

Supplementary Information

Single cell genome analysis supports a link between phagotrophy and primary plastid endosymbiosis

Debashish Bhattacharya^{1*}, Dana C. Price^{1,5}, Hwan Su Yoon^{2,5}, Eun Chan Yang³, Nicole J. Poulton³, Robert A. Andersen⁴, and Sushma Parankush Das¹

¹ Department of Ecology, Evolution and Natural Resources and Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, New Jersey, 08901, USA

² Department of Biological Sciences, Sungkyunkwan University, Suwon 440-746, Korea

³ Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine, 04575, USA

⁴ Friday Harbor Laboratories, University of Washington, Friday Harbor, WA 98250 USA

⁵ Equal contribution made by these authors.

* Corresponding author: Debashish Bhattacharya, Department of Ecology, Evolution and Natural Resources and Institute of Marine and Coastal Science, Rutgers University, New Brunswick, New Jersey, 08901, USA; phone: 732-932-8193, fax: 732-932-8746,

Email: bhattacharya@aesop.rutgers.edu

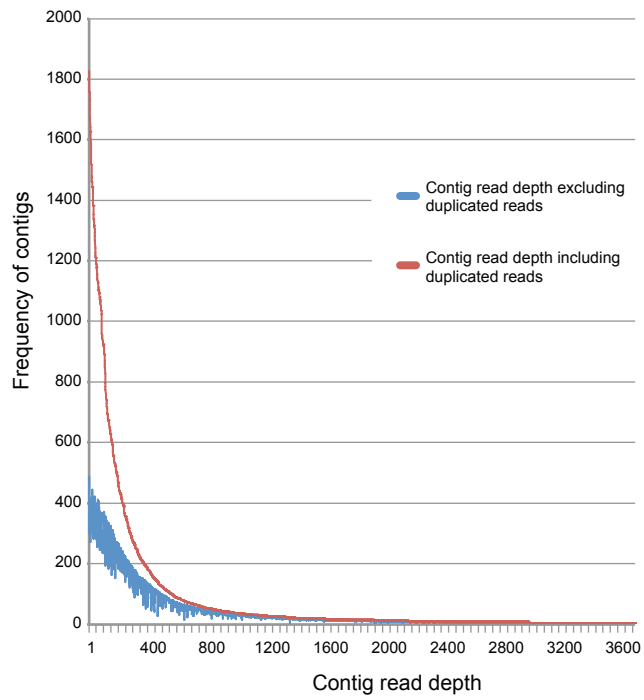


Fig. S1. Analysis of contig read depth using the 454 data. Normalization of reads (see Methods) provides a correction for WGA bias, in particular for abundant contigs that have lower genome coverage.

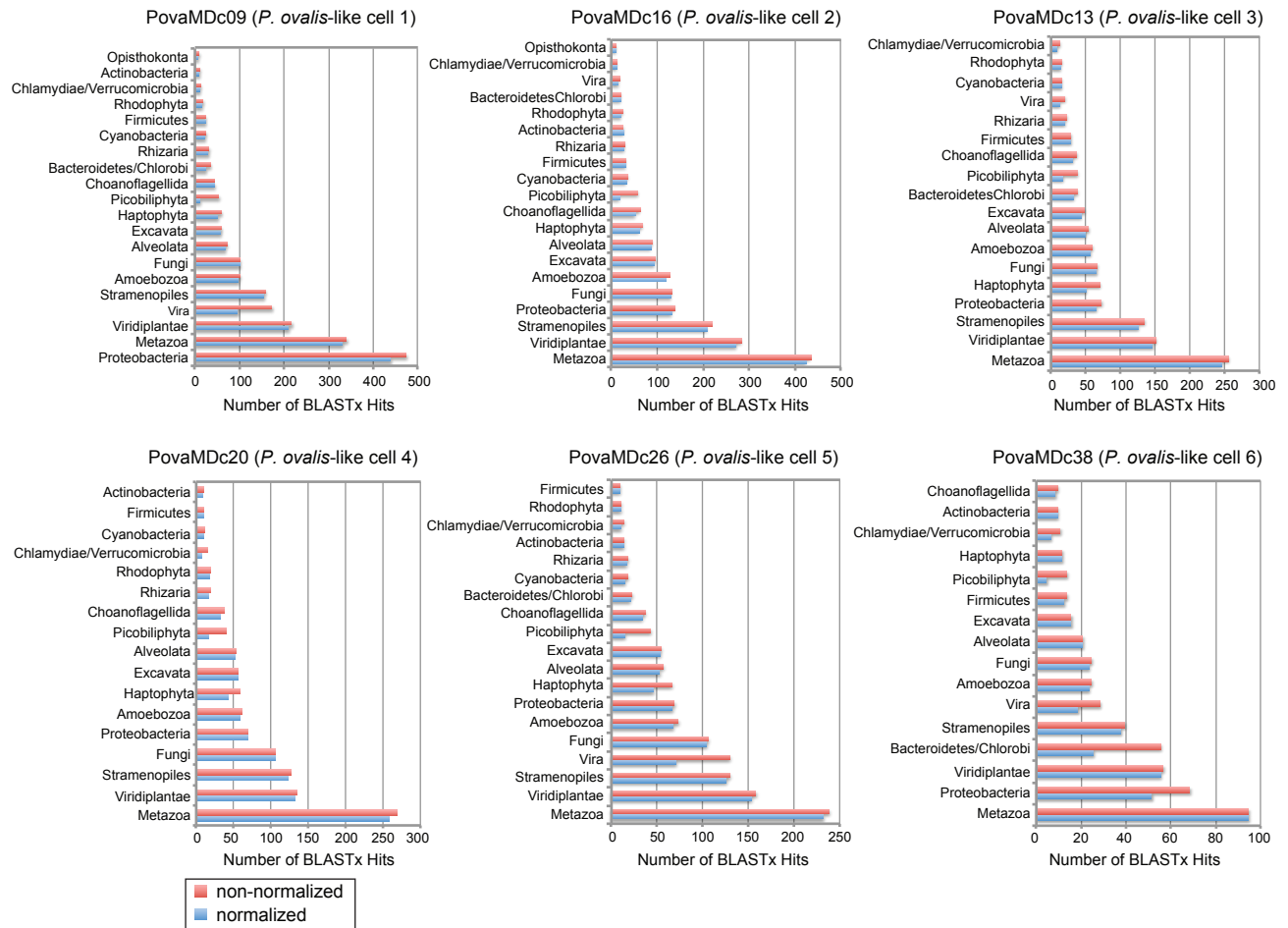


Fig. S2. Analysis of *P. ovalis*-like SAGs. Taxonomic distribution of the number of non-normalized and normalized (see Fig. S1) BLASTx hits (e-value $\leq 10^{-10}$) using the contigs from the six *P. ovalis*-like single cells for which we have 454 data.

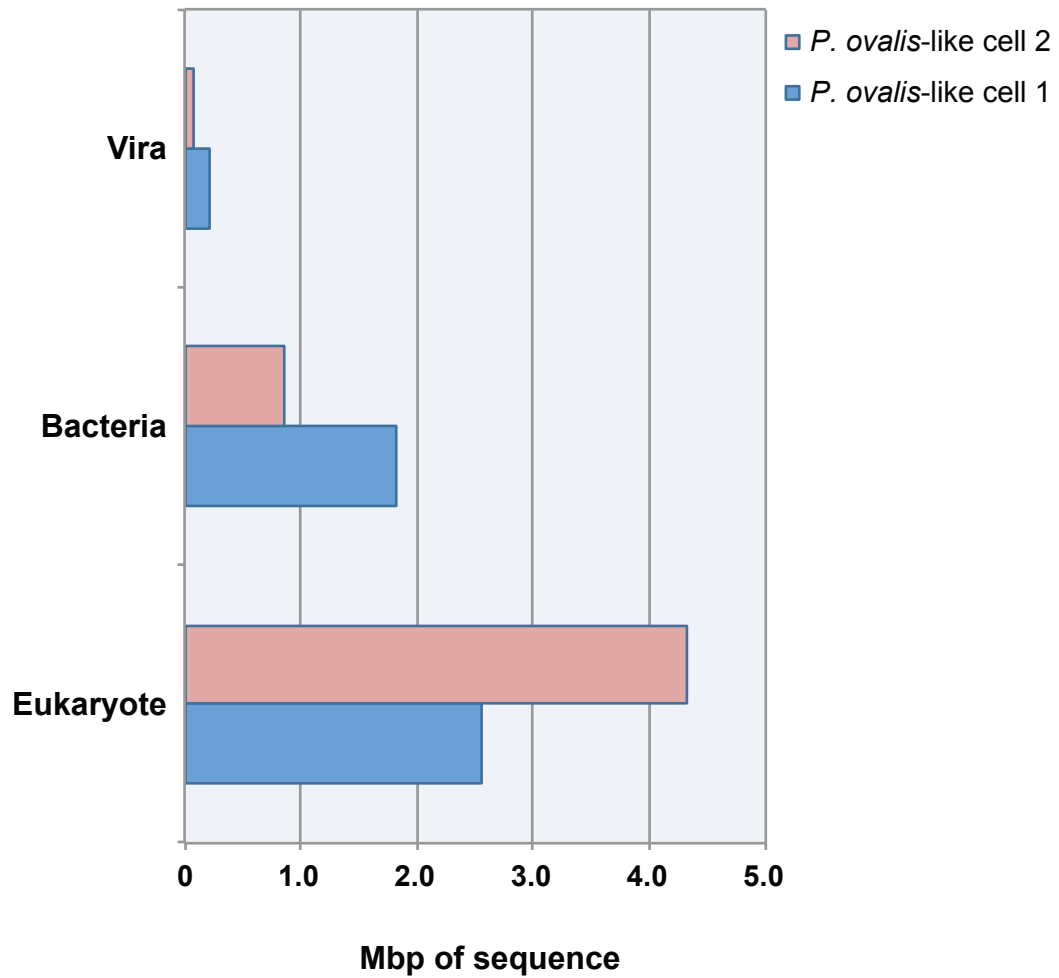


Fig. S3. Total megabases of DNA that encode proteins with BLASTx hits in the *P. ovalis*-like cell 1 and 2 joint assemblies. To estimate the amount of coding DNA in the joint assemblies, we determined the number of nucleotides encoded on all contigs that had significant BLASTx hits.

Fig. S4. Contig7191 from the *P. ovalis*-like cell 1 assembly of the 454+Illumina data that encodes proteins of α -cyanobacterial origin. The MAKER 2 web annotation server (<http://derringer.genetics.utah.edu/cgi-bin/MWAS/maker.cgi>) revealed 8 putative proteins that are shown below the contig sequence.

>Cell 1 Contig7191, Average coverage: 7,576.65

```
TATGCAATGTTTGATCGTGCGGTAATGAAGAAAAAAGATGGGCAGATTATTTGTTTAAAGAT
GGTCTATGATTGGACTGAATGATAAACTTCTTCAGCAGTATATTGAGTGGATTGCAAATCGT
CGTATGAAAGCAATTGGACTTAAACCAGTATATGATATTCGTGCAAATAGCAACCCTCTCCC
TGGACAGAGCATTGGATTTCTTCTAAAGGACTTCAAGTTGCTCCTCAAGAAACAGAAGTTGAA
TCTTATGTAGTTGGGGGTATCAAGCAAGATGTTAAAAAAGATACATTTGCTGGATTAAACTT
TAACTTTTATAACGCCCATTAGGGCGTTTTTTATTAAATAATTAATTATCCATTAATAATATGC
TTGAAGTAAAAATTTTAGGAGACAAAGTTCTTAAAGAAAAATCAAAGAGGGTGTGATAAAATT
GATGATAGCATTAAAAATATTTGTGCTTCTATGATTACTACTATGCTATCAAAAAATGCTGTTG
GTTTAGCAGCAAATCAAGTGGGAATTTTAAAGAGAATTATAATTGTTTTAGATGAAGACTCTC
CTAAAGTAATGATTAATCCAGAAATAATTGAATATTCTAGTGAAACTTGTAACTATGAAGAGG
GATGCTTAAGCATCCCCAATACACTTATGGATATTGAAAGACCAAAAATTTAAAATTAAGT
ATAGAACTTAAAAGGTAAGCCTTGTTACGAGACATATGACGGACTTACTGCTAGAATCATTG
AGCATGAGGTTGACCACTTAGATGGAATAACAATGGATGTTAGAAAAGAGGTAAAATGATTC
CAAAAATTATTTTCCAAGATTCAAATTATGATGAATGGTGTGAAGAACAATAACAGAATGCAT
ATAGATTAGCTGCTGAATATGATGAATTTATATTTGGTGACTATGATTATATTACTGACTGGAA
GCAAACCTTCTCTAATGACATAAATTAGTATCTCTAAATATACCAATAAGAATGTTGGTATATT
GATGAGTTGCGACTATCCCAATCCATGGTATTATTGTGATAATCCCTTTACCTCTGATGATATT
GGGGATTATTATGGGTTTGTACCTTATTGAAAATACCAATAACAATCGTAAATATATTGGT
AGAAAATATTTTGGTCTTTAGAAAACCTAAGGGAAAAACAAGAAAAGTAAAAACGGAATC
TGATTGGAAAAACTATTATGGGTCTTGTCCAGAACTTAAAGAAGATATTAATAATATGGAAG
AGGGATTTTLAGAAGGACTATCATCTCCCTCCACAAAACAAAAGGAAGAACAATTTTGAAG
AAACAAGACAACCTTCTCCTTAAACAACGTCCTCACAGAAGGACTTGACGATGGAACCCCAAGA
TACTACAATTCAAACATCCTCAGTAGATACTTCAGAAAAGATTATTATAGAAAATGTGATTGA
ACCTCTTCCTTATATCAGGGAATGGTCTATTGATAGAATTAATTTTTTGTAAAGTGATATTGAA
GACAATAAGGAAAAATTTCTGATGCTATTTCTATTGCTGAAGAATTTGATGAGTGGATCAAT
ATTACTGATGATGAGATAGAATGTCTTTTTATCAATAATTCTGATTGGGTTGAAAATGAAACC
GAGTGAATCATTGACAAAATCCTAAATAATCACTTATTATGTAAAATCCCTATTATGATAGG
GTAATCGTTATAAGTCTTTGACGTGACATTTAGAGCCGTGGGTAAGTCCCTGAGAAGGGGAA
ATTCTCCTTTACCTATACGGATGTAGAGTTCTATTAATAATAATGCTAAGAAACCTTTCAACT
GTAGCCGTATCTCTTTTAGGTGCGGTTGCGACATCAGCGGCAACATTGCCAACACCGGGTAT
GCTACTTCAAGTTCTACTCCACAAATCGCAATTTTAAACGGAAGAAAGCAAAACCCAGGAGAC
AACGACCAAGAGGTTGCTCCCGAAAAGCCAAAAGAAAAAAGGCTAGTTTGTAAGGATGTA
ATTCAAATGAAGCACGTACTCTAGAATTTCTTCAGAGTCGTGGAATTACTGATAGAAATGCC
TCGCAACCATTATGGGAAATATTCGTCAGGAATCAACATTTTTCCCTAATATCTGCGAAGGTG
GTGCTAGAGTTTCTACAACCTTGTAGAAAGTGGTGGGTATGGTCTAATCAATGGACTAATG
CTCCAAGGTTCTATGGGTTAGGGAAACACGCAGCACGTATCGGTGCTAATCCTTCTCCTACTAG
ATGCTCAACTTGATTATATGCTTCATGAGGGTGACTGGAAGATGATTGAACCTTATATGAAGA
CACCAGGCAAGACCATTAATCAATATATGCGTCTTGCTCAAAGGTGGATACGTTGGGGTCATC
ACGGAGCCAGAACCGATTATGCGTATAATTATGCGAATAAGTTAGTTCTAACTGAAGTGGGTT
CATAAAAATAAATAACAACCTAAGGAGTGCTTCGGCACTCCTTTTTTATAAATAACTAAAA
AGTATTTTTTAAATGAAAATAAAGAAATTCTGGGTCTTTATAATGCTTATGTTCAAATTCATA
CTGAACAACCTTGATGAGAATAGAGCAGCTGCAAGAAATCCTGAAGAGTATGAGCATAACAAA
GCAAAGAAAACGGATCGTAGAACTGCAGCAATGAATGATCCTCATAACAGGAATCAATTCTCC
TGCTTTTCGTAGATTTATGGCACAGCAAACCTGGTAGAAAATAAAGGAAGATTATGATGTTA
TGATCTCATCCTTTCATACCTTCTTTATGAAGGTTATGCTGAAACACCAGAAGCAGCAGAATC
CATTATGGTGAATATGAGTGAAGATTGGAGAGAGAGTATTGTTGAGGAACTTGAGCAACTTG
ATGAGATTATGAAACATCAACAAGAAAGAGATAGAATTGAAGGAATGAAAAGAAATCTTCT
AGGGCAAAAACATGGTCTTATGAACCACCAACAAAAGGAAAGCACGGTGAGATTGAGCGTCG
```

TCAGGCACAAAGACCACCTTCAAGATTGAGAAGTGGTAAACTTCCAAAGGAAAGAGAGGAGG
GTTGAGTTCACTTTCTAACCCATTCATAAGGTGCCTTATGGGCACCTTTTTTAAATGACTTGACT
TAAGTATCTAAAATCAATATAATTACCTTTGTTAAGGTTGAAGGGTTATATAAGATCAAAGCT
TATAGATATATAAGAAGATAAAGTCTTATGTATCAGTATAGTCTTAAAGATATTAGAGTAATT
GATGGTGATACGATTGAAGCATCAATTGACTTAGGATTTAATATCTTTCTAAAACAATCTATTC
GTCTAACAAATATTGATGCTCCTTCAATAAGAACACTAAATGAAGAAGTAAAACAATATGGA
TTTAGAGCAAAAAATAAGCTACAAGAGTATATTAACCTGGAACCGAAGTAATTGTAGCAAC
TCAAAATCCAAATAAGACTGAAAAGTTTGGAAAGAGTATTAGGTGAGGTTTATGTTCAAGGAC
ATAATCTAACCGCAAGTGAATACTTATTCGCAAATCAGTATGCTTGGGCATATGACCCATCAA
AAGACAAAGACCTATCAGAACTCGCACCGTTAAATGAGTAATCTACCAGAAGACCCAGAAGA
GATTATTGATTTAGTTGCTAAATCTCAGTATTTAAAAGTCAATACTGAGTTAGGTGAGATTAA
ATTAGAAGCACATAATGAAATTAATCTTCAACCAAAGGAACCTCCTTGGGAATGAACTTAA
GTTGGAAGAACTGGAGATATTACACCTACCATTACATTTGATACAAAAAACTAAGAGAAC
CTTGTA AAAATATTGACCTTGTGAAATAGTAGATAATGCTATCGAAGAGTATTTAAATGACT
AAGAAATCTAAGTTTAATATTTTTAAATTTCAATTTGGAGAAGTCAAAGAAGTATCTTTTCATT
ATGTAAGACTGGATTTCTTCTCAAAGGAAGTGTAGATGTTGTATCACTACTTCCATTTGTAGA
CCGTAAGTTTCTTTAACTTAGTTGATGAAATACAACCTAAGTTTGGTGTGATATTCTAAAC
GATTATATTATCAAAGATGACGAATTAACCTCCGTTTTCGTATAGATAGAACAATAGAGAAAGCA
CTTAGGGATTATGAGGATACCTGAATTTCTGGTGAGTTTCAGACTGTCTCCTAAAAGGTCTCTG
ATTGCTGCTCTAATTGCTTCTGGGATTGCCTTTCTCACTGTAAGGTGTGGTATCAGTCAGGAAG
ACCTTCTGAAGTTCTATAATGAACTCAGAAAGGAGATTAAGTTTAACTTACCTAATGAGAATA
ATGTTATCAATGAGATTGATAGACAACCTCAACGAAAGAATTAATAATGACCCAAAGTTACTTG
AATTCAAAGTTTCGTGGAGAAGTTGATGACGCAATCAGTCGTTATGAAAGAGAAGAATACCAA
AACCGTCAAATCAATATGAAAAACAAGAATATTTTGGAAAGAGATTAATAAACCCAAAGTACGA
TAATTTACAGAACTGATTGTGGAAAACGCAGTTTATTATGAGTTTGAGGATGGAACCTATGGG
TATTCGTGGTGCTTGGGTTTCTCCAGACCCACGGGAAATAGATTTAGAGCAATAAAAAGTGTTA
TATATAACATATCTTATTTTTATTGGAGATTATAATGTCTGAATCAGTAACTAAGCTTGTAGA
AGCATTGAAACTTGAAATCCGAAGATGAAAAGTTTGTAGGTGGTAATGGTGCTGCAGGAA
CCCGTGCTCGCAAGGCACTACAGGAAATCGCTAACTTGTAAAGAAAGGAGAGCAGAAATT
TCCGAAGAAAAGAACGCAAGAAAGGAGACCAAGGGGGCTTGACGGACCCGTGCCCTGTGGT
ATTCTATGAGGGTGGTTGAGAGACCACCAACTAAATCCTAAGTTTCTTAGGTCAAGGGAATGG
CCTCCCTGTGTCAGATAAGGTAAGGAAAGTAAAAGGAGAATGGGAACGCCTGTGGGATGATAC
CTCACCTGCCTTATCTACCGATAGAGGTTATGCCTCTGTTATATTCTTGAGATATAACACGCAT
AATCCATCAATACCCTCTATGCCTATCCTTCAAAAAATAATACTCGACAATTCTAGATGCGTC
GAGTATTATGAAGTGAAGCACAACCCAGAGGGTCTCTTGTCTCAGTAGCTCAGTTGGATAGAG
CATCTGCCTTCTAAGCAGTTGGTCGGGGTTCAAGTCCCTCCTGAGACGCCAGGGGAATTAGC
TCAGTTGGTAGAGCAGGGTCTTTGCAAGGCCAAGGTCAGGAGTTCAAGTCTCCTATTCTCCAC
TTGACAATCTAAGGTCTATGCCTTATGATTGTCTCATAACCACAAGGGGGTATAGCTTAATGGTT
AGAGCGCCTGCTTATAACGGGTTAGTCTGGGTTCAACTCCAGTATCCCTATTCCCACTTAT
GTTGGGTTTCTGCCCGTGTATCCCAATCGGCAGAGGAAGCGGACTTAAAATCCGTTCAAGTGC
GGGTTAGTTCAGCGGTAGAACGCTATCCTTCCAAGTTAGATGTCGTCGGTTCGATTCCGATA
CCCCGCTCTTGACTTTTTTCAAAAAAAGTCTTATAAATAAACACATAACTGATTAGAACAATG
ACCTACCCAATGCCACAAAACAGTTTAGTCTCCTTGATTGCCGCTATTGGCATATTGAGGGT
ACTCCCCTGTTTGCCTGTATGGTAGGTCAAATGTAAGACCGACCATAAAAAGCAAAAACACACA
GGGGAGAGAACTCAAAGTTTCCCTCCCCTTTTTCTTTGCCTTGTGACGGTTTCAAAAGTGTCCA
CCACCTACTCCCAAGAGACCAAACGGTGGTATTCTAATCAAGTGGTCGAGAGAGGACCACA
CGAACCTTGA AAAATTTAAACCTTAATGGGTCTGTA ACTCAACGGTAGAGTAACGGGCTTTTTAA
TCTGGAAGTTGTGGGTTCAAATCCACCAGACCCATTCCCGTTTTCGGGTATCCAGTATCCTGT
CTGGTGAACAATAACACAGGTGGGGATTTTTAGGTTATCCTCGTCGGGGTATTGTCGCTC
ACCCTAAACCTAATAACGAAGTTAAAGTTTCTGCGTTAAAGTCGTGAAACCTAGCAAGAGATA
ATTTCAAATGCGATTGGTGGGTTATGTCCACCTATTTGGGAGAGTGGCTACTGTTGGCAATA
CGTGTGGCAGCGGTCTGTA AAAACCGTTACATAGGAACCATCGGGGGTTCAATCCCTCCTCTC
TCACCTTGACCTATAGTGAAGTGGTCATCACGATACCCTGTCACGGTATTATCACGGTTCA
AATCCCGTTAGGGTTCGTCGTCCTACACTTATTCAGTTTTGTGTGTGGGCTGTAATTATAAACT
GGCATTCAGTATATTATATAATAATGATAAACTGAATGGGCTCTGTA ACTCAGTTGGTAGAGT

GACCCTATTATGAGGGGTAAGTCGCTGGTTCAAGTCCAGTCGGAGCAATTGGGAGAAATCCC
ACGCATGGTCTATCGTCTAGAGGCCAAGGACGCTGCTCTTTCACAGCAGAAACATGGGTTCA
AATCCCCTTAGGACTATATCCAGATGTAGCGCAGTTTGGCTAGCGCACCTGTTTTGGGAACAG
GGGGTCGGGGGTTCAAATCCCTCCATCTGGACCTTGGGAACATAGCTCCAATTGGTAGAGCAC
TTGATTGAAGATCAAGGTGTTATCGGTTCAAATCCGATTGTTCCACTTGACCTTTATGGTCAC
GGAGATTCACCTCCTCCGTTATGAAACTTCGGTTGGTTAGTGATAGGTTTTTCAGTCCGAAAAC
TGAAACTGGTAGGTTGTTGATATCGACTCCTACCTTGATGGAATCGTAGCTCAGTGGTAGAGC
ACTCGGCTGATAACCGAGCGGTCACAAGTTCAAATCTTGTCCGTTCCACTTGGGAGAGCACCT
TGGACGGTGGGTACGACGGTATCAGAGACGGTTCGATTCCGTCCCTTGGCAATGGTGGTTGCA
TTCCACCTTCTCCCTACTTGGAAAGATTGGCAGAGTGGTTAATGCAGCGGTTTGCTAAACCGTG
AGGGTAAAACCTCCGTTGGTTGATCCCAACATCTTCCGCCAGGTTGTGTAGTTTCAAGTGGTAG
AACACTTCTCTCATAAAGGAAGTCGTCGGTGGTTCAAATCCACCCACAGCCACTGTGACGTTAT
CCTAGTGGTTAAGGAACCAGTTTGTGAAACTGGGTAGATGGGTTCAATTCCCATAACGTCACCC
CGCCCTTATAGCTCAGTGGTAGAGCAACTCACTAGTAATGAGTAGGTCGTTGGTTCAAATCCG
ATTGAGGGCTCTTGTTCAGTAGCCAAGTGGTAAGGCGGAAGTCTGCAAACCTTCTAATCGTC
GGTTCGATTCCGACCTGAAACTCCAATGGGGAATCGTCTAAAGGCAGGACACCACCGCTTTGAC
CGTGTCTTATAGTTTCGACTCCTAGTTCCCCAGCCAGTCCGTGTAGCCCAATTGGAAGAGGC
ACTACCTTGAGGGGGTAGGTTGGAGGTTCAAATCCTCTCACGGACACCTTGTGGGTTGGT
GCAACTGGTAGCATCTTGGTCTCCAAAACCAAAGATCAGAGTTCAAATCTCTGGCCCTTCGCC
TATGCCCGTGTACTCCAATTGGTAGAGAGGGTGGACTTAGAATCCATACAGTGTAGGTTTCGAC
TCCTATCACGGGCATGTGCTCTCTTAACTCAGTTTGGATAGAGTGTGGCTACGAACCAAGA
AGTCGCAGGTTCAAATCCTGCAGAGAGTACCATTTTCATTATAAATCTAAATAGTAGGTGTAA
CAGGACAATGGAATTAGTAGAACCACATTCACAATCTGGTTCTGAACAGTTCTTATGAACC
ACTACATTTACAAAATTGGAAAAGAGCAGTTGTTCTTCTTTAAGGAGAAAGCAAGACTCAT
ATCTAAACGTGTTATACGTTTAGTCAATTATGTGAAAATTCCTTCAAACATGGAAAGGATGC
TTATCCTACAAGATCCTTGATTTACAAGAGGGATGATCATGAGTGTCAATACTGTGGGACTAC
AAAAGACCTAACACTTGACCATGTGATTCCCTCGTTCAAAGGTGGACAGGACACATGGGAAA
ACCTTGTTACCTGTTGCATCAAATGTAATTTAAAGAAAGGGGATAAACTTCTTTCAGAAACTA
ATATGACTTTGAAGAAAACCTCCAGAAGCACCATTTAATCGGGTGTATCTAGACCTTCAAAGA
GTAGAGTTCCGAATGGGTTGATTATGTAATTGGTTAAGGATGGTTTGGGGCTTGACGAATTC
AATCCTATAGAATATTATAAAACGGTGGTTGAGAGACCACCTATCGGGAGATTAACCTCAGCG
GTAGAGTTCTTGCCTTACAAGCAAGCAGTCACTGGTTTCGATTCCAGTATCTCCCACTTGACAAT
CGGGTATATATTATACCCTATGATTGTCTCATTGTGGGGAAGTGTAACGGTTGCACAGAAGTC
TCATAAGCTTCAGGTAGGTGGTTCAATTCCACCCCCGCTCCCAAATAACACACAAAACACACA
GGAGAAAACAAATGACGCCTTACGAACTACGCTTTGAAATTTTCAAGCAAGCATATGCTCATG
CTAATGATGAATATCTTGCATCATATAACATAGTTGATAATCATAATCAAATAATGAAAAGA
AATGGGATTATCCTACATTTCTTCTTATGAGAAAATTGAAGCACTTGCTGAAAAAATCAACA
ACTTTGTAAGTTCTAAATGATACAGTGGGGCGGAAACGCCCCCGTTAGTATTCCTTATAGCT
CAATTGGCAGAGCACGGAGCTGTTAACTCTGGGGTTCCTGGTTCGAGTCCAGGTGGGGGAGTT
GCGGAAAGTATCTTCCGCGAGTGGTGGTACTCACTACTCATTGCTCTCTAGCTCAGTGGT
AGAGCACTCGCCTTGAAGCGAGCGGTCATCGGTTCAAATCCGATGGGGGGCTTGACGAATA
CTTTACATTCGTCTATACTGTGACTGCAATTTAAACATTGCTAAATAAACTTCGTAGTTATAAT
TCTTAACAAAATATATGAAACTATTTCAACAACCTGATGCTCGCACCTGTTGCTGTTGGTCTGAT
TGCTCCCGCTGCGACTGCCGCGAAGTCAATATGGAAGGAGTAAATCAGTATTCAACTTCAGA
GCAGGTTACAAGTATCAACCAATTCTCTGATGTTCAACCTACCGATTGGGCATACCAAGCACT
CAGTAATCTTGTGAGCGTTACGGTTGTGTAGCAGGTTATCCTAACGGCACCTATGGTGGTGG
TAAGGCAATGACCCGTTATGAGGCAGCAGCACTCTTCAATGCTTGTCTTGACCGTGTAACCTGA
AGTCACCGATGAAATTCGTAAACTTCAGGAAGAATTCAAGGCAGAACTGATTGTTCTCCGTGC
TCGTGTAGAAGGTCTAGAAGTGAAGACTGCTGTTCTTGAAGCACAGCAATTCTCCACTACTAC
CAAATGAATGGTGAAGTAACTTTGTTCTTGGTGGTGTCTCTGGTTACGAAAGTCGTGGTAC
TAAGGCAGATGGAACCGCATTCAACTATGATGTTCTGCTGAACTTTGATACCTCATTTACTGGT
CGTGACCTACTCCGCACCCGTTCTCGTGCTGGTAACTTCAGTGAACCTCCTTTTGGTTCAAGTT
CTTCCCTGTTCAAACCTGGATAAGGCAGAAACCACTGGTGATACTGTAGAACTTGACCGTATCT
ACTATCAGTTCCTGTAGGTAACGATTTACCGTAACCGTTGGTCCAAAGGTTTCGTAACACCG
AATGGCTTGGGTTCTTCTGCTTATAAGTCTGAAATCCTTGACTTCTTCTACTGCTGGTGC
TCCAGGTGTTTATAACAAGGCAACTGGTGTGTTAGGTTTCAGTATGAAAGCAATGGTTT

CGTTGCTGGTCTGAACTATGTTGCTCAGGATGGTGATAGCACCGAAACTGGTGTATTCTCTTCT
GATGGTGCCCTGAACCTGATGGCTCAACTTGGTTATCGTCAAGATAATTGGGGTGTGGTGTG
GGTTATCGTTATGGTTCTGAAGGCACACGTCCTCGTACCTACAACGGTTTCCAAGGTGCTAAT
GGTACTCTGCTAGGCAATCAGGATTCCAATAGTGTTCGTAATGCTTATTGGCGTCCTACCG
AAACTGGTATCATTCTCAGTCAGTGTGGTTATGGTTATAATGCCGTAAGTGGTCGTGCTGG
TGATACTGACGCAACCGATTACAATCTTGGTTCGTTGGTCTTCAGTGGGATGACGCATTGCTT
GCTGGTAATGCTGCTGGTGTGCTGTGGGTGACCCCGCAAATGCCGAAGGTCTTGCTAACGAT
GCCCTGATGCTTGAGTTCTTCTACAAGTATCAAGTTACCGACAATATCAGCATCACTCCTGCTA
TCTTCTATGTAAGCAATGATGCTGACCAACGTGGTAATGCCTCTGAGTGGGGTGGTGTAAATCC
AAACCAAGTTCAAGTTCTGATTCCATCAGATATAAAATTGACCCACCGAAAGGTGGGTCTTTT
TAATGGTTAACCAAATCCTTAACCGTTTCTTAGTGGACTTAATTTAGGGTTTCTAGTAAACTTC
TTATGAAGTCGTTATTTTTTACTTCATTTTTTATGAAATTTAAACATCTTATTGCTATTGGACTT
GCTGCTGTTGCTCCTGCTGCCGCAATTTGCAAGTTCTACTATTAATGGTGTGGTGTCTTTCC
TGCCCAATTTATCAACGATGGTTCAGGATTATGCACGAACCTCTGGGAATAGGGTTAATTA
TCAGTCCGTTGGTCTGGTGTGTTGCTCAATTTATTGCGGGCACGGTTAACTTTGGAGCA
ACAGACGAACCAATCAAACCTCAGAAGCAGCAAGAGTAAAGCGTGGTGTGTTCAAATTC
TATGGTAGGTGAACAATGCTATTGCTTATAACAAGCCAGGATGCTCCCTGAAACTTACTCA
GAAGCAAACCTGTAGATATTTTCTCAGTTCGTTATTCGTGACTGGAAGCAGGTTGGATGTGCTGC
TGGTCTATGCGAGTTGTATATCGTTCTGATGGTCTGGAACCTTGCATTTACAAATTCT
CTTGATGCCTTTGGTGGTGGAGTCTGGTGTAGGCAAGTCAGTTAAGTGGCCTGTTGGTATTG
GTGAAAAGGAAATGAGGGTGTCTGGAACAATCAGAACAACCTCCTGGTGCAATTGGATAT
GTAAATACTGGATTTGTGAGAGCAAATAAACTTCAGGTAGCAGCACTTCAAATAAAGGCAGG
TAAGTTTATTCTTCTACTGCTGCTGCTGGTCTGCTGCTCTGAATGGTATTAACCTGGATGCA
AACCTTGTGGTGAATAATCCAATC

MAKER 2 Protein Predictions

>snap_masked-Contig7191-abinit-gene-0.0-mRNA-1 protein AED:1.00 eAED:1.00
MELVEPHSTILVLNSSYEPLHFTNWKRAVVLLFKEKARLISKRVIRLVNYVKIPFKHGKDAYPTRSL
IYKRDDHECQYCGTTKDLTLDHVIPRSKGGQDTWENLVTCCKCNLKKGDKLLSETNMTLKKTP
APFNRVYLDLQKSRVSEWVDYVIG

>snap_masked-Contig7191-abinit-gene-0.1-mRNA-1 protein AED:1.00 eAED:1.00
MRIPEFLVSFRLSPKRSLIAALIASGIAFLTVRCGISQEDLLKFYNELRKEIKFNLPNENNVINEIDRQL
NERINNDPKLLEFKVRGEVDDAISRYEREYQNRQINMKNKNILEEINKPKYDNLQKLIVENAVYY
EFEDGTMGIRGAWVSPDPREIDLEQ

>snap_masked-Contig7191-abinit-gene-0.2-mRNA-1 protein AED:1.00 eAED:1.00
MSESVTKLVEAFETWKSEDEKFGVGGNGAAGTRARKALQEIAKLVKERRAEISEEKNARKETKGA

>snap_masked-Contig7191-abinit-gene-0.3-mRNA-1 protein AED:1.00 eAED:1.00
MLRNLSTVAVSLLGAVATSAATLPTPGYATSSSTPQIAILTEESKTQETTTKEVAPEKPKERLVC
GCNSNEARTLEFLQSRGITDRNALATIMGNIRQESTFFPNICEGGARVSYNSCRSGGYGLIQWTNAP
RFYGLGKHAARIGANPSSSLDAQLDYMLHEGDWKMIEPYMKTTPGKTINQYMRLAQRWIRWGHG
ARTDYAYNYANKLVLTVVGS

>snap_masked-Contig7191-abinit-gene-0.4-mRNA-1 protein AED:1.00 eAED:1.00
MKFKHLIAIGLAAVAPAAAFASSTINGAGASFPAPIYQRFQDYARTSGNRVNYQSVGSGAGVRQ
FIAGTVNFGATDEPIKPEAARVVRGVVQIPMVGTTIAIAYNKPGCSLKLTKQQTVDIFSGRIRDWK
QVGCAGPMRVVYRSDGSGTTFAFTNSLDAFGGWSVPGVGSVKWVPGIGGKNEGVSQTIRTP
GKFIPTAAAGSAAALNGIKLDANLA

>snap_masked-Contig7191-abinit-gene-0.5-mRNA-1 protein AED:1.00 eAED:1.00
MKLFQQLMLAPVAVGLIAPAATAAELNMEGVNQYSTSEQVTSINQFSDVQPTDWAYQALSNLVE
RYGCVAGYPNGTYGGGKAMTRYEAAALFNACLDRVTEVTDEIRKLQEEFKAELIVLRARVEGLE
VKTAVLEAQQFSTTKLNGEVNVLGGVPGYESRGTAKADGTAFNYDVRLNFDTSFTGRDLLRTRL
RAGNFSELPGSSSLFKLKAETTGDVELDRIYYQFPVGNDFTVTVGPKVRNTEMAWVPSAYK

SEILDFSTAGAPGVYNKATGAGVGVQYESNGFVAGLNYVAQDGDSTETGVFSSDGALNLMAQL
GYRQDNWGVGVGYRYGSEGTRPRTYNGFQGANGLLNQDSNSVSVNAYWRPTETGIIPSVSAG
YGYNAVSGRAGD TDATDSQSWFVGLQWDDAFVAGNAAGVAVGQPANAEGLANDALMLEFFYK
YQVTDNISITPAIFYVSNADQRGNASEWGGVIQTKFKF

>snap_masked-Contig7191-abinit-gene-0.6-mRNA-1 protein AED:1.00 eAED:1.00
MIGLNDKLLQQYIEWIANRRMKAIGLKPVYDIRANSNPLPWTEHWISSKGLQVAPQETEVEVESYVV
GGIKQDVKKDTFAGFKL

>snap_masked-Contig7191-abinit-gene-0.7-mRNA-1 protein AED:1.00 eAED:1.00
MENKEILGLYNAYVQIHTEQLDENRAAARNPEEYEHNKAKKTD RR TAAMNDPHTGINS PAFRRF
MAQQTGRK LKEDYDVYDLILSYLLYEGYAETPEAAESIMVNMS EDWRESIVEELEQLDEIMKHQQ
ERDRIEGMKKKSSRAKTWSYEPPTKGKHGEIERRQAQRPPSRLRSGKLPKERE EG

Fig. S5. Contig13737 from the *P. ovalis*-like cell 1 assembly of the 454+Illumina data that encodes proteins of cyanophage origin. The MAKER 2 web annotation server (<http://derringer.genetics.utah.edu/cgi-bin/MWAS/maker.cgi>) revealed 13 putative proteins that are shown below the contig sequence.

>Contig13737 Average coverage: 4,537.08

```
AGGTGGAAGTGGGTCAGGAAGTCCGTAGCAAGTTATGATACATATCTAGCATCCTT
AGAGATGCTTTTTAAATCTTCTGCTTCAGATTCTGGTTCAGTTAGATTTTATGAAGATACTGAT
AATGGAACAAATTACGTTGCAATTGCAGCATCTGCTACTTTAGGATCTTCATACACATTAAGA
CTTCTGATGGAAGTGGAACTTTAGGACAAGTCTTCAAAGTGGTAAATTTATCT
TGGGTAACATCCAACCTTTATTATTAGTGATGGAACAAACACATCACAATTAATACAGATAAT
ACAATAACCTTTGTTTCTGGCGAGGGTATTGATGCTGTTGTTTCTCAAGATTTAGATACTATTA
CTATATCTGCAGAAGCAGGATCACCAACTAATGCAGGTATTGCTTCATATCAATCATCACAAT
TTTATTTTGTAATAACATATCAGGTTGGCGTTGTAAGTCTACTGATACTGTTAAAGGTCTAGC
ATCATTGATAATGGAGATTTTGTAGTTACAAATGGAACGTTACCTTAGCGGATAGTGCTAA
CGGTGCTGTTTAGCAATTTCTGGAACAGAAAATGAAGTAGAGGTTTCTAGAACAAATGGAAC
CGTACTGTTGGACTTCCAAATAATGTAAGTATTGGAATGACCTAACAGTTAATGGTAACTT
AAGAGTTGTTGGAAGTGCAGTACTTTTGAAGTAACTATAAGAGTAGAAGATAGACTAA
TTGAACTTGGTCTTGTGCTGGTACTGACGCAACACCCTGGGATTTGGGTATTGCTTT
TAACTATGGTGATGGAAGTCTAAGAAAGCTGGTGTCTTCTGGTTAGATAATCAGTATATTGG
TATTGCATCTGCTGTTAGCATATCAAATGATACTGGAACATCTGATGCTGATCCACAGATCAC
AATTTCCACTTTTCTCCAATCGTTGCTGATGGATTATACTTGGTGGATTAATGCTTCTGATT
TGGCAATAAATAGTTCAAATGAAGCAGTAAATCTTGTATTTGATGGAGGATCTTATTGATAAT
GAGTACTGAAATTAATCTTGAATATGGTGATGTTTTAAAAGCATAACCAGTCAAAATCTGGAGA
ACTCTTAACTCAATTAATAACTGCAGAGGCAAGACTAAATGCCTCTGCATCTCTTATTATGGA
GTTGAATAGACGAATAGAGGAATTAGAATCCGAAAAAGAAAAGCTTCAAAAATCTGTAAGTA
GATCAAAAAAATCATCAACTGATAATGTTGTTGATTACAATCAGTAATATTTGAGTTTTAATA
ATGGCAAATATATTTAAACCTAAGCGATCTAATGTACCATCTTCTATCCCAACAACAAGTAAT
CTTGTAGATGGTGAATTGGCAGTAAATCTGCTGATAAAAAAATATTTCTTCGTGATGGTGAA
GATATTATAGAAATATCCAATGCATCAACTTATATCTTATGCTTCAACAGCAGGTATAGCA
ACTTATGCTACTTCTAGTGGTATAGCAACTTCTGTTATAGGCGGCCTTGGTTCTATATCACAAC
TACAAGTTTCTGGTATTTCTACATTTACTGACGGTCCAGTTTTAATTGGGACTGCAACATCAAC
AGGAACTACATCACAATCACTTCAAGTTAATAGCGGTGCTTATGTTTCTGGAAATCTTGGAGT
TGGAAACCACAAATCCAACCTTATAAAGTTTCAAGTTGATGATACTGCTGCAACTGGTGCTGGTCT
TTTAGTTAGAGGAGGGGGTGCTGGATCCCAACTTGGTCTAGATTTGAAAAGAAATGTTGGTTCAAC
AGGATCATTGATTGATATTCTTGCCTCCAGTTCAAATCCACAAATTTCTTTTGGTGGATCTACA
ATAAACACTTTTACTGTTGGTGTAATATTAGTTTCAATTTCTGGTTCAATTTGAAATTTGCTGATA
GTAATAATATAGGTACAAATACAAGGTTTCTTATAAATCCAATTTGGTGATGTTGGTATCGGAA
CCACAAATCCAACAGCAAACTTGTGTTTTGGGTGATGTTAAAATAAGTGGAAATAACAACCTT
CTAGTGTTTTTGTATCTCTGGGGGGACTTATTCTGCACTTATTGACACAAAATCAGATGCTTC
CTTTGTAATGGGGCAAAATGATGCAATTTACACATTAGAAAACAATGATGATTTTTTAAGAGT
TCTTATTGAAAAACAAAGTGGTATAATTAATATTGGACAATTAATACTGCGTATATAAGTGG
AATTGAATTAAGACCAGGGACTACTGGTCTGTTAAACTACATCATGGTGGATCTGGAGATAA
TGTAAGATACAAACCCTGGTGTGGTGCAACAGTATTTGGAGATTTAGAAGTTTCTGGAAAT
TATAACTGCTTCTAATCTAACCAGGATATAAATACTTGTGATGATTTTCTTACAGTTTAAATGGT
TCTACAACCTCAATTTACAATTTCAAGTAGTTCTACAACTTTCTAAATAGTGAGATAACAAGT
GCTGCTAGATTATTAATTTCCGTTGGTGGTGTGTTTCAAGCACCAGACCCAACTCAGACAAAT
GGTACTATATAAGTGGTGGAACTGATTTATCAACAGACCCAATAAAAATTAATTTTGTAGAA
GCACCAAAATCTGGTCAATCATTTTTTGGAGTTGCTTATGGACTAACAAATAAGTCCATCACA
TCCTTTGTAACACCAGAACAATCAATCGCATATAGTATAGTATTTGGAGTCTAATTTAAAATG
GCAAGAAGAAGAGAAGAAAGATACCGATTTGATGCCTCAGAAAAGAAAATTTTTATTTTCAGA
TAGACATACTCAGAATGACATTCTATTAATTGTAATTTCTGTTGATGGTATTATCATTGCCAAT
CAATTAGATCCTGGTAAAGGATTTTCAACAACCTTATTTCCCACTACCAGTTGATAATGTCGATT
GGGAATATTCTACAGATGGATATACAGAGGTATTATTAGATTATGACACCACTGGAATGTCTG
```

ATACGGATGAACTTACAATCTTTATTGAAGAAGGACAAAAAGGATTA AAAAGTTCGTCCATATG
ATTTTGG AACGGATGCAATTGAAAGGCAACGTGTTGCAAATCCACAATCTACTAATTGACGCTG
ACTTTGAATATGGTCTCCAGAATACAAAGTGGCAATCACTGGGTTTGAATAGAAATATTCCAT
CGTCTTTGAGCTTCTGGTCTGCGCTATCTGTAGATGACATTGCATCTCAAGGTGCTGCAAC
ATATTCTACTATTAGAGTTACAGTTGCTACAGGAAGTGCCTGATGTTGGGTCTCCAGTATCA
ATTACAGGAACCAACCCACTTGC GGAAGGTTATTTGTTGTAATTGCAAATAACTCACCA
ACACTCTCTACCTTTGATATTCTAGCAAAAAGGTAGTGTTCCCTACTGGTCTATTTTCCAACTT
ATTCTGTTGTGAAAGAAGGTGGAACATATGGAGGAACTCCTATGGATGTTTCTTCTATGAGTG
GTGATGGTGC GGGAGGAAACGTA ACTATTA ACTTTGGAAGTCTCATGGTTTGATTCCAGGTT
CTCCTATTGCTGTTGTTGATACTACTGCTGGCACACAATCCCACGAAGGTAGATTCTTCGTTAC
TGAAGTTGTTGATGGAGATACTGTCATTTATGATGCAGGACAGACCGTAACAAGTGGTGCCAT
TACGACCACCAATATTTCTGTTTATGCAATCAATGATAGTTTTTATACACATAGACCATTTGAT
GGTGGTGTGTTGATGGGTCCACA ACTCCCTATTCACGGATTAGAAGCAAAGAGGCAAACAAA
AAGATATTTCCGTTATCAATCTGGTAAAGGTATTTGTTCTCGACGGGAACATTATTCAATCCA
GTTTTTGATATTTCAAACAGCAACATACTCTCACCAGATATTAGTGTAACA ACTCAAATTCCTC
ACGGAGTTCAAGAAGGTGCAACTGTAAA ACTTTATGGTATTACTTCAACTGGATAATAATGGAA
CTTATACAGTAAAAGCAATTACAAATGAATTTACATTTACTGTTGCTGGAGGAACTGCTC
CCACTGATGCTACTGCTGTCTTAGAAGCACAACTAGAATTGCTGTTACTGGATGGTCTGGTG
CTTCAATTCGTTGTGGAATTTTTGATGATTCAAACGGAATGTTCTGGGAGTATGACGGAAGAG
AACTTGCAAGTTGTAAGAAGATCTTCTACATTT CAGTTAGCTGGCGTTTGTCTTTATCAAATGG
ATCTCAGGCAGTTACTGGAACAAA CACAAGATTTACGGAACA ACTTAGAGTCGGTGATAAAA
TTGTAATTCGTGGTCAA ACTTATATCGTTTTCTACGATTACAAGTGATACAGCACTCTCAGTATC
TCCAGAAAATCGTGGTGTTACAGTTACAAATGTAAAACCAACAATTATAAGAGAACTTAGAG
TTCCACAAAGTAAATTCAATATGGATCGTATTGATGGA ACTGATACTCCATCAGGATATACTA
TTGATTTGGCAAAAATGCAGATGCTTGGAAATTCAGTATTCCTGGTATGGTGCTGGATTTATTGA
CTTCATGCTTCGTGGTCTGCTGGTGAATTCATCACGGTTCATCGTATGAAGAACAACAACATT
AACGATGAAGCATATATGAGAAGTGGAAACCTTCCAGCACGTTATGAAGTCTCTAATTATGCT
GCTTTGGATAAACTGGCATCAGCATCTGGAACATCTGGAGATCTTACACTTGCTGATGCTTCT
AGATTTCCA ACTGCATCCGCATCTTATCCAGAGTATGTGATGGTGACAAGTAATCAGTCTGGA
ACAATTTATCATGAAGTTATTGAATATACTGGAAAATCTGGTAATACATTA ACTGGA ACTACA
AGAGCAACTTCTTATACGCAATATCTTGCTGGTCTTCTAGAACATTTAGTGGCGCATCCTCAG
CACAAAATCATCCTGCAGGTTCAAGTGT CATTCTATTGAATACA ACTTGTGCTCCAACAATTC
ACACTGGGGTTCTGCTATAATTATGGATGGTGGATTTGATGAAGATAGTGGATACTTGTTTAA
TCTTTCCAGAACAAATGTGACTATAAGCGGTAATAGTTCCGTTGCGGTTTTATTCCGTCCAGCA
CCATCGGTTTCTGATACAATTC CAGGAGATTTGGGTGATAGAGAGGTTATTAATCGTTCTCAA
ATCACTCTTAAAGAACTTAGTATAAACAATAACTCCTCAAGAAACCTTGAAATTTCTGCGATT
ATTAATCCAAGTAATGTAGGTGCGGCAACTTGGCAGAACGCAAATACTACA ACTGTAGGTCTCT
GCGAATGTATTCCAACCATCGTTTCGCACAATATGCAACTACAGGTGGAACAATTCTTGGAGGA
ACAGCTCCTGCTGATGGTGAAATTTCTGTTTAGATACTTATCATCTTCGGGAACA AATATTTATG
ACTTGGGTGCAATTAAGAAGTTCAA AACTCTATTTTGGGAGGAGACAACACATATCCCGATG
GGCCTGAAGTTATTGTTTTCTACGTTACGAATAACAACGCACAATCAGCGTCAATTGACTTAC
TTCTTAGATGGACGGAGGCACAGGCATAAATGGCAGAATTATCAAGAGGTCAGATACTTAAAA
ATCCAGAAGTTCTTGGTATAGGAAATACTGGAGCAATAATTAGATTATTAGATGCTGATGAAA
ACTCACATATTGATATTAAGGTTCCCGATACTGTAGGGACTGCTTATACATTTACATTCCCTGA
TAGTGGTGGTACTACTGATTATTTAAAAACTGATGGAAGTGGAACTACAAGTTGGGCAAC
AGTATCTGCAAGCCCTGGTGGTTCTGATACTCAACTTCAATATAATAGTTCTGGTAGTTTTGCT
GGTGCTACAAATCTAACA ACTGATGGAATAACTTAAACCATCGGAGCACAAGGAGATATTAG
ATTAGCAGATTCTGATAGTAGTAACTATGTTGCTTTCCAGGCACCTGCAACTGTTGCTGCTGAT
AGACTTTATACACTCCCAGATACAATTGGAACGGCAGGTCAAGTATTAAGATTTCCACAAGA
ACTGATACTACAGCAACACTTGCTTGGCAGGATGATAA CACTGGTGGTGC GGGTAGCTCCCCA
GGTGGTGTGATACTTATGTT CAGTTAATGATGGAGGAACTACTTTTGGTGGTGTGCTGGA
TTCACATAACAACAAAACA ACTGATTCTGCAACACTTGTAGGTTCTCTTACTGTTGGTGGAGTTA
ATGCAACAACGGGAACCGACTATCAAATCAACAGCACATCAGTTCTAAATGCAACCACATTA
GGTCTGGTGTTTTAAACTCCTCACTCACTTCTGTCCGAACTCTTGGTGTCTGCAAGTTGATA
ATGTTAATGTGAATGGAAACGTAATATCTTCTACTAATACAAATGGTAATATTGATTTAGACC
CTAATGGA ACTGGTGTGTAAGAAGTCTTGGTGCGGCAGCACTAAGATTTATGGATAGTGATA

ATTCAAACATATGCCGCAATTGTTGCTCCATCATCAGTTACATCAGATTACACAATCACACTTCC
TGCTGCTGGTGGTGGCTGCGAATGAAATTCTCCAGTATGATGGTTCTCAAAATGCCTCATTGTT
TCCAATACAAGAACTCTTAATTTTGTAAATTGATGGTGGTGGTTCTGCGGTTACAACCTGGAGTA
AAGGGACATATTGTTCTTGATGGGGATTACACAGTTACGGGTTGGACAGTTATTGCCGACCAA
TCAGGAAGTATTGTAGTTGACGTAAACCGTGCTACTTATACAAACTTCCCAACAACCTGCTTCT
ATTGCTGGAACAGAACTTCCAACACTATCTGCGGCACAAAAAGCAGAAGACCTCACACTATCT
TCTTGGACAACACTACTTTCTGCTAGAGATGTATTGGAGTTTGAAGTTGATAGTGCTTCTACAG
TTGAAAGAGTTACAGTCGCACTCCGTTTAGTACCAAGATAATATTATGGCTATAACATATCAA
GCAATATCTACAGTTTCAATAAAGACAACACTACATCTAATGGCTCTCTAACACTCAACGCTCCT
GCTGGAGTAACGACTGATGACCTTTTAGTTGCTTGTATCTCATTTTCGTGGTGCAACAACCTCCT
CAATTCCTGCTGATTGGAATTTGATAGAAAGGACCACAGTAACTGGAAACATCAGCGGAAAC
AATACTAATGCAATTGCTTCTGGTCTGATGGCTTGGATTAAAAGAGGTGCAACAAACCCAAGT
TTTGTATTTGGTGCAGTAGCAGGAGAACAATTTCCAAACCTTGCTTTAGGTTATGTTGTAAGA
ATAGATGGTCAAGACCTTACAGACCCATTAGCAGGATCAAGTGTTAATACTTTAGCGAGTACT
AATACTACTGTAACAACCTGGTGGATATACTAGAAGTTTTTAACTGCTACTGCAGATGATTAT
ATAGAAATTATGTTATGTATGGGTGCTCAAGAAGTTACTTGGAGCACACAACAATATGAAACT
GGTCCAACCTGCGATGACTGAAATTGCTCAGCAAACATCTACAGCAGGTAGGGATGGTCTATT
GCTGTTGCTAGAACTAATGTATTGGGAGACACTACTGGTGGTTTTAGTGCTTTTCTTCTGCTG
CCCGTTCGCAATGCCCTTATGGTCGCATCTTTTGGTGGTCCGAGACCCGTTGCTGCTGGGTACTC
GTTCTCCACTTTCTTCTAGGGGGGGGGCTTGACGGGTCTTGTGCTCTGTGCTACTATAAATAAG
TGTTAAGGAATGGAAACAAACTGTAGCCTTCTTAATACTTACCTAGAAAACCTAAACGGAGT
AATTAACCTATGACCGCTTCAATCGCACAAAAGCGGCAGGTTGGTTCTTGGGAACAGTTCTGT
GAGTGGGTTACTTCCACCAACAACCGTCTTTATGTCGGTGGTTCGGAACCTTTATGATTCCCT
GCCTTCTTGCTGCTGCTATTTGCTTTATCGTTGCCTTCGTTGCTGCACCTCCTGTAGACATTGAT
GGCATTTCGTGAACCTGTTGCTGGTTCACTTCTCTACGGTAACAACATCATCTCTGGTGCTGTTG
TTCCAAGTTCCAATGCTATTGGACTTCACTTCTATCCCATCTGGGAAGCTGCTTCTCTGGATGA
ATGGCTCTACAACGGTGGTCCTTACCAACTCGTTGTGTTCCACTTCTGATTGGTGTGTTGCC
TATATGGGTTCGTGAGTGGGAACCTTTCATACCGTCTAGGTATGCGTCCTTGGATTTGTGTAGCAT
ATTCAGCACCTGTTGCTGCTGCTACTGCTGTGTTTCTTGTCTATCCTTTCCGGTCAAGGTTCTTTC
TCTGATGGTATGCCTCTGGGTATCAGTGGCACCTTCAACTATATGCTTGTATTCCAAGCAGAGC
ACAACATCCTGATGCATCCATTCCATATGCTTGGTGTAGCAGGTGTGTTCCGGTGGTTCAGTGT
TAGTGCGATGCACGGTCTCTGGTAACTTCTTCACTGGTTCGTGAAACCACTGAAACTGAGTC
CCAGAACTATGGTTACAAGTTCGGTCAAGAAGAAGAGACTTATAACATTGTTGCTGCTCACGG
GTATTTTGGTTCGCTTATCTTCCAATATGCCTCATTCAACAACCTCTCGTTCACTTCACTTCTTCC
TTGCTGCCTGGCCAGTTGTAGGCATCTGGTTCAGTCTTGGTGGTTTCTACGATGGCCTTCAA
TCTCAACGGTCTGAATTTTAAACCAGAGCATTATTGACTCACAAGGTTCGTGAATCAACTCTTG
GGCTGATGTTCTGAACCGTGCCTGGTCTGGGTATGGAAGTGATGCATGAGCGCAATGCTCACAA
CTTCCCTCTTGACCTTGCTGCTGCTGAGAACACTCCTGTTGCTCTCACTGCACCTGCAATCGGT
TGATAAAAACCTGAATAACTGATATAATTAAGAGGGTATAACAACCCTCTTTTTTATGTCTTAT
AATACTCAACCGAACCTATGCCCAATTGGGTATTTGGATGGGAATAGTTCTTATGATATTT
ACTATTTTATGTTTTGTTTTAATGACTTTAGGGATGATTTATTTCATAAGATTTATAGTATAATTT
CTAAATAATTTTTTACAATACTTAAACGATGCTTCTAGACTTAGCACACACGATTGCTGACTATA
CTATCTGTGGTGAAGGTAATGTATCAGAAAAGAAAAACAGAGGATACTCTTTTAATCAAAGCA
AGTGGTACAAGTCTTCATACTACTATCAGAAGATGATTTAGTTCTGGTGAATACTGATGCTGAA
CAACTAAAACCAGAACAAAAAAAACCAAGTATTGAAGTGCTTTTTTCATGCATGGATTATGAA
GCACTTCCCAGAAATTAATTATATTGCTCATAACATCCACCAAAGACTACACAGATACTTTG
CTCTCCTGCAGTCAATGACTTTGCCTGTCAGAGATGGTTCCCAGACCAGATTGTGAGAAACGG
TGTAGTGTCTGCCTTGTCCCTTATGCTCCTCCTGGTGTCTCGTTTATTGAAGAATGTGGATAAG
TATGTGGGTGAGTTTGTAGACCATTATAGATACTTTCTTAAGTTGATCCTTTTTGGAGAATCAG
GTATTATCACAGCATCACCTTACCAAAAAGGATTGTGCTGCTGCTACTTTGATGTGTGAAAAAT
CTGCAGAAATCTTTATTGGTGCTAAACTTCTTGGTGGTGTTCATTTCTTCCAGATGAAGAAAT
TGAGCATCTAGAAAACCTGTCCTGGTGAGCAGTATCGTCGCCGTATGTATTTTCGTCAAATAACC
TAGAGTATAAATACTGATTGACTTCTATGTTACGTTATGTTAAGATAAATATGAGAAATACAT
AGGAGGTTATGACTTCTTCAACACTTTCACAACCAATTCAACAGAGAGGATGGTTCGATGTCT
TGGATGACTGGCTTAAACGAGATCGTTTTGTATTTGTGGGTTGGTCTGGACTATTACTTTTTCC
CACTGCTTATTTGGCCCTTGGTGGCTGGCTTACTGGCACAACGTTTGTACGAGCTGGTACACC

CACGGGTTGGCGTCTAGTTATCTTGAGGGTGCTAATTCCTCACAGCAGCTGTTTCGACGCCTG
CAGATTCTATGGGTCATTCTTCTTCTACTTTGGGGTCCAGAGTCTCAGGGCGATATCGTCCG
CTGGTCCAACTTGGGGGACTCTGGCCTTTTGTGGCGCTCCACGGAAGCTTTGCCCTTATAGGG
TTCATGCTCCGCCAGTTTGAGATTGCTCGACTGGTAGGTATTAGACCTTATAATGCTATTGCGT
TCTCTGGTCTATCGCAGTATTCGTTTCTGTATTCTGATGTATCCTCTGGGACAATCCAGTTG
GTTCTTTGCTCCGTCCTTCGGGGTAGCAGCAATCTTCAGGTTCCCTTCTGTTTCTTCAGGGTTCC
ACAACTGGACTCTCAACCCCTCCATATGATGGGAGTTGCTGGTATACTAGGAGGTGCTCTAC
TCTGTGCTATTTCATGGTGCAACGGTTGAAAACACCCTCTACGAAGATGGAGAAAAGGCAAAT
ACCTTCAAAGCCTTTGAACCTACCCAAGAAGAGGAAACCTATTCAATGGTTACTGCGAATCGT
TACTGGTCTCAAATTTTTGGTATCGCTTTTAGTAATAAGCGTTGGCTCCATTTCTTTATGTTGTT
TGTACCTGTTATGGGTCTTTGGACATCTAGTATTGGGATTATTGGACTGGCTCTTAACCTTAGA
GCATACGATTTTCGTATCACAGGAGATTCATGCGGCAGAAGATCCGGAGTTTGAACTTTCTAC
ACGAAGAACATTCTGCTAAATGAAGGACTTCGTGCTTGGATGGCTCCAGTTGATCAACCTCAT
GAGAACTTTGTGTTCCCTGAGGAAGTTCTCCCGAGAGGGAAACGCACTCTAAAATAAATAAGG
GAGTTCTTATGAACTCCTTTTTTATGTTTCGCAACAATAGCACTGTTTATTTTTATCGGTCTTTT
TATGTTACACACTTTCACTATTTTTGATTATGAAACTTGACCCACTTAAAACCTGGATTCTATTGCG
TTGACTATGTTTATAACGACACAAAAGAAGTCGGCAGTTTATTTTTCAACTTGAATCCGCACAAG
AAGCACTTGTGAAGATGATGAGGAACGGTGTAGAATGTACTGGTATGAGGGAATGGAACA
AAACCCGAAATTATATCTATTAATAAATAAATAAAGTGAATCAATACAATTCTATAATGAGAA
ATAGATGTTGGGATTTTGTAAATGTCTTCATTTGCTAGAAAGTTACGGTGTGATAAAGCAATTAC
TAATGTTAAATTTACGAAATAGCACTACAGTGGTTCGATGATAAATAACTATACTTGTGATGT
TCATCTAGATAAATTTAGTTAAAGTCGATTCTTATTTTTAGACAAATATACGAAAATTGGGAAGA
GTAAGTTATGCGAGTTGGAATGATTGGATTGGGACGAATGGGAGAAGGAATGTCCCGTCGTA
TGATGAAAGCAGGTATTGAAGTATGGGGTTACAGGAGAAATCTTGATAAGGCAAATGAATCA
TTTGAAAGGGGTTATATTTTTGGATATGCTGCCGATATTGAAACTCTTGTTAGGGTAGTTAAAC
GAAATAAAAATGGAGAACAACCAGGAATCTCCAGATGGTAGTTCCTGCTGAAACAGTAGAG
GAGACGATTGATGAGTTACTACGATATTGTGGTGAAGGGGATATTATTATTGATCATGGCAAT
AGCAATTTTAAAGACAGTCGGAAGAGAGCAGAACGTCTTGCAAATTTGGGCATCCAATATAT
TGATTGTGGTACTAGTGGTGGTGTCTACGGTGTGGGTTCGTGGATACTGTCTTATGGTTGGGGG
TGGAGATACTGCAGTCACCACTTGTAAGAGTATTTTTGATGCCCTTGCCCCAGGGGTCGATGC
TGCCCCTAGAACTGCCCTGACGACCACCTGACCCCCGAGAGAGGGGTTGGTTGCACTGTGG
GGTCCAGGTGCTGGGCACCTTCGTTAAGATGGTACATAATGGTATTGAATATGGTATGATGCA
GGCATAACGCAGAAGGATTTAACATTCTGAAAAATGCAAACGCAGGTGCTAAGTATGTCCGTG
AAGGAGACGCAGAAGTTGCCCTATGGCAGACCCAGAATCTTATTGTTATGATATTAACGTTT
CTGAAGTTGCTGAGTTATGGCGTCGTGGTAGTGTGTTGGGTCTTGGTTACTTGACCTTACTGC
TAATGTGTTACGCAGCGACAGTGAGCTTGAACGATTCTCTGGAGGGGTATCCGACAGTGGTGA
GGGTCGCTGGACAGTTACTGCCGCTGTGGATTTGGGGGTTCCCGCTCCTGTATCACCCTGC
GCTCTATGAAAGATTTAATTTCTCGCGGTCTTGGTGTCTTCGCGTCCAAAATTCTAAATGGTATG
AGGTTTATGTTTGGTGGTTCATCACGTTTCGTAGCAACTCATTTACTCAAGGGTCTTCTGTCTTG
ACTACCTGACAAAACCTCTTGACTTCACTTGATTTTTCTGCTAAACTGAATCAGTCATCAACGA
CACTTATGGACATTCGGGATTTTAAAAATCAAGAGCATAAAGGTGAGTATTATGAACCTATG
AACAAAAGGTTGCTACCTACTGGGGGATCTTAAACAAAACCTTACAACAGTTTCTCCAGGAT
ATGGTATGGACCATTTGGACGGAAGGATTCTAAAGTTTCTTTCTGATATTACAAATACTGTAA
TGTATGAGACAACTGAAACATTTGAGAATCTTCATTCAGAATATAAGAAAGATAATGTTTTTG
TTTCTAAAACCCAAATCAAAGACTCTGTAAAAGAGGCACTGGAAGAATACAATTATTCTACAT
CACTAAAAGATTTTTCTTGGTGATTACATAATATTAGTATTTCTTAACATAATCACTTAAAAC
ATAAGAAACTCCCATAGATAAGTTAAGTTTCAATAACTATTATGAAGTTTCTTATCGTTTTATT
TGCGTCCTTATTTTTTGCACCTCCTTCTTGGGCATTAGACATTACTATGGGAAGTAATGGTAAC
TTGGTCTTTGAACCAGCAGATACTACAATTGCTGTTGGAGAAACAGTTCATTTTGTAAATGGT
ATGCTACCACCACATAATGTGATTGTTGAAGACCATCCAGAACTTTCTCACGATGGACTTTTGT
TCTCCCCTGGAGAAAGTTTTGATATTACATTCACAGAACCTGGAGATTACACTTATTGGTGGC
GTCCTCACAAAGGTGCTGGAATGGTTGCTACTATTCACGTCTCATAATGAATCACAATCACAA
CGACTATATTCATATGTTTGTATGCTGTCTCGCAATCGTTGGGGGAGTTGCCTCAATGATTTAT
GCCTATACTGTAATAAAAACAACAAAATCATACTAACTGAACTATGACTTATTCTATTACT
CTTCGTTCTCCCGATGGAACCTGAACAAGTTATTCAATGTCCTGAAGACCAATATATTCTTGAA
GCAGCAGACGAAGCTGGAGTTGACCTCCCTTATTCTTGTCTGCTGGTGCTTGTCTTCTTGTG

CGGGCAAGGTAATTAGTGGTGAAGTTGATAATGAGGAACAATCGTTCCTTGACGATGACCAA
CTTGCTGCTGGATTCACAATGCTCTGTGTTGCTTATCCAAAATCTGATTGTGTAATTGAAACTG
AACAGGAGGAAAATCTTTGAAAATCTTTAAAAAGCATCTCAAATCAAAAAACTCAAATCAAC
TTTGAGAACTCACGACCCTGTTGATAGTAGATTTAGGGAACCTGAAAATAGAGATAGAGGTG
AAAAGAAATAATTTTAGAGACCCGAAAGGGTCTTTTCTGTCTCTTGACTTCTAACCCCTTCT
GCTATAATAAGAACTCTCCGCAGACTTCAAAGATGGTTGGAAAACCTTGACCCAGAAGAACG
AGTGCTTCCCGAAACATTCGCACAAACCTCTGATGAACCTTATGTTTCGTCACGATTATAAACT
CCATTACACAACGAAGAAACCAGAAGTGTGATAACTTTGAAGACACTCAAAGGACTTGTT
TCAAACCTCAAAAACAATTTCTAGATTATGTTGAAGTGCTAGATAAGAAACAACCAAAGAAAA
AAGCAAAAAGGGTCTGAAATGCAAATTGATTGGATTAACCTCTCTGATGGGCATTTATTTGATT
TAGTTGGGTCTTAATTACGGAAAACCAAACAGACGGATGATTAGTTCAGAAAACACCTTATAA
ACTTGCCGAAATCTCCGTGAAACTTGGCCTGGAATATATAAGAAACCACACTCTCACAAAA
TGAAACTGAAATATCTTCTGGTATTACTACCCTACTTCTAACAGCATCTTGTACTCAAGAAAC
TGAGCCCTATGAGGTTTATAAGATTACCAAAGTAATACTAGAATTGATTGTAAGGGTAAGG
AATTTAACACGGAAAGACCACAGAAATTCGTGTGATTTCTGATATAGAAAATATCACAATCA
AAAACCTGTAAGTTAAAAGGTTCCATTCGGGTTTATGGTCTTGGAATGAATGGTGAAAATGAGG
AAGTAAAGAAATCTCCATAAGGCAGGCACACAGAAAAGAGCACAGGCAGCAGCACCTTCT
AATGTTCTTATTTCCAATATGAAGATTGAGAGTGTGAATAGAATTCCAGTTTATCTTGCTCTG
GTGTTACAAGAGTACTGTGGAAAACCTCTGAGTTTATTGGGACAACCTGATTCAACGGTCATCT
ATCTGGATGCCGAGAGTGCTTATAATACCATTTACAACAATACCTTTGATGTGAGTGGAACCT
TTACCCTTCGTCAGTTTAGAATTCGTGAAGTGATTGCCGTAGATGGTTCCGCATATAATACAAT
TTCTGGAACAAATTTAAGACCCGAGTAGGTGGTGGAGTATATCTCTATCGCAACTGTGGAGA
AGGTGGAACCGTCAGACACCAATCTCCTCAGCATAATCTAATTGAAAATAATGACTTTAATCT
ATCAGGATTACACATTAACAATTATGGTATCTGGTTAGGTGCTCGTAATGGAAATCGTTTCTA
CTGTAATGCCGATGAAGGTCATAAGTTCGGCAGTAGTGTAATGACCGTGATTTCCGAAACAG
CAACATCATTAGAGACAATAAGTTCAGGTTCTGACCGAACCATAAAGAATGATGGTGAAA
ATAATGTGATAAAATAAATAACTAATGTCCCTGTTTTGGAGTATTTGGTAATGGAAGAAAACC
TATTTCTGAAGAAACTCAAAGAATATTATACAAACACAAGATGCCGTAGAGCACGACCCA
AAAGCAGCACTACGAAATCCTGAGGGATACACACCAGTTGATGATGTAAGTATCACTACACC
AGTAGAAAATCACCAAGATTTGATGTTGGTGAAGTTTCCGCAACTTCTTCAGTTCCTACTCG
ACAAGTTTTGGATTTACCTCTGAAGTTTATGGTGGAGCTGGGGACTTTTTAATTGCTCCTGAT
GAAGAAGGAGTAGAAAATAATCCTCGGGGACTAGAATTAATTCATTAGCAGAAAACACAAA
TGTTAGGGTTCAAACCTATATCTAATCTTGGTGCTGGTTCTAATAATACTAACGATAATTCTGTG
GATAGTGAGACACCGGATGAAGTTATAGATGAAGTTCCTGTAGAACCAGCACCAAACCCCC
TGCTCCACCAGTTGAGCCTCCTGTGGACCCCCCGTTCCACCTGTTGAACCACCAGTGGACCC
ACCTGAGCCTCCAGTAGATCCTCCTGTGGACCCACCAGTTGATCCTCCTGTGGACCCACCAGT
TGATCCTCCTGTGGACCCACCAGTTGATCCTCCTGTGGACCCACCAGTTGATCCTCCTGAACCT
CCTGTTG

MAKER 2 Protein Predictions

>snap_masked-Contig13737-abinit-gene-0.0-mRNA-1 protein AED:1.00 eAED:1.00

MLLDAHTIADYTICGEGNVSERKTEDLLIKASGTSLHTLSEDDLVLVNTDAEQLKPEQKPSIEV
LFHAWIMKHFPEINYIAHTHPPKTTQILCSPAVNDFACQRWFPDQIVRNGVVSCLVPYAPPGARLL
KNVDKYVGEFVDHYRYFPKLILLENHGIITASPYQKDCAAATLMCEKSAEIFIGAKLLGGVQFLPD
EEIEHLENCPEQYRRRMVFVK

>snap_masked-Contig13737-abinit-gene-0.1-mRNA-1 protein AED:1.00 eAED:1.00

MANIFKPKRSNPSSIPTSNLVDGELAVNSADKKIFLRDGEDIIIEISNASNIISYASTAGIATYATSS
GIATSIVIGGLGSISQLQVSGISTFTDGPVLIGTATSTGTTSQLQVNSGAYVSGNLGVGTTNPTYKV
QIDDTAATGAGLLVRGGGAGSQLARFERNVSGTSLIDILASSNPQISFVSGTINTFTVGVNISSYS
GSFEIADSNNIGTNTRFLINPIGDVIGITNPTAKLDVLGDVKISGITSSVVFVSPGGTYSALIDTKSD
ASFVMGQNDIAIYTLNNDDFLRVLIEKQSGIINIGQLNTAYISGIELRPGTTGSKLHHGGSGDNVK
IQTTGVGATVFGDLEVSGIITASNLTEYKILDDISSQFNSTTQFTISSSSTNFLNSEITSAARLLISVG
GVVQAPDPTQTNGYYISGGTDLSTDPIKINFVEAPKSGQSFFGVAYGLTISPSQSFVTPEQSIAYSIVF
GV

>snap_masked-Contig13737-abinit-gene-0.2-mRNA-1 protein AED:1.00 eAED:1.00
MIGLGRMGEGMSRRMMKAGIEVWGYRRNLDKANESFERGYIFGYAADITLVRVVKRNKNGEQ
PGIFQMVVPAETVEETIDELLRYCGEGDIIHDHGNNSFKDSRKRAERLAKLGIQYIDCGTSGGVYGV
GRGYCLMVGGGDTAVTTCKSIFDALAPGVDAAPRTAPDDHLTPAERGWLHCGGPGAGHFVKMV
HNGIEYGMMQAYAEFNFILKNANAGAKYVREGDAEVAPMADPESYCYDINVSEVAELWRRGSV
VGSWLLDLTANVLRSDSELERFSGGVSDSGEGRWVTAAVDLGVPAVITTALYERFNSRGLGAF
ASKILNGMRFMFGGHHVR

>snap_masked-Contig13737-abinit-gene-0.3-mRNA-1 protein AED:1.00 eAED:1.00
MTASIAQKRQVGSWEQFCEWVTSTNNRLYVWFGTLMIPCLLAAICFIVAFVAAPPVDDIDGIREP
VAGSLLYGNNIISGAVVPSSNAIGLHFYPIWEAASLDEWLYNGGPPYQLVVFHFLIGVFAVMGREW
ELSYRLGMRPWICVAYSAPVAAATAVFLVYPFGQGSFSDGMPLGISGTFNYMLVFAQEHNILMHP
FHMLGVAGVFGGSLFSAMHGSLVTSSLVRETTETESQNYGYKFGQEEETYNIVAAHGYFGRLLIFQ
YASFNNSRSLHFFLAAWPVVGWFTALGVSTMAFNLNGLNFNQSIIDSQGRVINSWADVLNRAGL
GMEVMHERNAHNFPDLAAAENTPVALTAPAIG

>snap_masked-Contig13737-abinit-gene-0.4-mRNA-1 protein AED:1.00 eAED:1.00
MTYSITLRSPDGTEQVIQCPEDQYILEAADEAGVDLPYSCRAGACSSCAGKVISGEVDNEEQSFLD
DDQLAAGFTMLCVAYPKSDCVIETEQUEENL

>snap_masked-Contig13737-abinit-gene-0.5-mRNA-1 protein AED:1.00 eAED:1.00
MARRREERYRFDASEKKIFISDRHTQNDILLIVNSVDGIIIANQLDPGKGFSTTYFPLPVDNVDWEY
STDGYTEVLLDYDTTGMSDDELTIIEEGQKGLKVRPYDFGTDAIERQRVANPQSLIDADFEYGL
QNTKWQSLGLNRNIPSFELPGPALSVDIASQGAATYSTIRVTVATGTALDVGSPVSITGTTNPLA
EGLFVVIANNSTLSTFDILAKGSVPTGSIFTTYSVVKEGGTYGGTPMDVSSMSGDGAGGNVTINF
GSPHGLIPGSPIAVVDTTAGTQSHEGRFFVTEVVDGDTVIYDAGQTVTSGAITTTNISVYAINDSFYT
HRPFDGGVLMGPQLPIHGLEAKRQTKRYFRYQSGKILFSTGTLFNPVFDIQTATYSSPDISVTTQIP
HGVQEGATVKLYGITSTGYNGTYTVKAITNEFTFTVAAGGTAPTDATAVLEAQPRIAVTGWGAS
IRCGIFDDSNMGFWYDGRRELAVRRSSTFQLAGVCSLSNGSQAVTGTNTRFTEQLRVGDKIVIRG
QTYIVSTITSDTALSVPENRGVTVTNVKPTIIRELVRVQSKFNMDRIDGTDTPSGYTIDLAKMQML
GIQYSWYGAGFIDFMLRGPAGEFITVHRMKNNNINDEAYMRSGNLPARYEVSNYAALDKLASAS
GTSGDLTLADASRFPTASASYPEYVMVTSNQSGTIYHEVIEYTGKSGNTLTGTTTRATSYTQYLAGA
SRTFSGASSAQNHPAGSSVILLNTTCAPTISHWGSAIMDGGFDEDSGYLFNLSRTNVTISGNSSVAV
LFRPAPSVDITPGDLGDREVINRSQITLKELSIINNSSRNLEISAIINPSNVGAATWQNANTTTVGP
ANVFQPSFAQYATTGGTILGGTAPADGEILFRYLSSSGTNIYDLGAIKEVQNSILGGDNTYPDGPVEI
VFYVTNNAQSASIDLLLRWTEAQA

>snap_masked-Contig13737-abinit-gene-0.6-mRNA-1 protein AED:1.00 eAED:1.00
MGSNGNLVFEPADTTIAVGETVHFVNGMLPPHNVIVEDHPELSHDGLLFSPGESFDITFTEPGDYTY
WCGPHKGAGMVATIHVS

>snap_masked-Contig13737-abinit-gene-0.7-mRNA-1 protein AED:1.00 eAED:1.00
MTSSTLSQPIQQRGWFDVDDWLKRDRFVFGWVSGLLLFPTAYLALGGWLTGTTFTVTSWYTHGL
ASSYLEGANFLTAAVSTPADSMGHSLLLLWGPEQGDIVRFWQLGGLWPFVALHGSFALIGFMLR
QFEIARLVGIRPYNAIAFSGPVAVFVSFVFLMYPLGQSSWFFAFSFGVAAIFRLLFLQGPHNWTNPF
HMMGVAGILGGALLCAIHGATVENTLYEDGEKANTFKAFEPTQEEETYSMVTANRYWSQIFGIAF
SNKRWLHFFMLFVPMGLWTSSIGIHLALNLRAYDFVVSQEIHAEDPEFETFYTKNILLNEGLRA
WMAPVDQPHENFVFPPEVLPRGNAL

>snap_masked-Contig13737-abinit-gene-0.8-mRNA-1 protein AED:1.00 eAED:1.00
MEENLFLEETQKNIIQTQDAVEHDPKAALRNPEGYTPVDDVSITTPVEKSPRFVDGVEVSATSSVPT
QVLDLPLEVYGGAGDFLIAPDEEGVENNPRGLELIPLAETPNVRVQTISNLGAGSNNNDNSVDSE
TPDEVIDEVPVEPAPNPPAPPVEPPVDPPVPPVEPPVDPPPEPPVDPPVDPPVDPPVDPPVDPPV
DPPVDPPVDPPPEPP

>snap_masked-Contig13737-abinit-gene-0.9-mRNA-1 protein AED:1.00 eAED:1.00
MAITYQAISTVSIKTTTNSGSLTLNAPAGVTDDLLVACISFRGATTPSIPADWNLIERTTVTGNISG

NNTNAIASGLMAWIKRGATNPSFVFGAVAGEQFPNLALGYVVRIDGQDLTDPLAGSSVNTLASTN
TTVTTGGYTRSFLTATADDYIEIMLCMGAQEVTWSTQQYETGPTAMTEIAQQTSTAGR DGSIAVA
RTNVLGDTTGGFSAVSSAARRNALMVASFSGGPRPVAAGYSFSTFF

>snap_masked-Contig13737-abinit-gene-0.10-mRNA-1 protein AED:1.00 eAED:1.00
MKLKYLLVLLPLLLTASCTQETEPYEVHKITKSNTRIDCKGKEFNNGKTTEIRVISDIENITIKNCKL
KGSIRVYGLGMNGENEEVKKSSHKAGHTERAQAAAPSNVLISNMKIESVNRIPVYLAPGVTRVTV
ENSEFIGTTDSTVIYLD AESAYNTIYNNTFDVSGNF TLRQFRIREVIAVDGSAYNTISGNKFKTAGV
GGVYLYRNCGEGGTVRHQSPQHNL IENNDFNLSGLHINNYGIWLGARNGNRFYCN ADEGHKFGS
SVNDRDFANSNIIRDNKFSGSDRTIKNDGENNVIK

>snap_masked-Contig13737-abinit-gene-0.11-mRNA-1 protein AED:1.00 eAED:1.00
SGSVRFYEDTDNGTNYVAIAASATLGSSYTLRLPDGTGLGQVLQTDGSDNTITFVSGEGIDAVVS
QDLDTITISAEAGSPTNAGLASFDNGDFVVTNGNVTLADSANGAVLAISGTENEVEVSRTNGTVTV
GLPNNVTIGNDLTVNGNLRVVGSAVTFETETIRVEDRLIELGLVAGATDANTTWDLGIAFN YGDG
TAKKAGVFWLDNQYIGIASAVSISNDTGTSDADPQITISTFAPIVADGLYLGGFNASDLAINSSNEA
VNLVFDGGSY

>snap_masked-Contig13737-abinit-gene-0.12-mRNA-1 protein AED:1.00 eAED:1.00
MVGKLDPEERVLPEFAQTSDEPYVRHDYKLHYTTKKPEVFDNFEDTQRTWFQTPKQFLDYVEV
LDKKQPKKKAKGF

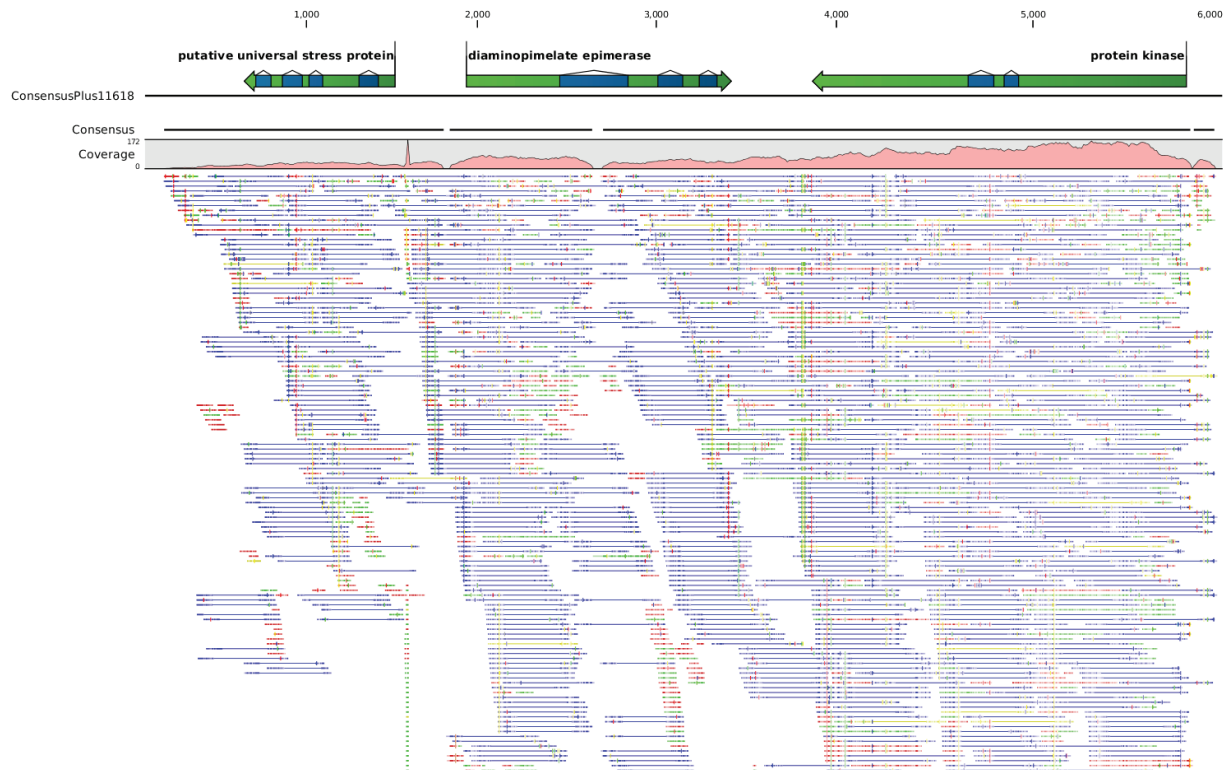


Fig. S6. Partial view of paired-end Illumina reads that span contig ConsensusPlus1618. The complete set of reads was too numerous to include in this figure. Note that many paired-end reads span the region of zero coverage upstream of and in the intron encoded by the gene DAP epimerase.

TAAAGAAAGTTCTACAGACAATGAGCACAAGGAAATTTTCTACTGCTTTTTACGAACTTAATG
TGCAGACTTGGGAACGCGGGGCTGGTCTTACACAAGCGTGCGGCACAGGTGCATGTGCTTCAT
TTGTTGCTGCATGCTTGCTGGGCCTGTGCCCTCGAGAAACCATTGAAACGAAGACTGAGATAC
AAGTCTCAAGTATTCAAAATTCCTCAGACGGGACAATGGGTACAGCAGTGCATGTTCTTGGGG
GCACGTTGTTTATCAGGTGGACAGAGGCAGACAAACTGTTGATGACTGGACCAGCAAATTTG
TCTTTAAAGGAATACTACCATAGCTTTCGAAGGATGAGGCAAGCCAAATAGCCCAACATCAA
CTTACAGAAATCAACCAATAAAACAAACCAAAAAGTCGTACGCCGTGCTAGAACTAAACATCG
CTTTGACTTATATTGAAAAAAAAAAGATGATTTTCACCAATAAAGGTCAAGCGTCCTTTTTA
AAATAATACAAAACCTTTTTGTAAATGCAAAGGTATTGGTACCTTACTACCATTTCATTTAC
ATGCAATAACTTCTTCTTCCAAACATTGTATCAGATGACGGAGAGGATAGGCCTAGTGGCGGC
CTTAAGCGCAAGATCATGCTGCGGGTATGCAACAATCCGTCTCCATTGGGAAGCTGACACCTT
TCAACTCTACTACTGTGTACTGAAAATTGAAAATAGTTGACAATATAATACAGATAAAGTTTA
TTTTGATGAAGAGAACGAGCCCCTATAACAAGTTCTCTTCAAGCTTCCGCAGCCGGGCGAGCG
GAGAGGGCCCCGTATTGATAGTGTGATCAAATCTGGGCTCGGTCCATCCAGATTACGATTCT
CAAGGAGCTGTTCTTCTCCTCTTCGCTCAGGTTTCAGAGCGAAATGCAATTTAGACGACTTCCC
TTGTTAAGTGAATCTTTTCCAGTTTACGAGGCTTTGTGAATGAGCGTTCACACTGTGAGGTT
GTTTGTGAGCCTGACAGGGGCTCCTGGAGGAAGAAAGCAAAGTTCGCATAGCTTTAGAAGAGTC
TTCGTTACTACCATTTGAAACAGCAACGGAATTATCATCTAAGGACAGTTGAATTGGACTTGA
GGTTCACCAGGTATTGTGGAACGTATAGCGGAGGTAAATTTTACTGTAAGTACTTATCTAT
ATCAGAGTGAAGATGTGACGAATCTGCTTCTCGCACATACGAGTAGTCTCGGAAATTCGATTC
GACTCGAAGCCAACGAACACATTCCGCATGATTGTGCGGACCTCTGCGACGTGAGCTCCCTG
ATTATGCGGCTTCCGACAATCGGAATTGGAAACACCAGATGCTTCTTCTACTTGTATCGCCAAA
CTGTTGAGCAACCTGGGCGGAACTTTCAATAGCTGGCAGAGCTCCAGTGCTAGATAGTACAAT
TGGTTTCGGGAGATAGGGAGGAGATCCCTGTTTTTGCAGCACTTGATCCAGTCTAAGCCGGT
AAAGAAGCGGTGTGCCTTTATTTCCGAAGCACCGCCTTTTGCCTCAAGGCGACTGGATGGATC
GCGTACAAGAATACACTAAAAGACATATATTTGAATGTAAAACAAACTGACCAAAAATACAT
TACAACAAACAGAATCAATTGAAGCTAGCTATAATATTGTGCGGATCACAATTCTTATTCTGAA
GAATCCAGACCTGCAGGATATCTTTTGCCTGCTGTGAAAGGTCTCGGGGAAGATCTAGAGGA
GCCTTCACATTTTGGACAAAGGAAAATGAAAATATTTTGTGACGGACGAAGTATGATATCAAAG
CGTAACTGACCGAAGTTATTTTATGGAAAAGTTCTGTTTCGCGTACGTGAGTACCAAGGAGGAA
GGCCGCACAACATCTCGTAGAGCAAGGCCCAAAGCTCCACCAGTCACTTCCAAGCCGTACC
CCTGTGCGCCGGATCACTTACAGGAGCCAGGTAATCCGGAGTACCACAGAACTGAAAGCTTTGG
AGTCAGAGTGCTCCAAACTTTTCGAAAGACCAAAAATCTGTGAGCCGAATATGTCCAGTTTCAT
CAAGAAGAATGTTCTCCGGTTTTAGATCTCTGTACATGATTCCGGCGGTATGCAAGTACTCGA
GCGCCAATAATATTTCCGCTGCGTAAAAGCGGGCGGCATCTTACAGAGAACCGGCGTTGCTTGG
AGAGATGGCGGAACAATTCGCTCCACAGCAATACTCAAGAACTAAGCAAAGGTTATCCCGT
GTTTGGAAACGCAAAATGCAGTTTGACAATAAAGGGGTGGTCAAAGGTAGCGAGTAACCACT
CTCTGTCTTTGCGTATAAGACTTGCTTCTTGTGTTACAATGCTGTCTTTGGAAATAATTTTCATGG
CAAAAAGGCGACCGCAGCTGTTCTCTCGCACTAACATTACTTTTCCCGTGTCTCCGGTTCCGAG
GACGGTCAAAGGTACAAAATTTTCGAGAGTACATCACAGGCTGCTCTTTCTTAATTCTTTACG
CGAGTCACATCAGACAACAAAGGATCCTTCTCTGCCCCAGAAGCACAATCACTTGTCTCACAC
CCCAGAGCAATTTTAGAATAAACTACAATACCACCGCTATCAAAAAGTACGCTGCAGCGATG
AAAATACCAATCACCATGATGCTCTGGACTGGCTGACTGACTGGATACCCTAGGGGTAAACAC
TGAAAGTGGACAATTGTCTTCGCGTGAAAAGAAAAATATGGCGGAGGACGAAGGTGACAAAA
GTGTGGTCAAGTTACAAAAGCAGTCACGTGTTATGAAAGTGGTAGAGATGTGAGATAGTCTCC
TTCTTACCACATCCTTCTTCTTCTCACCTCCTCCTCCTCCT

AUGUSTUS and Manually Validated Protein Predictions

>Putative universal stress protein

MVCVDGSPHSMKAFDSACYWKRPNDFLVVHAVETVKPVPLVGGPGPIMLPVSEYTATNEALKK
EAQKLMARLGDKCAKKGMDHVMVVLTTGNPKVAVVQYAQDKKVDMIFVGRGLGAVLSQIIS
FLTVNATCLSFDTNRTLKLYPLPLMRCRVRKGRKLLKAPD

>DAP epimerase

MILYSQISCINFHNKKVLFGILTA AVPKDSKTRSSTCDTHVRHVSEGQPYCLSEGGQRRTDYSVSK
VLDKMSNFKKKRSFQSRSLAPYKSIQPEIIFQKYEGCGNDFIICDFVDLPPKSLREIQQVDIRKRA
SFLCDNRNGIGADGILALSPLDTRSDARMHILNSDGSSPEMCGNGLRCLVKFLMSTRGPISKRHWT
IDTPAGIISAWPQDNGLISVDMGSPRLCMEDVPCRPTTSIPDCCSQDESWKSRVMLHFPLEIDIRPVL
ETSSAFTSSFYSSCCNTTTHAPTLRATTTITTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTSTATTVMGTSNGLKTSS
TKGLESPSSSHGAVSARTDRSLESVSFPPCHTGFFRATAIGMGNPHVIIPTESRISLEDLRKIGPLEN
HSAFPQKTNVHFLQVTKTSRSNKTNVASTVKIDDRYWAEMFGLKNVKKVLQTMSTRKFSTAFYE
LNVQTWERGAGLTQACGTGACASFVAACLGLCPRETIETKTEIQVSSI

>Putative protein kinase

FVPLTVLGTGSTGKVMLVRENSCGRLFAMKIISKDSIVSKKQVLYAKTERWLLATFDHPPFIVKLHF
AFQTRDNLCLVLEYCCGGELFRHLSKQRRFSEDAARFYAAEILLALEYLHRRIMYRDLKPENILL
DETGHIRLTDFGLSKSLEHSDSKAFSFCGTPEYLAPEVIRRQGYGLGSDWWSFGALLYEMLCGLPP
WYSRTRTELFHKIILVRDPSSRLGAKGGASEIKAHRFFTGLDWDQVLQKQGSPPYLP

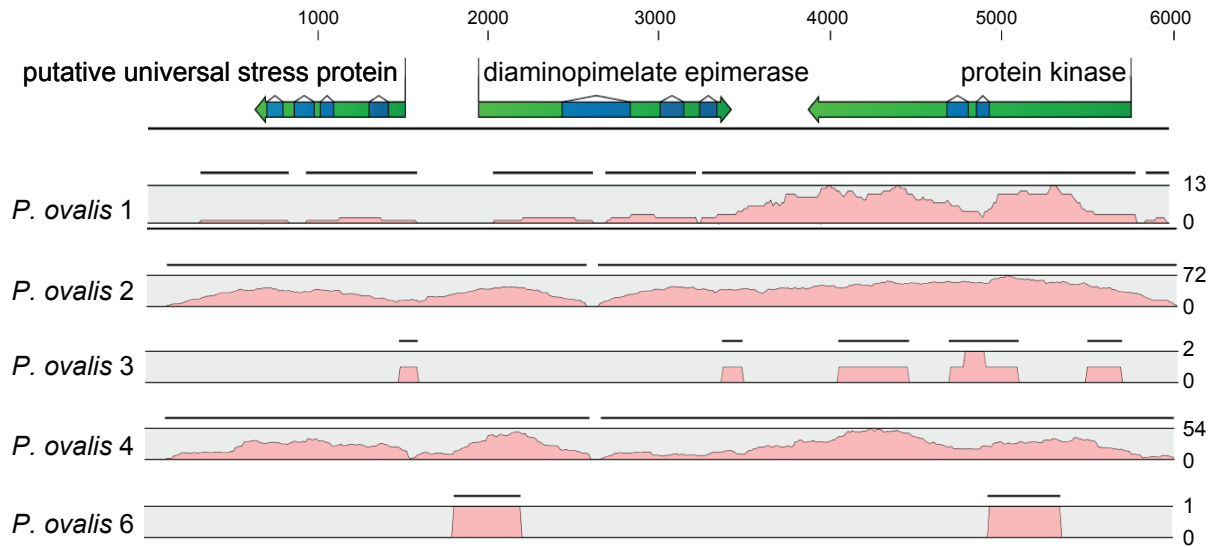


Fig. S8. Mapping of 454 reads from the *P. ovalis*-like SAG data. Each SAG dataset of 454 reads was used to map against the *P. ovalis*-like ConsensusPlus1618 that encodes a DAP epimerase gene of α -cyanobacterial origin. The other two genes on this contig are of amoebal (eukaryotic) provenance. The coverage across the contig is shown for SAG datasets from which we recovered matches.

Fig. S9. Contig11624 from *P. ovalis*-like cell 2 that encodes proteins of eukaryotic and cyanobacterial origins. The putative protein predictions are shown below the contig.

>Rhizaria-Paulinella_ovalis-like_Cell2_Contig11624 Average coverage: 298.96

```
AGTTAATATTGATTGGCCACGAACGCACAAAAGCAAAGACATTAAGGTTACTGATTGACTTCCT
TTATAAATCGAAAAGATGTTTCTCTTGTAGGTGAACTTCCAAGAACGTTTCGCTTGAAGTGGTTT
CTCAAATCATAACATCTTATTATGTCAGAATGGATTTCTGCACCTGATAGCTTTTTAGAAAAGAT
GGTGACTGGGTTAGACACATCAAGGTCAATCTTGTGGAATCATTGCAAAAAGGCATATGTTGT
CATGCGTTTTTATTTTTCTTGAAGGGCGATCTCCACCGATTAATTATTCAAGCTTCGGCGACTT
GGACTCCTGACAAAATGTCGATTCATTTTCTTATCGCACAAATATTTGACTTCAGACTG
AAAGTAAGAAAAAAGTTCTGTCTATTTTCATGGAGAAAACGAAATCAAATGGAGTGGTGGCCT
AGGGTCTTGAGGTAATGTGATAGGACGAATTCGATATTCGCACTTCATATGATTCTTAGG
GAGACCCTATGATCAACATCAAGAAAATACAACCCGAAAATTCACGGTATACCGTAAAGGTG
ATTGATACCCAGATATATCATTCAATATTGTAGGATTGATGATCTGGATTCTGATACGCAGCG
AACCGTGAAGAAAATGATGGTAATTGTTTTGGAAGTGGGTGGGATGAGAGTGTCCATTCACA
CTATCGATTTCTCACAGTTTGATCAAAGACAAAAGGAGCTGGGCCTTCCAACCTCAGATGAGG
TCGCTCTATCGCTTTGCCATGCTCATGCAGATAATATCTGAATCAGTTTTCTGACAGAGTACTG
AAAGACTGTTTGTGTTTTTGTGCTCTCTATCATATGTTTTGTGATATCTTATTTTTCGTCTT
CTTCCAAGAAAGATCAAGCTCACTTGTGTGTGATAGCAACAAAAGCATGAGATTCTAAAGCG
ATTTATGGAGCAGGTCAAGTAAACTGATGTTCTGAGATCGTCTTCTTATTATCTGCTGCTT
TATCTTCATAGCATCCCGAGATGGACTTTTCAAATGCTAAAATGGCTGAGAGACGAAAAGAAT
ATTTTCAGTTTTTCATGTCCAAGTTCCTTGTGTTAGCAGATACAAATATCGCGTTGTGGTATCCCGG
AGAGGTCAGTCGGTATGCATTTCCGGTCCAGAGTAAGGGGATCCTGGTCTGTCATTTTAGAGATT
CGGACTTCGCAAGTGGGCCTGTGTATAACAAGTATCTATATCTTTTTGTTGGAGCGATATGCAT
AGACATAATGTTGTACCCTTGAATTTGCTCTCAATCCATACAAAGCAGGTTGACAACTGGCA
TCTAAAATCATGTTTCGTGTAGTTTTGAGTGATTCTACACTGGGAATTGAATAATTTTGCACAT
TTGTTTCGAAAGCTTTTACTTTCTAGTAAAAAGGAAACAATGGTAGTCTTTCAATGTATGTGA
ACTTTTTTCCAATTTTCGCTGTAAGTCAATATACTTCGAGGGATTCTTCAAACCTCATCAAAA
AAATCCTTTTTTCGACATTGGGAATATGATACGTCGATTGATTGAGCATTCTTGTACTTACCT
TTTAAGAAGAAACGTTATTACTTTAATATGAAGCGACTCTGTTTTTGTCTACTTTACGAATATA
CAGTAGTATTTGACTCAGTACACTAAATTTAGCAGACAGAACTCTCGTCTCCTTTAGGCGG
GACTACAATGACCCGTTTGGGTTGGTTCGAGCTTATGAGAAATCCCTAACTCGGATTCCAATAC
TACTCGAGTCAACTCCCTCTCCTGACCTAACTTATCTTTAGAAATTATAAGGGAACCAAGCTTT
TTGCCCTTGACCTGAACGACCACAAAGTCTTAACGGATTTCTCAAGATGGAGTTCTTTCTCTG
CTGTGGATACACGGGGCCAAAGTTGCTCATGCACAGATTGACTATGCGTCTCGCAGATACCAG
CGCCCTGTCTTTTGAAGTTGTAGATAACAGAGAGATGCTCAAAAAGTTCAGCAGAAATATGAG
GAGCAAAGGGGAGCTAACATAACAAGAAGCGATCGAAGGGCCTCAACGAAGACAACCTGAAGA
ATGTAACCTGCAAAGCAAAGACCAAGTGTGAATTTACTTTGAAAGGGTAGGATACCTAGGA
TCTTGCTGAATTGAGTCTTTTAGGACATTCGAATATTTTCATCAAAAAGCGCGATTGCTACGTTGA
AATTTTTCAATTTCAATTTGTTCTGACACCTGCTTTACTGCATCCATCCTTGCCAAGGCAACCTC
GGAGTCGGTTTCAGAGAGATCAGTTGGACTGCAAGCAGAGTAGATTGCTTCAACAGACTTCTT
ATCCTTTGACAAAATGGATTTTTCTGTTCACTTGCAGATGTTCAAAGCGGAGGACCCACAG
AATCAACGAGGGGTGTGCGGTGGAAGAGGGGTTTTGGGGGGACCATTCCAGAGATTTTTCT
GGGGGAGCCTTGAATAAGATGGCCAGTCGACAAAACATCTGCTCCATAAGCATTACAGCATCT
TCTGGATTACGCCATTATACTTTGACTTAGACATTTTTTCGAAAACCTTGACCACTTTCCCTTC
ATATTGTGCAAGCTCTGATCTCTTAACGATCCGGTTGGAAGTTGCGTCTCGAAAGGTTTCTCCA
AGGACAAGTCCTTGACTAAGAAGGCAATGGAACGGTTCGGTTGAGCAAGATGTCCTTTTGA
AAACAAGAAACGTGAAATGAATCTCGAATATAGCAGATGGAGAATGGCATGTTCAACTCCGC
CAATATAGACGTCAACTGGCATCCAGTGATCGAGGCTACAGTAGAGGCAAACGAGCATTACA
CACATGAAAATGGGACTTCCACAGACGTACGTTACCTTTTCTTTTCATAAGACATTTGTCTCG
GCATAGGGGCGGACAGAGATAACGCAAAAAGTACCATGAAGAATCCACAAACGTGTCTAAG
GTATCAGTATCTTTTCGGCCATACCGCCACAACCTGGGCACTTGACTTCTCTCCACGGACGAG
CTTCTTCAGATAGCAATCCCTGCTTTCCTAACTAAGAGAACGAAACTCCTTGAATTGAATGAA
ACGAGTATGGTCTGGGTACCTGAAGATCTGGCAGCAGGACAGGCAGCTCTTCTTGGTACAG
```

GGACCACGCCACACGAGTCACAATGTATCATAGGAATCGGTGCACCCCAATATCTCTGTGCGAG
AAACAAGCCAATCCTGCAAACGAAAATCAATTTTTGTCTTGCCAATCTTGGAAGATAGAGAT
TTTCCGAAATGATATCCCGTGCGGATGCAGTGCTAAGTCTGTAAATCTGCGCTATTGATCAT
CCATCCCTCTCCAGAATAACACTGCTTCTCATTCTTTCTTTGGTTTCTTTGGAGTGCTCTGACTAA
GATCTTGTTCAAGTATCACAACTTCAAGAACTGAAAGATCGAAATGTTGGGCGAAAATATGGT
CTCGAGTATCATGTGCAGGAACTCCCATGACCGCACCAGTCCCATATTGAGGTAAGACATAGT
CTGCGACGAAAATAGGCAGATCACGCTCTAGCCATGGATGCGTTGCAGTGACTCCCAAATTA
CTCCACTTTTCTCAGTACTACTTACGCGCTGAATATCTGTTTGAGAGTCAACTTTCTTTTGTAT
CGATCAAGATCACGAAGGCAGGCTTTTGAAAAGCCGCCAAAGAAACGCAGGATTTGATAGAAT
TGGATGCTCTGGAGCCAAAGCCTTGACACGTATGTTGGCAAAGAAAATATATATTTATATAAG
AAAATATCACAGAAGAACTCACCAAAAAGGATACCCCATACAGAGTCTCAGGTCGTGATGTA
AAAATTCAAGCGGCTATAATATAAAAAATTAAGATGTCTGCTGACTGATATCCAAATTAAG
GACCATGTCAAAGCCAGGATTTTGAAACGACAAAATTGCACCCTGCAAAAAATCGATAACTTT
ATATACCTAACAAATGCAACTAGGAAACCTTTGATAGACCAATCCAATTCGTTGCAAGCATT
CACTTCCTCGGGCCACTCTTTCAATTCATCTAAGCCATGGAGAAGATTGCTTGCCATCGATCGA
ACATCAAAAAACCATTGCTTAAGGATTCTCTTCTCGACCAAGCACCTATGTCAGGACTGTGG
CAGTCTTCGTGATTTATAACTAGTCTCACCTGATCGCCAGGAGCGGCCTTCGGCGTCAACTTG
TTCATTAGCCAAAACAGTTTTATCTACTGGGTCCCAATTGACCTTTGCAGCCTTGATACGCT
AGACCGCTCTTAAAAAGTTCTAAAAATAGCCATTGCGTCCAATGATAGTACTCAGGATCGCAA
GTTGTGACTTCTAATTCCCAGTCGAAAGAAAGTCCCAGAGAATTCATTTGTTTCTTCATGGAG
GAAATGTTCTTCTGAGTCCATTCAGCTGGATCGACTCCTATGTATTTAATCCCACATACGATGC
AGTCTTTCCTGTGGCTAGGATATTATTGGACCTCTTTCCAAGGCCGATTCTCGGCTGGTAATC
CAAATGCATCCCATCCCATAGGATGCAATACATCAAAGCCCTACATGGGGACGTGCCAATCAT
GTCATATATATAGCAGCCTCGGATTTGTGTGTGTGCTCAGATATATATTTTACTCGCAAAAAGC
AGCCATAGGCAAACCTGAAGCCGGTAGTATCTCGCGAGCGAGTCACTATAGAAGCAGGGTCA
TGCCGCCATTGAAATGGAAAGAGAAGACTGGTAAACTACCTGATAGTGTAGACGCGCACATG
GCCATATGGAGTTCACCGGAAGGATATGGAAACATCGATAGAATATACTTCTTTCCTGCGCC
TATGCCAACAGCATGTTTTTAATAATCCTTCAACGGAAAAGGCTCGCGCACGCAATTACCCT
GATTTGAGTTCTCAGAACCTTTCGAAGTGGCCTTCAGAAAACAATGGATGTTGATTGCAACAA
CAACA
ACAGCTTGCTTAAGTTTCCAAATCCTCTGCCATTTAGATTCCACTTGGGAAGGAATGTAAGTA
CCCGACTGCAAAAACGCAGCAAAACATTTTCAAGTAAATGTCAATATCCGCAGGTGCACGTCGA
GGCGTACAGTACCAAAGAAAGCCACCGCCTTGACGAAATCATGGGGCACCACGCACGATAAAA
GTCAATCTTTTGTGCTTCTGCGAAGATATTACAATGATACATGTTAAGATTCTGTTTCGGAA
CTCTATCACATAGACATCTAGACATCCCTTCTTGAAGAGACACAAATTTTCTAAATAATCCGG
TCGACGTCCAATGACCTATACCTTCAAGATCCTTCAATTAGACTACTTTACGCCTTCAATTGACT
TCATTGCCTTCAATAACCCTTCCATTAGATTCCATTATTCCCAAGCCTTTCCTTCCCATC
CCGCTCCAAAAGCTAAAATGAGTACATGCAGCGCTCTCCGCTCTGCATTTCCGCGGCAAC
CAGAAAGCTGCTTTCGCGAGCATAATTTGTGCTCATGTTTCTCAACTGCAAGATACCTTTCT
ATGCGACGACCAGGCTGGATATCAGATTTTGAAGAATGATACTTCCGCGAGGGAAGTCTGCGAG
ACGTCGCCTAAGAAACGCGGATACTGTATCTATTTTCCCGACATATCCTCATGAGCATGTATT
GAATATACAAACAGTTGTGGATTTTATACGCGCAAGGGTACGATCGATCTTTCCTGTGTTTAT
CTATTTAGATTTATTATCAGATCTATTTTGTAGGCGGTTGAGTAAGTTATGGTGTTTAGGAAAA
GTTTATTCATCTCCCTTGGATTCCGCTCACTTTCACATAGGGACTCAGTGATGTCAAAGTATAT
GAAACAAATTATTGCTGTAAGTACTGGGTTATAGCTACTGCAGCTTCTTCAAGAGCCGCGTAT
GCGCATGGACATAAACTCGCGAAGTACGGCAAGCAAGGGAGATCCCGTTCACGTATAAATTT
AGAAGGAGAACGTCGAGACGACTGGTTTATAGTTGATCTGCAATTTTGTGTTGGTGCACATCTT
TGTACAGGAAAAGCGAAGTGAACATTGGCCTCAAATTTTCGCACCATTGCCCTGACGATAAACT
GATCCCCATGACAGAAGAAGGAACCTTGGAAATTTCTGATGTTCAAGACTCTCAAGATGTATC
TGCGTCTTCGAAGTCTTCTTGACATTATTGCAAAAAGCAGAAAAATCGATTGCACTCCATGACT
TAGGCTCAAGAATTATATGGGTGGGACGGTTCATCTTTTACTCTTGTGTTGTCAGCTGCTGTT
GTTGTTGTTGAAATAGTCACTTACCTGCCGCTTGCGCCGCGTCCAATATCTTCTCTGTTGTC
AACCTGGAAGCAACTACTACCAATAAGCCTAAGCGTGGTTAAGCCAGTATACCAAGTTACAG
AACATATCAAATTTGGGATAATGCTCCGTTTTTATCGAGGGGAAGGCAAGTACACAGCCAAA
AAGCAGACCGCTTACATAACAAGACGACAGAACATCAGATTGAAACAATGCTCTTGTCTTCTG
TCGTTGGGGAGGCCAGTCACACAACCAAGAAGCAGACAGCTTCCATAAGGCTACAGAACATA

TCCGGGATATGATGGAAGCAATTCTCCCGTTCCTTGTTGAGGGGAAGAGCAGTCGCAAAACCA
AAAAGCAGTCCGCTTACATTAAGTTACCGTGCTACCTACCAGTAGCAGTTCTGAAATTCTTTTC
AACTGAAAATGGTTAATTCTTAAATTATGGGGGGTTATGGTTTAAACTTACACCCCCAGTA
TTTCTTGTATATAAAATTGAATTTAGTCCCCAGTGCACCAAGAAAAATGACAATACGGTAATC
GATAAAACCGCATTACTATTACAGATATTGGGGCCTTTTATGAAATCTGGCAGTCTCGTCAG
TTGAAAACCTTTTCAAAGTGAAGAAAAAGTCTCAGGAAAAGTTATGTTATCGATACGATTTCG
CACTGGGCAAAGAGCTCTGGATAATAAAGAGCGAAAAATAACTCAAGAGTCTTCCGAGTTGA
GTATGGCTCACCCCTTAATAGCGATCATCGTATCGCCTACCTCCACCTCCACCGCTAGACCTCC
CGCTGTCGTAGTATCTTGAGCGGTCTCGACCGTCTCGACCGTCTCGACCGTCTCGACGGTCATC
GTAGCCACCACCTCGATCGCCGCTCGACCACCACCTCGATCGCCGCTCGACCACCACCTCG
ATCACCACCTCGATCGCGGCCTAGCAAATTGGAAAAGACAAATGTTCCACTCTCGACATGCAGC
AGCTCCGACTACGAAACAACGAGGCCCTAGCTTCGAAAATTGTGTGGGCTGATTATTGAAAGT
GATAGATTAACCTCAGCGCGCACACGAGAGAGAAAAGAAAGAGAGA

AUGUSTUS Protein Predictions

>Putative nuclear migration proteins (nudC)

VNFQERSLEVVSQIIHLIMSEWISAPDSFLEKMTVGLDTSRIDDLSDTQRTVKKMMVIVLEVGGM
RVSIHTIDFSQFDQRQKELGLPTSDEVALSLCHAHADNI

>Putative leucyl-tRNA synthetase

AAIEVVIEVVVEAAIEVVVEAAIEVVATMTVETVETVETVETAQDTTITAGGLAVEVEVGDITMIAIK
GCLLLGCVTGLPNDKKQEHCNLMFCRLVMLIGSSCFQVDNREDIGRGARRQKTSKTQIHLESLEH
QEIPRSTINQSSRRSPSKFIRERDLPLPYFASLPCAYAALEEAVAITQRRLADFPREVSFSKSDIQ
PGRRIERYLAVENHEQQNMLGESSFLVAGESAERRALHSGTYIPSQVESKWQRIWKLKQAVKVFD
HFVVVVVVVVVVVVVVVVVAINIHCFLKATSKGSENSNQGKKYILSMFPYPSGELHMGHVRVYTIS
DSLARYYRLQGFVLPMPGWDAFGLPAENAALERGVDPAEWTQKNISSMKKQMNSLGLSFDWE
LEVTTCDPEYYHWTQWLFLELFKSLAYHKAAKVNWDVDPDKTVLANEQVDAEGRSWSDELKE
WPEEVKCLQRNWIGLSKALAPEHPILSNPAFLGRLSKACLRDLDRYKKKVDSQTDIQRVSSTKSG
VNLGVTATHPWLERDLPIFVADYVLPQYGTGAVMGVPAHDTRDHIFAQHFDLSVLEVVIPEQDLS
QSTPKPKENEKQCYSGEGWMINSADLTGLSTASARDIISENLYLAKIGKTKIDFRLQDWLVSQR
YWGAPIPMIHCDSCGVVPPPREELPVLLPDLQGLLSEEARPWREVKCPCSGGMAERDITDITFV
DSSCLDHWMPVDVYIGGVEHAILHLLYSRFISRFLFSKGHLAQPEPFHCLLSQGLVLGETFRDATSN
RIVKRSELAQYEGKVVKVFEKMSKSKYNGVNPEDAVNAYGADVCLQGPRRWLQRLWLLAFEHL
ASEQKNPFLSKDKKSVEAIYSACSPDLDSEVALARMDAVKQVLHSSVVFVEALRSLLVMLA
PFAPHISAELFEHLSVIYNFKRQGAGICETHSQSVHEQLWPRVSTAEKELHLEKSVKTLVVVQVKG
KKLGLIISKDKLGQERELTRVVLESELGISHKLDQPKRVIVVPPKGGREFCLLNLVY

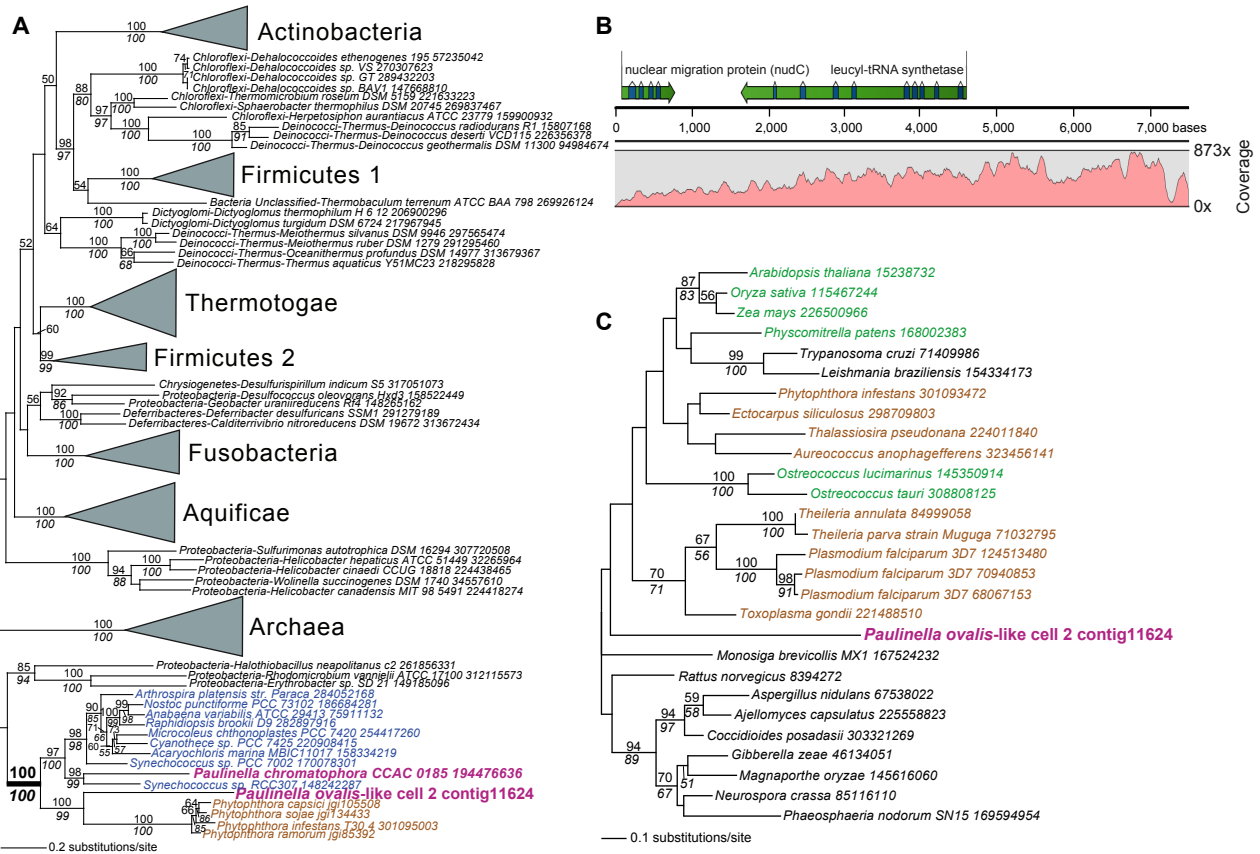


Fig. S10. Ancient cyanobacterial HGT found in the *P. ovalis*-like cell 2 SAG data. **A**, Maximum likelihood (RAxML, WAG + Γ + F model) phylogeny of leucyl-tRNA synthetase proteins. **B**, Intron distribution and coverage of *P. ovalis*-like cell 2 genome contig11624 that encodes two proteins. **C**, Maximum likelihood (RAxML, WAG + Γ + F model) phylogeny of nuclear migration proteins (nudC). In Figures 5A and 5C, Cyanobacteria are in blue text, other Bacteria are in black text, the chromatophore (plastid) and the *P. ovalis*-like cell data are in magenta text, Viridiplantae is in green text, and chromalveolates in brown text. RAxML and PhyML bootstrap values (100 replicates) are shown above and below the branches, respectively (only those > 50% are shown). The unit of branch length is the number of substitutions per site. The NCBI gi numbers (when available) are shown after each taxon name.

Table S1. Assembly output for the six *Paulinella ovalis*-like cells (SAGs) studied in our work. The 454 reads were assembled using the native Roche GS De Novo Assembler Software V2.5 beta.

	PovaMDc09	PovaMDc16	PovaMDc13	PovaMDc20	PovaMDc26	PovaMDc38
	Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
# of reads	781K	684K	727K	734K	608K	682K
# Mbp sequenced	308	247	180	190	180	221
Large Contigs:						
# of contigs	2084	2275	977	1126	1182	399
Avg. contig size	2329	2227	1989	2031	2086	2219
N50	2589	2456	2100	2163	2259	2370
Largest contig	29673	34453	14463	15081	26168	20823
All Contigs:						
# of contigs	6338	6865	4486	4766	4294	2067
# of bases	6783369	7204965	3493618	3821037	3882433	1516467
# BLASTx hits	2101	1954	1194	1147	1306	540

Table S& The number of protein sequences in the database that was used for the BLASTx and phylogenomic analysis (based on phyla).

Grouping	Species/Strain	RefSeq	JGI	EST ¹	Independent ²	Total
Archaea	121	225,867	0	0	0	225,867
Bact-Actinobacteria	282	994,292	0	0	0	994,292
Bact-Aquificae	10	19,325	0	0	0	19,325
Bact-BacteroidetesChlorobi	141	450,081	0	0	0	450,081
Bact-ChlamydiaeVerrucomicrobia	38	79,759	0	0	0	79,759
Bact-Chloroflexi	15	52,585	0	0	0	52,585
Bact-Cyanobacteria	68	225,555	0	0	0	225,555
Bact-Deferribacteres	2	5,338	0	0	0	5,338
PROKARYOTES Bact-Deinococci	12	26,191	0	0	0	26,191
Bact-Dictyoglomi	2	3,656	0	0	0	3,656
Bact-Elusimicrobia	2	2,305	0	0	0	2,305
Bact-Environmental	2	408	0	0	0	408
Bact-FibrobacteresAcidobacteria	6	28,629	0	0	0	28,629
Bact-Firmicutes	759	2,099,809	0	0	0	2,099,809
Bact-Fusobacteria	25	59,335	0	0	0	59,335
Bact-Gemmatimonadetes	1	3,935	0	0	0	3,935
Bact-Nitrospirae	3	6,366	0	0	0	6,366
Bact-Planctomycetes	6	36,794	0	0	0	36,794
Bact-Proteobacteria	1239	4,251,165	0	0	0	4,251,165
Bact-Spirochaetes	44	72,342	0	0	0	72,342
Bact-Synergistetes	6	13,162	0	0	0	13,162
Bact-Tenericutes	55	32,455	0	0	0	32,455
Bact-Thermotogae	11	20,807	0	0	0	20,807
Bact-Unclassified	9	17,518	0	0	0	17,518
Alveolata	70	167,836	0	584,904	0	752,740
Amoebozoa	22	30,550	12,410	138,624	0	181,584
Cryptophyta	8	1,419	0	40,320	0	41,739
Excavata	30	134,643	0	443,424	0	578,067
Haptophyta	5	140	39,124	56,868	0	96,132
Opisthokonta-Choanoflagellida	4	9,203	0	74,886	0	84,089
Opisthokonta-Fungi	186	569,377	212,456	132,168	0	914,001
Opisthokonta-Metazoa	2120	1,067,024	140,855	30,108	0	1,237,987
Opisthokonta-Others	4	0	0	46,494	0	46,494
Plantae-Glaucophyta	3	149	0	57,696	0	57,845
Plantae-Rhodophyta	23	1,168	0	331,482	28,975	361,625
Plantae-Viridiplantae	228	385,435	114,102	114,294	0	613,831
Rhizaria	5	1,211	0	29,112	0	30,323
Stramenopiles	47	41,980	81,762	96,078	0	219,820
Vira	2475	84,202	0	0	0	84,202
Others	39	1,062	0	0	0	1,062
Total	8,128	11,223,078	600,709	2,176,458	28,975	14,029,220

¹ The actual numbers of EST contigs are the numbers in this column divided by 6 due to six-frame translations.

² These data represent protein models from *Cyanidioschyzon merolae* and *Calliarthron tuberculosis*.

Contig1481	Stamenocleis-Fragilariaops_cnf2012968	2,00E-01
Contig1482	Stamenocleis-Fragilariaops_cnf2012968	1.80E-09
Contig1218	Stamenocleis-Fragilariaops_cnf2012458	1.10E-32
Contig5713	Stamenocleis-Fragilariaops_cnf2012458	1.70E-40
Contig6841	Stamenocleis-Fragilariaops_cnf2012158	5.70E-22
Contig4465	Stamenocleis-Fragilariaops_cnf2012178	3.10E-33
Contig2182	Stamenocleis-Fragilariaops_cnf2012473	2.70E-45
Contig4680	Stamenocleis-Fragilariaops_cnf2012473	7.80E-07
Contig2588	Stamenocleis-Fragilariaops_cnf2012473	1.60E-41
Contig4398	Stamenocleis-Fragilariaops_cnf2012918	4.00E-24
Contig1497	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g210118213	1.10E-31
Contig6222	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	4.80E-50
Contig1086	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	4.50E-50
Contig1761	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	2.00E-09
Contig2995	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	1.10E-11
Contig5294	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	4.30E-26
Contig4813	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	4.40E-28
Contig4799	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	7.90E-07
Contig5414	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	4.40E-28
Contig5308	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	3.50E-07
Contig5414	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011943	6.30E-24
Contig7256	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21011946	6.50E-07
Contig6230	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012454	5.70E-05
Contig4638	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012454	6.70E-07
Contig771	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012496	3.10E-13
Contig1521	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012496	1.10E-11
Contig4095	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012497	5.90E-12
Contig1033	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012497	3.40E-07
Contig1026	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012531	1.40E-07
Contig239	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012531	3.20E-05
Contig9295	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012802	1.74E-14
Contig14619	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012878	2.80E-09
Contig4714	Stamenocleis-Paeodactylum_Incunatum_CCP_1055_1_g21012878	1.30E-18
Contig1053	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.50E-16
Contig706	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.10E-27
Contig583	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.70E-25
Contig10378	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-10
Contig3807	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.50E-12
Contig4084	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.60E-11
Contig1285	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.80E-27
Contig3594	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.30E-47
Contig3626	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-06
Contig4703	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.80E-14
Contig6823	Stamenocleis-Phytophthora_brasicae_asp14491616_1	8.80E-14
Contig9311	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.00E-21
Contig352	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-12
Contig1470	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-12
Contig1452	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.90E-16
Contig4714	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-11
Contig1540	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.10E-13
Contig1017	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.10E-13
Contig5298	Stamenocleis-Phytophthora_brasicae_asp14491616_1	6.30E-12
Contig1764	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-14
Contig859	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.20E-49
Contig790	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.40E-28
Contig526	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.90E-12
Contig2841	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-32
Contig3165	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.70E-13
Contig469	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-06
Contig732	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.40E-06
Contig110	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.00E-06
Contig983	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-10
Contig1805	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.30E-27
Contig1467	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.30E-27
Contig1930	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.60E-13
Contig143	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.0E-07
Contig1963	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.30E-02
Contig137	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.30E-11
Contig1260	Stamenocleis-Phytophthora_brasicae_asp14491616_1	0
Contig1060	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.40E-09
Contig1884	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.60E-15
Contig4901	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-37
Contig1035	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.40E-13
Contig1215	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.60E-17
Contig4504	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.30E-22
Contig720	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.80E-32
Contig1426	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.80E-09
Contig347	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-02
Contig573	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.00E-07
Contig1030	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.10E-02
Contig1304	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.80E-34
Contig350	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.90E-27
Contig6198	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.90E-27
Contig1724	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.30E-08
Contig1405	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-12
Contig4794	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.10E-10
Contig4148	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.10E-15
Contig6645	Stamenocleis-Phytophthora_brasicae_asp14491616_1	8.30E-20
Contig4124	Stamenocleis-Phytophthora_brasicae_asp14491616_1	8.30E-20
Contig10074	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-20
Contig1048	Stamenocleis-Phytophthora_brasicae_asp14491616_1	6.80E-23
Contig8412	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.60E-23
Contig3481	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.60E-14
Contig1014	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-12
Contig1047	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.50E-14
Contig4834	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-12
Contig1826	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.60E-07
Contig1404	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.50E-15
Contig2080	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-12
Contig1241	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-53
Contig721	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.00E-09
Contig2676	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.00E-09
Contig9219	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.30E-10
Contig6812	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.10E-15
Contig3672	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.70E-21
Contig265	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.70E-21
Contig1037	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-32
Contig1945	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.40E-09
Contig1581	Stamenocleis-Phytophthora_brasicae_asp14491616_1	6.00E-10
Contig4752	Stamenocleis-Phytophthora_brasicae_asp14491616_1	6.10E-11
Contig10031	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-08
Contig1126	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-08
Contig147	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.00E-12
Contig792	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.00E-12
Contig3966	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-12
Contig1873	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.70E-14
Contig10326	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.80E-50
Contig5437	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.40E-03
Contig13671	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-23
Contig567	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-09
Contig10377	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.70E-34
Contig2229	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-17
Contig10546	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.10E-13
Contig3618	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-31
Contig6688	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.50E-12
Contig4571	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.80E-17
Contig309	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.50E-09
Contig419	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-07
Contig704	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.70E-06
Contig217	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-27
Contig26	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-27
Contig1794	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.70E-16
Contig2934	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.90E-05
Contig10430	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.70E-12
Contig147	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.70E-12
Contig954	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-08
Contig9211	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.70E-37
Contig296	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.30E-15
Contig5712	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.80E-15
Contig3417	Stamenocleis-Phytophthora_brasicae_asp14491616_1	8.20E-34
Contig5044	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.50E-06
Contig1320	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-20
Contig206	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-18
Contig8187	Stamenocleis-Phytophthora_brasicae_asp14491616_1	8.30E-31
Contig2229	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-16
Contig10397	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.40E-06
Contig1070	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.00E-14
Contig1171	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-19
Contig1363	Stamenocleis-Phytophthora_brasicae_asp14491616_1	9.40E-07
Contig1016	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.50E-31
Contig225	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-87
Contig11711	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-50
Contig10	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.90E-10
Contig1251	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.20E-21
Contig9713	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.80E-19
Contig427	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.00E-62
Contig416	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.30E-40
Contig4042	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.40E-35
Contig701	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-08
Contig936	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.70E-11
Contig2929	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-08
Contig6408	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-17
Contig2924	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.30E-46
Contig13187	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.10E-59
Contig2522	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.80E-08
Contig4585	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.90E-30
Contig3000	Stamenocleis-Phytophthora_brasicae_asp14491616_1	5.30E-16
Contig571	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.00E-11
Contig446	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-21
Contig5471	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-21
Contig4513	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-14
Contig1953	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-14
Contig9528	Stamenocleis-Phytophthora_brasicae_asp14491616_1	7.80E-06
Contig1348	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.50E-34
Contig1046	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.80E-06
Contig1396	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.70E-07
Contig115	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.50E-26
Contig1436	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.70E-16
Contig1365	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.40E-14
Contig383	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-11
Contig4122	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.10E-40
Contig3064	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.80E-54
Contig3847	Stamenocleis-Phytophthora_brasicae_asp14491616_1	4.20E-53
Contig497	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.10E-19
Contig187	Stamenocleis-Phytophthora_brasicae_asp14491616_1	3.70E-25
Contig4614	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.20E-12
Contig3878	Stamenocleis-Phytophthora_brasicae_asp14491616_1	8.00E-10
Contig1106	Stamenocleis-Phytophthora_brasicae_asp14491616_1	2.00E-03
Contig520	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.00E-53
Contig2678	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.50E-09
Contig473	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.80E-19
Contig79	Stamenocleis-Phytophthora_brasicae_asp14491616_1	1.80E-19
Contig734	Stamenocleis-Sargassum_binderi_asp12453168_2	2.10E-07
Contig1231	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299207	3.70E-13
Contig1790	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	2.40E-12
Contig148	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	7.70E-13
Contig9499	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	5.10E-52
Contig1117	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	3.70E-13
Contig611	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	1.30E-08
Contig14919	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	1.30E-40
Contig1347	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	1.60E-27
Contig7187	Stamenocleis-Thalassiosira_pseudonana_CCP1135_1_g210299243	1.60E-06

Table S4. Normalized BLAST top hits to contigs derived from the 454 + Illumina assembly of *P. ovalis* like cell 2 genome data.

Contig	Accession	Accession	Accession	Accession
Contig1147	Adriactinobacterium_ovalis_DSM_1521	4.802.16	XZF16566 arabinoside synthase [Arabinobacterium oviforme] DSM 15212	
Contig2111	Actinobacterium-Aeromonitium_haemolyticum_DSM_20955	4.906.20	YF020684 ATP synthase F1 subunit beta [Actinobacterium haemolyticum DSM 20955]	
Contig2109	Actinobacterium-Aeromonitium_haemolyticum_IC1_18119	2.842.09	YF024615 hypothetical protein_Ae147 [Actinobacterium IC1]	
Contig5827	Actinobacterium-Bacteroides_cavaema_DSM_12233	1.937.07	YF0281241 Phenylyl-CoA dioxygenase [Bacteroides cavaema DSM 12233]	
Contig8155	Actinobacterium-Bacteroides_jinnahii_82_210920	1.181.12	YF035316 sarA/warE/warF/warG-like protein [Bacteroides jinnahii 82]	
Contig4264	Actinobacterium-Calothrix_aestiva_DSM_44920	6.102.07	YF0112049 hypothetical protein_Can_2189 [Calothrix aestiva DSM 44920]	
Contig4168	Actinobacterium-Calothrix_aestiva_DSM_44920	6.102.10	YF011664 phosphatase [Calothrix aestiva DSM 44920]	
Contig4173	Actinobacterium-Calothrix_aestiva_DSM_44920	6.102.21	YF011727 ribonuclease H1 subunit [Calothrix aestiva DSM 44920]	
Contig11592	Actinobacterium-Calothrix_aestiva_DSM_4324	4.208.20	YF01049912 calyxin [Actinobacterium DSM 4324]	
Contig3493	Actinobacterium-Calothrix_aestiva_DSM_4324	4.208.20	YF01049912 calyxin [Actinobacterium DSM 4324]	
Contig262	Actinobacterium-Kinetococcus_industrialis_3653201	6.536.00	YF0116181 multidomain endonuclease [Kinetococcus industrialis 3653201]	
Contig222	Actinobacterium-Kinetococcus_industrialis_3653201	4.485.00	YF0116181 multidomain endonuclease [Kinetococcus industrialis 3653201]	
Contig1486	Actinobacterium-Kribbella_favosa_DSM_17326_2324	1.258.19	YF01301347 flavin-containing monooxygenase PPO [Kribbella favosa DSM 17326]	
Contig284	Actinobacterium-Micromonospora_aureofaciens_ATT_C2	6.482.25	YF01038848 adenine/methylphosphatase [Micromonospora aureofaciens ATT_C2]	
Contig1193	Actinobacterium-Mycobacterium_tarimense_ATT_C1347	1.536.00	XZF047712 cAMP-binding domain-containing protein [Mycobacterium tarimense ATT_C1347]	
Contig1183	Actinobacterium-Mycobacterium_tarimense_ATT_C1347	6.026.10	XZF047712 cAMP-binding domain-containing protein [Mycobacterium tarimense ATT_C1347]	
Contig11203	Actinobacterium-Mycobacterium_sp._HCS_G107870	6.116.00	YF015748 RNA domain-containing protein [Mycobacterium sp. HCS]	
Contig2100	Actinobacterium-Mycobacterium_sp._HCS_G107870	6.116.00	YF015748 RNA domain-containing protein [Mycobacterium sp. HCS]	
Contig7626	Actinobacterium-Mycobacterium_sp._HCS_G107870	4.538.09	XZF03908 X-ray diffracting lipopitidase-like protein [Mycobacterium sp. HCS]	
Contig1077	Actinobacterium-Nocardium_multivulgare_DSM_442	3.026.12	YF01030823 flavin monooxygenase/NADH oxidase [Nocardium multivulgare DSM 442]	
Contig4708	Actinobacterium-Nocardium_multivulgare_DSM_442	6.116.00	YF01030823 flavin monooxygenase/NADH oxidase [Nocardium multivulgare DSM 442]	
Contig4794	Actinobacterium-Nocardium_multivulgare_DSM_442	7.292.50	XZF013400 agrinine dehydratase [Nocardium multivulgare DSM 442]	
Contig4982	Actinobacterium-Nocardium_multivulgare_DSM_442	7.292.50	XZF013400 agrinine dehydratase [Nocardium multivulgare DSM 442]	
Contig950	Actinobacterium-Rhodococcus_opacus_84_G223628	2.908.30	YF0127004 oxidoreductase [Rhodococcus opacus 84]	
Contig2128	Actinobacterium-Rhodococcus_opacus_DSM_594	2.912.1	YF0144517 formaldehyde dehydrogenase [Rhodococcus opacus DSM 594]	
Contig1617	Actinobacterium-Rhodococcus_opacus_DSM_594	2.912.1	YF0144517 formaldehyde dehydrogenase [Rhodococcus opacus DSM 594]	
Contig12352	Actinobacterium-Rhodococcus_opacus_DSM_594	2.292.11	YF0145799 pyridoxamine 5-phosphate oxidase-like protein [Rhodococcus opacus DSM 594]	
Contig474	Actinobacterium-Rhodococcus_opacus_DSM_594	2.912.1	YF0145799 pyridoxamine 5-phosphate oxidase-like protein [Rhodococcus opacus DSM 594]	
Contig1050	Actinobacterium-Sanguibacter_bekdadii_DSM_10542	2.026.00	YF01314440 signal transduction tyrosine kinase [Sanguibacter bekdadii DSM 10542]	
Contig1631	Actinobacterium-Sanguibacter_bekdadii_DSM_10542	2.026.00	YF01314440 signal transduction tyrosine kinase [Sanguibacter bekdadii DSM 10542]	
Contig1631	Actinobacterium-Sanguibacter_bekdadii_DSM_10542	1.258.10	YF0121077 hypothetical protein_SAV_802 [Sanguibacter bekdadii DSM 10542]	
Contig919	Actinobacterium-Sanguibacter_bekdadii_DSM_10542	1.258.10	YF0121077 hypothetical protein_SAV_802 [Sanguibacter bekdadii DSM 10542]	
Contig1845	Actinobacterium-Streptomyces_ghanemensis_ATT_C4	3.476.39	XZF057588 serine protease [Streptomyces ghanemensis ATT_C4]	
Contig1627	Actinobacterium-Streptomyces_ghanemensis_ATT_C4	3.476.39	XZF057588 serine protease [Streptomyces ghanemensis ATT_C4]	
Contig1646	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	4.288.10	XZF072628 N50R family peptidase [Streptomyces hygroscopicus ATT_C2]	
Contig1649	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	4.288.10	XZF072628 N50R family peptidase [Streptomyces hygroscopicus ATT_C2]	
Contig11819	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	1.258.10	YF0176123 predicted protein [Streptomyces sp. CI]	
Contig8075	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	1.258.10	YF0176123 predicted protein [Streptomyces sp. CI]	
Contig775	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	1.116.14	XZF013400 hypothetical protein_Scs_6553 [Streptomyces hygroscopicus DSM 4301]	
Contig1276	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	1.116.14	XZF013400 hypothetical protein_Scs_6553 [Streptomyces hygroscopicus DSM 4301]	
Contig1209	Actinobacterium-Streptomyces_hygroscopicus_ATT_C2	1.258.25	XZF0130769 protein [Streptomyces hygroscopicus DSM 4301]	
Contig424	Aveolidea-Aveolidea_calyculata_espig01805104_1	6.842.11	XFX45561 avc147879 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig5006	Aveolidea-Aveolidea_calyculata_espig01805104_2	5.028.25	XFX45999 avc154522 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig5005	Aveolidea-Aveolidea_calyculata_espig01805104_3	8.708.10	XFX45251 avc152030 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig278	Aveolidea-Aveolidea_calyculata_espig01805104_4	1.138.30	XFX45831 avc166474 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig4773	Aveolidea-Aveolidea_calyculata_espig01805104_5	1.138.30	XFX45831 avc166474 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig4773	Aveolidea-Aveolidea_calyculata_espig01805104_6	1.138.30	XFX45831 avc166474 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig4773	Aveolidea-Aveolidea_calyculata_espig01805104_7	1.138.30	XFX45831 avc166474 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig234	Aveolidea-Aveolidea_calyculata_espig01805104_8	4.338.22	XFX43888 avc187740 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig234	Aveolidea-Aveolidea_calyculata_espig01805104_9	4.338.22	XFX43888 avc187740 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig1036	Aveolidea-Aveolidea_calyculata_espig01805104_10	7.482.21	XFX451588 avc194790 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig1036	Aveolidea-Aveolidea_calyculata_espig01805104_11	7.482.21	XFX451588 avc194790 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig151	Aveolidea-Aveolidea_calyculata_espig01805104_12	1.536.29	XFX45685 avc201433 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig151	Aveolidea-Aveolidea_calyculata_espig01805104_13	1.536.29	XFX45685 avc201433 Aveolidea calyculata ACCYV expressed sequence tags Aveolidea calyculata cDNA, mRNA sequence	
Contig4642	Aveolidea-Aveolidea_minutum_espig20784138_2	6.798.12	XFW297438 Avn0011001P028122_phn111-vc Aveolidea minutum normalized EST library Aveolidea minutum cDNA, mRNA sequence	
Contig959	Aveolidea-Aveolidea_minutum_espig20784138_3	6.798.12	XFW297438 Avn0011001P028122_phn111-vc Aveolidea minutum normalized EST library Aveolidea minutum cDNA, mRNA sequence	
Contig1607	Aveolidea-Aveolidea_minutum_espig20784138_4	3.088.11	XFW89048 Avn0011001P0142028_phn111-vc Aveolidea minutum normalized EST library Aveolidea minutum cDNA, mRNA sequence	
Contig1608	Aveolidea-Aveolidea_minutum_espig20784138_5	3.088.11	XFW89048 Avn0011001P0142028_phn111-vc Aveolidea minutum normalized EST library Aveolidea minutum cDNA, mRNA sequence	
Contig1289	Aveolidea-Aveolidea_minutum_espig20784138_6	3.088.11	XFW89048 Avn0011001P0142028_phn111-vc Aveolidea minutum normalized EST library Aveolidea minutum cDNA, mRNA sequence	
Contig1289	Aveolidea-Aveolidea_minutum_espig20784138_7	3.088.11	XFW89048 Avn0011001P0142028_phn111-vc Aveolidea minutum normalized EST library Aveolidea minutum cDNA, mRNA sequence	
Contig171	Aveolidea-Babaia_bovis_728_G1560819_1	2.322.22	XZF025684 avc1141210 Babaia bovis EST library [Babaia bovis ATCC 6255], mRNA sequence	
Contig171	Aveolidea-Babaia_bovis_728_G1560819_2	2.322.22	XZF025684 avc1141210 Babaia bovis EST library [Babaia bovis ATCC 6255], mRNA sequence	
Contig1059	Aveolidea-Babaia_bovis_728_G1560819_3	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1059	Aveolidea-Babaia_bovis_728_G1560819_4	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_5	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_6	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_7	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_8	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_9	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_10	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_11	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_12	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_13	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_14	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_15	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_16	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_17	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_18	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_19	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_20	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_21	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_22	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_23	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_24	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_25	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_26	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_27	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_28	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_29	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_30	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_31	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_32	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_33	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_34	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_35	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_36	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_37	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_38	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_39	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_40	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_41	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_42	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_43	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_44	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_45	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_46	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_47	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_48	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_49	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_50	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_51	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_52	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_53	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_54	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_55	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_56	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_57	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_58	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_59	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_60	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_61	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_62	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_63	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_64	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_65	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_66	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_67	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_68	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_69	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_70	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_71	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_72	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_73	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_74	3.958.10	XZF0102484 AvfAAc AvfAA family (Babaia bovis)	
Contig1529	Aveolidea-Babaia_bovis_728_G1560819_75	3.958.10	XZF	

Contig13785 Vridplante-Volvo_carter1_f_naganema_g30281 6.30E-09 >XP_002947840 hypothetical protein VOLCADRAFT_88148 [Volvo cartier f_naganema]
Contig13804 Vridplante-Volvo_carter1_f_naganema_g30281 6.30E-09 >XP_002947868 hypothetical protein VOLCADRAFT_88144 [Volvo cartier f_naganema]
Contig13807 Vridplante-Volvo_carter1_f_naganema_g30281 8.82E-06 >XP_002949059 hypothetical protein VOLCADRAFT_39977 [Volvo cartier f_naganema]
Contig13817a Vridplante-Volvo_carter1_f_naganema_g30281 1.78E-15 >XP_002949912 hypothetical protein VOLCADRAFT_32474 [Volvo cartier f_naganema]
Contig14046 Vridplante-Volvo_carter1_f_naganema_g30281 2.26E-27 >XP_002949262 hypothetical protein VOLCADRAFT_39216 [Volvo cartier f_naganema]
Contig14048 Vridplante-Volvo_carter1_f_naganema_g30281 6.78E-25 >XP_002950039 hypothetical protein VOLCADRAFT_39216 [Volvo cartier f_naganema]
Contig14049 Vridplante-Volvo_carter1_f_naganema_g30281 6.78E-25 >XP_002950039 hypothetical protein VOLCADRAFT_39218 [Volvo cartier f_naganema]
Contig14078 Vridplante-Volvo_carter1_f_naganema_g30281 3.68E-14 >XP_002950289 hypothetical protein VOLCADRAFT_39270 [Volvo cartier f_naganema]
Contig1412 Vridplante-Volvo_carter1_f_naganema_g30281 4.95E-18 >XP_002950289 hypothetical protein VOLCADRAFT_39288 [Volvo cartier f_naganema]
Contig14214 Vridplante-Volvo_carter1_f_naganema_g30281 8.76E-07 >XP_002950328 hypothetical protein VOLCADRAFT_39288 [Volvo cartier f_naganema]
Contig14464 Vridplante-Volvo_carter1_f_naganema_g30281 3.78E-07 >XP_002950429 hypothetical protein VOLCADRAFT_39292 [Volvo cartier f_naganema]
Contig14600 Vridplante-Volvo_carter1_f_naganema_g30281 7.30E-07 >XP_002950559 hypothetical protein VOLCADRAFT_39278 [Volvo cartier f_naganema]
Contig14412 Vridplante-Volvo_carter1_f_naganema_g30281 2.18E-08 >XP_002950834 hypothetical protein VOLCADRAFT_39390 [Volvo cartier f_naganema]
Contig14707 Vridplante-Volvo_carter1_f_naganema_g30281 4.00E-07 >XP_002950941 hypothetical protein VOLCADRAFT_39492 [Volvo cartier f_naganema]
Contig14907 Vridplante-Volvo_carter1_f_naganema_g30281 3.89E-21 >XP_002951091 hypothetical protein VOLCADRAFT_44554 [Volvo cartier f_naganema]
Contig14909 Vridplante-Volvo_carter1_f_naganema_g30281 6.20E-21 >XP_002951155 hypothetical protein VOLCADRAFT_39497 [Volvo cartier f_naganema]
Contig14949 Vridplante-Volvo_carter1_f_naganema_g30281 5.96E-26 >XP_002951204 hypothetical protein VOLCADRAFT_44554 [Volvo cartier f_naganema]
Contig14965 Vridplante-Volvo_carter1_f_naganema_g30281 5.30E-06 >XP_002951491 hypothetical protein VOLCADRAFT_39247 [Volvo cartier f_naganema]
Contig15149 Vridplante-Volvo_carter1_f_naganema_g30281 5.56E-25 >XP_002951638 hypothetical protein VOLCADRAFT_32106 [Volvo cartier f_naganema]
Contig13885 Vridplante-Volvo_carter1_f_naganema_g30284 2.90E-08 >XP_002953254 hypothetical protein VOLCADRAFT_96131 [Volvo cartier f_naganema]
Contig14790 Vridplante-Volvo_carter1_f_naganema_g30284 1.60E-26 >XP_002953346 hypothetical protein VOLCADRAFT_39287 [Volvo cartier f_naganema]
Contig17271 Vridplante-Volvo_carter1_f_naganema_g30286 6.70E-23 >XP_002955667 hypothetical protein VOLCADRAFT_67197 [Volvo cartier f_naganema]
Contig12644 Vridplante-Volvo_carter1_f_naganema_g30286 7.70E-55 >XP_002956829 hypothetical protein VOLCADRAFT_67893 [Volvo cartier f_naganema]
Contig16003 Vridplante-Volvo_carter1_f_naganema_g30286 4.50E-26 >XP_002957233 hypothetical protein VOLCADRAFT_63997 [Volvo cartier f_naganema]
Contig14912 Vridplante-Volvo_carter1_f_naganema_g30286 1.20E-46 >XP_002958029 hypothetical protein VOLCADRAFT_39362 [Volvo cartier f_naganema]
Contig14915 Vridplante-Volvo_carter1_f_naganema_g30286 2.30E-11 >XP_002959726 hypothetical protein VOLCADRAFT_31081 [Volvo cartier f_naganema]
Contig14624 Vridplante-Zea_mays_g21227370 6.30E-38 >NP_001130129 hypothetical protein LOC00191216 [Zea mays]
Contig14953 Vridplante-Zea_mays_g21227370 2.70E-15 >NP_001137026 hypothetical protein LOC00217721 [Zea mays]
Contig15343 Vridplante-Zea_mays_g21238271 1.90E-14 >NP_001137122 hypothetical protein LOC00217502 [Zea mays]
Contig14769 Vridplante-Zea_mays_g21249274 1.30E-10 >NP_001148518 hypothetical protein C_sulfur_2 [Zea mays]
Contig11500 Vridplante-Zea_mays_g21249274 2.50E-08 >NP_001148509 hypothetical protein LOC00279203 [Zea mays]
Contig14926 Vridplante-Zea_mays_g21249274 1.60E-08 >NP_001149093 RNA binding motif protein, Y-sister 2 [Zea mays]
Contig11922 Vridplante-Zea_mays_g212607544 1.10E-10 >NP_001152493 LOC100286123 [Zea mays]
Contig15463 Vridplante-Zea_mays_g212607500 5.10E-07 >NP_001144303 hypothetical protein LOC00173195 [Zea mays]
Contig1042 Vridplante-Zea_mays_g212610036 2.40E-40 >NP_001147911 ATP-dependent RNA helicase DRX23 [Zea mays]
Contig15066 Vridplante-Zea_mays_g212609087 1.60E-05 >NP_001150313 hypothetical protein LOC00308420 [Zea mays]
Contig14111 Vridplante-Zea_mays_g213331561 1.80E-19 >NP_001169685 hypothetical protein LOC00383566 [Zea mays]
Contig14862 Vridplante-Zea_mays_g213331544 5.10E-15 >NP_001169316 hypothetical protein LOC00383205 [Zea mays]
Contig1382 Vridplante-Zea_mays_g213333855 5.90E-14 >NP_001167960 hypothetical protein LOC00384853 [Zea mays]
Contig14952 Vridplante-Zea_mays_g213333820 1.60E-52 >NP_001169686 hypothetical protein LOC00383379 [Zea mays]
Contig2005 Vridplante-Zea_mays_g30808266 4.50E-17 >NP_001181817 hypothetical protein LOC100501566 [Zea mays]