall organization | | 3 |
| GO:0003676 | | nucleic acid binding | 7 | **Slimmed molecular function GO** | | | | |
| GO:0003824 | | catalytic activity | 1 | GO:0000166 | | nucleotide binding | | 6 |
| GO:0004803 | | transposase activity | 2 | GO:0001882 | | nucleoside binding | | 1 |
| GO:0005215 | | transporter activity | 3 | GO:0003674 | | molecular function | | 4 |
| GO:0005488 | | binding | 5 | GO:0003676 | | nucleic acid binding | | 6 |
| GO:0016491 | | oxidoreductase activity | 1 | GO:0004871 | | signal transducer activity | | 1 |
| GO:0016740 | | transferase activity | 3 | GO:0005198 | | structural molecule activity | | 1 |
| GO:0016787 | | hydrolase activity | 1 | GO:0005215 | | transporter activity | | 8 |
| GO:0016829 | | lyase activity | 1 | GO:0005488 | | binding | | 2 |
| GO:0016853 | | isomerase activity | 1 | GO:0009055 | | electron carrier activity | | 1 |
| GO:0016874 | | ligase activity | 2 | GO:0016491 | | oxidoreductase activity | | 9 |
| GO:0043167 | | ion binding | 7 | GO:0016597 | | amino acid binding | | 1 |
| GO:0046906 | | tetrapyrrole binding | 1 | GO:0016740 | | transferase activity | | 12 |
| GO:0048037 | | cofactor binding | 1 | GO:0016787 | | hydrolase activity | | 8 |
| GO:0051540 | | metal cluster binding | 1 | GO:0016829 | | lyase activity | | 4 |
| **Slimmed cellular component GO** | | | | GO:0016853 | | isomerase activity | | 1 |
| GO:0005575 | cellular component | | 2 | GO:0016874 | | ligase activity | | 1 |
| GO:0005576 | extracellular region | | 1 | GO:0019842 | | vitamin binding | | 1 |
| GO:0005618 | cell wall | | 1 | GO:0030246 | | carbohydrate binding | | 1 |
| GO:0005622 | intracellular | | 2 | GO:0033218 | | amide binding | | 1 |
| GO:0016020 | membrane | | 5 | GO:0043167 | | ion binding | | 11 |
| GO:0030312 | external encapsulating structure | | 1 | GO:0046906 | | tetrapyrrole binding | | 1 |
| GO:0030313 | cell envelope | | 1 | GO:0048037 | | cofactor binding | | 5 |
| GO:0042597 | periplasmic space | | 1 | GO:0051540 | | metal cluster binding | | 1 |
| GO:0043234 | protein complex | | 1 | **Slimmed cellular component GO** | | | | |
| GO:0044464 | cell part | | 6 | GO:0005575 | | cellular component | 16 | |
|  |  | |  | GO:0005576 | | extracellular region | 2 | |
|  |  | |  | GO:0005618 | | cell wall | 2 | |
|  |  | |  | GO:0005622 | | intracellular | 6 | |
|  |  | |  | GO:0005773 | | vacuole | 1 | |
|  |  | |  | GO:0016020 | | membrane | 21 | |
|  |  | |  | GO:0031982 | | vesicle | 1 | |
|  |  | |  | GO:0042597 | | periplasmic space | 2 | |
|  |  | |  | GO:0043226 | | organelle | 1 | |
|  |  | |  | GO:0043229 | | intracellular organelle | 1 | |
|  |  | |  | GO:0043232 | | intracellular non-membrane-bounded organelle | 1 | |
|  |  | |  | GO:0044464 | | cell part | 24 | |

**Supplementary Table 5.** Antimicrobial resistance genes, drug targets, transporters and virulence factors found in *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809 genomes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Associated process | Hit in *B. firmus* I-1582 genome (PATRIC ID) | Hit in *Bacillus* sp. ZZV12-4809 genome (PATRIC ID) | Gene | Product |
| Antibiotic resistance | fig|1399.11.peg.2101 | fig|1399.9.peg.5136 | *alr* | Alanine racemase (EC 5.1.1.1) |
|  | fig|1399.11.peg.1993 | fig|1399.9.peg.332 | *ant* | Aminoglycoside 6-nucleotidyltransferase |
|  | fig|1399.11.peg.957 | fig|1399.9.peg.3833 | *bceB* | Bacitracin export permease protein BceB |
|  | fig|1399.11.peg.958 | - | *bceB* | Bacitracin export permease protein BceB |
|  | fig|1399.11.peg.961 | fig|1399.9.peg.3836 | *bceR* | Two-component response regulator BceR |
|  | fig|1399.11.peg.960 | fig|1399.9.peg.3835 | *bceS* | Two-component sensor histidine kinase BceS |
|  | fig|1399.11.peg.1958 | fig|1399.9.peg.5147 | *ddl* | D-alanine--D-alanine ligase (EC 6.3.2.4) |
|  | fig|1399.11.peg.2091 | - | *ddl* | D-alanine--D-alanine ligase (EC 6.3.2.4) |
|  | fig|1399.11.peg.4941 | - | *ddl* | D-alanine--D-alanine ligase (EC 6.3.2.4) |
|  | fig|1399.11.peg.492 | fig|1399.9.peg.967 | *dxr* | 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) |
|  | fig|1399.11.peg.2131 | fig|1399.9.peg.4666 | *EF-G* | Translation elongation factor G |
|  | fig|1399.11.peg.2132 | fig|1399.9.peg.4667 | *EF-Tu* | Translation elongation factor Tu |
|  | fig|1399.11.peg.3214 | fig|1399.9.peg.517 | *fabK*-like | Putative FabK-like enoyl-[acyl-carrier-protein] reductase |
|  | fig|1399.11.peg.1413 | fig|1399.9.peg.25 | *fabL* | Enoyl-[acyl-carrier-protein] reductase [NADPH] (EC 1.3.1.104) |
|  | fig|1399.11.peg.220 | fig|1399.9.peg.2358 | *folA* | Dihydrofolate reductase (EC 1.5.1.3) |
|  | fig|1399.11.peg.4613 | fig|1399.9.peg.4833 | *folP* | Dihydropteroate synthase (EC 2.5.1.15) |
|  | fig|1399.11.peg.1290 | fig|1399.9.peg.2321 | *gdpD* | Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46) |
|  | fig|1399.11.peg.1393 | fig|1399.9.peg.44 | *gdpD* | Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46) |
|  | fig|1399.11.peg.3781 | fig|1399.9.peg.4584 | *gdpD* | Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46) |
|  | fig|1399.11.peg.4205 | fig|1399.9.peg.5065 | *gdpD* | Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46) |
|  | fig|1399.11.peg.259 | fig|1399.9.peg.544 | *gdpD* | Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46) |
|  | fig|1399.11.peg.2677 | fig|1399.9.peg.2082 | *gidB* | 16S rRNA (guanine(527)-N(7))-methyltransferase (EC 2.1.1.170) |
|  | fig|1399.11.peg.2692 | fig|1399.9.peg.2096 | *gyrB* | DNA gyrase subunit B (EC 5.99.1.3) |
|  | fig|1399.11.peg.3910 | fig|1399.9.peg.4278 | *inhA* | Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) |
|  | fig|1399.11.peg.1700 | fig|1399.9.peg.847 | *ileS* | Isoleucyl-tRNA synthetase (EC 6.1.1.5) |
|  | fig|1399.11.peg.3124 | fig|1399.9.peg.4312 | *kasA* | 3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.179) |
|  | fig|1399.11.peg.709 | fig|1399.9.peg.220 | *liaF* | Membrane protein LiaF(VraT) |
|  | fig|1399.11.peg.4115 | fig|1399.9.peg.218 | *liaR* | Response regulator LiaR(VraR) |
|  | fig|1399.11.peg.707 | fig|1399.9.peg.3169 | *liaR* | Response regulator LiaR(VraR) |
|  | fig|1399.11.peg.4116 | fig|1399.9.peg.219 | *liaS* | Sensor histidine kinase LiaS(VraS) |
|  | fig|1399.11.peg.708 | fig|1399.9.peg.3168 | *liaS* | Sensor histidine kinase LiaS(VraS) |
|  | fig|1399.11.peg.879 | fig|1399.9.peg.1590 | *murA* | UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7) |
|  | fig|1399.11.peg.820 | fig|1399.9.peg.1896 | *murA* | UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7) |
|  | fig|1399.11.peg.450 | fig|1399.9.peg.1009 | *pgsA* | CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5) |
|  | fig|1399.11.peg.877 | fig|1399.9.peg.1894 | *rho* | Transcription termination factor Rho |
|  | fig|1399.11.peg.2126 | fig|1399.9.peg.4661 | *rpoB* | DNA-directed RNA polymerase beta subunit (EC 2.7.7.6) |
|  | fig|1399.11.peg.2127 | fig|1399.9.peg.4662 | *rpoC* | DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6) |
|  | fig|1399.11.peg.2133 | fig|1399.9.peg.4668 | *S10p* | SSU ribosomal protein S10p (S20e) |
|  | fig|1399.11.peg.2129 | fig|1399.9.peg.4664 | *S12p* | SSU ribosomal protein S12p (S23e) |
|  | fig|1399.11.peg.4771 | fig|1399.9.peg.3399 | *ykkCD* | Broad-specificity multidrug efflux pump YkkC |
| Drug target | fig|1399.11.peg.3581 | fig|1399.9.peg.2136 | *gap* | NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) |
|  | fig|1399.11.peg.856 | fig|1399.9.peg.1875 | *glyA* | Serine hydroxymethyltransferase (EC 2.1.2.1) |
|  | fig|1399.11.peg.3584 | fig|1399.9.peg.2133 | *gpmI* | 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.12) |
|  | fig|1399.11.peg.2693 | fig|1399.9.peg.2097 | *gyrA* | DNA gyrase subunit A (EC 5.99.1.3) |
|  | fig|1399.11.peg.1770 | fig|1399.9.peg.1821 | *icd* | Isocitrate dehydrogenase [NADP] (EC 1.1.1.42) |
|  | fig|1399.11.peg.3461 | fig|1399.9.peg.4741 | *luxS* | S-ribosylhomocysteine lyase (EC 4.4.1.21) |
|  | fig|1399.11.peg.5015 | fig|1399.9.peg.5134 | *ndoA* | Programmed cell death toxin YdcE |
|  | fig|1399.11.peg.2305 | fig|1399.9.peg.3694 | *parE* | DNA topoisomerase IV subunit B (EC 5.99.1.3) |
|  | fig|1399.11.peg.6 | fig|1399.9.peg.4200 | *pdp* | Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2) |
|  | fig|1399.11.peg.1763 | fig|1399.9.peg.1828 | *pfkA* | 6-phosphofructokinase (EC 2.7.1.11) |
|  | fig|1399.11.peg.855 | fig|1399.9.peg.1874 | *upp* | Uracil phosphoribosyltransferase (EC 2.4.2.9) |
| Transporter | fig|1399.11.peg.959 | fig|1399.9.peg.3834 | *bceA* | Bacitracin export ATP-binding protein BceA |
|  | fig|1399.11.peg.3432 | fig|1399.9.peg.2011 | *Bsu4004* | Gluconate permease, Bsu4004 homolog |
|  | fig|1399.11.peg.788 | fig|1399.9.peg.300 | *daacs* | Na+/H+-dicarboxylate symporter |
|  | fig|1399.11.peg.614 | fig|1399.9.peg.2692 | *dnaK* | Chaperone protein DnaK |
|  | fig|1399.11.peg.2057 | fig|1399.9.peg.5553 | *dppC* | Dipeptide ABC transporter, permease protein DppC (TC 3.A.1.5.2) |
|  | fig|1399.11.peg.1724 | fig|1399.9.peg.825 | *ftsW* | Peptidoglycan glycosyltransferase FtsW (EC 2.4.1.129) |
|  | fig|1399.11.peg.2523 | fig|1399.9.peg.1718 | *mreB* | Rod shape-determining protein MreB |
|  | fig|1399.11.peg.99 | fig|1399.9.peg.2474 | *qcrB* | Menaquinone-cytochrome c reductase |
|  | fig|1399.11.peg.2811 | fig|1399.9.peg.1767 | *sdhFP* | Succinate dehydrogenase flavoprotein subunit (EC 1.3.5.1) |
|  | fig|1399.11.peg.2812 | fig|1399.9.peg.1766 | *sdhIS* | Succinate dehydrogenase iron-sulfur protein (EC 1.3.5.1) |
|  | fig|1399.11.peg.1441 | fig|1399.9.peg.2809 | *spoiiia-e* | Stage III sporulation protein AC |
|  | fig|1399.11.peg.1442 | fig|1399.9.peg.2810 | *spoiiia-e* | Stage III sporulation protein AD |
| Human virulence factor | fig|1399.11.peg.3576 | fig|1399.9.peg.2142 | *clpP* | ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92) |
|  | fig|1399.11.peg.4856 | fig|1399.9.peg.923 | *codY* | GTP-sensing transcriptional pleiotropic repressor CodY |
|  | fig|1399.11.peg.2296 | fig|1399.9.peg.3271 | *EF1623* | Ethanolamine utilization protein similar to PduA/PduJ |
|  | - | fig|1399.9.peg.3684 | *EF1623* | Ethanolamine utilization protein similar to PduA/PduJ |
|  | fig|1399.11.peg.1 | fig|1399.9.peg.4194 | *fur* | Ferric uptake regulation protein FUR |
|  | fig|1399.11.peg.2659 | fig|1399.9.peg.2065 | *purA* | Adenylosuccinate synthetase (EC 6.3.4.4) |
|  | fig|1399.11.peg.2844 | fig|1399.9.peg.5326 | *purB* | Adenylosuccinate lyase (EC 4.3.2.2) |
| **Genes only present in one strain** | |  |  |  |
| Antibiotic resistance | - | fig|1399.9.peg.3896 | *bcrC* | Undecaprenyl-diphosphatase BcrC (EC 3.6.1.27) |
|  | - | fig|1399.9.peg.545 | *blaA* | Class A beta-lactamase (EC 3.5.2.6) |
|  | - | fig|1399.9.peg.2997 | *blaD* | Class D beta-lactamase (EC 3.5.2.6) |
|  | - | fig|1399.9.peg.2937 | *catA9* | Chloramphenicol O-acetyltransferase (EC 2.3.1.28) |
|  | - | fig|1399.9.peg.4474 | *fosB* | Fosfomycin resistance protein FosB |
|  | - | fig|1399.9.peg.3004 | *rlmA (II)* | 23S rRNA (guanine(748)-N(1))-methyltransferase (EC 2.1.1.188) |
|  | - | fig|1399.9.peg.2628 | *vanS* | Vancomycin (or other glycopeptides) histidine kinase VanS |
|  | - | fig|1399.9.peg.2629 | *vanR* | Vancomycin (or other glycopeptides) response regulator VanR |
|  | - | fig|1399.9.peg.2627 | *vanXY*-unclassified | D-Ala-D-Ala dipeptidase/carboxypeptidase (EC 3.4.16.4)(EC 3.4.13.22) |
|  | - | fig|1399.9.peg.2966 | *vatD* | Streptogramin O-acetyltransferase |
|  | - | fig|1399.9.peg.2646 | *vgbA* | Streptogramin B lyase (EC 4.2.99.-) |
| Drug target | fig|1399.11.peg.4112 | - | *hemH* | Coproporphyrin ferrochelatase (EC 4.99.1.9) |
| Transporter | fig|1399.11.peg.2036 | - | - | Putative gap-family peptidoglycolipid addressing protein |
|  | - | fig|1399.9.peg.5394 | *aapA* | Amino-acid permease AapA |
| Human virulence factor | - | fig|1399.9.peg.1510 | *capB* | Poly-gamma-glutamate synthase subunit PgsB/CapB (EC 6.3.2.-) |
|  | - | fig|1399.9.peg.1509 | *capC* | Poly-gamma-glutamate synthase subunit PgsC/CapC (EC 6.3.2.-) |

**Supplementary Table 6.** BLASTP hits for homologous sequences of known virulent proteases (GenBank accessions) found in the analyzed genomes using sequence similarity. The parameters used were: at least 30% sequence identity over 60% of both protein lengths with an E-value of 10−5.

| Query accession (GenBank) | | Genomic hit  (PATRIC ID) | Alignment interval (aa) | Identity (%) | Query cover (%) | Subject cover (%) | E-value |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***B. firmus* I-1582** | | | | | | | |
| ABD96101.1 | fig|1399.11.peg.2959 | | 321 | 32 | 66 | 82 | 6,00E-32 |
| fig|1399.11.peg.4886 | | 461 | 30 | 84 | 72 | 9,00E-32 |
| fig|1399.11.peg.5033 | | 287 | 34 | 65 | 89 | 8,00E-30 |
| fig|1399.11.peg.743 | | 423 | 34 | 69 | 73 | 3,00E-29 |
| CAD20578.1 | fig|1399.11.peg.743 | | 423 | 36 | 75 | 74 | 4,00E-32 |
| fig|1399.11.peg.5033 | | 287 | 33 | 62 | 83 | 4,00E-28 |
| ADK74313.2 | fig|1399.11.peg.5033 | | 287 | 32 | 62 | 83 | 3,00E-28 |
| fig|1399.11.peg.743 | | 423 | 30 | 82 | 85 | 5,00E-28 |
| AAM93666.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| AAT71322.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| AAU01968.1 | fig|1399.11.peg.5033 | | 287 | 33 | 65 | 86 | 2,00E-28 |
| fig|1399.11.peg.4886 | | 461 | 33 | 71 | 60 | 5,00E-28 |
| CAA63841.1 | fig|1399.11.peg.4746 | | 449 | 34 | 62 | 64 | 1,00E-26 |
| ACS66684.1 | fig|1399.11.peg.743 | | 423 | 34 | 75 | 77 | 1,00E-30 |
| ABL74286.1 | fig|1399.11.peg.4746 | | 449 | 34 | 62 | 64 | 2,00E-28 |
| fig|1399.11.peg.743 | | 423 | 31 | 78 | 83 | 1,00E-26 |
| ABF72192.1 | fig|1399.11.peg.2959 | | 321 | 35 | 62 | 80 | 8,00E-32 |
| fig|1399.11.peg.4746 | | 449 | 36 | 62 | 64 | 1,00E-29 |
| AAX54901.1 | fig|1399.11.peg.4746 | | 449 | 34 | 62 | 64 | 1,00E-26 |
| AAX54903.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| AAW21809.2 | fig|1399.11.peg.4886 | | 461 | 30 | 84 | 77 | 1,00E-28 |
| fig|1399.11.peg.4746 | | 449 | 34 | 62 | 64 | 1,00E-26 |
| AEP02888.1 | fig|1399.11.peg.743 | | 423 | 30 | 73 | 81 | 4,00E-21 |
| ABX79998.1 | fig|1399.11.peg.743 | | 423 | 34 | 73 | 73 | 1,00E-29 |
| ABX79999.1 | fig|1399.11.peg.2959 | | 321 | 33 | 66 | 82 | 3,00E-30 |
| ADA70794.1 | fig|1399.11.peg.2959 | | 321 | 33 | 89 | 81 | 1,00E-32 |
| fig|1399.11.peg.5033 | | 287 | 35 | 75 | 74 | 2,00E-29 |
| ABO32256.1 | fig|1399.11.peg.2959 | | 321 | 32 | 65 | 81 | 4,00E-31 |
| fig|1399.11.peg.743 | | 423 | 32 | 74 | 78 | 1,00E-29 |
| fig|1399.11.peg.4886 | | 461 | 30 | 84 | 72 | 1,00E-28 |
| AAV30845.1 | fig|1399.11.peg.2959 | | 321 | 48 | 65 | 78 | 5,00E-68 |
| fig|1399.11.peg.4886 | | 461 | 42 | 71 | 59 | 6,00E-55 |
| fig|1399.11.peg.5033 | | 287 | 42 | 68 | 90 | 1,00E-54 |
| fig|1399.11.peg.743 | | 423 | 39 | 65 | 65 | 4,00E-39 |
| fig|1399.11.peg.4746 | | 449 | 31 | 92 | 89 | 6,00E-33 |
| AAV30844.1 | fig|1399.11.peg.362 | | 555 | 38 | 95 | 93 | 1,00E-95 |
| AFZ77001.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| ABI93801.1 | fig|1399.11.peg.2959 | | 321 | 49 | 65 | 78 | 8,00E-70 |
| fig|1399.11.peg.5033 | | 287 | 43 | 68 | 90 | 1,00E-55 |
| fig|1399.11.peg.743 | | 423 | 40 | 66 | 66 | 8,00E-42 |
| fig|1399.11.peg.4746 | | 449 | 31 | 88 | 86 | 1,00E-34 |
| ABI93803.1 | fig|1399.11.peg.362 | | 555 | 38 | 95 | 93 | 1,00E-96 |
| AAU81559.2 | fig|1399.11.peg.2959 | | 321 | 49 | 65 | 78 | 1,00E-70 |
| fig|1399.11.peg.5033 | | 287 | 43 | 68 | 90 | 1,00E-56 |
| fig|1399.11.peg.743 | | 423 | 39 | 65 | 65 | 2,00E-41 |
| fig|1399.11.peg.4746 | | 449 | 31 | 92 | 89 | 4,00E-34 |
| ABI93802.1 | fig|1399.11.peg.362 | | 555 | 38 | 95 | 93 | 6,00E-95 |
| AFX62372.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| AFX62373.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| CAC80694.2 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG09329.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG10090.1 | fig|1399.11.peg.2959 | | 321 | 98 | 100 | 100 | 0 |
|  | fig|1399.11.peg.5033 | | 287 | 42 | 71 | 77 | 1,00E-47 |
|  | fig|1399.11.peg.4746 | | 449 | 36 | 81 | 64 | 6,00E-29 |
| EWG10233.1 | fig|1399.11.peg.1042 | | 215 | 100 | 100 | 100 | 2,00E-152 |
| EWG10594.1 | fig|1399.11.peg.2616 | | 424 | 96 | 100 | 100 | 0 |
| EWG10706.1 | fig|1399.11.peg.1458 | | 281 | 96 | 100 | 100 | 0 |
| EWG11446.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG12155.1 | fig|1399.11.peg.849 | | 755 | 95 | 100 | 100 | 0 |
| EWG12508.1 | fig|1399.11.peg.4742 | | 430 | 98 | 100 | 100 | 0 |
| EWG12667.1 | fig|1399.11.peg.1281 | | 1445 | 96 | 99 | 99 | 0 |
| EWG12917.1 | fig|1399.11.peg.3383 | | 655 | 98 | 100 | 100 | 0 |
| EWG12971.1 | fig|1399.11.peg.1997 | | 293 | 98 | 100 | 100 | 0 |
| EWG13047.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG13067.1 | fig|1399.11.peg.743 | | 423 | 97 | 100 | 100 | 0 |
|  | fig|1399.11.peg.5033 | | 287 | 36 | 69 | 93 | 5,00E-41 |
|  | fig|1399.11.peg.4886 | | 461 | 32 | 85 | 73 | 1,00E-39 |
|  | fig|1399.11.peg.2959 | | 321 | 36 | 61 | 74 | 3,00E-36 |
|  | fig|1399.11.peg.4746 | | 449 | 35 | 67 | 65 | 9,00E-31 |
| EWG13080.1 | fig|1399.11.peg.731 | | 796 | 96 | 100 | 100 | 0 |
| ***Bacillus* sp. ZZV12-4809** | | | | | | | |
| ABD96101.1 | fig|1399.9.peg.252 | | 419 | 33 | 70 | 74 | 3,00E-32 |
| fig|1399.9.peg.3746 | | 321 | 32 | 66 | 82 | 9,00E-32 |
| fig|1399.9.peg.4004 | | 461 | 30 | 84 | 72 | 1,00E-30 |
| CAD20578.1 | fig|1399.9.peg.252 | | 419 | 33 | 82 | 84 | 3,00E-36 |
| fig|1399.9.peg.4004 | | 461 | 31 | 78 | 65 | 5,00E-27 |
| AAT71322.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| AAU01968.1 | fig|1399.9.peg.252 | | 419 | 37 | 60 | 61 | 7,00E-31 |
| fig|1399.9.peg.4004 | | 461 | 33 | 71 | 60 | 4,00E-28 |
| AAX54903.1 | fig|1399.9.peg.252 | | 419 | 32 | 78 | 82 | 8,00E-28 |
| CAA63841.1 | fig|1399.9.peg.252 | | 419 | 32 | 72 | 77 | 1,00E-27 |
| ACS66684.1 | fig|1399.9.peg.252 | | 419 | 33 | 76 | 78 | 1,00E-33 |
| ABL74286.1 | fig|1399.9.peg.252 | | 419 | 31 | 78 | 82 | 2,00E-28 |
| ABF72192.1 | fig|1399.9.peg.3746 | | 321 | 35 | 62 | 80 | 1,00E-31 |
| AAX54901.1 | fig|1399.9.peg.252 | | 419 | 35 | 60 | 64 | 3,00E-28 |
| AAW21809.2 | fig|1399.9.peg.4004 | | 461 | 32 | 68 | 61 | 2,00E-28 |
| AAM93666.1 | fig|1399.9.peg.252 | | 419 | 32 | 69 | 77 | 1,00E-26 |
| AEP02888.1 | fig|1399.9.peg.252 | | 419 | 30 | 73 | 80 | 5,00E-22 |
| ABX79998.1 | fig|1399.9.peg.252 | | 419 | 31 | 80 | 83 | 2,00E-32 |
| fig|1399.9.peg.4004 | | 461 | 32 | 66 | 58 | 7,00E-25 |
| ABX79999.1 | fig|1399.9.peg.252 | | 419 | 33 | 70 | 74 | 2,00E-31 |
| fig|1399.9.peg.3746 | | 321 | 33 | 66 | 82 | 3,00E-30 |
| fig|1399.9.peg.4004 | | 461 | 30 | 84 | 72 | 4,00E-27 |
| ADA70794.1 | fig|1399.9.peg.252 | | 419 | 39 | 82 | 61 | 2,00E-33 |
| fig|1399.9.peg.3746 | | 321 | 33 | 89 | 81 | 2,00E-33 |
| fig|1399.9.peg.3481 | | 312 | 35 | 69 | 67 | 2,00E-24 |
| ABO32256.1 | fig|1399.9.peg.252 | | 419 | 33 | 74 | 77 | 5,00E-33 |
| fig|1399.9.peg.3746 | | 321 | 33 | 65 | 81 | 1,00E-31 |
| AAV30845.1 | fig|1399.9.peg.3746 | | 321 | 50 | 65 | 78 | 1,00E-68 |
| fig|1399.9.peg.4004 | | 461 | 37 | 87 | 75 | 1,00E-55 |
| fig|1399.9.peg.252 | | 419 | 38 | 73 | 73 | 1,00E-40 |
| fig|1399.9.peg.3481 | | 312 | 40 | 61 | 78 | 1,00E-39 |
| fig|1399.9.peg.4924 | | 446 | 30 | 98 | 99 | 1,00E-27 |
| AAV30844.1 | fig|1399.9.peg.3622 | | 555 | 40 | 95 | 93 | 1,00E-97 |
| fig|1399.9.peg.1988 | | 571 | 37 | 89 | 90 | 5,00E-77 |
| AFZ77001.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| ABI93801.1 | fig|1399.9.peg.3746 | | 321 | 50 | 65 | 78 | 2,00E-70 |
| fig|1399.9.peg.4004 | | 461 | 38 | 87 | 75 | 6,00E-57 |
| fig|1399.9.peg.252 | | 419 | 39 | 74 | 74 | 1,00E-43 |
| fig|1399.9.peg.3481 | | 312 | 40 | 61 | 78 | 2,00E-41 |
| fig|1399.9.peg.4924 | | 446 | 30 | 98 | 99 | 3,00E-28 |
| ABI93803.1 | fig|1399.9.peg.3622 | | 555 | 40 | 95 | 93 | 2,00E-98 |
| fig|1399.9.peg.1988 | | 571 | 36 | 89 | 90 | 3,00E-77 |
| AAU81559.2 | fig|1399.9.peg.3746 | | 321 | 50 | 65 | 78 | 4,00E-71 |
| fig|1399.9.peg.4004 | | 461 | 38 | 87 | 75 | 1,00E-57 |
| fig|1399.9.peg.252 | | 419 | 39 | 73 | 73 | 6,00E-43 |
| fig|1399.9.peg.3481 | | 312 | 41 | 61 | 78 | 8,00E-42 |
| fig|1399.9.peg.4924 | | 446 | 30 | 98 | 99 | 6,00E-28 |
| ABI93802.1 | fig|1399.9.peg.3622 | | 555 | 40 | 95 | 93 | 3,00E-97 |
| fig|1399.9.peg.1988 | | 571 | 37 | 89 | 90 | 1,00E-77 |
| AFX62372.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| AFX62373.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| CAC80694.2 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG09329.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG10090.1 | fig|1399.9.peg.3746 | | 321 | 93 | 100 | 100 | 0 |
|  | fig|1399.9.peg.3481 | | 312 | 48 | 91 | 96 | 3,00E-78 |
|  | fig|1399.9.peg.4004 | | 461 | 43 | 71 | 48 | 3,00E-48 |
|  | fig|1399.9.peg.252 | | 419 | 34 | 81 | 66 | 4,00E-37 |
| EWG10233.1 | fig|1399.9.peg.3368 | | 215 | 94 | 100 | 100 | 3,00E-135 |
| EWG10594.1 | fig|1399.9.peg.5541 | | 424 | 92 | 100 | 100 | 0 |
| EWG10706.1 | fig|1399.9.peg.2826 | | 281 | 95 | 100 | 100 | 0 |
| EWG11446.1 | 0 | | 0 | 0 | 0 | 0 | 0 |
| EWG12155.1 | fig|1399.9.peg.1620 | | 754 | 93 | 99 | 99 | 0 |
| EWG12508.1 | fig|1399.9.peg.3186 | | 430 | 97 | 100 | 100 | 0 |
| EWG12667.1 | fig|1399.9.peg.4576 | | 1446 | 94 | 100 | 100 | 0 |
| EWG12917.1 | fig|1399.9.peg.402 | | 655 | 96 | 100 | 100 | 0 |
| EWG12971.1 | fig|1399.9.peg.342 | | 293 | 96 | 100 | 100 | 0 |
| EWG13047.1 | fig|1399.9.peg.272 | | 608 | 92 | 100 | 100 | 0 |
| EWG13067.1 | fig|1399.9.peg.252 | | 419 | 95 | 100 | 100 | 0 |
|  | fig|1399.9.peg.4004 | | 461 | 32 | 85 | 72 | 2,00E-39 |
|  | fig|1399.9.peg.3746 | | 321 | 37 | 61 | 74 | 2,00E-36 |
|  | fig|1399.9.peg.4924 | | 446 | 30 | 68 | 70 | 5,00E-16 |
| EWG13080.1 | fig|1399.9.peg.240 | | 795 | 93 | 100 | 100 | 0 |

**Supplementary Table 7.** Unique BLASTP hits (from Supplementary Table 6) for putative nematode-virulent proteases in the *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809 genomes.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Bacillus* strain | Protein sequence PATRIC ID | Annotation | Length (aa) | Conserved family domain | Comment |
| I-1582 | fig|1399.11.peg.2959 | Major intracellular serine protease precursor (EC 3.4.21.-) | 321 | Peptidase S8 | Sep1 (Geng *et al*. 2016) |
| I-1582 | fig|1399.11.peg.4886 | Protease | 461 | Peptidase S8 |  |
| I-1582 | fig|1399.11.peg.5033 | Protease | 287 | Peptidase S8 |  |
| I-1582 | fig|1399.11.peg.743 | Subtilisin precursor (EC 3.4.21.62) | 423 | Peptidase S8 |  |
| I-1582 | fig|1399.11.peg.4746 | Serine protease AprX | 449 | Peptidase S8 |  |
| I-1582 | fig|1399.11.peg.362 | Zinc metalloproteinase aureolysin (EC 3.4.24.29) | 555 | Peptidase M4 |  |
| I-1582 | fig|1399.11.peg.2616 | Ankyrin repeat protein, putative | 424 | Peptidase M48 |  |
| I-1582 | fig|1399.11.peg.1458 | RNA binding methyltransferase FtsJ like | 281 | Haemolysin A/rRNA methyltransferase TlyA |  |
| I-1582 | fig|1399.11.peg.849 | hypothetical protein | 755 | Peptidase S8 |  |
| I-1582 | fig|1399.11.peg.4742 | Uncharacterized metalloprotease YhfN | 430 | CAAX prenyl protease 1, Peptidase M48 |  |
| I-1582 | fig|1399.11.peg.1281 | Bacillopeptidase F | 1445 | Peptidase S8, Peptidase M6-like (only one domain) |  |
| I-1582 | fig|1399.11.peg.3383 | Hypothetical protein | 655 | Peptidase M6-like |  |
| I-1582 | fig|1399.11.peg.1997 | Protease HtpX | 293 | Peptidase M48 |  |
| I-1582 | fig|1399.11.peg.731 | Immune inhibitor A metalloprotease | 796 | Peptidase M6 |  |
| I-1582 | fig|1399.11.peg.1042 | FIG01964566: Predicted membrane protein, hemolysin III homolog | 215 | Haemolysin III |  |
| ZZV12-4809 | fig|1399.9.peg.252 | Subtilisin precursor (EC 3.4.21.62) | 419 | Peptidase S8 |  |
| ZZV12-4809 | fig|1399.9.peg.3746 | Major intracellular serine protease precursor (EC 3.4.21.-) | 321 | Peptidase S8 | Sep1 (Geng *et al*. 2016) |
| ZZV12-4809 | fig|1399.9.peg.4004 | Protease | 461 | Peptidase S8 |  |
| ZZV12-4809 | fig|1399.9.peg.3481 | Hypothetical protein | 312 | Peptidase S8 |  |
| ZZV12-4809 | fig|1399.9.peg.4924 | Hypothetical protein | 446 | Peptidase S8 |  |
| ZZV12-4809 | fig|1399.9.peg.3622 | Zinc metalloproteinase aureolysin (EC 3.4.24.29) | 555 | Peptidase M4 |  |
| ZZV12-4809 | fig|1399.9.peg.1988 | Zinc metalloproteinase precursor (EC 3.4.24.29) / aureolysin | 571 | Peptidase M4 |  |
| ZZV12-4809 | fig|1399.9.peg.5541 | ankyrin repeat protein, putative | 424 | Peptidase M48 |  |
| ZZV12-4809 | fig|1399.9.peg.2826 | RNA binding methyltransferase FtsJ like | 281 | Haemolysin A/rRNA methyltransferase TlyA |  |
| ZZV12-4809 | fig|1399.9.peg.1620 | hypothetical protein | 754 | Peptidase S8 |  |
| ZZV12-4809 | fig|1399.9.peg.3186 | Uncharacterized metalloprotease YhfN | 430 | CAAX prenyl protease 1, Peptidase M48 |  |
| ZZV12-4809 | fig|1399.9.peg.4576 | bacillopeptidase F | 1446 | Peptidase S8, Peptidase M6-like (only one domain) |  |
| ZZV12-4809 | fig|1399.9.peg.402 | hypothetical protein | 655 | Peptidase M6-like |  |
| ZZV12-4809 | fig|1399.9.peg.342 | Protease HtpX | 293 | Peptidase M48 |  |
| ZZV12-4809 | fig|1399.9.peg.272 | putative secreted protease | 608 | Peptidase M6-like |  |
| ZZV12-4809 | fig|1399.9.peg.240 | Immune inhibitor A metalloprotease | 795 | Peptidase M6 |  |
| ZZV12-4809 | fig|1399.9.peg.3368 | FIG01964566: Predicted membrane protein, hemolysin III homolog | 215 | Haemolysin III |  |

**Supplementary Table 8.** Analysis of the putative chitinase sequences found in the genomes of *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809, with results from NCBI CDD for domain and functional annotations.

|  |  |  |
| --- | --- | --- |
| Query sequence | Protein domain(s)   * Superfamily * Family * Aligned interval (aa); E-value * Accessions(a) (b) | Short domain description |
| fig|1399.9.peg.2658| | * GH18\_chitinase-like superfamily; * GH18\_chitinase; * 355–595; 2.7e-86; * cl10447; cd06548 | GH18 (glycosyl hydrolase, family 18) type II chitinases hydrolyze chitin, an abundant polymer of N-acetylglucosamine, which is a major component of the cell wall of fungi and the exoskeleton of arthropods. |
| * FN3; * 137–204; 2.3e-09; * cl21522; smart00060 | Fibronectin type 3 domain; one of three types of internal repeat within the plasma protein, fibronectin. |
| fig|1399.9.peg.3724| | * GH18\_chitinase-like superfamily; * GH18\_3CO4\_chitinase; * 66–367; 1.4e-44; * cl10447; cd06545 | The *Bacteroides thetaiotaomicron* protein represented by pdb structure 3CO4 is an uncharacterized bacterial member of the family 18 glycosyl hydrolases with homologs found in *Flavobacterium*, *Stigmatella*, and *Pseudomonas*. |
| fig|1399.9.peg.3666| | * GH18\_chitinase-like superfamily; * GH18\_CFLE\_spore\_hydrolase; * 259–563; 2.3e-94; * cl10447; cd02874 | Cortical fragment-lytic enzyme (CFLE) is a peptidoglycan hydrolase involved in bacterial endospore germination. CFLE degrades a spore-specific peptidoglycan constituent called muramic-acid delta-lactam that comprises the outer cortex. |
| * Cu\_amine\_oxidN1 superfamily; * Cu\_amine\_oxidN1; * 67–139; 2.5e-05; * cl06736; pfam07833 | Cu-amine oxidases catalyze the oxidative deamination of primary amines to the corresponding aldehydes, while reducing molecular oxygen to hydrogen peroxide. |
| fig|1399.11.peg.2278| | * GH18\_chitinase-like superfamily; * GH18\_CFLE\_spore\_hydrolase; * 259–563; 1.5e-90; * cl10447; cd02874 | See description for cd02874 |
| * Cu\_amine\_oxidN1 superfamily; * Cu\_amine\_oxidN1; * 67–132; 2.8e-04; * cl06736; pfam07833 | See description for pfam07833 |
| fig|1399.11.peg.4380| | * GH18\_chitinase-like superfamily; * GH18\_3CO4\_chitinase; * 27–113; 5.8e-08; * cl10447; cd06545 | Truncated GH18\_3CO4 domain; see description for cd06545 |

(a) superfamily cluster accession from NCBI CDD or SMART database

(b) position-specific scoring matrix (PSSM) / protein families collection (Pfam)

**Supplementary Table 9.** All predicted/putative secondary metabolite clusters (antiSMASH analysis) in the genomes of *Bacillus firmus* I-1582 and *Bacillus* sp. ZZV12-4809.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cluster | Type | Genome location (contig) | From | To | Cluster size (bp) | Most similar homologous cluster | | | Most similar known cluster | | | | Additional information |
| ID | Similar genes (%) | | MIBiG BGC-ID (a) | Biosynthetic gene cluster name | Similar genes (%) | |
| B. firmus I-1582 | | | | | | | | | | | | | |
| Cluster 1 | T3pks | NODE\_1\_length\_277315\_cov\_80.3941 | 118902 | 159984 | 41083 | ACWD01000030\_c1 | 100 | - | | - | - |  | |
| Cluster 2 | Cf-putative | NODE\_5\_length\_109936\_cov\_111.444 | 59668 | 71895 | 12228 | LNDH01000009\_c1 | 17 | - | | - | - |  | |
| Cluster 3 | Cf-putative | NODE\_5\_length\_109936\_cov\_111.444 | 81712 | 90212 | 8501 | JXLS01000066\_c1 | 14 | - | | - | - |  | |
| Cluster 4 | Cf-putative | NODE\_6\_length\_100661\_cov\_78.5669 | 31203 | 51370 | 20168 | AJLS01000136\_c1 | 22 | - | | - | - |  | |
| Cluster 5 | Cf-saccharide | NODE\_8\_length\_96798\_cov\_86.1395 | 36954 | 60070 | 23117 | CP012600\_c7 | 41 | BGC0000792\_c1 | | Polysaccharide | 23 | Most similar subcluster: Avermectin oleandrose (AB032523.1); 25% genes similar | |
| Cluster 6 | Cf-putative | NODE\_10\_length\_80921\_cov\_93.6957 | 30248 | 37659 | 7412 | - | - | - | | - | - |  | |
| Cluster 7 | Cf-putative | NODE\_11\_length\_78866\_cov\_87.7244 | 61634 | 66414 | 4781 | LIUR01000001\_c6 | 9 | - | | - | - |  | |
| Cluster 8 | Cf-putative | NODE\_11\_length\_78866\_cov\_87.7244 | 68453 | 78745 | 10293 | JXRP01000019\_c1 | 16 | BGC0001090\_c1 | | Bacillomycin | 20 |  | |
| Cluster 9 | Cf-putative | NODE\_13\_length\_76437\_cov\_86.9749 | 11988 | 20711 | 8724 | LMEZ01000001\_c4 | 12 | - | | - | - |  | |
| Cluster 10 | Cf-putative | NODE\_17\_length\_67099\_cov\_88.4517 | 53963 | 65379 | 11417 | CP000922\_c1 | 7 | - | | - | - |  | |
| Cluster 11 | Lanthipeptide | NODE\_18\_length\_65837\_cov\_99.7641 | 1 | 31154 | 31154 | ALXI01000173\_c1 | 16 | - | | - | - | 1: leader/core peptide, putative Class I(b): VKENMFDLDVQVLKAASGEQPD-IKDhaKLFCDhbPGCVDQDhbDhbKDhbDhaCCLKCKDhbDhaLCNDhaNMDhaCW;  Contig border cluster/ possibly incomplete | |
| Cluster 12 | Cf-putative | NODE\_22\_length\_57374\_cov\_81.5008 | 195 | 7969 | 7775 | JOTM01000019\_c1 | 11 | - | | - | - |  | |
| Cluster 13 | Cf-putative | NODE\_23\_length\_56314\_cov\_78.3736 | 130 | 6866 | 6737 | LFZU01000001\_c3 | 7 | - | | - | - |  | |
| Cluster 14 | Cf-putative | NODE\_24\_length\_55898\_cov\_90.2335 | 45720 | 49389 | 3670 | JRFO01000005\_c3 | 6 | - | | - | - |  | |
| Cluster 15 | Siderophore | NODE\_27\_length\_50215\_cov\_102.313 | 28262 | 41766 | 13505 | ACWD01000072\_c1 | 32 | BGC0000942\_c1 | | Petrobactin | 33 |  | |
| Cluster 16 | Cf-putative | NODE\_34\_length\_41890\_cov\_81.5435 | 17486 | 22549 | 5064 | CP011110\_c6 | 9 | - | | - | - |  | |
| Cluster 17 | Cf-putative | NODE\_37\_length\_36602\_cov\_74.6194 | 14931 | 31635 | 16705 | CCDJ010000002\_c1 | 13 | - | | - | - |  | |
| Cluster 18 | Cf-putative | NODE\_38\_length\_36313\_cov\_78.9992 | 13347 | 19121 | 5775 | CP012600\_c3 | 7 | - | | - | - |  | |
| Cluster 19 | Cf-fatty acid | NODE\_39\_length\_35727\_cov\_91.3177 | 7647 | 29866 | 22220 | LMHK01000003\_c1 | 18 | - | | - | - |  | |
| Cluster 20 | Cf-putative | NODE\_46\_length\_32799\_cov\_95.9456 | 23845 | 31329 | 7485 | FR773526\_c1 | 7 | BGC0000796\_c1 | | S-layer glycan | 9 |  | |
| Cluster 21 | Cf-fatty acid | NODE\_54\_length\_27181\_cov\_79.2117 | 7316 | 27181 | 19866 | ACWD01000043\_c1 | 26 | - | | - | - | Contig border cluster/ possibly incomplete | |
| Cluster 22 | Cf-putative | NODE\_59\_length\_26108\_cov\_60.5532 | 17902 | 23858 | 5957 | AYJU01000001\_c1 | 3 | - | | - | - |  | |
| Cluster 23 | Cf-putative | NODE\_72\_length\_21863\_cov\_64.9777 | 7045 | 17103 | 10059 | CP009684\_c15 | 7 | - | | - | - |  | |
| Cluster 24 | Cf-saccharide | NODE\_85\_length\_16505\_cov\_73.2168 | 1 | 16505 | 16505 | AKKV01000020\_c1 | 10 | BGC0000796\_c1 | | S-layer glycan | 9 | Contig border cluster/ possibly incomplete | |
| Cluster 25 | Terpene | NODE\_107\_length\_9159\_cov\_73.9875 | 1 | 9159 | 9159 | JWJE02000036\_c1 | 12 | - | | - | - | Contig border cluster/ possibly incomplete | |
| Cluster 26 | Cf-putative | NODE\_124\_length\_5182\_cov\_46.6323 | 151 | 4689 | 4539 | AAJM01000005\_c1 | 12 | - | | - | - |  | |
| Cluster 27 | Cf-putative | NODE\_129\_length\_4428\_cov\_39.9781 | 156 | 4241 | 4086 | JXAJ01000001\_c2 | 5 | BGC0000796\_c1 | | S-layer glycan | 14 |  | |
| Cluster 28 | Cf-saccharide | NODE\_149\_length\_2065\_cov\_34.1775 | 1 | 2065 | 2065 | AKKV01000020\_c1 | 5 | - | | - | - | Contig border cluster/ possibly incomplete | |
| Bacillus sp. ZZV12-4809 | | | | | | | | | | | | | |
| Cluster 1 | Cf-putative | NODE\_1\_length\_587391\_cov\_36.5028 | 30683 | 34225 | 3543 | CP002293\_c3 | 8 | - | | - | - | |  |
| Cluster 2 | Cf-putative | NODE\_1\_length\_587391\_cov\_36.5028 | 52240 | 58244 | 6005 | - | - | - | | - | - | |  |
| Cluster 3 | Cf-putative | NODE\_1\_length\_587391\_cov\_36.5028 | 335867 | 355358 | 19492 | CP009283\_c3 | 12 | BGC0000796\_c1 | | S-layer glycan | 9 | |  |
| Cluster 4 | Cf-fatty acid | NODE\_1\_length\_587391\_cov\_36.5028 | 368340 | 389299 | 20960 | LMRU01000001\_c2 | 7 | - | | - | - | |  |
| Cluster 5 | Cf-putative | NODE\_1\_length\_587391\_cov\_36.5028 | 393370 | 403603 | 10234 | LILC01000013\_c1 | 4 | - | | - | - | |  |
| Cluster 6 | Cf-putative | NODE\_1\_length\_587391\_cov\_36.5028 | 408391 | 430591 | 22201 | ACNI01000155\_c1 | 15 | - | | - | - | |  |
| Cluster 7 | Cf-putative | NODE\_3\_length\_302206\_cov\_33.0507 | 87988 | 98474 | 10487 | CP011862\_c4 | 11 | - | | - | - | |  |
| Cluster 8 | Cf-putative | NODE\_3\_length\_302206\_cov\_33.0507 | 145921 | 153991 | 8071 | CP003187\_c5 | 3 | - | | - | - | |  |
| Cluster 9 | Cf-putative | NODE\_3\_length\_302206\_cov\_33.0507 | 181742 | 188871 | 7130 | - | - | - | | - | - | |  |
| Cluster 10 | Bacteriocin-Terpene | NODE\_3\_length\_302206\_cov\_33.0507 | 261850 | 302206 | 40357 | ACWD01000043\_c1 | 34 | - | | - | - | | Contig border cluster/ possibly incomplete |
| Cluster 11 | Cf-saccharide | NODE\_4\_length\_264972\_cov\_39.112 | 1 | 15292 | 15292 | AKKV01000020\_c1 | 10 | BGC0000764\_c1 | | Exo-polysaccharide | 11 | | Contig border cluster/ possibly incomplete |
| Cluster 12 | Cf-putative | NODE\_4\_length\_264972\_cov\_39.112 | 36695 | 48754 | 12060 | APRS01000013\_c1 | 16 | - | | - | - | |  |
| Cluster 13 | Cf-putative | NODE\_4\_length\_264972\_cov\_39.112 | 91951 | 98483 | 6533 | CP001982\_c4 | 5 | - | | - | - | |  |
| Cluster 14 | Cf-putative | NODE\_6\_length\_213038\_cov\_39.4491 | 30523 | 39350 | 8828 | JXLS01000066\_c1 | 14 | - | | - | - | |  |
| Cluster 15 | Cf-putative | NODE\_6\_length\_213038\_cov\_39.4491 | 93280 | 102681 | 9402 | JWIR02000021\_c1 | 3 | - | | - | - | |  |
| Cluster 16 | Lanthipeptide | NODE\_7\_length\_208029\_cov\_35.6017 | 1 | 30103 | 30103 | CYHN01000001\_c1 | 10 | - | | - | - | | 1: leader/core peptide, putative Class II: MSYSKRERIFLFYTGGCEMESVEAA-YIAGIIDDhaEGDhbNHDhaNQEDhbCE  2: leader/core peptide, putative Class II(b): MSSEEIILSWKNPESRQKVENHPSGTSFNELEFDEMLSVCGG-DhaGDIKPEADhbPDhbIALDhbDhaALVDhaAPADhaFVGDhaYLVDhbKVFCK;  Contig border cluster/ possibly incomplete |
| Cluster 17 | Cf-saccharide | NODE\_7\_length\_208029\_cov\_35.6017 | 164747 | 208029 | 43283 | AGCN01000032\_c1 | 15 | BGC0000796\_c1 | | S-layer glycan | 14 | | Contig border cluster/ possibly incomplete |
| Cluster 18 | T3pks | NODE\_8\_length\_182807\_cov\_34.5978 | 68207 | 109289 | 41083 | ACWD01000030\_c1 | 100 | - | | - | - | |  |
| Cluster 19 | T1pks-Nrps | NODE\_9\_length\_179541\_cov\_32.701 | 14907 | 84725 | 69819 | BAUY01000012\_c1 | 22 | BGC0001033\_c1 | | Paenilamicin | 35 | | Monomers prediction: (ala-ala) + (mal) + (asn) + (ala-ala-asp) |
| Cluster 20 | Cf-putative | NODE\_11\_length\_167313\_cov\_30.447 | 849 | 35915 | 35067 | JPVN01000015\_c1 | 10 | - | | - | - | |  |
| Cluster 21 | Cf-putative | NODE\_11\_length\_167313\_cov\_30.447 | 154150 | 167150 | 13001 | GG705060\_c1 | 22 | BGC0001098\_c1 | | Iturin | 22 | |  |
| Cluster 22 | Cf-putative | NODE\_13\_length\_152528\_cov\_30.811 | 6906 | 22107 | 15202 | AJLS01000136\_c1 | 22 | - | | - | - | |  |
| Cluster 23 | Cf-putative | NODE\_13\_length\_152528\_cov\_30.811 | 95028 | 115354 | 20327 | JXCK01000020\_c1 | 7 | BGC0000796\_c1 | | S-layer glycan | 14 | |  |
| Cluster 24 | Bacteriocin | NODE\_14\_length\_120695\_cov\_43.258 | 32400 | 43224 | 10825 | JPNZ01000001\_c3 | 2 | - | | - | - | |  |
| Cluster 25 | Cf-putative | NODE\_14\_length\_120695\_cov\_43.258 | 80731 | 97705 | 16975 | CP012159\_c24 | 2 | - | | - | - | |  |
| Cluster 26 | Cf-putative | NODE\_16\_length\_103040\_cov\_34.006 | 1844 | 9606 | 7763 | JOTM01000019\_c1 | 11 | - | | - | - | |  |
| Cluster 27 | Cf-putative | NODE\_17\_length\_99667\_cov\_31.965 | 11023 | 26338 | 15316 | LJIZ01000030\_c1 | 21 | - | | - | - | |  |
| Cluster 28 | Cf-putative | NODE\_17\_length\_99667\_cov\_31.965 | 27811 | 35159 | 7349 | APVL01000007\_c2 | 15 | - | | - | - | |  |
| Cluster 29 | Siderophore | NODE\_17\_length\_99667\_cov\_31.965 | 37193 | 52651 | 15459 | APVL01000007\_c2 | 26 | - | | - | - | |  |
| Cluster 30 | Cf-putative | NODE\_21\_length\_78140\_cov\_35.3167 | 10802 | 16787 | 5986 | JXRP01000019\_c1 | 11 | - | | - | - | |  |
| Cluster 31 | Cf-putative | NODE\_21\_length\_78140\_cov\_35.3167 | 18830 | 29282 | 10453 | JXRP01000019\_c1 | 16 | BGC0001098\_c1 | | Iturin | 22 | |  |
| Cluster 32 | Cf-fatty acid | NODE\_22\_length\_77980\_cov\_35.5981 | 41420 | 63659 | 22240 | ATLO01000004\_c1 | 9 | - | | - | - | |  |
| Cluster 33 | Siderophore | NODE\_23\_length\_72740\_cov\_39.2002 | 14967 | 28480 | 13514 | ACWD01000072\_c1 | 32 | BGC0000942\_c1 | | Petrobactin | 33 | |  |
| Cluster 34 | Cf-putative | NODE\_25\_length\_67830\_cov\_32.8632 | 44362 | 49391 | 5030 | CP011110\_c6 | 9 | - | | - | - | |  |
| Cluster 35 | Cf-saccharide | NODE\_39\_length\_28297\_cov\_35.2256 | 1 | 19719 | 19719 | CP012600\_c7 | 11 | - | | - | - | | Contig border cluster/ possibly incomplete |
| Cluster 36 | Cf-putative | NODE\_44\_length\_25567\_cov\_39.5311 | 13617 | 25433 | 11817 | LFJO01000002\_c1 | 14 | - | | - | - | |  |

(a) Minimum Information about a Biosynthetic Gene cluster (MIBiG) database of Biosynthetic Gene Clusters (BGC)

(b) Dha: Didehydroalanine; Dhb: Didehydrobutyrin

**REFERENCES**

Chun, J., A. Oren, A. Ventosa, H. Christensen, D. R. Arahal *et al.*, 2018 Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68: 461-466.

Geng, C., X. Nie, Z. Tang, Y. Zhang, J. Lin *et al.*, 2016 A novel serine protease, Sep1, from *Bacillus firmus* DS-1 has nematicidal activity and degrades multiple intestinal-associated nematode proteins. Sci. Rep. 6: 25012.

Lee, I. *et al.*, 2017 ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. Int. J. Syst. Evol. Microbiol.67: 2053-2057.

Richter, M., and R. Rosselló-Móra, 2009 Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U.S.A. 106: 19126-19131.

Richter, M., R. Rosselló-Móra, F. Oliver Glöckner, and J. Peplies, 2016 JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32: 929-931.

Teeling, H., J. Waldmann, T. Lombardot, M. Bauer, and F. O. Glöckner, 2004 TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. BMC Bioinformatics 5: 163.

Yoon, S.-H. *et al.*, 2017 Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int. J. Syst. Evol. Microbiol. 67: 1613-1617.