MrGreen annotation notes.

Made from the PECAAN output with extra notes entered into DNAMaster while doing some QC.

CDS 41 - 379

/gene="1"

/function="hypothetical function"

/product="hypothetical function"

/locus tag="SEA\_MRGREEN\_1"

/note=Original Glimmer call @bp 41 has strength 13.15; Genemark calls start at 41

/note=

/note=SSC: Start = 41, Stop = 379. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 2.055 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 339 bp is not the longest possible ORF. GAP: 0 bp. ST: SS=NA. F: hypothetical function. FS: PHDBLAST= PhageName= Azizam, ProteinNumber= 1, Function= function unknown, EValue= 7.0E-60. NCBIBLAST= PhageName= hypothetical protein JTF52\_gp01 [Microbacterium phage Azizam] >gb|QDF18028.1| hypothetical protein BELTHELAS\_1 [Microbacterium phage Belthelas] >gb|QGJ97461.1| hypothetical protein SEA\_AZIZAM\_1 [Microbacterium phage Azizam] >gb|QJD53899.1| hypothetical protein SEA\_MCSHIE\_1 [Microbacterium phage McShie], Coverage= 100.0, SubjectRange= 1:112, QueryRange= 1:112, EValue= 8.93118E-74. HHPRED= . CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 376 - 1089

/gene="2"

/function="hypothetical function"

/product="hypothetical function"

/locus tag="SEA\_MRGREEN\_2"

/note=Original Glimmer call @bp 361 has strength 8.71; Genemark calls start at 376

/note=

/note=SSC: Start = 376, Stop = 1089. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.332 is not the highest start score. SCS: Start is called by Glimmer and is not called by Genemark. LO: 714 bp is not the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: hypothetical function. FS: PHDBLAST= PhageName= Scamander, ProteinNumber= 2, Function= function unknown, EValue= 1.0E-129. NCBIBLAST= PhageName= hypothetical protein SEA\_SCAMANDER\_2 [Microbacterium phage Scamander] >gb|QBZ72505.1| hypothetical protein SEA\_RHYSAND\_2 [Microbacterium phage Rhysand] >gb|QDF15887.1| hypothetical protein SEA\_LAVIMO\_2 [Microbacterium phage LaviMo] >gb|QFP94572.1| hypothetical protein SEA\_VANISIUS\_2 [Microbacterium phage Vanisius] >gb|QJD53925.1| hypothetical protein SEA\_OTWOR\_2 [Microbacterium phage Otwor] >gb|QJD54065.1| hypothetical protein SEA\_DANNO\_2 [Microbacterium phage Danno], Coverage= 97.9339, SubjectRange= 1:235, QueryRange= 1:242, EValue= 2.94961E-156. HHPRED= . CDD= .

/note=

/note=Gap: -4 checked all sources

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 1092 - 2552

/gene="3"

/function="terminase"

/product="terminase"

/locus tag="SEA\_MRGREEN\_3"

/note=Original Glimmer call @bp 1092 has strength 10.85; Genemark calls start at 1092

/note=

/note=SSC: Start = 1092, Stop = 2552. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 2.482 is not the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 1461 bp is the longest possible ORF. GAP: 2 bp. ST: SS=NA. F: terminase. FS: PHDBLAST= PhageName= Azizam, ProteinNumber= 3, Function= terminase, EValue= 0.0. NCBIBLAST= PhageName= terminase [Microbacterium phage Azizam] >gb|QGJ97463.1| terminase [Microbacterium phage Azizam], Coverage= 100.0, SubjectRange= 1:486, QueryRange= 1:486, EValue= 0.0. HHPRED= Accession= PF03354.16, Description= Terminase\_1 ; Phage Terminase, Probability= 100.0. Coverage= 88.0658, SubjectRange= 1:467, QueryRange= 1:479. CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 2650 - 3804

/gene="4"

/function="portal protein"

/product="portal protein"

/locus tag="SEA\_MRGREEN\_4"

/note=Original Glimmer call @bp 2650 has strength 10.21; Genemark calls start at 2746

/note=

/note=SSC: Start = 2650, Stop = 3804. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 1.264 is not the highest start score. SCS: Start is called by Glimmer and is not called by Genemark. LO: 1155 bp is the longest possible ORF. GAP: 97 bp. ST: SS=NA. F: portal protein. FS: PHDBLAST= PhageName= Bri160, ProteinNumber= 4, Function= portal protein, EValue= 0.0. NCBIBLAST= PhageName= portal protein [Microbacterium phage Bri160] >gb|QLF83184.1| portal protein [Microbacterium phage Bri160], Coverage= 100.0, SubjectRange= 1:384, QueryRange= 1:384, EValue= 0.0. HHPRED= Accession= 6TE9\_B, Description= Phage portal protein, HK97 family;

/note=Gap: 98

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 3801 - 5384

/gene="5"

/function="major capsid and protease fusion protein"

/product="major capsid and protease fusion protein"

/locus tag="SEA\_MRGREEN\_5"

/note=Original Glimmer call @bp 3801 has strength 14.08; Genemark calls start at 3801

/note=

/note=SSC: Start = 3801, Stop = 5384. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 3.097 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 1584 bp is not the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: major capsid and protease fusion protein. FS: PHDBLAST= PhageName= Quaker, ProteinNumber= 5, Function= major capsid and protease fusion protein, EValue= 0.0. NCBIBLAST= PhageName= major capsid and protease fusion protein [Microbacterium phage Quaker] >gb|AXC35363.1| major capsid and protease fusion protein [Microbacterium phage KayPaulus] >gb|QKY78771.1| major capsid and protease fusion protein [Microbacterium phage Livingwater] >gb|QOI67281.1| major capsid and protease fusion protein [Microbacterium phage Sippinontea] >gb|QRI45125.1| major capsid and protease fusion protein [Microbacterium phage Wolfpack] >gb|AXC35996.1| major capsid and protease fusion protein [Microbacterium phage Quaker], Coverage= 100.0, SubjectRange= 1:527, QueryRange= 1:527, EValue= 0.0. HHPRED= Accession= 6TSU\_T4, Description= Major capsid protein Rcc01687;

/note=This start allows for a 4bp overlap.

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 5388 - 5741

/gene="6"

/function="head-to-tail adaptor"

/product="head-to-tail adaptor"

/locus tag="SEA\_MRGREEN\_6"

/note=Original Glimmer call @bp 5388 has strength 17.77; Genemark calls start at 5388

/note=

/note=SSC: Start = 5388, Stop = 5741. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.672 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 354 bp is the longest possible ORF. GAP: 3 bp. ST: SS=NA. F: head-to-tail adaptor. FS: PHDBLAST= PhageName= Belthelas, ProteinNumber= 6, Function= head-to-tail adaptor, EValue= 5.0E-60. NCBIBLAST= PhageName= head-to-tail adaptor [Microbacterium phage Belthelas], Coverage= 100.0, SubjectRange= 1:117, QueryRange= 1:117, EValue= 8.87778E-79. HHPRED= Accession= 5A21\_C, Description= 15 PROTEIN; VIRAL PROTEIN, VIRAL INFECTION, TAILED BACTERIOPHAGE, SIPHOVIRIDAE, SPP1, VIRAL ASSEMBLY, HEAD-TO-TAIL INTERFACE, DNA GATEKEEPER, ALLOSTERIC MECHANISM; 7.2A {BACILLUS PHAGE SPP1}, Probability= 98.9. Coverage= 88.8889, SubjectRange= 1:102, QueryRange= 1:116. CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 5738 - 6118

/gene="7"

/function="tail terminator"

/product="tail terminator"

/locus tag="SEA\_MRGREEN\_7"

/note=Original Glimmer call @bp 5738 has strength 15.38; Genemark calls start at 5738

/note=

/note=SSC: Start = 5738, Stop = 6118. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 1.515 is not the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 381 bp is the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: tail terminator. FS: PHDBLAST= PhageName= Quaker, ProteinNumber= 7, Function= tail terminator, EValue= 2.0E-65. NCBIBLAST= PhageName= tail terminator [Microbacterium phage Quaker] >gb|AXC35365.1| tail terminator [Microbacterium phage KayPaulus] >gb|QDF18034.1| minor tail protein [Microbacterium phage Belthelas] >gb|QKY78773.1| tail terminator [Microbacterium phage Livingwater] >gb|QOI67283.1| tail terminator [Microbacterium phage Sippinontea] >gb|QRI45127.1| tail terminator [Microbacterium phage Wolfpack], Coverage= 100.0, SubjectRange= 1:126, QueryRange= 1:126, EValue= 2.95531E-84. HHPRED= . CDD= .

/note=

/note=This start site allows for a 4 bp overlap.

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 6157 - 6591

/gene="8"

/function="major tail protein"

/product="major tail protein"

/locus tag="SEA\_MRGREEN\_8"

/note=Original Glimmer call @bp 6157 has strength 12.9; Genemark calls start at 6157

/note=

/note=SSC: Start = 6157, Stop = 6591. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 1.529 is not the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 435 bp is the longest possible ORF. GAP: 38 bp. ST: SS=NA. F: major tail protein. FS: PHDBLAST= PhageName= Bri160, ProteinNumber= 8, Function= major tail protein, EValue= 2.0E-81. NCBIBLAST= PhageName= major tail protein [Microbacterium phage Bri160] >ref|YP\_009996769.1| major tail protein [Microbacterium phage Quaker] >gb|AXC35366.1| major tail protein [Microbacterium phage KayPaulus] >gb|QDF18035.1| major tail protein [Microbacterium phage Belthelas] >gb|QKY78774.1| major tail protein [Microbacterium phage Livingwater] >gb|QOI67284.1| major tail protein [Microbacterium phage Sippinontea] >gb|QRI45128.1| major tail protein [Microbacterium phage Wolfpack], Coverage= 100.0, SubjectRange= 1:144, QueryRange= 1:144, EValue= 7.97173E-100. HHPRED= Accession= 2K4Q\_A, Description= Major tail protein V; gpV, Bacteriophage Lambda, Major tail protein, VIRAL PROTEIN; NMR {Enterobacteria phage lambda}, Probability= 98.9. Coverage= 97.9167, SubjectRange= 8:156, QueryRange= 8:142. CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 6604 - 6987

/gene="9"

/function="hypothetical function"

/product="hypothetical function"

/locus tag="SEA\_MRGREEN\_9"

/note=Original Glimmer call @bp 6604 has strength 15.37; Genemark calls start at 6604

/note=

/note=SSC: Start = 6604, Stop = 6987. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.25 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 384 bp is the longest possible ORF. GAP: 12 bp. ST: SS=NA. F: hypothetical function. FS: PHDBLAST= PhageName= BurtonThePup, ProteinNumber= 9, Function= function unknown, EValue= 1.0E-66. NCBIBLAST= PhageName= hypothetical protein JTF51\_gp09 [Microbacterium phage PaoPu] >ref|YP\_009996619.1| hypothetical protein JTF52\_gp09 [Microbacterium phage Azizam] >ref|YP\_009996644.1| hypothetical protein JTF53\_gp09 [Microbacterium phage Bri160] >ref|YP\_009996669.1| hypothetical protein JTF54\_gp09 [Microbacterium phage Kaijohn] >ref|YP\_009996695.1| hypothetical protein JTF55\_gp09 [Microbacterium phage Nobel] >ref|YP\_009996770.1| minor tail protein [Microbacterium phage Quaker] >gb|AVR56186.1| hypothetical protein SEA\_BURTONTHEPUP\_9 [Microbacterium phage BurtonThePup] >gb|AXC35367.1| hypothetical protein SEA\_KAYPAULUS\_9 [Microbacterium phage KayPaulus] >gb|AXH66495.1| hypothetical protein SEA\_SCAMANDER\_9 [Microbacterium phage Scamander] >gb|AXQ64358.1| minor tail protein [Microbacterium phage Minima] , Coverage= 100.0, SubjectRange= 1:127, QueryRange= 1:127, EValue= 2.16294E-85. HHPRED= . CDD= .

/note=

/note=Gap: 13 While the gene has no function according to NCBI, the rest of the EE cluster says that gene 9 is a minor tail protein. upstream (gene 8): major tail protein, downstream (gene 10): tail assembly chaperone, BurtonThePup. checked all sources

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 7001 - 7321

/gene="10"

/function="tail assembly chaperone"

/product="tail assembly chaperone"

/locus tag="SEA\_MRGREEN\_10"

/note=Original Glimmer call @bp 7001 has strength 18.36; Genemark calls start at 7001

/note=

/note=SSC: Start = 7001, Stop = 7321. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.485 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 321 bp is the longest possible ORF. GAP: 13 bp. ST: SS=NA. F: tail assembly chaperone. FS: PHDBLAST= PhageName= Bri160, ProteinNumber= 10, Function= tail assembly chaperone, EValue= 3.0E-53. NCBIBLAST= PhageName= tail assembly chaperone [Microbacterium phage Bri160] >gb|QLF83170.1| tail assembly chaperone [Microbacterium phage Bri160], Coverage= 100.0, SubjectRange= 1:106, QueryRange= 1:106, EValue= 5.34832E-68. HHPRED= . CDD= .

/note=

/note=Gap: 14

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS join(7001. .7285;7285. .7437)

/gene="11"

/function="tail assembly chaperone"

/product="tail assembly chaperone"

/locus tag="SEA\_MRGREEN\_11"

/note=Genemark calls start at 7348

/note=

/note=SSC: Start = 7348, Stop = 7437. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 1.375 is not the highest start score. SCS: Start is not called by Glimmer and is called by Genemark. LO: 90 bp is not the longest possible ORF. GAP: 26 bp. ST: SS=NA. F: tail assembly chaperone. FS: PHDBLAST= PhageName= McShie, ProteinNumber= 11, Function= tail assembly chaperone, EValue= 2.0E-9. NCBIBLAST= PhageName= tail assembly chaperone [Microbacterium phage McShie], Coverage= 100.0, SubjectRange= 114:142, QueryRange= 114:29, EValue= 2.39041E-10. HHPRED= . CDD= .

/note=

/note=Gap: 27. All members of this pham (61) are currently drafts. For all members of the EE cluster, gene 11 has the same start codon as gene 10, except for this pham, who’s gene 11 does not overlap with gene 10 at all.

/note=

/note=All nondraft genes in this pham move the start of gene 11 to the start of gene 10. This would make the start of this gene 7001.

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 7553 - 9658

/gene="12"

/function="tape measure protein"

/product="tape measure protein"

/locus tag="SEA\_MRGREEN\_12"

/note=Original Glimmer call @bp 7553 has strength 13.34; Genemark calls start at 7553

/note=

/note=SSC: Start = 7553, Stop = 9658. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.755 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 2106 bp is the longest possible ORF. GAP: 115 bp. ST: SS=NA. F: tape measure protein. FS: PHDBLAST= PhageName= Quaker, ProteinNumber= 12, Function= tape measure protein, EValue= 0.0. NCBIBLAST= PhageName= tape measure protein [Microbacterium phage Quaker] >gb|AXC36003.1| tape measure protein [Microbacterium phage Quaker] >gb|QKY78778.1| tape measure protein [Microbacterium phage Livingwater], Coverage= 100.0, SubjectRange= 1:701, QueryRange= 1:701, EValue= 0.0. HHPRED= Accession= 6V8I\_CF, Description= Tape Measure Protein, gp57; phage tail, tail tip, tape measure protein, VIRAL PROTEIN; 3.7A {Staphylococcus virus 80alpha}, Probability= 99.9. Coverage= 32.097, SubjectRange= 859:1088, QueryRange= 859:644. CDD= .

/note=

/note=SIF- TMHMM 10 transmembrane regions identified via TMHMM and SOSUI

CDS 9655 - 10614

/gene="13"

/function="minor tail protein"

/product="minor tail protein"

/locus tag="SEA\_MRGREEN\_13"

/note=Original Glimmer call @bp 9655 has strength 9.0; Genemark calls start at 9655

/note=

/note=SSC: Start = 9655, Stop = 10614. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 2.005 is not the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 960 bp is the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: minor tail protein. FS: PHDBLAST= PhageName= Azizam, ProteinNumber= 13, Function= minor tail protein, EValue= 1.0E-178. NCBIBLAST= PhageName= minor tail protein [Microbacterium phage Azizam] >gb|QGJ97472.1| minor tail protein [Microbacterium phage Azizam] >gb|QNJ56055.1| minor tail protein [Microbacterium phage JRok], Coverage= 100.0, SubjectRange= 1:319, QueryRange= 1:319, EValue= 0.0. HHPRED= . CDD= .

/note=

/note=Gap: -3

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 10614 - 12674

/gene="14"

/function="minor tail protein"

/product="minor tail protein"

/locus tag="SEA\_MRGREEN\_14"

/note=Original Glimmer call @bp 10614 has strength 8.71; Genemark calls start at 10614

/note=

/note=SSC: Start = 10614, Stop = 12674. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.807 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 2061 bp is the longest possible ORF. GAP: -1 bp. ST: SS=NA. F: minor tail protein. FS: PHDBLAST= PhageName= Bri160, ProteinNumber= 14, Function= minor tail protein, EValue= 0.0. NCBIBLAST= PhageName= minor tail protein [Microbacterium phage Bri160] >gb|QLF83174.1| minor tail protein [Microbacterium phage Bri160], Coverage= 100.0, SubjectRange= 1:686, QueryRange= 1:686, EValue= 0.0. HHPRED= . CDD= .

/note=

/note=Gap:0

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 12676 - 13254

/gene="15"

/function="minor tail protein"

/product="minor tail protein"

/locus tag="SEA\_MRGREEN\_15"

/note=Original Glimmer call @bp 12676 has strength 10.9; Genemark calls start at 12676

/note=

/note=SSC: Start = 12676, Stop = 13254. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.49 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 579 bp is the longest possible ORF. GAP: 1 bp. ST: SS=NA. F: minor tail protein. FS: PHDBLAST= PhageName= Kaijohn, ProteinNumber= 15, Function= minor tail protein, EValue= 1.0E-100. NCBIBLAST= PhageName= minor tail protein [Microbacterium phage Kaijohn] >gb|QGH78535.1| minor tail protein [Microbacterium phage Kaijohn], Coverage= 100.0, SubjectRange= 1:192, QueryRange= 1:192, EValue= 1.45775E-126. HHPRED= . CDD= .

/note=

/note=Gap:2

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 13300 - 13566

/gene="16"

/function="hypothetical function"

/product="hypothetical function"

/locus tag="SEA\_MRGREEN\_16"

/note=Original Glimmer call @bp 13300 has strength 17.9; Genemark calls start at 13300

/note=

/note=SSC: Start = 13300, Stop = 13566. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 1.529 is not the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 267 bp is the longest possible ORF. GAP: 45 bp. ST: SS=NA. F: hypothetical function. FS: PHDBLAST= PhageName= KayPaulus, ProteinNumber= 16, Function= function unknown, EValue= 2.0E-35. NCBIBLAST= PhageName= hypothetical protein JTF52\_gp16 [Microbacterium phage Azizam] >gb|QGJ97475.1| hypothetical protein SEA\_AZIZAM\_16 [Microbacterium phage Azizam], Coverage= 100.0, SubjectRange= 1:87, QueryRange= 1:88, EValue= 7.53633E-27. HHPRED= . CDD= .

/note=

/note=Similarly to EE phage Kaijohn, there are three minor tail proteins upstream with an endolysin downstream.

/note=

/note=Gap:46 checked all sources

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 13591 - 14283

/gene="17"

/function="endolysin"

/product="endolysin"

/locus tag="SEA\_MRGREEN\_17"

/note=Original Glimmer call @bp 13591 has strength 11.9; Genemark calls start at 13591

/note=

/note=SSC: Start = 13591, Stop = 14283. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 3.342 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 693 bp is the longest possible ORF. GAP: 24 bp. ST: SS=NA. F: endolysin. FS: PHDBLAST= PhageName= Quaker, ProteinNumber= 17, Function= endolysin, EValue= 1.0E-136. NCBIBLAST= PhageName= endolysin [Microbacterium phage Quaker] >gb|AXC36008.1| endolysin [Microbacterium phage Quaker] >gb|QKY78783.1| lysin A [Microbacterium phage Livingwater], Coverage= 100.0, SubjectRange= 1:230, QueryRange= 1:230, EValue= 3.34307E-163. HHPRED= . CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 14280 - 14516

/gene="18"

/function="membrane protein"

/product="membrane protein"

/locus tag="SEA\_MRGREEN\_18"

/note=Original Glimmer call @bp 14250 has strength 8.13; Genemark calls start at 14283

/note=

/note=SSC: Start = 14280, Stop = 14516. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 2.15 is the highest start score. SCS: Start is not called by Glimmer and is not called by Genemark. LO: 237 bp is not the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: membrane protein. FS: PHDBLAST= . NCBIBLAST= PhageName= hypothetical protein JTF58\_gp18 [Microbacterium phage Quaker] >gb|AXC35376.2| hypothetical protein SEA\_KAYPAULUS\_18 [Microbacterium phage KayPaulus] >gb|QKY78784.1| hypothetical protein SEA\_LIVINGWATER\_18 [Microbacterium phage Livingwater] >gb|QRI45138.1| membrane protein [Microbacterium phage Wolfpack] >gb|QWS68300.1| membrane protein [Microbacterium phage Concrete] >gb|AXC36009.2| hypothetical protein SEA\_QUAKER\_18 [Microbacterium phage Quaker], Coverage= 100.0, SubjectRange= 1:78, QueryRange= 1:78, EValue= 3.47187E-50. HHPRED= . CDD= .

/note=

/note=One transmembrane protein was found in the gene and I listed it as a membrane protein because the transmembrane protein was found on both SOSUI and TMHMM.

/note=

/note=Start 14280 allows for a 4bp overlap with the previous gene.

/note=

/note=SIF-TMHMM: 1 transmembrane domain identified using TMHMM and SOSUI.

CDS 14513 - 14737

/gene="19"

/function="membrane protein"

/product=" membrane protein "

/locus tag="SEA\_MRGREEN\_19"

/note=Original Glimmer call @bp 14513 has strength 10.6; Genemark calls start at 14513

/note=

/note=SSC: Start = 14513, Stop = 14737. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.843 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 225 bp is the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: hypothetical function. FS: PHDBLAST= PhageName= Azizam, ProteinNumber= 19, Function= function unknown, EValue= 4.0E-36. NCBIBLAST= PhageName= hypothetical protein JTF52\_gp19 [Microbacterium phage Azizam] >ref|YP\_009996654.1| hypothetical protein JTF53\_gp19 [Microbacterium phage Bri160] >ref|YP\_009996780.1| hypothetical protein JTF58\_gp19 [Microbacterium phage Quaker] >gb|AXC34902.1| hypothetical protein SEA\_VITULAELIGANS\_19 [Microbacterium phage VitulaEligans] >gb|AXC35377.1| hypothetical protein SEA\_KAYPAULUS\_19 [Microbacterium phage KayPaulus] >gb|QDF18046.1| hypothetical protein BELTHELAS\_19 [Microbacterium phage Belthelas] >gb|QJD53917.1| hypothetical protein SEA\_MCSHIE\_19 [Microbacterium phage McShie] >gb|QKY78785.1| hypothetical protein SEA\_LIVINGWATER\_19 [Microbacterium phage Livingwater] >gb|QOI67295.1| membrane protein [Microbacterium phage Sippinontea] >gb|QPL14311.1| hypothetical protein SEA\_JERKY\_19 [Microbacterium phage Jerky] , Coverage= 100.0, SubjectRange= 1:74, QueryRange= 1:74, EValue= 2.27443E-43. HHPRED= . CDD= .

/note=

/note=Gap: -3. upstream and Lsr2-like DNA binding protein, helix turn helix DNA binding protein downstream seen in KaiJohn, KayPaulus, and Wolfpack. checked all sources

/note=

/note=We are not sure if this gene should be identified as NA or NI.

/note=

/note=SIF-TMHMM: 1 transmembrane domain identified using TMHMM and SOSUI.

CDS complement (14806 - 15018)

/gene="20"

/function="Lsr2-like DNA bridging protein"

/product="Lsr2-like DNA bridging protein"

/locus tag="SEA\_MRGREEN\_20"

/note=Original Glimmer call @bp 15018 has strength 11.6; Genemark calls start at 15018

/note=

/note=SSC: Start = 15018, Stop = 14806. (Reverse). CP: Does contain all GeneMarkHost capacity. SD: ZScore 2.655 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 213 bp is the longest possible ORF. GAP: 2 bp. ST: SS=NA. F: Lsr2-like DNA bridging protein. FS: PHDBLAST= . NCBIBLAST= PhageName= Lsr2-like DNA bridging protein [Microbacterium phage Quaker] >gb|AXC35378.1| Lsr2-like DNA bridging protein [Microbacterium phage KayPaulus] >gb|QKY78786.1| Lsr2-like DNA bridging protein [Microbacterium phage Livingwater] >gb|QOI67296.1| Lsr2-like DNA bridging protein [Microbacterium phage Sippinontea] >gb|QRI45140.1| Lsr2-like DNA bridging protein [Microbacterium phage Wolfpack] >gb|AXC36011.1| Lsr2-like DNA bridging protein [Microbacterium phage Quaker], Coverage= 100.0, SubjectRange= 1:70, QueryRange= 1:70, EValue= 7.75731E-42. HHPRED= Accession= 4E1P\_B, Description= Protein lsr2; anti-parallel beta sheet, dimer, DNA BINDING PROTEIN; 1.728A {Mycobacterium tuberculosis}, Probability= 99.9. Coverage= 84.2857, SubjectRange= 1:59, QueryRange= 1:60. CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS complement (15021 - 15527)

/gene="21"

/function="helix-turn-helix DNA binding domain"

/product="helix-turn-helix DNA binding domain"

/locus tag="SEA\_MRGREEN\_21"

/note=Original Glimmer call @bp 15464 has strength 12.08; Genemark calls start at 15527

/note=

/note=SSC: Start = 15527, Stop = 15021. (Reverse). CP: Does contain all GeneMarkHost capacity. SD: ZScore 0.494 is not the highest start score. SCS: Start is not called by Glimmer and is called by Genemark. LO: 507 bp is not the longest possible ORF. GAP: 81 bp. ST: SS=NA. F: helix-turn-helix DNA binding domain. FS: PHDBLAST= PhageName= Quaker, ProteinNumber= 21, Function= helix-turn-helix DNA binding protein, EValue= 4.0E-91. NCBIBLAST= PhageName= helix-turn-helix DNA binding protein [Microbacterium phage Quaker] >gb|AXC36016.1| helix-turn-helix DNA binding protein [Microbacterium phage Quaker], Coverage= 100.0, SubjectRange= 5:172, QueryRange= 5:168, EValue= 5.85324E-118. HHPRED= Accession= 4LHF\_A, Description= Regulatory protein cox; helix-turn-helix, DNA binding, VIRAL PROTEIN; 2.401A {Enterobacteria phage P2}, Probability= 98.7. Coverage= 32.7381, SubjectRange= 6:69, QueryRange= 6:87. CDD= .

/note=

/note=I moved the start because there was evidence of a helix-turn-helix upstream of my gene. The start called includes this coding potential and provides the helix-turn-helix function. A majority of phage in this cluster also moved their start.

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS complement (15609 - 15836)

/gene="22"

/function="helix-turn-helix DNA binding domain"

/product="helix-turn-helix DNA binding domain"

/locus tag="SEA\_MRGREEN\_22"

/note=Original Glimmer call @bp 15836 has strength 6.41; Genemark calls start at 15836

/note=

/note=SSC: Start = 15836, Stop = 15609. (Reverse). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 0.93 is the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 228 bp is the longest possible ORF. GAP: 509 bp. ST: SS=NA. F: helix-turn-helix DNA binding domain. FS: PHDBLAST= PhageName= Belthelas, ProteinNumber= 22, Function= helix-turn-helix DNA binding protein, EValue= 4.0E-36. NCBIBLAST= PhageName= helix-turn-helix DNA binding protein [Microbacterium phage Belthelas], Coverage= 100.0, SubjectRange= 1:75, QueryRange= 1:75, EValue= 4.1949E-45. HHPRED= . CDD= .

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 16346 - 16735

/gene="23"

/function="hypothetical function"

/product="hypothetical function"

/locus tag="SEA\_MRGREEN\_23"

/note=Original Glimmer call @bp 16346 has strength 6.2; Genemark calls start at 16553

/note=

/note=SSC: Start = 16346, Stop = 16735. (Forward). CP: Does not contain all GeneMarkHost capacity. SD: ZScore 1.961 is the highest start score. SCS: Start is called by Glimmer and is not called by Genemark. LO: 390 bp is the longest possible ORF. GAP: 509 bp. ST: SS=NA. F: hypothetical function. FS: PHDBLAST= PhageName= Quaker, ProteinNumber= 23, Function= function unknown, EValue= 3.0E-66. NCBIBLAST= PhageName= hypothetical protein JTF58\_gp23 [Microbacterium phage Quaker] >gb|AXC36013.1| hypothetical protein SEA\_QUAKER\_23 [Microbacterium phage Quaker] >gb|QKY78789.1| hypothetical protein SEA\_LIVINGWATER\_23 [Microbacterium phage Livingwater], Coverage= 100.0, SubjectRange= 1:129, QueryRange= 1:129, EValue= 1.6627E-83. HHPRED= . CDD= .

/note=

/note=Gap: 510. upstream (gene 22): helix-turn-helix DNA binding protein, downstream (gene 24): MerR-like helix-turn-helix DNA binding domain protein, Quaker and McShie. checked all sources

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 16833 - 17051

/gene="24"

/function="helix-turn-helix DNA binding domain, MerR-like"

/product="helix-turn-helix DNA binding domain, MerR-like"

/locus tag="SEA\_MRGREEN\_24"

/note=Original Glimmer call @bp 16833 has strength 9.14; Genemark calls start at 16842

/note=

/note=SSC: Start = 16833, Stop = 17051. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 1.748 is not the highest start score. SCS: Start is called by Glimmer and is not called by Genemark. LO: 219 bp is not the longest possible ORF. GAP: 97 bp. ST: SS=NA. F: helix-turn-helix DNA binding domain, MerR-like. FS: PHDBLAST= . NCBIBLAST= PhageName= MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage KayPaulus] >gb|AXH66510.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Scamander] >gb|QBZ72527.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Rhysand] >gb|QDF15909.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage LaviMo] >gb|QDF16584.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Hulk] >gb|QFP94594.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Vanisius] >gb|QFP97688.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage TeddyBoy] >gb|QJD53947.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Otwor] >gb|QJD54087.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Danno] >gb|QOI67300.1| MerR-like helix-turn-helix DNA binding domain protein [Microbacterium phage Sippinontea] , Coverage= 100.0, SubjectRange= 1:72, QueryRange= 1:72, EValue= 4.94861E-44. HHPRED= Accession= 4LHF\_A, Description= Regulatory protein cox; helix-turn-helix, DNA binding, VIRAL PROTEIN; 2.401A {Enterobacteria phage P2}, Probability= 99.1. Coverage= 75.0, SubjectRange= 6:68, QueryRange= 6:68. CDD= .

/note=

/note=Gap: 98

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

CDS 17048 - 17347

/gene="25"

/function="HNH endonuclease"

/product="HNH endonuclease"

/locus tag="SEA\_MRGREEN\_25"

/note=Original Glimmer call @bp 17048 has strength 5.67; Genemark calls start at 17048

/note=

/note=SSC: Start = 17048, Stop = 17347. (Forward). CP: Does contain all GeneMarkHost capacity. SD: ZScore 1.411 is not the highest start score. SCS: Start is called by Glimmer and is called by Genemark. LO: 300 bp is not the longest possible ORF. GAP: -4 bp. ST: SS=NA. F: HNH endonuclease. FS: PHDBLAST= . NCBIBLAST= PhageName= HNH endonuclease [Microbacterium phage Bri160] >ref|YP\_009996686.1| HNH endonuclease [Microbacterium phage Kaijohn] >ref|YP\_009996786.1| HNH endonuclease [Microbacterium phage Quaker] >gb|AXC35383.1| HNH endonuclease [Microbacterium phage KayPaulus] >gb|QJD53923.1| HNH endonuclease [Microbacterium phage McShie] >gb|QKY78791.1| HNH endonuclease [Microbacterium phage Livingwater] >gb|QOI67301.1| HNH endonuclease [Microbacterium phage Sippinontea] >gb|QRI45145.1| HNH endonuclease [Microbacterium phage Wolfpack], Coverage= 100.0, SubjectRange= 1:99, QueryRange= 1:99, EValue= 3.85423E-68. HHPRED= . CDD= .

/note=

/note=There is a 4bp overlap with the previous gene at the start 17048.

/note=

/note=SIF-TMHMM: 0 transmembrane domains identified using TMHMM and SOSUI.

BASE COUNT 2936 a 6082 c 5870 g 2533 t

ORIGIN