Templeton et al. Fig. S2.

Evidence of intron presence in *Ascogregarina* genes of bacterial affinity in by BLAST search. Three examples are given below with only top hits listed. The presence of introns are indicated by the gaps between aligned fragments and sometimes by the stop codons within aligned sequences.

A) Top hit in BLAST search for contig 2623:

```
>ref|ZP_01101351.1| phospho-2-dehydro-3-deoxyheptonate aldolase, phe-sensitive [gamma
proteobacterium KT 71]
gb|EAQ99452.1| phospho-2-dehydro-3-deoxyheptonate aldolase, phe-
sensitive [Congregibacter  litoralis KT71] Length=352

Score = 35.8 bits (81),  Expect(3) = 3e-29
 Identities = 16/26 (61%), Positives = 20/26 (76%), Gaps = 0/26 (0%)  Frame = -1

Query  1002  QLATGSTRIMGVMIESNLYEGGQNLI  925
             Q+A GS  I G+MIESNL EG QN++
Sbjct  288   QVARGSRSIFGIMIESNLVEGRQNVV  313

Score = 79.7 bits (195),  Expect(3) = 3e-29
 Identities = 39/54 (72%), Positives = 42/54 (77%), Gaps = 0/54 (0%)  Frame = -1

Query  1605  GAIGARTSECQLHRELASGLSMPVGFKNGTTGDVQVAVDACVSVKHPHTFLGVS  1444
             GAIGART+E Q HRELASGLS PVGFKN T GDVQVA+DA  S   PH FL V+
Sbjct  160   GAIGARTTESQTHRELASGLSCPVGFKNATDGDVQVAIDAINSASQPHRFLSVT  213

Score = 58.9 bits (141),  Expect(3) = 3e-29
 Identities = 36/112 (32%), Positives = 51/112 (45%), Gaps = 23/112 (20%)  Frame = -2

Query  1421  RFF*ITKQGLPAICHSTV*ASSQIHRS*NSPYSMYAIH*GNPDCHIILRGASSGPNYSEA  1242
             RF  +TKQG  AI  +T                      GN DCHIILRG    PNY
Sbjct  208   RFLSVTKQGHSAIFQTT----------------------GNADCHIILRGGKH-PNYDMF  244

Query  1241  HIEAAATLLERGGCSARLVVDFSHGNSQKVRATSVITCFRITKDRKMSARQL  1086
              ++ A+ +L + G  AR+++D SH NS+K+ A  V     +       +R +
Sbjct  245   SVDDASAMLSKSGLPARIMIDASHANSRKIPARQVDVAMDVASQVARGSRSI  296
```

B) Top hit in BLAST search for contig 1226:

```
>ref|YP_132940.1| short chain dehydrogenase [Photobacterium profundum SS9]
 emb|CAG23140.1| hypothetical dehydrogenase [Photobacterium profundum SS9] Length=275

Score = 81.3 bits (199),  Expect(2) = 2e-17
 Identities = 52/164 (31%), Positives = 81/164 (49%), Gaps = 37/164 (22%)  Frame = +2

Query  671   KGCKVTVSAKSTQEQPNLPGTIYSVAKEVEALGVEALPFKV*CP*CRTDRTHGRCHRREM  850
             +G  + ++AKS Q  P L GTI+SVA+EVE  G  AL  K+
Sbjct  29    QGANIVIAAKSDQPHPKLSGTIHSVAQEVEEAGGHALAIKL-------------------  69

Query  851   GPVRXXXXXXXXXXXXXX*XT*RQLGVWFGVVLCFWCCRVDIMVANAGALWWKDMVDTPM  1030
               VR   +   +          +Q    FG         +D++V NA A+    + DT
Sbjct  70    -DVRDEEAVNLAM---------QQAFETFG--------SIDVLVNNASAIVLTRLQDTDT  111
```

```
Query  1031  RRYDLINDVNARATFACAKAVLPHMLRQKHGHIITMSPPLSFFS  1162
             +R+DLIN +N R +F C+KA +P++ + ++ HIIT+SPP++   S
Sbjct  112   KRFDLINSINVRGSFVCSKAAIPYLKKSRNPHIITLSPPINLAS  155

Score = 33.5 bits (75),  Expect(2) = 2e-17
 Identities = 16/27 (59%), Positives = 18/27 (66%), Gaps = 0/27 (0%)  Frame = +1

Query  469  LAGQVAIITGASRGIGREVALAYAAAG  549
            L G    ITG+SRGIGRE+AL  A  G
Sbjct  4    LTGVTVFITGSSRGIGREIALICAQQG  30
```

C) Top hit in BLAST search for contig 3330:

```
>ref|ZP_01308644.1| Propionyl-CoA carboxylase [Oceanobacter sp. RED65]
 gb|EAT10723.1| Propionyl-CoA carboxylase [Oceanobacter sp. RED65] Length=535

Score =  140 bits (353),  Expect(2) = 1e-54
 Identities = 76/163 (46%), Positives = 95/163 (58%), Gaps = 29/163 (17%)  Frame = +3

Query  474  VGSEAERSGIAKHGAKLVTAVSCAAVPKITVIVGASFGAGNYAMCGRAYDPRFLFSWPNA  653
            VGS+ E+ GIAKHGAK+V AV+CA VPK T+++G SFGAGNY MCGRAYDP F++ WPNA
Sbjct  381  VGSKYEQEGIAKHGAKMVNAVACAKVPKFTMVIGGSFGAGNYGMCGRAYDPNFMWMWPNA  440

Query  654  RVGVMGGQQAAGMPSLCTNMALAV*MF*WKWSAVASPNNMKKIPRN*RLP*ISAE*VSPL  833
            R+ VMGG+QAAG+ +L T  A        K     S  +
Sbjct  441  RISVMGGEQAAGVMALITKGAKE------KRGEPFSEQD---------------------  473

Query  834  HLSRGF*AEVTAVYEQQSYAEYGSARLWDDAVIQPKQTRQVRG  962
                 F A +   YEQ+ +  Y SARLWDD VI P Q+R+V G
Sbjct  474  --EAAFKAPLLEKYEQEGHPYYASARLWDDGVIDPAQSREVLG  514

Score = 99.8 bits (247),  Expect(2) = 1e-54
 Identities = 54/124 (43%), Positives = 78/124 (62%), Gaps = 13/124 (10%)  Frame = +2

Query  65   EVAEIIEPPRDCEEELDLIAPLGFQWKVSGVKGLLSRLLDGSHFAEFKPEFGTSLVTGWG  244
            +VA+++ P    EE   L+ P   + K   V  +++R++DGS F E+K  FG SLV G+
Sbjct  275  DVADVVAPLYPAEEIYGLV-PSDLK-KPMDVHEIIARIVDGSEFDEYKARFGNSLVCGFA  332

Query  245  KIMGHPVGGLKYIHNKSVGILANSGVLTCDAAVKGTHFIQLCCSRQTPLLFLQNITGFMV  424
            +I G+PVG           I+AN+G+L  ++A KG +FI+LCC R  PL+FLQNITGFMV
Sbjct  333  RIYGYPVG-----------IIANAGILFSESAQKGANFIELCCQRNIPLVFLQNITGFMV  381
```