

Supplementary text 2:

Examples of intergenic hits into unannotated ORFs

Spacer 1:

>JOE01000009.1_62635_14_spacer_63457_32

GACCAGGCGTCCCCGCTGACGCTGATCGCGGT

Hit coordinates:

ContigID: LGDA01000250

Start: 45153

Stop: 45184

Intergenic sequence:

>gi|925387131|gb|LGDA01000250.1|:44655-45964 Streptomyces sp. WM6368 P399contig36.1, whole genome shotgun sequence

TAAGACATGCGGTGTGGCCGGGCACATGGATGCCCGGCATCCGACAGGGTCCTGCGAGCTGCCGCGTG
AGCACAGTGCAGCAGGACAGGCGGAGTCTGGCGGGGAACGGGAGTCGGCTGCGGCATACCCCGGAGGACG
ACTAAGTACAGCTGGATGGTTGCCGGCCACGCCGTCGGTGGACCATCCGACGGGCGGTCTCTGCTGCTGC
GGCGGCGTGGCCAGGAGCAGGTGCGGGTCTCCAGCCGGAACGGCGCGTACCGCCCCGGCTCCGGCT
CGTACCGGCGATACGGGGCGCTCTCGTAGACGACGGTGGTGTTCGGGCAGTCTCCGCGACCCCGAGCAG
CGCTTGGCGCTCGGCGTCTCGGCGGCAAGGCTCCAGCGCAACTTCGCCGCGGTCCACTGGGCGGCGTAC
TCGCAGTGGTAACCGCTGGCCGGCGGAATCCACTGCGCCGGATCCTTGTGCGCCTTCGACCGGTTGACT
TCGCGGTACCCGCGATCAGCGTCAGCGGGGACGCCTGGTTCGTCGTACGCCTCACGCCGGACCGCGTC
CCACGCGAAGCTGCCCGAGTCGTGGACTTCGGCCAAAGGTACGAAATGGTCTACATCAAGGCCCGCGGGC
TCCGTCAACAACGACGTCGCCGTAAGGGCTCAGCCGCGATCCACCCGACATCTTGACGCCCGCAGCGACGA
CCGGGGCCTCGACGGCCTCGGACAGGATGACTTCCCTGCGGGTATCACAGCCGTGGTTCGCGTTCAGCCC
GCGGTTCCAGTGCTTGTACAGGTCCCACTTGTAAACCTCCCGCTGCTCATCAGCGACGGGGATCCGGTTCG
ATGGCCTCGAACAGCGGCAGCGGCGCACGGACTCCGGGAGCCCCAGCGATACCGGACGAGGCGGGGACGC
CCGCTGGGTGCCGGCCGCGTGCGCCGGGTCGGTGGCGAGCAGGGGCAGGGCAGCGAGAGCGAGCGC
GGGAGGACCACGCTGCAGCAGGTTCTTGATCACGCAGTGGGTTGTAGCGGTGCCCCAGCGGGACAGGGCC
TGGGGCCTCAGCTACCCGACAGGCCGCGGAATCACCGCAGAAACGATCAGCGGCTGGGCCGACCGGC
AGCCGGTCAACGCTCTCAGCGCGAGATCATCTTTGGCCCGTGGGCAACGGCTCGGTCCCGGGCAACCC
ACACGCATCATCGGGAGTACCAGCAGATTCCAGCGCACGGCAGTTGCAACATCATGGCGTCAGGTGCCT
CGAGTCAGCTGCACATGACGGTCTTTGCGACGTCTCCAGACGAGGGCCG

Top 10 BLASTX hits: into HNH endonuclease

Fields: query id, subject id, subject ids, subject length, s. start, s. end, evalue, query seq, subject seq, q. start, q. end, score, subject sci names, subject title

10 hits found

Spacer 2:

>JQ0001000126.1_6927_10067_28_spacer_9569_35
AGTCTATTTTGCCTGCTAACGGTTTCAGCATCTC

Hit coordinates:

ContigID: AOUX01000192

Start: 3802

Stop: 3836

Intergenic sequence:

>gi|456831403|gb|AOUX01000192.1|:3205-4402 *Leptospira interrogans* serovar Naam str. Naam ctg718000007680, whole genome shotgun sequence

```
TTAGCTCCGATCTGTCCAACCTTTTTAGGGCGTTTTTTTTAATAGTTTTTTTGGTCTGAATTACGTTGGT
CGCGCGATCAGGATAATATTTTCAACCAAATCAACCGTATACGGTGGTAAGACTTTTGACTTCTCGTCC
GTAATTTCAATTTCTTGCGAATCGATCCTTTAATGGTAGCAAAAGGCAATTCAACAGTTTTTTTATTTA
CAAATATTAATTTTTATTATTCAGAATATATTTACGTATTTGTTGTCCAAGCTTTCTTTCAAAAAC
TAATCTCGTTCAGTATTTGTAATTTAGAACGAACAGTTTCAACTCCGCGTTTTTTGTCCCCGTCGATG
CGTTTTAGATCGGATTCGATCTCGATTAATCTTTGAAGTCGTTATTAAGATCCTCTAATGTTTTAGGCG
GTCGGGTTGAACCTCATCGACAGTCTCCTTATTGATTTTTAATCTACTTGCCTCGTTGTCGACTAATA
GTAATTTACGTTTCCCTTTTGGGACCTTTGTTCCCTTTTTTTCGATCGCCTCAATAACAAGGAGCGATTT
AGGTCCTGGGATTAATCTGATTGTTTTATTGTAATCGGAGATGCTGAAACCGTTAGCAGTGCAAAATAGA
CTCATACGGCTATATAATTTGCCTATATGCGCCTTCAAGTTCTTCTCGGTGGCCTCCGAAAGTAATCTTT
CTGACTTTTGGGTTTGTGTTGCTCATTGTACATCTATAAATTATAATTACTACTATGGTTAGGCATAAT
AAACGTATTATAAAAAGCTCAAGCATCTAGATCCCAGAAAATATAAAACACAGTTACCCAAATAGCAGCA
GCTAAAAAGATCAAGAATGCTTGTAAAAAGAGATACGCTGTCATAACTTTACAACCCCAAGTGGATT
AACGGAACCCCTGTGATGATGAAGGGCTATAGCGATCAATGAGGCGACGCACTTAAGGCTTGTTTTTTT
GCGTCGGACCGATCCATGTGGAGCTCGTCACCATCACCATCTTTCTGAAAAAATAGATCTGAATGCTCTA
TAGTTTGTAAAGCACGTCTGAAAAGATCATATATCGGCCTCCACCAAACCTGTTTATTGTTCCAAATC
TTAACTGCGTCCGCAATTGATTGCGCAATATTGTGGTTACGCGACAAGAGTGACATCTAACGGCATAGG
CCTTGTTTC
```

Top 10 BLASTX hits: into hypothetical protein

```
# Fields: query id, subject id, subject ids, subject length, s. start, s. end, evalue, query seq, subject seq, q. start, q. end, score, subject sci names, subject title
# 10 hits found
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|446021850|ref|WP_000099705.1| gi|446021850|ref|WP_000099705.1| 176 1 145 8.09e-94
MSSTRPPKTLLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MSSTRPPKTLLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL 436
2 735 Leptospira interrogans hypothetical protein [Leptospira interrogans]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|488052634|ref|WP_002124031.1| gi|488052634|ref|WP_002124031.1| 176 1 145 3.59e-93
MSSTRPPKTLLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
```

MSSTRPPKLTLEDLNKQLRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL 436
2 730 Leptospira interrogans Gam-like protein [Leptospira interrogans]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|516461609|ref|WP_017850447.1| gi|516461609|ref|WP_017850447.1| 176 1 145 8.23e-87
MSSTRPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MSLTQPPKLTLEDLNKRLQRLIKIDSDLKRIDGDKNAEVEVTRSKFTDTERDLVFERESLDKQVREYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKTALKSWTGAEL 436
2 688 Leptospira interrogans hypothetical protein [Leptospira interrogans]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|490915838|ref|WP_004777741.1| gi|490915838|ref|WP_004777741.1| 176 1 145 1.75e-79
MSSTRPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MSSTQPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNTEVESVRSRFDTERELVFEKENLDKQVREYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKTLPPTYVDLVEKYYSRPNVVIQTKKTIKKAALKNWTNAEL 436
2 640 Leptospira kirschneri Gam-like protein [Leptospira kirschneri]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|487897388|ref|WP_001970854.1| gi|487897388|ref|WP_001970854.1| 178 1 147 1.00e-69
LSMSSTRPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MSTTQNPQPKLTLEDLNKMQRLVEIESDLKRIEAGEKNSEVESVRSRFDTERDLVFEKEELDKQVRDFVQNKDTLFAHRKTVELPFATIKRIDSQEIEITDEKSKELPPYSVDLIEKFYPERANNAIQIKKTVKKTALKSWTDAEL 442
2 575 Leptospira interrogans Gam-like protein [Leptospira interrogans]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|1001638526|ref|WP_061216057.1| gi|1001638526|ref|WP_061216057.1| 176 1 145 4.86e-69
MSSTRPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MNAAQTPKLTLEDLNARLQRLVEIEANLKRIGGDKNTEVESVRSRFDTERDLVFEKEQLDRQVREFVLQNKDTLFVHRKTIELPFATIKRIDSQEIEITDEKSKTLPYSVDLIEKFYPERASDAIQIKKSIKKAALKSWTDAEL 436
2 571 Leptospira santarosai Gam-like protein [Leptospira santarosai]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|490641537|ref|WP_004506532.1| gi|490641537|ref|WP_004506532.1| 178 1 147 1.55e-67
LSMSSTRPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MNRAPNPQPKLTLEDLNKMQRLIEIKSDLKRIEAGEKNTEVESIRSRFVGVGERDLVFEKEELDKQVREFVQNKDTLFVHRKTIELPFATIKRIDSQEIEITDEKSKTLPYSVDLIEKFYPERANNAIQIKKSVKKTALKSWTDAEL 442
2 561 Leptospira weilii Gam-like protein [Leptospira weilii]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|515129273|ref|WP_016758123.1| gi|515129273|ref|WP_016758123.1| 177 8 146 1.70e-65
PKTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
PKSLEDLNIRMQRLEIESDLKRIEAGEKNTEVEDTRSHFVEVERDLVFEKERLDKQIRAFVNMENKNTLFVHRKTIELPFATIKRIDSQEIEITDEKSKTLPYSVDLIEKFYPERANNAIQIKKSVKKTALKSWTDAEL 418 2
547 Leptospira MULTISPECIES: hypothetical protein [Leptospira]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|516465861|ref|WP_017854324.1| gi|516465861|ref|WP_017854324.1| 103 1 103 6.90e-64
MSKQTQKSERLLSEATEKKLKAHIGKLYSRMSLFCANGFSISDYNKTIIRLIPGPKSLLVIEAIEKKGTKVPGKRRKLLLDNERASRLKINKETVDEFNPTA
MSKQTQKSERLLSEATEKKLKAHIGKLYSRMSLFCANGFSISDYNKTIIRLIPGPKSLLVIEAIEKKGTKVPGKRRKLLLDNERASRLKINKETVDEFNPTA 725 417 529 Leptospira
interrogans hypothetical protein [Leptospira interrogans]
gi|456831403|gb|AOUX01000192.1|:3205-4402 gi|516474782|ref|WP_017863226.1| gi|516474782|ref|WP_017863226.1| 176 1 145 1.48e-61
MSSTRPPKLTLEDLNKRLQRLIEIESDLKRIDGDKNAEVEVTRSKFTNTERELVFERESLDKQIRKYILNKNLIFVNKKTVELPFATIKRIDSQEIEITDEKSKVLPPYTVDLVEKYYPDRATNVIQTKKTIKKNALKSWTDAEL
MNAAQTPKLTLEDLNALHQLRLVEIEANLKRIDGDKNTEVESVRSRFDTERDLVFEKEQLDRQVREFVLQNKDTLFVHRKTIELPFATIKRIDSQEIEITDEKSKTLPYSVDLIEKFYPERVSDAIQIKKLIKTTLKSWTDAEL 436
2 521 Leptospira santarosai hypothetical protein [Leptospira santarosai]

Spacer 3:

>HG916826.1_842637_846204_26_spacer_844192_32
TACATCAGCAAGGACGGCCCGGACATGATCCC

Hit coordinates:

ContigID: LMMB01000004

Start: 84179

Stop: 84210

Intergenic sequence:

>gi|944790866|gb|LMMB01000004.1|:84017-84790 Pseudomonas sp. Leaf83 contig_12, whole genome shotgun sequence
TGTACAAGCCAGTATTTCCGGGGTTC TAGCCGGATCGCTTCTCAGATGGGAGCGTTTTGGGAACACTTT
GGGAATGGTGGAACGAAAAAGGCAGCCTGCGAGGGCTGCCTTTTGTCTGTTTCTAGGGCCTGGATTCACAG
CTTTAGCGCGTCTCCAGGAGCGGGATCATGTCCGGGCCGCTCTTGCTGATGTACTTGCCGTAATGACGT
TCGAGCATTGCGATTGTCTGTGCTTCCATACGGCTGACGCAGGCTGGGCAGAGGCAC TTGTAACCGCCG
CTGTTGGTGGTCATGGCTGGGGCTTCTGTGCAGGGCAGGATGGCGCCCACTTCGGCGTGGCGCGCGGC
GATGTGCTGGCGGTGGGCGGTGATGATGCTGGCCAGCTCGGCTTCGTCGATTTGCCCGTCTTCCAGGGCC
TTGAGGATCATCGCTCGACCTTGCCCGCTTTACCGCCGTGCCCATGGCGCGGGCCAGCAGGTCAATGT
TGTCCTGTCTGTGTCAGCTCGGGCAGGGGACGTAACCGCCGTTGTAGAGGCTGGTGAGGTAGTCGGGCAGA
TAGCTGGTGCCGGCGACCTTCTCCAGCTGGCGCAC TTGTTCTCGTCCAGCGGGCGGTTGCCGGGGTTCT
CGTACAGCTTGTGTCGAAGTCTGTTGAGGTCCAGCCCCGGGCAGGTGGCGGCGCATTCGCGGCCCCCTGG
GAAGGCGCCCCGGTGAGGGGGAATGCGAGGGCGATATCGGCAAGGCCGTGCTGGTGGGCGATCCGTTGTG
AATC

Top 10 BLASTX hits: into hypothetical protein

Fields: query id, subject id, subject ids, subject length, s. start, s. end, evalue, query seq, subject seq, q. start, q. end, score, subject sci names, subject title
20 hits found

Query ID	Subject ID	Subject IDs	Subject Length	s. start	s. end	evalue	Query Seq	Subject Seq	q. start	q. end	score	Subject Sci Names	Subject Title
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 980973412 ref WP_059392489.1	gi 980973412 ref WP_059392489.1	156	83	156	1.47e-63	LPELTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASII TAHRQHIAARHAEV GAILALHRKPQP	MPEPADTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASII TAHRQHIAARHAEV GAILALHRKPQP	515	294	355	Pseudomonas toyotomiensis	hypothetical protein [Pseudomonas toyotomiensis]
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 980973412 ref WP_059392489.1	gi 980973412 ref WP_059392489.1	156	18	85	1.47e-63	GAFPGGRECAATCPGLDLKQFDNKL YENPGNRPLTDEQVRQLEKVAGTSYLPDYLTSLYNGVYVPCPS	AAFPGGRECAATCLGLDLKQFDNKL YENPGHRPLTDEQVRQLEKVAGTSYLPDYITGLYNGVVFAMPE	709	506	327	Pseudomonas toyotomiensis	hypothetical protein [Pseudomonas toyotomiensis]
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 545127478 ref WP_021488921.1	gi 545127478 ref WP_021488921.1	155	82	155	2.10e-63	LPELTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASII TAHRQHIAARHAEV GAILALHRKPQP	MPELTDNIDLLARAMGTAVKRGKVDAMILRALEDGQIDEAELASII TAHRQHIAARHAEV GAILALHRKPQP	515	294	367	Pseudomonas mendocina	hypothetical protein [Pseudomonas mendocina]
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 545127478 ref WP_021488921.1	gi 545127478 ref WP_021488921.1	155	18	84	2.10e-63	GAFPGGRECAATCPGLDLKQFDNKL YENPGNRPLTDEQVRQLEKVAGTSYLPDYLTSLYNGVYVPCPS	AAFPGGRECAA-CLGLDLKQFDNKL YENPGHRPLTDEQVRQLEKVAGT TYLPDYLTDLYNGVVFAMPE	709	506	313	Pseudomonas mendocina	hypothetical protein [Pseudomonas mendocina]
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 917193888 ref WP_051800600.1	gi 917193888 ref WP_051800600.1	156	83	156	2.14e-62	LPELTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASII TAHRQHIAARHAEV GAILALHRKPQP	MPEPADTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASII TAHRQHIAARHAEV GAILALHRKPQP	515	294	355	Pseudomonas oleovorans	hypothetical protein [Pseudomonas oleovorans]
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 917193888 ref WP_051800600.1	gi 917193888 ref WP_051800600.1	156	18	85	2.14e-62	GAFPGGRECAATCPGLDLKQFDNKL YENPGNRPLTDEQVRQLEKVAGTSYLPDYLTSLYNGVYVPCPS	AAFPGGRECAAACLGLDLKQFDNKL YENPGHRPLTDEQVRQLEKVAGTSYLPDYITGLYNGVVFAMPE	709	506	317	Pseudomonas oleovorans	hypothetical protein [Pseudomonas oleovorans]

gi 944790866 gb LMMB01000004.1 :84017-84790	gi 835622920 ref WP_047589232.1	gi 835622920 ref WP_047589232.1	156	83	156	8.82e-60
LPELTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASIIIAHRQHIAARHAEVGAIALHRKPQP	MPELADTDNIDLLARAMGTTIKRGTVDAEILKALEDGEISEAELASIIAAHRQHIAARHAEVGAIALHRKPQP					515
294 331 Pseudomonas mendocina	hypothetical protein [Pseudomonas mendocina]					
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 835622920 ref WP_047589232.1	gi 835622920 ref WP_047589232.1	156	18	84	8.82e-60
GAFPGGRECAATCPGLDLKQFDNKLLENPGRPLTDEQVRQLEKQVAGTSYLPDYLTSLYNGVYVPCP	AAFPGGRECAATCLGLDLKQFDNKLLENPGRPLTDEQVLQLEKQVAGTSYLPDYISGLYNGVYVAMP					709
509 318 Pseudomonas mendocina	hypothetical protein [Pseudomonas mendocina]					
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 503480628 ref WP_013715289.1	gi 503480628 ref WP_013715289.1	159	18	84	3.53e-53
GAFPGGRECAATCPGLDLKQFDNKLLENPGRPLTDEQVRQLEKQVAGTSYLPDYLTSLYNGVYVPCP	AAFPGGRECAATWGLDLKQFDNKLLENPGRPLTDEQVLQLEKQVAGTSYLPDYISGLYNGVYVAMP					709
509 313 Pseudomonas mendocina	hypothetical protein [Pseudomonas mendocina]					
gi 944790866 gb LMMB01000004.1 :84017-84790	gi 503480628 ref WP_013715289.1	gi 503480628 ref WP_013715289.1	159	83	159	3.53e-53
LPELTDNIDLLARAMGTAVKRGKVDAMILKALEDGQIDEAELASIIIAHRQHIAARHAEVGAIALHRKPQP*PP	MPELAELDNIDLLERAMTTTIKRGTVDAMILTALKDGEINEAELASIIAHRQHMAARHAEVSSILALHSKRQEPKP		515	285	279	Pseudomonas mendocina
protein [Pseudomonas mendocina]						hypothetical

Spacer 4:

>LFQC01000004.1_87106_87459_1_spacer_87136_35
GTAACAGATGATATTAATTCAAAGAATTAACCT

Hit coordinates:

ContigID: AOSX01000029

Start: 620311

Stop: 620345

Intergenic sequence:

>gi|727535586|gb|AOSX01000029.2|:619946-620968 Clostridium botulinum Af84 Contig_29, whole genome shotgun sequence

```
AAATTCGGATAGAAACATAGATATTTCTCATATAGATTTTTTTGATTAATTATATGCACTGATGCTATGA
ATAAAAACAAAAAATTCTAAAAGTTTTCCAACCGCATTAAATTTTAAAACCTATCCTAAGCATTGCAA
TATAAGGCTTAAGATAGGTTTTGTTTATGTTTTCTGAAAAATCACAAATGGCTGGAAAAATTTTTTTAA
TCATTGTAAAATCAAATGCATATGCTATTCCTTAAAATTAAGGAATGGCTAATTTACTATGGATGAACAT
TAACATAGGATGATTTAAATTTGGAAAAGCTTACTAGATTTCTTACCTATATCGGGTTGAACAATAACAT
GAGATGCATTTAAATAAGTTAATTCCTTTGAATTTAATATCATCTGTTACACTGAACAATAACATAAGAT
GTATTATTA AAAATGGTAGTTGTAGCTTATGTTATAGCTATTATTTTTCAATATTTAGAGGAATTTTTT
ACATATGTAGAATATTTGTATATAGTCCCTTTAATTCATTATCAAATCTTCTTGAATAAAAAGGAGCCT
GTGTATGCAGTGGGGGCTCCTTTTTTATATTATTTGAAGGAATTTTTAACATTTATAGAATATTAATA
TAACGGCTTCCCAATAGGTTGATTATAGACCTCTTGCTTAAGCAAAAAGAACCCCAATAAAAAGGGGTT
CTTTTTGTATGGATTTATATTGGTGCTATGTATTGGTCTATTTTTACTATATCCAGATGTAGAAAAGT
TAATCAGTAAGAAATATATAAGATGTAAATTTTTAAATGTATTTCTATAATGAGAGTACCTTATTTTTTA
TTTTTTTGATAAAAATTTAAGAAAATTAGCAATTTATTAATAAAAATTTAAAATAATTATTTAATTTCA
TTTTAAAGTTGTATCAGAATGTATAAAAAGTAAAAGATGTAGGATAAACAAAGAAGGAATTGAATTA
TTGTAAGAATATAAATATTATGGTAAAAATATAAAATAAATA
```

Top 10 BLASTX hits: (1 hit found) into hypothetical protein

Fields: query id, subject id, subject ids, subject length, s. start, s. end, evalue, query seq, subject seq, q. start, q. end, score, subject sci names, subject title

1 hits found

gi 727535586 gb AOSX01000029.2 :619946-620968	gi 1119667520 ref WP_072587154.1	gi 1119667520 ref WP_072587154.1	80	49
79	2.91e-07VTLNNNIRCIKNGSCSLCYSYYFSIFRGIF LKLNNNIRCIKNGSCSLCYSYYFSLFRRI	396	488	136
	hypothetical protein [Clostridium botulinum]			Clostridium botulinum