To the annotation team:

Two Providence College students worked this summer on the annotation of Milly, a K2 cluster phage discovered by a student at Georgia State College and University. Providence College is interested in the K2 cluster phages as we have discovered a phage, ZoeJ, on our campus that is a member of this subcluster. We requested that we be allowed to collaborate with the Georgia team on the annotation of Milly, and our request was granted. Earlier this summer, Debbie Jacobs-Sera sent us the DNA Master file of the Georgia team’s annotations.

The Providence College team, consisting of myself, Ellen Wheeler and Nicole Cullen, annotated Milly without viewing the Georgia State College and University results beforehand. Below, we have noted the discrepancies between the Providence College team and the Georgia team. We hope that our results are clearly presented and will allow the annotation team to make the final call regarding gene assignments.

Also, I have submitted a csv file with all three of our names on it, but I feel that the Georgia team should get top billing for the GenBank submission. They were the ones who discovered the phage, purified it, and isolated the DNA.

Thank you very much.

Sincerely,

Kathleen Cornely, Providence College

DISCREPANCIES IN NOTES FOR MILLY BETWEEN PROVIDENCE COLLEGE AND GEORGIA STATE COLLEGE AND UNIVERSITY (discrepancies are noted in red)

Some overall notes:

* The Georgia team does not check functionality using HHPred, Hatfull map or Phamerator (only BLAST results were used).
* The Providence College team’s BLAST results include phage ZoeJ; Team Georgia’s does not.
* The Providence College team lists the SD score in the notes while Team Georgia notes only whether or not the SD score was the highest score.
* In the DNA Master file submitted by Providence College, all genes have been re-BLASTed and all ORFs are valid.

**Gene 1**

SSC: Start: 88 Stop: 222 (FWD) CP: ORF does not include all coding potential shown on GeneMark-smeg output. SD: 299, not best score. SCS: agrees with Glimmer call, but disagrees with GeneMark call. Original Glimmer call @bp 88 has strength 16.95; GeneMark calls start at 1. Gap: N/A. BLAST: gp1 of ZoeJ and TM4; q1:s1 and 100% aligned for both. LO: 135 bp; longest possible ORF. F: putative zinc peptidase. FS: Blast.

**Gene 2**

SSC: Start 224 Stop: 436 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 210, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 224 has strength 9.18. Gap: 1 bp. BLAST: gp2 of ZoeJ and TM4; q1:s1 and 100% alignment for both. LO: 213 bp; not longest ORF. F: hypothetical protein.

**Gene 3:**

SSC: Start: 426 Stop: 674 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 567, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 426 has strength 9.07. Overlap: 11 bp overlap with previous gene. BLAST: gp3 of ZoeJ and TM4; q1:s1, 100% identical and 100% similarity for both, 100% alignment for ZoeJ. LO: 249 bp; not longest possible ORF. F: ATP binding protein/DNA binding protein. FS: HHpred.

**Gene 4**

SCC: Start: 652 Stop: 2076 (FWD). CP: ORF includes all of the coding potential shown on GeneMark-smeg output. SD: 399, not highest score. SCS: Disagrees with GeneMark predictions, agrees with Glimmer. Original Glimmer call @bp 652 has strength 11.34; GeneMark calls start at 754. Gap: 23 bp overlap with previous gene. BLAST: terminase of ZoeJ; q1:s1, terminase gp4 of TM4, q4:s5. LO: 1323, not longest possible ORF. F: terminase. FS: BLAST.

**Gene 5**

SSC: Start: 2088 Stop: 3593 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 378, not highest score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 2088 has strength 16.78. Gap: 11 bp gap with previous gene. Blast: gp5 of TM4 q17:s17; gp5 of ZoeJ q22:s22. LO: 1506 bp; longest possible ORF. F: phage portal protein. FS: Phamerator and BLAST.

**Gene 6**

SSC: Start: 3574 Stop: 5094 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 672, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 3574 has strength 12.58. Overlap: 20 bp overlap with previous gene. Blast: gp73 of 33D, q1:s1; gp6 of TM4; q1:s1, 100% aligned, 95.48% similar, capsid maturation protease of ZoeJ, q1:s1. LO: 1521 bp; longest possible ORF. F: capsid maturation protease. FS: BLAST

**Gene 7**

SSC: Start: 5091 Stop: 5276 (FWD). CP: ORF includes almost all coding potential shown on GeneMark-smeg output (a very small amount at the beginning might be missed). SD: 546, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 5091 has strength 9.65. Overlap: 4 bp overlap with previous gene. Blast: gp7 of ZoeJ and TM4; q1:s1 for both. LO: 186 bp; longest possible ORF. F: hypothetical protein.

**Gene 8**

SSC: Start: 5349 Stop: 5891 (FWD). CP: ORF includes all coding potential shown in GeneMark-smeg output SD: 525, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 5349 has strength 10.23. Gap: 72 bp gap with previous gene. Blast: Scaffolding protein of ZoeJ; q34:s32, 97.74% similarity, 96.24% identity; gp8 of TM4; q68:s73. LO: 543 bp; longest possible ORF. F: Scaffolding protein. FS: BLAST.

**Gene 9**

SSC: Start: 5947 Stop: 6864 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 819, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 5947 has strength 17.69. Gap: 55 bp gap with the previous gene. Blast: major head subunit of 33D; q1:s1; gp9 of TM4; q1:s1; major capsid protein of ZoeJ and Validus; q1:s1 for both; gp11 of Larva; q1:s1. LO: 918 bp, longest possible ORF. F: major capsid protein. FS: Phamerator, BLAST and HHpred.

**Gene 10**

SSC: Start: 6976 Stop: 7362 (FWD). CP: ORF includes all potential shown on GeneMark-smeg output. SD: 494, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 6976 has strength 6.96. The earlier call at 6907 does not have any good BLAST hits and the coding potential missed by the earlier start is minimal, if any. Gap: 11 bp gap with previous gene. Blast: gp10 of both TM4 and ZoeJ; q1:s1 for both, 100% similarity with TM4. LO: 387 bp; not longest possible ORF. F: hypothetical protein.

**Gene 11**

SSC: Start: 7362 Stop: 7712 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 525, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 7362 has strength 8.17. Overlap: 1 bp overlap with previous gene. Blast: gp11 of both TM4 and ZoeJ; q1:s1 for both, 99.14% similarity and 98.28% identity with TM4. LO: 351 bp; longest possible gene. F: hypothetical protein.

**Gene 12**

SSC: Start: 7693 Stop: 7965 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 399, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 7693 has strength 13.17. Overlap: 20 bp overlap with previous gene. Blast: gp66 of 33D, q1:s1; gp12 of both TM4 and ZoeJ; q8:s1, 93.98% similar and 90.36% identical with TM4. LO: 273 bp; longest possible ORF. F: hypothetical protein.

**Gene 13**

SSC: Start: 7962 Stop: 8381 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 651, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 7962 has strength 10.53. Overlap: 4 bp overlap with previous gene. Blast: gp13 of both TM4 and ZoeJ; q1:s1, 99.28% similarity and 98.56% identical with TM4. LO: 420 bp; not the longest possible ORF. F: hypothetical protein.

**Gene 14**

SSC: Start: 8532 Stop: 9146 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 504, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 8532 has strength 18.25. Gap: 150 bp gap with previous gene. Blast: gp14 of TM4; q1:s1, major tail subunit of ZoeJ and 33D; q1:s1 for both. LO: 615 bp; longest possible ORF. F: major tail subunit. FS: BLAST.

**Gene 15**

SSC: Start: 9253 Stop: 9702 (FWD). CP: ORF includes nearly all coding potential shown on GeneMark-smeg output (a small amount is missed at the beginning, but earlier start has no q1:s1 BLAST matches). SD: 481, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 9253 has strength 13.09. Gap: 106 bp with previous gene. Blast: tail assembly chaperone of ZoeJ; q1:s1; gp15 of TM4, q26:s25. LO: 450 bp, not longest ORF. F: tail assembly chaperone. FS: BLAST.

**Gene 16**

SSC: Start: 9253 Stop: 10115 (FWD). CP: includes all coding potential show on GeneMark-smeg output. SD: N/A SCS: does not agree with either Glimmer or GeneMark predictions. Original Glimmer call @bp 9732 has strength 9.58. Gap: no gap with preceding gene (frameshift). Blast: tail assembly chaperone of ZoeJ; q1:s1. LO: 864, longest possible ORF. F: tail assembly chaperone. FS: BLAST.

Note that the Georgia team did not call the frameshift.

**Gene 17**

SSC: Start: 10116 Stop: 13784 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 567, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 10116 has strength 13.65. Gap: 0 bp with previous gene. Blast: tapemeasure protein of both ZoeJ and TM4; q1:s1. LO: 3669 bp, longest possible ORF. F: tapemeasure. FS: BLAST and Hatfull map, Phamerator.

**Gene 18**

SSC: Start: 13889 Stop: 15016 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 441, not best score. SCS: Agrees with GeneMark but not Glimmer. Original Glimmer call @bp 13898 has strength 12.20; GeneMark calls start at 13889. Gap: 104 bp gap with previous gene. Blast: gp18 of TM4; q1:s1, minor tail protein of ZoeJ, q1:s1. LO: 1128 bp, longest possible ORF. F: distal tail protein. FS: BLAST and HHpred.

**Gene 19**

SSC: Start: 15013 Stop: 16782 (FWD). CP: ORF does not include all coding potential shown on GeneMark-Smeg output. SD: 420, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 15013 has strength 11.14. Overlap: 4 bp overlap with previous gene. Blast: minor tail protein of ZoeJ; q1:s1, gp19 of TM4; q1:s1, gp24 of Fionnbharth; q1:s1, gp22 of Pixie; q1:s1. LO: 1770 bp; not longest possible ORF. F: minor tail subunit. FS: BLAST and HHPred.

**Gene 20**

SCS: Start: 16782 Stop: 17246 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 672, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 16782 has strength 13.35. Gap: 1 bp overlap. Blast: minor tail protein of ZoeJ; q1:s1, gp20 of TM4; q1:s1. LO: 465, not longest possible ORF. F: minor tail protein. FS: BLAST.

**Gene 21**

SCS: Start: 17339 Stop: 18418 (FWD). CP: ORF includes all possible coding potential shown on GeneMark-Smeg. SD: 651, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 17339 has strength 15.84. Gap: 92 bp with previous gene. Blast: minor tail protein of both ZoeJ and Validus; q1:s1, gp21 of TM4; q1:s1, gp26 of Fionnbharth; q7:s7, gp24 of Pixie; q1:s1. LO: 1080 bp; longest possible ORF. F: minor tail protein FS: BLAST.

**Gene 22**

SSC: Start: 18423 Stop: 18677 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 510, best score. SCS: Agrees with Glimmer but not GeneMark predictions. Original Glimmer call @bp 18423 has strength 10.81; GeneMark calls start at 18411. Gap: 4 bp gap with previous gene. Blast: minor tail protein of ZoeJ; q1:s1, 97% identical and similar, gp22 of TM4; q1:s1, gp27 of Fionnbharth; q1:s1, gp26 of BarrelRoll; q1:s1, minor tail protein of Emerson; q1:s1. LO: 255, not longest possible ORF. F: minor tail protein. FS: BLAST.

**Gene23**

SSC: Start: 18661 Stop: 21000 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 567, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 18661 has strength 14.25. Overlap: 17 bp overlap with previous gene. Blast: minor tail protein of ZoeJ; q1:s1, gp26 of Pixie; q1:s1, gp27 of MacnCheese; q1:s1, gp23 of TM4; q1:s1. LO: 2340 bp, longest possible ORF. F: glycosidase, hydrolase. FS: HHPred.

**Gene24**

SSC: Start: 21001 Stop: 21282 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 651, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 21001 has strength 17.53. Gap: 0 bp gap with previous gene. Blast: minor tail protein of ZoeJ; q1:s1, 96.74% similarity, gp27 of Pixie; q1:s1, gp28 of MacnCheese; q1:s1, gp24 of TM4; q1:s1. LO: 282 bp; longest possible ORF. F: minor tail protein. FS: BLAST.

**Gene25**

SSC: Start: 21282 Stop: 22277 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 210, not best score. SCS: Agrees with Glimmer, disagrees with GeneMark. Original Glimmer call @bp 21282 has strength 11.71; GeneMark calls start at 21363. Later start excludes coding potential. Gap: 1 bp overlap. Blast: gp25 of TM4; q1:s1, minor tail protein of ZoeJ; q1:s1. LO: 996 bp, not longest possible ORF. F: minor tail protein. FS: BLAST.

**Gene26**

SSC: Start: 22292 Stop: 22531 (FWD). CP: ORF leaves out a miniscule part of the coding potential shown on GeneMark-smeg output. SD: 714, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 22292 has strength 11.02. Gap: 14 bp gap with previous gene. Blast: gp26 of TM4; q4:s5, 90.54% similarity, gp26 of ZoeJ; q4:s6, gp9 of AnnaL29; q1:s1, gp6 of Goose; q1:s1. LO: 240 bp; not longest possible ORF. F: hypothetical protein.

**Gene27**

SSC: Start: 22685 Stop: 23053 (FWD). CP: ORF includes all but a miniscule amount of coding potential shown on GeneMark-smeg output. SD: 357, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 22685 has strength 16.86. Gap: 153 bp gap with previous gene. Blast: gp27 of ZoeJ; q1:s1, 93.39% similarity; gp27 of TM4; q1:s1; gp29 of CrimD and BarrelRoll, q1:s1 for both. LO: 369 bp; not longest possible ORF. F: hypothetical protein.

**Gene28**

SSC: Start: 23083 Stop: 24774 (FDW).CP: ORF includes all but a tiny fraction of beginning coding potential shown on GeneMark-smeg output. SD: 504, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 23083 has strength 10.18. Gap: 29 bp gap with previous gene. Blast: Lysin A of ZoeJ; q1:s1; Lysin A of Validus, q2:s1. LO: 1692 bp; longest possible ORF. Function: Lysin A. FS: BLAST, HHpred, Phamerator.

**Gene29**

SSC: Start: 24780 Stop: 25982 (FWD). CP: ORF neglects a small amount of coding potential shown on GeneMark-smeg output. SD: 378, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 24780 has strength 13.91. Gap: 5 bp gap with previous gene. Blast: Lysin B of ZoeJ; q1:s1, 92.29% similarity, gp30 of TM4; q1:s1, Lysin B of 33D; q1:s1. LO: 1203 bp; second longest possible ORF. F: Lysin B. FS: BLAST, Phamerator and HHPred.

**Gene30**

SSC: Start: 26001 Stop: 26387 (FDW). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 416, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 26001 has strength 15.56. Gap: 18 bp gap with previous gene. Blast: gp31 of TM4; q1:s1, 96.09% similarity and 95.31% identity, holin of ZoeJ; q1:s1, 91.91% similarity. LO: 387 bp; longest possible ORF. F: Holin. FS: BLAST.

**Gene 31**

SSC: Start: 26470 Stop: 26889 (FWD). CP: ORF includes all but a small amount of coding potential shown on GeneMark-smeg output. SD: 300, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 26470 has strength 9.95. Gap: 82 bp gap with previous gene. Blast: gp31 of ZoeJ; q1:s1, 100% similarity and 98.56% identical, gp33 of TM4; q1:s1, 98.54% similarity, gp45 of 33D; q1:s1. LO: 420 bp, longest possible ORF. F: hypothetical protein.

**Gene 32**

SSC: Start: 26879 Stop: 27370 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 588, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 26879 has strength 12.30. Overlap: 11 bp overlap with previous gene. Blast: gp32 of ZoeJ; q1:s1, 98.16% similarity and 92.01% identical, gp34 of TM4; q1:s1, 98.16% similarity and 90.80% identical, gp29 of Arturo; q1:s1, gp30 of Peaches; q1:s1. LO: 492 bp; longest possible ORF. F: hypothetical protein.

**Gene 33**

SSC: Start: 27367 Stop: 27615 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 315, not best score. SCS: Agrees with Glimmer but not GeneMark predictions. Original Glimmer call @bp 27367 has strength 5.05; GeneMark calls start at 27421. Overlap: 4 bp overlap with previous gene. Blast: gp35 of TM4; q1:s1, 100% similarity and 98.78% identical, gp33 of ZoeJ; q1:s1, 96.34% similar and 95.12% identical. LO: 249 bp; not longest possible ORF. F: hypothetical protein.

**Gene 34**

SSC: Start: 27602 Stop: 28786 (FWD). CP: ORF leaves out a small portion of the possible coding potential shown on GeneMark-smeg output. SD: 315, not best score. SCS: Agrees with Glimmer predictions, because includes more of the coding potential possible and has a higher SD, does not agree with GeneMark. Original Glimmer call @bp 27602 has strength 13.68; GeneMark calls start at 27635. Overlap: 14 bp overlap with previous gene. Blast: gp36 of TM4; q1:s1, 99.75% similarity and 99.24% identical, gp34 of ZoeJ; q1:s1, 99.75% similarity and 98.73% identical, gp36 of Anaya; q1:s1, gp35 of CrimD; q1:s1, gp35 of Adephagia; q1:s1. LO: 1185 bp; longest possible ORF. F: DNA repair protein FS: HHpred.

**Gene 35**

SSC: Start: 28944 Stop: 29279 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 357, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 28944 has strength 17.02. Gap: 157 bp gap with previous gene. Blast: gp58 of DS6A; q1:s11. LO: 336 bp, longest possible ORF. F: transcription factor. FS: HHpred.

**Gene 36**

SSC: Start: 29410 Stop: 29814 (FWD). CP: ORF includes all the coding potential shown on GeneMark-Smeg output. SD: 546, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 29410 has strength 15.41. Gap: 130 bp gap with previous gene. Blast: putative transcription factor of ZoeJ; q1:s1, 98.51% similarity; gp38 of TM4; q1:s1, 97.01% similarity and 96.27% identical. LO: 405 bp; longest possible ORF. F: putative transcription factor, DNA-binding protein. FS: DNA Master, HHpred.

**Gene 37**

SSC: Start: 29858 Stop: 30196 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SCS: Agrees with both Glimmer and GeneMark. Original Glimmer call @bp 29858 has strength 11.51 Gap: 43 bp gap with previous gene. Blast: gp37 of ZoeJ; q27:s23, 93.02% similarity. LO: 339 bp; second longest possible ORF. F: hypothetical protein.

**Gene 38**

SSC: Start: 30315 Stop: 31019 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 462, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 30315 has strength 10.83. Gap: 118 bp gap with previous gene. Blast: gp38 of ZoeJ; q14:14s, 97.74% similarity and 96.38% identical, gp40 of TM4; q14:s14, 96.38% similarity. LO: 705 bp; longest possible ORF. F: hypothetical protein.

**Gene 39**

SSC: Start: 31470 Stop: 31087 (REV). CP: ORF includes all but a small part of the coding potential shown on GeneMark-Smeg output. SD: 294, second best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 31470 has strength 6.74.Gap: 67 bp gap with previous gene. Blast: gp93 of TM4; q1:s1, gp39 of ZoeJ; q1:s1, 100% aligned for both. LO: 384 bp; longest possible ORF. F: hypothetical protein.

**Gene 40**

SSC: Start: 32030 Stop: 31467 (REV). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 32030 has strength 13.54. Overlap: 4 bp overlap with previous gene. Blast: queuine tRNA-ribosyltransferase of ZoeJ; q1:s1, 94.89% similarity and 90.34% identical. LO: 564 bp; longest possible ORF. F: queuine tRNA ribosyltransferase. FS: BLAST.

**Gene 41**

SSC: Start: 32168 Stop: 32866 (FWD). CP: ORF includes most of coding potential shown on GeneMark- smeg output. SD: 462, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 32168 has strength 14.03. Gap: 137 bp gap with previous gene. Blast: gp41 of ZoeJ; q1:s1, 95.26% similarity and 94.83% identical. LO: 699 bp; not longest possible ORF. F: replication protein FS: HHpred.

**Gene 42**

SSC: Start: 32988 Stop: 33833 (FWD). CP: ORF does not include all coding potential shown of GeneMark-Smeg output, small part of 5' end cut off. SD: 462, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 32988 has strength 11.29. Gap: 121 bp gap with previous gene. BLAST: gp42 of ZoeJ; q1:s1, 98.89% identical, 99.26% similar; gp42 of MacnCheese, 11:s1, gp86 of Larva, q1:s1. LO: 846 bp; not longest possible ORF. F: Queuine tRNA-ribosyltransferase protein. FS: HHpred.

**Gene 43**

SSC: Start: 34290 Stop: 35390 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 399, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 34290 has strength 16.50. Gap: 456 bp gap with previous gene. Blast: integrase of ZoeJ; q1:s1, 100% aligned, 100% similarity and 100% identity; integrase of Validius; q1:s1; gp43 of Anaya; q1:s1; gp41 of Angelica; q1:s1; gp41 of CrimD; q1:s1. LO: 1101 bp; longest possible ORF. F: integrase FS: DNA Master, HHpred.

**Gene 44**

SSC: Start: 35416 Stop: 35658 (FWD). CP: ORF does not include all coding potential shown on GeneMark-smeg output, cuts off tiny part of 5' end. SD: 357, not best score. SCS: Does not agree with both Glimmer and GeneMark predictions. Original Glimmer call @bp 35401 has strength 9.99. Gap: 25 bp gap with previous gene. BLAST: hypothetical protein of *Mycobacterium parascrofulaceum*; q1:s1; hypothetical protein of *Mycobacterium columbiense*; q1:s1 LO: 243, not the longest possible ORF. F: Hypothetical protein.

**Gene 45**

SSC: Start: 36565 Stop: 35717 (REV). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 525, best score. SCS: Agrees with Glimmer, but not GeneMark (later start excludes coding potential). Original Glimmer call @bp 36565 has strength 11.93; GeneMark calls start at 36451. Gap: 58 bp gap with previous gene. LO: 849 bp, longest ORF. Blast: gp 44 of ZoeJ; q11:s12. Function: hypothetical protein.

**Gene 46**

SSC: Start: 37024 Stop: 36653 (REV). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 378, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 37024 has strength 7.63. Gap: 87bp gap with previous gene. BLAST: Immunity repressor of ZoeJ; q1:s1, 96.91% identical, 97.94% similar; gp43 of BarrelRoll; q1:s1, 84.62% similar, gp 43 of CrimD and Angelica and gp45 of Anaya, q1:s1 for all three. LO: 372, not longest possible ORF. F: Transcriptional repressor protein. FS: BLAST and HHPred.

**Gene 47**

SSC: Start: 37198 Stop: 37515 (FWD). CP: ORF includes all coding potential identified by GeneMark-smeg output. SD: 441, best score. SCS: Does not agree with either GeneMark or Glimmer. Original GeneMark call @bp 37207, not called by Glimmer. Gap: 173 bp gap with previous gene. BLAST: gp42 of TM4; q3:s13, Cro-like protein of ZoeJ, q1:s1, gp44 of Angelica, q4:s1; LO: 318, not longest possible ORF; F: Cro-like protein, transcription regulator protein. FS: BLAST and HHpred.

**Gene 48**

SSC: Start 37512 Stop: 37769 (FWD). CP: ORF does not include all coding potential shown on GeneMark-Smeg output. SD: 294, not best possible score. SCS: Agrees with GeneMark predictions; does not agree with Glimmer prediction. Original Glimmer call @bp 37641 has strength 4.91; GeneMark calls start at 37512. Gap: 2 bp gap with previous gene. BLAST: Xis of ZoeJ; q1:s1, 97.65% identical, 98.82% similar; gp43 of TM4, q1:s1, 96.47% identical, 96.47% similar, gp45 of Angelica and BarrelRoll, q1:s1 for both, HTH DNA binding protein of Validus, q2:s5. LO: 258, longest possible ORF. F: Regulatory protein. FS: BLAST.

**Gene 49**

SSC: Start: 37774 Stop: 37995 (FWD). CP: ORF does not include all coding potential shown on GeneMark-smeg output. SD: 525, best possible score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 37774 has strength 12.77. Gap: 4 bp gap with previous gene. BLAST: gp44 of TM4, q1:s1, 82.19% identical, 86.30% similar; gp48 of ZoeJ, q1:s1, 86.30% similar. LO: 222, not longest possible ORF. F: hypothetical protein.

**Gene 50**

SSC: Start: 38071 Stop: 38289 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 403, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 38071 has strength 5.48. Gap: 75 bp gap with previous gene. Blast: gp49 of ZoeJ; q1:s1, gp45 of TM4; q1:s1. LO: 219 bp; longest possible ORF. F: hypothetical protein.

**Gene 51**

SSC: Start: 38286 Stop: 38567 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 462, best possible score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 38286 has strength 9.99. Overlap: 4 bp overlap with previous gene. BLAST: gp46 of TM4, q1:s1, 94.62% identical, 96.77% similar; gp50 of ZoeJ, q1:s1, 93.55% identical, 96.77% similar; gp33 of 33D, q1:s1, 93.55% identical, 95.70% similar. LO: not longest possible ORF. F: Hypothetical protein.

**Gene 52**

SSC: Start: 38581 Stop: 38790 (FWD). CP: ORF leaves out a tiny fraction of the coding potential shown on GeneMark-smeg output. SD: 756, best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 38581 has strength 13.83. Gap: 13 bp gap with previous gene. Blast: gp47 of TM4; q1:s1, 98.78% similarity and 94.12% identity; gp51 of ZoeJ; q1:s1, 100% aligned. LO: 210 bp; longest possible ORF. F: hypothetical protein.

**Gene 53**

SSC: Start: 38795 Stop: 39535 (FWD). CP: Includes all coding potential shown on GeneMark-smeg output. SD: 336, not best possible score. SCS: Agrees with GeneMark but not Glimmer. Original Glimmer call @bp 38846 has strength 13.06; GeneMark calls start at 38795. Gap: 4 bp gap with previous gene. BLAST: gp48 of TM4, gp52 of ZoeJ, q1:s1 for both. LO: 741, not longest possible ORF. F: DNA binding protein. FS: HHPred.

**Gene 54**

SSC: Start: 39535 Stop: 39765 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 756, best score. SCS: Agrees with both Glimmer and GeneMark. Original Glimmer call @bp 39535 has strength 10.62. Gap: 1 bp overlap with previous gene. Blast: gp49 of TM4; q1:s1, 100% alignment, similarity, and identity; WhiB-like transcriptional regulator of ZoeJ; q1:s1, 100% similarity and 98.68% identity. LO: 231 bp, second longest possible ORF. F: transcriptional protein FS: Blast.

**Gene 55**

SSC: Start 39762 Stop: 40520 (FWD). CP: ORF does not include all coding potential shown on GeneMark-smeg output. SD: 462, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 39762 has strength 11.80. Overlap: 4 bp overlap with previous gene. BLAST: gp54 of ZoeJ, q1:s1, 81.82% identical, 95.15% similar; gp50 of TM4, q1:s1, 91.33% identical, 94.87% similar. LO: 759, not longest possible ORF. F: hydrolase, glycosidase. FS: HHpred.

**Gene 56**

SSC: Start: 40517 Stop: 41149 (FWD). CP: ORF includes all coding potential shown in GeneMark-smeg output. SD: 567, best score. SCS: Agrees with both GeneMark and Glimmer predictions. Original Glimmer call @bp 40517 has strength 15.86. Overlap: 4 bp overlap with previous gene. BLAST: gp52 of TM4, q1:s1, 99.00% identical, 100.00% similar; gp55 of ZoeJ, q1:s1, 88.50% identical, 94.00% similar; gp111 of Henry, q1:s1. LO: 633, longest possible ORF. F: Hypothetical protein.

**Gene 57**

SSC: Start: 41142 Stop: 41537 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 420, best possible score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 41142 has strength 10.71. Overlap: 8 bp overlap with previous gene. BLAST: gp56 of ZoeJ, q1:s1, 99.24% identical, 100.00% similar. LO: longest possible ORF. F: Hypothetical protein. FS: BLAST.

**Gene 58**

SSC: Start: 41530 Stop: 41808 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 483, best possible score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 41530 has strength 15.19. Overlap: 8 bp overlap with previous gene. BLAST: gp57 of ZoeJ, q1:s1, 100.00% identical, 100.00% similar; gp54 of TM4, q1:s1, 96.74% identical, 98.73% similar; gp60 of Fionnbharth, q1:s1. LO: 279, longest possible ORF. F: Hypothetical protein. FS: BLAST.

**Gene 59**

SSC: Start 41805 Stop: 41921 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 294, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 41805 has strength 17.15. Gap: 4 bp overlap. BLAST: gp58 of ZoeJ, q1:s1, 97.37% identical, 97.37% similar. LO: 117, not longest possible ORF. F: Hypothetical protein.

**Gene 60**

SSC: Start: 42004 Stop: 42537 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 325, not best score. SCS: Disagrees with both Glimmer and GeneMark predictions, chose start because better q1:s1 results. Original Glimmer call @bp 42019 has strength 13.98. Gap: 82 bp gap with previous gene. BLAST: DNA Q of ZoeJ, q1:s1, 97.18% identical, 98.87% similar; gp57 of TM4, q1:s1, 97.74% identical, 97.74% similar; gp59 of MacnCheese, q1:s1, gp56 of Pixie, q1:s1; LO: 534 bp, not longest ORF. F: DNA Q-like protein, exonuclease. FS: HHpred, BLAST.

**Gene 61**

SSC: Start: 42534 Stop: 43358 (FWD). CP: ORF does not include all coding potential shown by GeneMark-smeg output. SD: 315, not best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 42534 has strength 11.64. Gap: 4 bp overlap with previous gene. BLAST: hydrolase of ZoeJ, q1:s1, 93.80% identical, 95.35% similar; gp58 of TM4, q1:s1, 89.19% identical, 93.80% similar. LO: 825, not longest possible ORF. F: hydrolase protein. FS: HHpred, BLAST.

**Gene 62**

SSC: Start: 43355 Stop: 44212 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 525, best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 43355 has strength 13.89. Gap: 4 bp overlap with previous gene. BLAST: exonuclease of ZoeJ, q1:s1, 97.54% identical, 98.95% similar; gp59 of TM4, q1:s1, 95.09% identical, 97.54% similar; gp20 of 33D, q1:s1, 94.74% identical, 97.19% similar. LO: 858, longest possible ORF. F: exonuclease protein. FS: HHpred, BLAST.

**Gene 63**

SSC: Start: 44209 Stop: 44385 (FWD). CP: ORF does not include all coding potential shown by GeneMark-smeg output, a small part of the 5' end is excluded. SD: 336, not best score. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 44209 has strength 16.55. Overlap: 4 bp overlap with previous gene. BLAST: gp60 of TM4, q1:s1, 96.55% identical, 98.25% similar; gp62 of ZoeJ, q1:s1, 89.66% identical, 92.98% similar. LO: 177, not longest possible ORF. F: hypothetical protein.

**Gene 64**

SSC: Start: 44382 Stop: 44600 (FWD). CP: ORF includes all coding potenital on GeneMark-smeg output. SD: 336, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 44382 has strength 15.40. Overlap: 4 bp overlap with previous gene. BLAST: gp63 of ZoeJ, q1:s1, 100.00% identical, 100.00% similar; gp61 of TM4, q1:s1, gp59 of FIonnbharth, q1:s1. LO: 219, not longest ORF. F: toxin. FS: HHpred.

**Gene 65**

SSC: Start: 44691 Stop: 45011 (FWD). CP: ORF includes all coding potential on GeneMark-Smeg output. SD: 403, not best score but best BLAST results. SCS: Agrees with Glimmer and DNA Master predictions. Original Glimmer call @bp 44691 has strength 16.88. Gap: 90 bp gap with previous gene. BLAST: gp62 of TM4, q1:s1, 96.23% identical, 98.11% similar; gp64 of ZoeJ, q1:s1, 89.62% similar. LO: 321, longest ORF. F: hypothetical protein.

**Gene 66**

SSC: Start: 45008 Stop: 45217 (FWD). CP: ORF excludes a small amount of beginning coding potential on GeneMark-smeg output. SD: 504, not best score, best score had no BLAST results. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 45008 has strength 16.82. Overlap: 4 bp overlap with previous gene. BLAST: gp65 of ZoeJ, q1:s1, 94.20% identical, 97.10% similar. LO: 210, not longest ORF. F: hypothetical protein.

**Gene 67**

SSC: Start: 45272 Stop: 45391 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 336, best and only score. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 45272 has strength 11.15. Gap: 54 bp gap with previous gene. BLAST: gp65 of TM4, q1:s1, 97.44% identical, 100.00% similar; gp66 of ZoeJ, q1:s1, 94.87% identical, 100.00% similar; gp69 of Validus, q1:s1, gp64 of Angelica, q1:s1. LO: 120 bp, longest possible ORF. F: hypothetical protein.

**Gene 68**

SSC: Start: 45460 Stop: 45978 (FWD). CP: ORF does not include all coding potential on GeneMark-smeg output, a small part of the 5' end is excluded. SD: 450, not best score. SCS: Agrees with GeneMark predictions but not Glimmer, GeneMark start position yielded better BLAST results. Original Glimmer call @bp 45433 has strength 17.04; GeneMark calls start at 45460. Gap: 68 bp gap with previous gene. BLAST: hypothetical protein of ZoeJ, q1:s1, 98.84% identical, 99.42% similar; hypothetical protein of TM4, 95.00% aligned, 98.26% identical, 99.42% similar; gp66 of TM4, q1:s1, 98.26% identical, 99.42% similar; gp66 of Pixie, q1:s1, gp 69 of MacnCheese, q1:s1; hypothetical protein of Validius, q1:s1; gp68 of Anaya, q1:s1. LO: 519, not longest possible ORF. F: hypothetical protein.

**Gene 69**

SSC: Start: 46066 Stop: 46311 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 609, best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 46066 has strength 15.52. Gap: 87 bp gap with previous gene. BLAST: Nrd-H like glutaredoxin gp67 of ZoeJ, q1:s1, 100.00% identical, 100.00% similar, gp67 of TM4, q1:s1. LO: 246, not longest ORF. F: NrdH-like glutaredoxin. FS: HHpred, BLAST.

**Gene 70**

SSC: Start: 46311 Stop: 46478. CP: ORF does not include all coding potential shown on GeneMark-smeg output, small part of 5' end is excluded. SD: 336, not best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 46311 has strength 18.47. Overlap: 1 bp overlap with previous gene. BLAST: gp69 of ZoeJ, q1:s1, 98.18% identical, 100.00% similar; gp68 of TM4, q1:s1, 96.36% identical, 96.36% similar. LO: 168, not longest ORF. F: hypothetical protein.

**Gene 71**

SSC: Start: 46475 Stop: 46819 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 273, not best score but good BLAST results. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 46475 has strength 11.05. Gap: 4 bp overlap. BLAST: HNH endonuclease of ZoeJ, q1:s1, 92.86% identical, 97.32% similar; gp69 of TM4, q1:s1. LO: 345, longest possible ORF. F: HNH endonuclease protein. FS: HHPred, BLAST.

**Gene 72**

SSC: Start: 46832 Stop: 49378 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 231, not best score but better BLAST results. SCS: Disagrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 46865 has strength 13.37. Gap: 4 bp gap with previous gene. BLAST: DNA primase/helicase of ZoeJ, q1:s1, 92.92% identical, 95.63% similar. LO: 2547, longest ORF. F: DNA replication protein, primase/helicase. FS: HHpred, BLAST.

**Gene 73 PC**:

SSC: Start: 49375 Stop: 49839 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 525, best score. SCS: Disagrees with both Glimmer and GeneMark predictions; not called by either. The gene was added to minimize the gap and to cover coding potential. Gap: 4 bp overlap with previous gene. BLAST: gp72 of ZoeJ, q44:s43. LO: 465, not longest ORF. F: hypothetical protein.

**Gene 74 PC, Gene 73 GCSU**

SSC: Start: 49842 Stop: 50477 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 651, best score. SCS: Disagrees with Glimmer predictions, agrees with GeneMark predictions, GeneMark start position has better BLAST results. Original Glimmer call @bp 49836 has strength 13.38; GeneMark calls start at 49842. Gap: 2 bp gap with previous gene. BLAST: RusA of ZoeJ, q1:s1, 98.44% identical, 99.48% similar; gp72 of TM4, q1:s1, 96.35% identical, 98.96% similar. LO: 636, not longest ORF. F: RusA protein. FS: HHpred, BLAST.

**Gene 75 PC, Gene 74 GCSU**

SSC: Start: 50477 Stop: 50788 (FWD). CP: ORF excludes a small amount of coding potential shown on GeneMark-smeg output. SD: 588, best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 50477 has strength 14.79. Overlap: 1 bp overlap with previous gene. BLAST: gp74 of ZoeJ, q1:s1, 99.03% identical, 100.00% similar. LO: 312, not longest ORF. F: hypothetical protein.

**Gene 76 PC, Gene 75 GCSU**

SSC: Start: 50785 Stop: 51114 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 420, not best score but better scores exclude large amounts of coding potential. SCS: Disagrees with both Glimmer and GeneMark predictions, new start position has better BLAST results. Original Glimmer call @bp 50812 has strength 5.60. Overlap: 4 bp overlap with previous gene. BLAST: gp74 of TM4, q1:s1, 93.58% identical, 95.41% similar; gp75 of ZoeJ, q1:s1, 91.74% identical, 94.50% similar. LO: 330, longest ORF. F: hypothetical protein.

**Gene 77 PC, Gene 76 GCSU**

SSC: Start: 51107 Stop: 52009 (FWD). CP: ORF excludes a bit of coding potential shown on GeneMark-smeg output. SD: 651, best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 51107 has strength 9.29. Overlap: 8 bp overlap. BLAST: gp76 of ZoeJ, q1:s1, 97.33% identical, 99.00% similar; gp75 of TM4, q1:s1, 93.73% identical, 96.70% similar. LO: 903, longest possible ORF. F: transcriptional regulator. FS: HHPred.

**Gene 78 PC, Gene 77 GCSU**

SSC: Start: 52006 Stop: 52467 (FWD). CP: ORF includes all but a small amount of coding potential on GeneMark-smeg output. SD: 546, not best score. SCS: Disagrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 52078 has strength 16.99. Gap: 4 bp gap with previous gene. BLAST: gp76 of TM4, q1:s1; gp77 of ZoeJ, q1:s1. LO: 462, not longest ORF. F: hypothetical protein.

**Gene 79 PC, Gene 78 GCSU**

SSC: Start: 52464 Stop: 52733 (FWD). CP: ORF excludes a tiny fraction of coding potential shown on GeneMark-smeg output. SD: 399, not best score but better BLAST results. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 52464 has strength 9.23. Gap: 4 bp overlap with previous gene. BLAST: gp77 of TM4, q1:s1, 96.63% identical, 97.75% similar; gp78 of ZoeJ, q1:s1, 98.88% identical, 98.88% similar. LO: 270, not longest ORF. F: hypothetical protein.

**Gene 80 PC, Not called by GCSU**

SSC: Start: 52756 Stop: 52893 (FWD). CP: ORF does not include all coding potential shown on GeneMark-smeg output, small part of 5' end is excluded. SD: 450, not best score. SCS: Agrees with Glimmer predictions. Original Glimmer call @bp 52756 has strength 13.29. Gap: 22 bp gap. BLAST: no Blast results at this start position so start position was chosen that includes most of the coding potential with minimal overlap with previous gene. LO: 138, not longest possible ORF. F: no Blast results and no hits on HHPred. (The Georgia State College team elected to delete this gene, which is certainly reasonable given the lack of BLAST results. The Providence College team decided not to delete the gene because there was significant coding potential at this position.)

**Gene 81 PC/Gene 79 GCSU**

SSC: Start: 52951 Stop: 53262 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 504, best score. SCS: Does not agree with Glimmer or GeneMark predictions, new start position includes more coding potential and has good BLAST results. Original Glimmer call @bp 53050 has strength 15.43. Gap: 156 bp gap with previous gene. BLAST: gp80 of ZoeJ, q1:s1, 100.00% identical, 100.00% similar; gp79 of TM4, q1:s1, 98.06% identical, 98.06% similar; hypothetical protein of TM4, 100.00% identical, 100.00% similar. LO: 312, not longest ORF. F: hypothetical protein.

**Gene 82 PC/Gene 80 GCSU**

SSC: Start: 53316 Stop: 53885 (FWD). CP: ORF includes all but a small amount of coding potential shown on GeneMark-smeg output. SD: 377, not best score but best BLAST results. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 53316 has strength 9.79. Gap: 53 bp gap with previous gene. BLAST: gp80 of TM4, q2:s3, 99.00% aligned; SprT-like protein of ZoeJ, q2:s3, 99.00% aligned. LO: 570, longest possible ORF. F: SprT-like protein, hydrolase protein. FS: BLAST, HHPred.

**Gene 83 PC/Gene 81 GCSU**

SSC: Start: 53925 Stop: 54704 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 462, not best score. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 53925 has strength 13.21. Gap: 39 bp gap with previous gene. BLAST: gp82 of TM4, q1:s1, 96.71% identical, 98.35% similar; gp 82 of ZoeJ, q1:s1, 90.61% identical, 91.84% similar. LO: 780, longest possible ORF. F: hypothetical protein.

**Gene 84 PC/Gene 82 GCSU**

SSC: Start: 54701 Stop: 54961 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 399, not best score but has good blast results and includes all coding potential. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 54701 has strength 10.21. Gap: 4 bp overlap with previous gene. BLAST: gp83 of ZoeJ, q1:s1, 98.84% identical, 98.84% similar; gp83 of TM4, q1:s1, 96.47% identical, 96.47% similar; gp74 of Larva, q1:s1. LO: 261, longest possible ORF. F: hypothetical protein.

**Gene 85 PC/Gene 83 GCSU**

SSC: Start: 54955 Stop: 55176 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 441, not best score but best BLAST results. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 54955 has strength 13.74. Overlap: 7 bp overlap with previous gene. BLAST: gp84 of TM4, q1:s1, 100.00% identical, 100.00% similar; gp84 of ZoeJ, q1:s1, 98.44% identical, 98.44% similar; gp79 of ChrisnMitch, q1:s1, 100.00% aligned. LO: 222, not longest ORF. F: hypothetical protein.

**Gene 86 PC/Gene 84 GCSU**

SSC: Start: 55310 Stop: 55798 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 462, best score. SCS: Agrees with Glimmer and GeneMark predictions. Original Glimmer call @bp 55310 has strength 15.94. Gap: 133 bp gap with previous gene. BLAST: gp85 of ZoeJ, q1:s1, 87.14% identical, 95.71% similar, gp85 of TM4, q1:s1. LO: 489, longest ORF. F: hypothetical protein.

**Gene 87 PC/Gene 85 GCSU**

SSC: Start: 55795 Stop: 56034 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 195, worst score. SCS: Agrees with Glimmer, but not GeneMark. Original Glimmer call @bp 55795 has strength 11.45; GeneMark calls start at 55798. Blast: no blast results. Gap: 4 bp overlap with previous gene. LO: 240 bp, longest ORF. F: hypothetical protein.

**Gene 88 PC/Gene 86 GCSU**

SSC: Start: 56157 Stop: 56420 (FWD). CP: ORF includes most of coding potential, the ORF stops slightly short of coding potential end. SD: 462, best score. SCS: Agrees with both Glimmer and GeneMark-Smeg predictions. Original Glimmer call @bp 56157 has strength 17.48. Gap: 122 bp gap with previous gene. Blast: gp86 of ZoeJ; q1:s1; 100% similarity and alignment, gp95 of Pixie; q1:s1; 97.7% aligned. LO: 264 bp, longest ORF. F: hypothetical protein.

**Gene 89 PC/Gene 87 GCSU**

SSC: Start: 56544 Stop: 56873 (FWD). CP: ORF includes all coding potential shown on GeneMark-smeg output. SD: 378, not best score. SCS: Disagrees with both Glimmer and GeneMark. This ORF includes all coding potential and had better blast values. Original Glimmer call @bp 56640 has strength 16.29. Blast: gp87 of TM4; q1:s1; gp 87 of ZoeJ, q2:s22. Gap: 123 bp gap with previous gene. LO: 330 bp, longest ORF. F: hypothetical protein.

**Gene 90 PC/Gene 88 GCSU**

SSC: Start: 57047 Stop: 57283 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 357, best score. SCS: Agrees with GeneMark, but not Glimmer (Glimmer call excludes coding potential). Original Glimmer call @bp 57092 has strength 6.96; GeneMark calls start at 57047. Blast: gp89 of ZoeJ; q1:s1; alignment 100%, similarity and identity are both 94.87%, gp89 of TM4, q1:s1. Gap: 173 bp gap with previous gene. LO: 237 bp, longest ORF. F: hypothetical protein.

**Gene 91 PC/Gene 89 GCSU**

SCS: Start: 57296 Stop: 57592 (FWD). CP: ORF includes all but a small amount of coding potential shown on GeneMark-smeg output. SD: 315, not best score. SCS: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 57296 has strength 18.21. Blast: gp90 of TM4; q1:s1, gp90 of ZoeJ; q1:s1. Gap: 12 bp gap with previous gene. LO: 297 bp, longest possible ORF. F: endonuclease. FS: HHpred.

**Gene 92 PC/Gene 90 GCSU**

SCS: Start: 57605 Stop: 57775 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 399, best score. SSC: Agrees with both Glimmer and GeneMark predictions. Original Glimmer call @bp 57605 has strength 10.24 Blast: gp91 of ZoeJ; q1:s1; 92.86% similarity and 91.07% identical, gp91 of TM4; q1:s1. Gap: 12 bp gap with previous gene. LO: 171 bp, longest ORF. F: hypothetical protein.

**Gene 93 PC/Gene 91 GCSU**

SCS: Start: 57772 Stop: 58050 (FWD). CP: ORF includes all coding potential shown on GeneMark-Smeg output. SD: 378, not best score. SSC: Agrees with GeneMark, but not with Glimmer predictions. Original Glimmer call @bp 57769 has strength 9.37; GeneMark calls start at 57772. Blast: HNH endonuclease of ZoeJ; q1:s1, 95.65% similar and 94.57% identical, gp100 of Pixie; q1:s1, gp99 of MacnCheese; q1:s1. Gap: 4 bp overlap with previous gene. LO: 279 bp, not longest ORF. F: HNH endonuclease FS: Blast and HHpred.