TrixiePhattel Merged Notes for all Genes

|  |
| --- |
| Gene Information Stop 636 |
| Start called at 385 in order to capture coding potential. BLAST shows a full coverage alignment with hypothetical protein from clustermate Lewando. HHPred shows similarity for various enzymes, but with probability below 90%. Match with peptidoglycan endolysin was with uncharacterized region. |
| Function: Hypothetical Protein |
|  |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 636 |
| Glimmer | 385 |
| GeneMark | not called |
| Length(s) | 252 |
| Overlap(s) | n/a |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | n/a |
| Best Z-score? | 424 z=2.065 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~390 |
| Staterator Screenshot    Gene: TrixiePhattel\_1 Start: 385, Stop: 636, Start Num: 3 Candidate Starts for TrixiePhattel\_1: (Start: 3 @385 has 1 MA's), (6, 424), (8, 493), (9, 529), (10, 571),  Start 3: • Found in 7 of 7 ( 100.0% ) of genes in pham • Manual Annotations of this start: 1 of 1 • Called 57.1% of time when present • Phage (with cluster) where this start called: Arzan\_2 (FI), Tenney120\_1 (AU6), TrixiePhattel\_1 (AU6), Uzumaki\_1 (AU6),  <http://phages.wustl.edu/starterator/Pham198720Report.pdf> | |
| Blast Screen Shot | |
| Function  No HHPred predicted functions above 90%. High percentage with Peptidoglycan DL-endopeptidase | |

|  |
| --- |
| Gene Information |
| Start called at 385 in order to capture coding potential. BLAST shows a full coverage alignment with hypothetical protein from clustermate Lewando. HHPred shows similarity for various enzymes, but with probability below 90%. Match with peptidoglycan endolysin was with uncharacterized region. |
| Function: Hypothetical Protein |
|  |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 636 |
| Glimmer | 385 |
| GeneMark | not called |
| Length(s) | 252 |
| Overlap(s) | n/a |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | n/a |
| Best Z-score? | 424 z=2.065 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~390 |
| Staterator Screenshot    Gene: TrixiePhattel\_1 Start: 385, Stop: 636, Start Num: 3 Candidate Starts for TrixiePhattel\_1: (Start: 3 @385 has 1 MA's), (6, 424), (8, 493), (9, 529), (10, 571),  Start 3: • Found in 7 of 7 ( 100.0% ) of genes in pham • Manual Annotations of this start: 1 of 1 • Called 57.1% of time when present • Phage (with cluster) where this start called: Arzan\_2 (FI), Tenney120\_1 (AU6), TrixiePhattel\_1 (AU6), Uzumaki\_1 (AU6),  <http://phages.wustl.edu/starterator/Pham198720Report.pdf> | |
| Blast Screen Shot | |
| Function  No HHPred predicted functions above 90%. High percentage with Peptidoglycan DL-endopeptidase | |

|  |
| --- |
| Gene Information Stop 1051 |
| Start called at 665, due to high Z score, agreement with Starterator data and BLAST alignment with protein from from clustermate Zenia |
| Function: Hypothetical Protein |
| No BLAST alignments with proteins with known function and no high probability hits with HHPRed. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 1051 |
| Glimmer | 665 |
| GeneMark | 665 |
| Length(s) | 111 |
| Overlap(s) | 28 gap |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 656 ( |
| Best Z-score? | 665 z=2.468 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~700 |
| Staterator Screenshot    Gene: TrixiePhattel\_2 Start: 665, Stop: 1051, Start Num: 27 Candidate Starts for TrixiePhattel\_2: (24, 656), (Start: 27 @665 has 49 MA's), (55, 869), (65, 947), (68, 974), (72, 998),  Start 27: • Found in 72 of 211 ( 34.1% ) of genes in pham • Manual Annotations of this start: 49 of 167 • Called 100.0% of time when present • Phage (with cluster) where this start called: Acai\_1 (AU1), AreFloNak\_1 (AU1), Argan\_1 (AU6), BarbieDoll\_2 (AU6), Biscute\_1 (AU6), Boog\_2 (AW), Bouchard\_1 (AU2), Breylor17\_1 (AU1), BronxBay\_2 (AW), Brunswick\_1 (AU1), CapnMurica\_2(AU1), CastorTray\_2 (AU1), Caterpillar\_2 (AU4), Chlochlo\_1 (AU2), Darby\_1 (AU1), DevitoJr\_1 (AU1), Djungelskog\_2 (AW), DoctorPepper\_2 (AW), Egad\_2 (AW), ElephantMan\_1 (AU1), GantcherGoblin\_1 (AU6), Giantsbane\_1 (AU2), Gordon\_1 (AU1), HerbBucket\_2 (AW), Ingrid\_1 (AU3), Inked\_2 (AU), Issa\_1 (AU1), KevinMinion\_1 (AU6), Kinny\_2 (AU6), Leathea\_2 (AU6), Lewando\_2 (AU6), LilHuddy\_1 (AU2), Linda\_2 (AW), Loretta\_1 (AU3), LucySwiss\_1 (AU1), Makai\_1 (AU5), MediumFry\_2 (AU4), Michelle\_2 (AW), MiniBagel\_2 (AU1), MrAaronian\_2 (AW), Natasha\_1 (AW), Navi1117\_2 (AU6), Nightmare\_1 (AU1), Niktson\_1 (AU1), Nivinsha\_1 (AU1), Phaby\_1 (AU2), Phaila\_2 (AU6), Powelldog\_1 (AW), ProfFrink\_2 (AW), Raunak\_2 (AW), Renaldo\_2 (AU6), RustyBoy\_1 (AW), Salk\_2 (AW), ScienceWizSam\_2 (AU1), Shepard\_1 (AU2), Shiba\_2 (AW), Sloopyjoe\_2 (AW), Sporto\_1 (AW), StarLord\_2 (AW), Stayer\_2 (AW), Synepsis\_1 (AU1), Tatanka\_1 (AU1), Teacup\_2 (AU1), Tenney120\_2 (AU6), Tenno\_2 (AU1), Tokki\_1 (AU2), TrixiePhattel\_2 (AU6), Truckee\_1 (AU5), Trustiboi\_1 (AU1), Uzumaki\_2 (AU6), Zeina\_2 (AU6), Zippen\_2 (AU3), | |
| BLAST | |

|  |
| --- |
| Gene Information Stop 1451 |
| Start of 1041 predicted by Glimmer and Genemark. Corresponds to most frequently called start on Staterator and captures coding potential. 100% alignment with BLAST of clustermate. |
| Function |
| HNH Endonuclease |
| HHPRed predicts with 98.5 confidence with Geobacillus virus HNH Endonuclease. HNH motif present within 30 AA’s located within conserved domain. |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 1451 |
| Glimmer | 1041 |
| GeneMark | 1041 |
| Length(s) | 528 |
| Overlap(s) | 11 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1041 (11 overlap) |
| Best Z-score? | 1041 (tied with 1032)=2.636 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~1050 |
| Staterator Screenshot  Gene: TrixiePhattel\_3 Start: 1041, Stop: 1451, Start Num: 36 Candidate Starts for TrixiePhattel\_3: (16, 963), (19, 978), (21, 990), (Start: 24 @1002 has 3 MA's), (Start: 33 @1032 has 1 MA's), (Start: 36 @1041 has 58 MA's), (62, 1281), (63, 1290), (72, 1398)  http://phages.wustl.edu/starterator/Pham85049Report.pdf | |
| BLAST | |
| HHPred | |

|  |
| --- |
| Gene Information Stop 2447 |
| Start 1,473 |
| Start 1,473 has the best Z score, and has 48 Manual Annotations on Starterator with a 100% call rate when present. BLAST alignment with cluster member Uzumaki shows 100% coverage starting with the 1st amino acid. |
| Function: endolysin |
| BLAST shows strong alignment with endolysins and Lysin A. Phagesdb Function Frequency shows multiple counts of endolysin and Lysin A. HHPred has 99%+ calls with peptidoglycan hydrolysing proteins and conserved domain database shows a conserved peptidoglycan binding domain. Endolysin called as there is no evidence of other endolysins in the genome. |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction |  |
| Stop | 2447 |
| Glimmer | 1473 |
| GeneMark | not called |
| Length(s) | 975 |
| Overlap(s) | 21 gap |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | #1 1473 Z= 3.297 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~1530 |
| Staterator Screenshot    Start 17: • Found in 69 of 73 ( 94.5% ) of genes in pham • Manual Annotations of this start: 48 of 52 • Called 100.0% of time when present • Phage (with cluster) where this start called: Acai\_3 (AU1), AreFloNak\_3 (AU1), Argan\_3 (AU6), BarbieDoll\_4 (AU6), Biscute\_4 (AU6), Boog\_4 (AW), BouleyBill\_24 (EA4), Breylor17\_3 (AU1), BronxBay\_4 (AW), Brunswick\_3 (AU1), Bustleton\_24 (EA4), CastorTray\_4 (AU1), Caterpillar\_4 (AU4), Cheeto1\_25 (EA9), ChickenKing\_24 (EA9), Darby\_3 (AU1), DevitoJr\_3 (AU1), Djungelskog\_4 (AW), DoctorPepper\_4 (AW), Egad\_4 (AW), ElephantMan\_3 (AU1), GaeCeo\_25 (EA9), GantcherGoblin\_3 (AU6), Gordon\_4 (AU1), HerbBucket\_4 (AW), Ingrid\_3 (AU3), Jingles\_24 (EA4), Kauala\_24 (EA4), KevinMinion\_3 (AU6), Kinny\_4 (AU6), Koji\_24 (EA4), Leathea\_4 (AU6), Lewando\_4 (AU6), Linda\_4 (AW), Loretta\_3 (AU3), Makai\_3 (AU5), MediumFry\_4 (AU4), Michelle\_4 (AW), MrAaronian\_4 (AW), Natasha\_3 (AW), Navi1117\_4 (AU6), Niktson\_3 (AU1), Nivinsha\_3 (AU1), Phaila\_4 (AU6), Pherbot\_24 (EA4), Powelldog\_3 (AW), PrincePhergus\_24 (EA4), ProfFrink\_4 (AW), Raunak\_4 (AW), Renaldo\_5 (AU6), RenegadeRaider\_25 (EA4), RustyBoy\_3 (AW), Salk\_4 (AW), ScienceWizSam\_4 (AU1), Shiba\_4 (AW), Sinatra\_24 (EA4), Sloopyjoe\_4 (AW), Sporto\_3 (AW), StarLord\_4 (AW), Stayer\_4 (AW), Synepsis\_3 (AU1), Tatanka\_3 (AU1), Tenney120\_4 (AU6), Theresita\_24 (EA7), TrixiePhattel\_4 (AU6), Truckee\_3 (AU5), Trustiboi\_3 (AU1), Uzumaki\_4 (AU6), Zeina\_4 (AU6),  Gene: TrixiePhattel\_4 Start: 1473, Stop: 2447, Start Num: 17 Candidate Starts for TrixiePhattel\_4: (Start: 17 @1473 has 48 MA's), (19, 1488), (31, 1740), (40, 1944), (43, 1974), (49, 2064),  http://phages.wustl.edu/starterator/Pham227423Report.pdf | |
| BLAST | |
| HHpred    Conserved Domain | |

|  |
| --- |
| Gene Information Stop 2610 |
| Start changed to 2448 to align with Starterator data and BLASTs with cluster members. Genemark coding potential was not captured by original called start. |
| Function: membrane protein |
| No BLAST alignments with proteins of known function and no significant HHPred results. Deep TMHMM shows transmembrane domain, but there is no indication of a more specific function. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction |  |
| Stop | 2610 |
| Glimmer | 2460 |
| GeneMark | 2460 |
| Length(s) | 153 |
| Overlap(s) | -12 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? |  |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~2480 |
|  | |
| Starterator screenshot:  Gene: TrixiePhattel\_5 Start: 2460, Stop: 2612, Start Num: 3 Candidate Starts for TrixiePhattel\_5: (Start: 1 @2448 has 5 MA's), (3, 2460), (4, 2469), (7, 2505), (8, 2544), (10, 2550), (13, 2565), (14, 2592),  Start 1: • Found in 14 of 15 ( 93.3% ) of genes in pham • Manual Annotations of this start: 5 of 5 • Called 92.9% of time when present • Phage (with cluster) where this start called: Argan\_4 (AU6), BarbieDoll\_5 (AU6), Biscute\_5 (AU6), GantcherGoblin\_4 (AU6), KevinMinion\_4 (AU6), Kinny\_5 (AU6), Lewando\_5 (AU6), Navi1117\_5 (AU6), Phaila\_5 (AU6), Renaldo\_6 (AU6), Tenney120\_5 (AU6), Uzumaki\_5 (AU6), Zeina\_5 (AU6),  <http://phages.wustl.edu/starterator/Pham12059Report.pdf> | |
| BLAST | |
| Alignment? | |
|  | |
| HHPred | |

|  |
| --- |
| Gene Information Stop 2976 |
| Start 2605 |
| Start 2605 was called by both Glimmer and GeneMark and has an 8 base overlap with the previous gene. Coding potential starts almost immediately at this start codon. Starterator start 20 is consistently called for AU6 phages and has 93 manual annotations and is called 99.2% of the time when present. BLAST aligns with the first amino acid of best hits. |
| Function: Hypothetical protein |
| Despite the PECAAN HHpred showing a 90.1% probability with gp6 from phage HK suggesting head-to-tail adaptor, further alignment with HHpred shows only 80% probability with gp6 and only with a short portion of the protein. |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 2976 |
| Glimmer | 2605 |
| GeneMark | 2605 |
| Length(s) | 372 |
| Overlap(s) | 8 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 8 |
| Best Z-score? | 1.981 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~2610 |
| Staterator Screenshot | |
|  | |
| Gene: TrixiePhattel\_6 Start: 2605, Stop: 2976, Start Num: 20 Candidate Starts for TrixiePhattel\_6: (Start: 17 @2599 has 32 MA's), (Start: 20 @2605 has 93 MA's), (33, 2689), (39, 2731), (41, 2752), (51, 2815), (53, 2824), (54, 2830), (56, 2842), (60, 2884), (69, 2929), (71, 2938),  Start 20: • Found in 124 of 218 ( 56.9% ) of genes in pham • Manual Annotations of this start: 93 of 169 • Called 99.2% of time when present • Phage (with cluster) where this start called: Aflac\_15 (DJ), AlainaMarie\_15 (DJ), Annihilus\_5 (BI2), Arataki\_11 (DJ), Argan\_5 (AU6), Artorias\_11 (DJ), Arzan\_6 (FI), Avazak\_11 (DJ), BarbieDoll\_6 (AU6), BetterYeti\_17 (DJ), Bilo\_5 (BI2), Bing\_5 (BI5), BluerMoon\_15 (DJ), Burley\_15 (DJ), Captrips\_11 (DJ), Capybara\_11 (DJ), CherryBlossom\_6 (BI1), Chikenjars\_15 (DJ), Citrus\_14 (DJ), ClubPenguin\_4 (BI7), Conley\_11 (DJ), CricKo\_5 (BI4), Crocheter\_14 (DJ), Dancer\_13 (DJ), Demure\_12 (DJ), Doxi13\_5 (BI2), DrGrey\_5 (BI1), Duffington\_15 (DJ), Eddiemania\_14 (DJ), EndAve\_15 (DJ), Esketit\_6 (BI1), Fazel\_12 (DJ), FidgetOrca\_6 (BI1), Figliar\_15 (DJ), FortCran\_11 (DJ), FrodoSwaggins\_5 (BI1), FromIndy\_5 (BI4), GantcherGoblin\_5 (AU6), Gill\_11 (DJ), GirlPower\_5 (BI1), GoblinVoyage\_5 (BI2), Gravy\_11 (DJ), Gusicorn\_17 (DJ), HazuAndZazu\_5 (BI1), Hortihabitatio\_15 (DJ), Hoshi\_6 (BI1), HotFries\_4 (BI2), Hydrus\_14 (DJ), IceWarrior\_6 (BI1), Indigenous\_6 (BI1), Jaylociraptor\_6 (BI1), Jazzy4900\_7 (FI), Jodelie19\_15 (DJ), JulesRay\_13 (DJ), Kardashian\_5 (BI6), Kenosha\_12 (DJ), Kerry\_11 (DJ), KevinMinion\_5 (AU6), Kinny\_6 (AU6), Leathea\_6 (AU6), Lewando\_6 (AU6), LibertyBell\_5 (BI3), Madamato\_6 (BI1), Madvan\_15 (DJ), Magel\_11 (DJ), MajinBuu\_16 (DJ), Maya\_6 (BI1), Meibysrarus\_6 (BI1), Miek\_5 (BI4), Moozy\_5 (BI2), Mossy\_12 (DJ), Mysticmoon\_13 (DJ), Nadmeg\_15 (DJ), Namo\_6 (BI1), Navi1117\_6 (AU6), Nithya\_15 (DJ), NoPickles\_11 (DJ),  Odesza\_11 (DJ), OhMyWard\_15 (DJ), OlgasClover\_15 (DJ), OlympicHelado\_6 (BI1), Pepy6\_012 (CC), Perkunas\_13 (DJ), Petito\_12 (DJ), Phaila\_6 (AU6), Phepper\_15 (DJ), Pherobrine\_12 (DJ), PherryCruz\_4 (BI2), Poco6\_014 (CC), Popy\_5 (BI1), Rainydai\_5 (BI4), RavenPuff\_4 (BI2), Regigigas\_5 (BI4), Renaldo\_7 (AU6), Rickmore\_14 (DJ), Rima\_6 (BI1), Roney\_11 (DJ), RosaAsantewaa\_5 (BI2), Runhaar\_15 (DJ), Scap1\_5 (BI2), Schwartz33\_15 (DJ), Secretariat\_11 (DJ), SendItCS\_5 (BI4), ShaggyRogers\_18 (DJ), SheRa\_5 (BI2), Soshi\_5 (BI1), Spectropatronm\_6 (BI1), Sunny4976\_6 (FI), TaidaOne\_6 (BI1), Talon44\_13 (DJ), Tanis\_11 (DJ), TenaciousP\_15 (DJ), Tenney120\_6 (AU6), Thiqqums\_5 (BI4), TieDye\_5 (BI1), TonyStarch\_6 (BI1), TrixiePhattel\_6 (AU6), Trufflozitus\_14 (DJ), Untouchable\_14 (DJ), Uzumaki\_6 (AU6), Vardy\_15 (DJ), Zeina\_6 (AU6), Zeph\_14 (DJ)  Gene: TrixiePhattel\_6 Start: 2605, Stop: 2976, Start Num: 20 Candidate Starts for TrixiePhattel\_6: (Start: 17 @2599 has 32 MA's), (Start: 20 @2605 has 93 MA's), (33, 2689), (39, 2731), (41, 2752), (51, 2815), (53, 2824), (54, 2830), (56, 2842), (60, 2884), (69, 2929), (71, 2938),  http://phages.wustl.edu/starterator/Pham174562Report.pdf | |
| BLAST | |
| HHpred | |

|  |
| --- |
| Gene Information Stop 3449 |
| Start 2949 |
| Start 2949 was called by Glimmer and captures coding potential that start 2997 does not. The 28 bp overlap with the previous gene, but within the ~30 range of legitimate overlaps. This start also aligns with the first amino acid of BLAST results with cluster members. |
| Function: Hypothetical protein |
| No significant BLAST or HHpred results with proteins or domains of known function. |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 3449 |
| Glimmer | 2949 |
| GeneMark | 2997 |
| Length(s) | 501 |
| Overlap(s) | 28 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 28 overlap |
| Best Z-score? | 2, Z=2.642, 2949 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~2950 |
|  | |
| Staterator Screenshot | |
|  | |
| Stop 4 has lower called manual annotations (53.1) vs Stop 7 (61.1) but occurs more commonly in AU6 cluster phages. Stop 4 captures coding potential missed by Stop 7 | |
| BLAST | |
| HHpred | |

|  |
| --- |
| Gene Information Stop 5199 |
| Start 3466 |
| Start 3466 was called by both Glimmer and GeneMark and captures the coding potential. This start is highly conserved across multiple clusters with 156 manual annotations. This gives alignment with the first amino acid of almost all BLAST hits. |
| Function: terminase |
| Shows 100% probability with high coverage of HK97 terminase large subunit. AU6 phages including TrixiePhattel don’t show signs of a second terminase, so the function of “terminase” is assigned. In Siphoviridae, synteny has the terminase gene on the left side of the genome before the portal protein, which aligns in this instance. |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 5199 |
| Glimmer | 3466 |
| GeneMark | 3466 |
| Length(s) | 1734 |
| Overlap(s) | -16 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 3.053 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~3470 |
| BLAST | |
| Starterator screenshot    Gene: TrixiePhattel\_8 Start: 3466, Stop: 5199, Start Num: 9 Candidate Starts for TrixiePhattel\_8: (Start: 6 @3460 has 2 MA's), (Start: 9 @3466 has 156 MA's), (19, 3565), (33, 3766), (42, 3850), (49, 3928), (50, 3946), (51, 3949), (54, 4018), (61, 4120), (64, 4144), (72, 4234), (78, 4306), (82, 4372), (84, 4405), (86, 4435), (91, 4477), (93, 4534), (95, 4603), (96, 4615), (104, 4729), (106, 4747), (116, 4816), (117, 4819), (125, 4912), (130, 4999), (132, 5023), (136, 5047), (138, 5059), (139, 5071), (144, 5149), (145, 5152),  Start 9: • Found in 213 of 218 ( 97.7% ) of genes in pham • Manual Annotations of this start: 156 of 169 • Called 93.9% of time when present • Phage (with cluster) where this start called: Acai\_7 (AU1), Aflac\_17 (DJ), AlainaMarie\_17 (DJ), Annihilus\_7 (BI2), Arataki\_13 (DJ), Arcadia\_8 (AM), AreFloNak\_7 (AU1), Argan\_7 (AU6), Artorias\_13 (DJ), Arzan\_8 (FI), Avazak\_13 (DJ), BarbieDoll\_8 (AU6), BetterYeti\_19 (DJ), Bilo\_7 (BI2), Bing\_7 (BI5), Biscute\_8 (AU6), BluerMoon\_17 (DJ), Boog\_7 (AW), Bouchard\_7 (AU2), Breylor17\_7 (AU1), Brunswick\_7 (AU1), Burley\_17 (DJ), Camille\_10 (EL), CapnMurica\_7 (AU1), Captrips\_13 (DJ), Capybara\_13 (DJ), CastorTray\_8 (AU1), Caterpillar\_8 (AU4), Cheesy\_8 (AM), CherryBlossom\_8 (BI1), Chikenjars\_17 (DJ), Chlochlo\_7 (AU2), Circum\_8 (AM), Citrus\_16 (DJ), ClubPenguin\_6 (BI7), Conley\_13 (DJ), Correa\_8 (AM), CricKo\_7 (BI4), Crocheter\_16 (DJ), Damascus\_10 (EL), Darby\_7 (AU1), Demure\_14 (DJ), DevitoJr\_7 (AU1), DizzyRudy\_10 (EL), Djungelskog\_8 (AW), Doxi13\_7 (BI2), DrGrey\_7 (BI1), Duffington\_17 (DJ), Dynamite\_8 (AM), Eddiemania\_16 (DJ), Eevee\_12 (JA), Egad\_8 (AW), ElephantMan\_6 (AU1), Elsa\_8 (AM), Elver\_15 (FK), EndAve\_17 (DJ), Esketit\_8 (BI1), Fazel\_14 (DJ), FidgetOrca\_8 (BI1), Figliar\_17 (DJ), FortCran\_13 (DJ), FrodoSwaggins\_7 (BI1), FromIndy\_7 (BI4), Gandionco\_15 (FK), GantcherGoblin\_7 (AU6), Giantsbane\_7 (AU2), Gill\_13 (DJ), GirlPower\_7 (BI1), GoCrazy\_7 (AM), GoblinVoyage\_7 (BI2), Gordon\_7 (AU1), Gravy\_13 (DJ), Gusicorn\_19 (DJ), Hankly\_7 (AM), HazuAndZazu\_7 (BI1), Heisenberger\_8 (AM), HerbBucket\_7 (AW), Hortihabitatio\_17 (DJ), Hoshi\_8 (BI1), HotFries\_6 (BI2), Hydrus\_16 (DJ), IceWarrior\_8 (BI1), Indigenous\_8 (BI1), Ingrid\_7 (AU3), Inked\_8 (AU), Issa\_7 (AU1), JEGGS\_8 (AM), Jaylociraptor\_8 (BI1), Jazzy4900\_9 (FI), Jodelie19\_17 (DJ), JoyLin\_12 (JA), JulesRay\_15 (DJ), Kardashian\_7 (BI6), Kardesai\_7 (AM), KeaneyLin\_7 (AM), Kenosha\_14 (DJ), Kerry\_13 (DJ), KevinMinion\_7 (AU6), Kinny\_8 (AU6), Leathea\_8 (AU6), Lewando\_8 (AU6), LibertyBell\_7 (BI3), LilHuddy\_7 (AU2), Linda\_8 (AW), Loretta\_7 (AU3), LucySwiss\_7 (AU1), Madamato\_8 (BI1), Madvan\_17 (DJ), Magel\_13 (DJ), MajinBuu\_18 (DJ), Makai\_7 (AU5), Maya\_8 (BI1), MediumFry\_8 (AU4), Meibysrarus\_8 (BI1), Miek\_7 (BI4), MiniBagel\_7 (AU1), Moozy\_7 (BI2), Mossy\_14 (DJ), MrAaronian\_8 (AW), Mudcat\_8 (AM), Mysticmoon\_15 (DJ), Nadmeg\_17 (DJ), Namo\_8 (BI1), NapoleonB\_8 (AM), Nason\_8 (AM), Natasha\_6 (AW), Navi1117\_8 (AU6), Nightmare\_7 (AU1), Niktson\_6 (AU1), Nithya\_17 (DJ), Nivinsha\_7 (AU1), NoPickles\_13 (DJ), Odesza\_13 (DJ), OhMyWard\_17 (DJ), OlgasClover\_17 (DJ), OlympicHelado\_8 (BI1), Paella\_16 (FK), Pepy6\_014 (CC), Perkunas\_15 (DJ), Petito\_14 (DJ), Phaby\_7 (AU2), Phaila\_8 (AU6), Phepper\_17 (DJ), Pherobrine\_14 (DJ), PherryCruz\_6 (BI2), Pippin15\_11 (JA), Poco6\_016 (CC), Popy\_7 (BI1), Powelldog\_7 (AW), ProfFrink\_8 (AW), Qui\_16 (FK), Raunak\_8 (AW), RavenPuff\_6 (BI2), Renaldo\_9 (AU6), Rickmore\_16 (DJ), Rima\_8 (BI1), Roney\_13 (DJ), RosaAsantewaa\_7 (BI2), Runhaar\_17 (DJ), RustyBoy\_6 (AW), Salk\_8 (AW), Scap1\_7 (BI2), Schwartz33\_17 (DJ), ScienceWizSam\_8 (AU1), Secretariat\_13 (DJ), SendItCS\_7 (BI4), ShaggyRogers\_20 (DJ), SheRa\_7 (BI2), Shepard\_7 (AU2), Shiba\_8 (AW), Soshi\_7 (BI1), Spectropatronm\_8 (BI1), Sporto\_7 (AW), StarLord\_8 (AW), Sunny4976\_8 (FI), Synepsis\_6 (AU1), TaidaOne\_8 (BI1), Talon44\_15 (DJ), Tanis\_13 (DJ), Tatanka\_6 (AU1), Teacup\_7 (AU1), TenaciousP\_17 (DJ), Tenney120\_8 (AU6), Tenno\_8 (AU1), Thiqqums\_7 (BI4), TieDye\_7 (BI1), Tokki\_7 (AU2), TonyStarch\_8 (BI1), Tribby\_8 (AM), TrixiePhattel\_8 (AU6), Truckee\_7 (AU5), Trufflozitus\_16 (DJ), Trustiboi\_7 (AU1), Untouchable\_16 (DJ), Uzumaki\_8 (AU6), Vardy\_17 (DJ), Xenomorph\_7 (AM), Zeina\_8 (AU6), Zeph\_16 (DJ), Zippen\_8 (AU3),  http://phages.wustl.edu/starterator/Pham710Report.pdf | |
|  | |
| HHpred    Conserved Domain | |

|  |
| --- |
| Gene Information |
| Start 5391 Kept, originally called by glimmer and genemark. Coding potential was captured and there was a full coverage alignment with phage in the same cluster (Zenia). BLAST shows a 100% alignment, HHpred shows low probability of other protein function |
| Function:  Hypothetical Protein |
| -Terminase before this hypothetical protein with other genes in pham. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 6053 |
| Glimmer | 5391 |
| GeneMark | 5391 |
| Length(s) | 663 |
| Overlap(s) | Gap -191 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -191 |
| Best Z-score? | 3.286 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~5410 |
| Staterator Screenshot  <http://phages.wustl.edu/starterator/Pham86733Report.pdf> | |
|  | |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **Similar genes surrounding Zenia , terminase right before this hypothetical protein.** |
| **BLASTp (Identified function)**  **100% coverage with Zenia** |
| **HHpred** |
| **Conserved Domain: No data** |
| **DeepTMHMM: No data** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 6603 |
| Start 6055 kept, originally called by glimmer and genemark. Coding potential captured, and full coverage with other phage in cluster (Zenia). BLAST showed 100% alignment. HHpred shows 99% probability score with spermidine acetyltransferase. |
| Function- Acetyltransferase |
| -Same arrangement of genes as Zenia, also shown several times within AU6 cluster |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 6603 |
| Glimmer | 6055 |
| GeneMark | 6055 |
| Length(s) | 549 |
| Overlap(s) | -1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | #1 6055 Z=2.048 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~6050 |
| Staterator Screenshot | |
| Gene: TrixiePhattel\_10 Start: 6055, Stop: 6603, Start Num: 3 Candidate Starts for TrixiePhattel\_10: (Start: 3 @6055 has 4 MA's), (9, 6100), (11, 6148), (18, 6262), (21, 6331), (27, 6424), (36, 6586),  [Pham86901Report.pdf](http://phages.wustl.edu/starterator/Pham86901Report.pdf) | |

|  |
| --- |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **Same arrangement as Zenia, also shown several times within AU6 cluster even though it is typically found in the “Middle” of the phage genome.** |
| **BLASTp (Identified function)**  **100% coverage with zenia** |
| **HHpred**    **99% probability score with spermidine acetyltransferase** |
| **Conserved Domain**    **Conserved domain with acetyltransferase** |
| **DeepTMHMM (No results)** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 7939 |
| Start 6632 kept, called by glimmer and genemark. Coding potential captured. Full alignment coverage with phage in the same cluster (Uzumaki). BLAST shows 100% alignment, and starterator agrees with a high number of manual annotations. HHpred shows 100% probability score with Escherichia phage portal protein |
| Function: Portal protein |
| Same arrangement of surrounding genes with Uzumaki. Portal proteins typically found on the left hand of phage genomes. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 7939 |
| Glimmer | 6632 |
| GeneMark | 6632 |
| Length(s) | 2308 |
| Overlap(s) | -28 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -28 (gap) |
| Best Z-score? | 3.125 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~6635 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham701Report.pdf> | |
| Blast | |
|  | |

|  |
| --- |
| Function |
| Phamerator  **Context and synteny**  Same arrangement of surrounding genes with Uzumaki. Portal proteins typically found on the left hand of phage genomes. |
| **Blastp (Identified function?)**  100% aligned with Gantchergoblin with an identified function of portal protein |
| HHPred    100% probability score with Escherichia phage portal protein |
| Conserved Domain  Conserved domain with phage portal protein |
| DeepTMHMM  No results (function not consistent with transmembrane domain) |
| Other Evidence? |

|  |
| --- |
| Gene Information Stop 8664 |
| Start at 7939 called by both Glimmer and Genemark.  Coding potential fully captured according to the Genemark Gene Activity Sheet. Starterator report shows manual annotation of gene in clustermates by other research teams.  Blast Results show alignment with other clustermates Lewando, Uzumaki, Zeina for Hypothetical Protein. HHPred shows no significant results for any matches above 90%. |
| Function: Hypothetical Protein |
|  |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 8664 |
| Glimmer | 7939 |
| GeneMark | 7939 |
| Length(s) | 726 |
| Overlap(s) | 1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 1, Z=1.628, 7939 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~7950 |
| Staterator Screenshot | |
| [Starterator](http://phages.wustl.edu/starterator/Pham2636Report.pdf)  Gene: TrixiePhattel\_12 Start: 7939, Stop: 8664, Start Num: 10  Candidate Starts for TrixiePhattel\_12:  (Start: 10 @7939 has 5 MA's), (26, 8206), (30, 8236), (34, 8296), (37, 8335), (50, 8479), (52, 8524),(55, 8554), (61, 8656),    Start 10:  • Found in 17 of 63 ( 27.0% ) of genes in pham  • Manual Annotations of this start: 5 of 42  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Argan\_11 (AU6), Arzan\_10 (FI),  BarbieDoll\_16 (AU6), Biscute\_16 (AU6), GantcherGoblin\_11 (AU6), Jazzy4900\_11  (FI), KevinMinion\_11 (AU6), Kinny\_15 (AU6), Lewando\_15 (AU6), Navi1117\_13  (AU6), Phaila\_11 (AU6), Renaldo\_17 (AU6), Sunny4976\_10 (FI), Tenney120\_13  (AU6), TrixiePhattel\_12 (AU6), Uzumaki\_12 (AU6), Zeina\_12 (AU6),  Gene: TrixiePhattel\_12 Start: 7939, Stop: 8664, Start Num: 10  Candidate Starts for TrixiePhattel\_12:  (Start: 10 @7939 has 5 MA's), (26, 8206), (30, 8236), (34, 8296), (37, 8335), (50, 8479), (52, 8524),(55, 8554), (61, 8656), | |
| Blast Screenshot | |
| Function:  No HHPred Results higher than 90%. Best match | |

|  |
| --- |
| Gene Information Stop 10848 |
| Start at 8686 called by both Glimmer and Genemark. Coding potential fully captured according to the Gene Coding Report. Starterator shows a large amount of manual annotations of gene in clustermates by other research teams. Blast results show high match for major capsid and protease fusion protein, HHPred also shows high probability of Major Capsid Protein. |
| Function: Major Capsid and Protease Fusion Protein |
| Gene shows best alignment with Zeina and Uzumaki. Tenny and other clustermates also possess this gene but with misalignment from Trixie\_Phattel |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 10848 |
| Glimmer | 8686 |
| GeneMark | 8686 |
| Length(s) | 2163 |
| Overlap(s) | -21 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -21 |
| Best Z-score? | 1, Z=3.053, 17 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~8700 |
| Staterator Screenshot | |
| [Starterator](http://phages.wustl.edu/starterator/Pham708Report.pdf)  Gene: TrixiePhattel\_13 Start: 8686, Stop: 10848, Start Num: 3  Candidate Starts for TrixiePhattel\_13:  (Start: 3 @8686 has 138 MA's), (7, 8752), (10, 8791), (25, 8932), (28, 8959), (90, 9412), (96, 9454),(99, 9487), (111, 9595), (145, 9976), (148, 10030), (149, 10051), (152, 10075), (154, 10084), (160,10195), (163, 10210), (164, 10216), (177, 10330), (186, 10429), (194, 10468), (197, 10513), (211,10675), (216, 10735),  Start 3:  • Found in 183 of 218 ( 83.9% ) of genes in pham  • Manual Annotations of this start: 139 of 168  • Called 98.9% of time when present  • Phage (with cluster) where this start called: Acai\_13 (AU1), Aflac\_22 (DJ),  AlainaMarie\_22 (DJ), Annihilus\_9 (BI2), Arataki\_20 (DJ), AreFloNak\_13 (AU1),  Argan\_12 (AU6), Artorias\_19 (DJ), Arzan\_11 (FI), Avazak\_19 (DJ), BarbieDoll\_17  (AU6), BetterYeti\_24 (DJ), Bilo\_9 (BI2), Bing\_14 (BI5), Biscute\_17 (AU6),  BluerMoon\_22 (DJ), Boog\_12 (AW), Bouchard\_14 (AU2), Breylor17\_14 (AU1),  BronxBay\_13 (AW), Brunswick\_13 (AU1), Burley\_22 (DJ), Camille\_14 (EL),  CapnMurica\_13 (AU1), Captrips\_19 (DJ), Capybara\_19 (DJ), CastorTray\_15 (AU1),  Caterpillar\_13 (AU4), CherryBlossom\_14 (BI1), Chikenjars\_22 (DJ), Chlochlo\_14  (AU2), Citrus\_21 (DJ), ClubPenguin\_12 (BI7), Conley\_19 (DJ), Count\_19 (EL),  CricKo\_11 (BI4), Crocheter\_21 (DJ), Damascus\_14 (EL), Dancer\_21 (DJ), Darby\_13  (AU1), Demure\_21 (DJ), DevitoJr\_13 (AU1), DizzyRudy\_14 (EL), Djungelskog\_13  (AW), DoctorPepper\_13 (AW), Doxi13\_9 (BI2), DrGrey\_14 (BI1), Duffington\_22 (DJ),  Eddiemania\_23 (DJ), Egad\_13 (AW), ElephantMan\_13 (AU1), EndAve\_22 (DJ),  Esketit\_14 (BI1), Fazel\_20 (DJ), FidgetOrca\_14 (BI1), Figliar\_22 (DJ), FortCran\_19  (DJ), FrodoSwaggins\_14 (BI1), FromIndy\_12 (BI4), GantcherGoblin\_12 (AU6),  Giantsbane\_15 (AU2), Gill\_19 (DJ), GirlPower\_13 (BI1), GoblinVoyage\_9 (BI2),  Gordon\_13 (AU1), Gravy\_19 (DJ), Gusicorn\_24 (DJ), HazuAndZazu\_14 (BI1),  HerbBucket\_12 (AW), Hortihabitatio\_23 (DJ), Hoshi\_14 (BI1), HotFries\_8 (BI2),  Hydrus\_23 (DJ), IceWarrior\_14 (BI1), Indigenous\_14 (BI1), Ingrid\_14 (AU3),  Inked\_15 (AU), Issa\_13 (AU1), Jaylociraptor\_14 (BI1), Jodelie19\_22 (DJ),  JulesRay\_21 (DJ), Kardashian\_12 (BI6), Kenosha\_21 (DJ), Kerry\_19 (DJ),  KevinMinion\_12 (AU6), Kinny\_16 (AU6), Leathea\_13 (AU6), Lewando\_16 (AU6),  LibertyBell\_12 (BI3), LilHuddy\_14 (AU2), Linda\_13 (AW), Loretta\_14 (AU3),  LucySwiss\_13 (AU1), Madamato\_14 (BI1), Madvan\_22 (DJ), Magel\_19 (DJ),  MajinBuu\_23 (DJ), Makai\_14 (AU5), Maya\_14 (BI1), MediumFry\_13 (AU4),  Michelle\_13 (AW), Miek\_11 (BI4), MiniBagel\_13 (AU1), Moozy\_9 (BI2), Mossy\_23  (DJ), MrAaronian\_13 (AW), Mysticmoon\_22 (DJ), Nadmeg\_22 (DJ), Namo\_14 (BI1),  Natasha\_13 (AW), Navi1117\_14 (AU6), Nightmare\_14 (AU1), Niktson\_13 (AU1),  Nithya\_22 (DJ), Nivinsha\_13 (AU1), NoPickles\_19 (DJ), Odesza\_19 (DJ),  OhMyWard\_22 (DJ), OlgasClover\_22 (DJ), OlympicHelado\_14 (BI1), Pepy6\_025  (CC), Perkunas\_21 (DJ), Petito\_20 (DJ), Phaby\_14 (AU2), Phaila\_12 (AU6),  Phepper\_22 (DJ), Pherobrine\_21 (DJ), PherryCruz\_8 (BI2), Poco6\_026 (CC),  Popy\_14 (BI1), Powelldog\_15 (AW), ProfFrink\_13 (AW), Rainydai\_11 (BI4),  Raunak\_13 (AW), RavenPuff\_8 (BI2), Regigigas\_12 (BI4), Renaldo\_18 (AU6),  Rickmore\_20 (DJ), Rima\_14 (BI1), Roney\_19 (DJ), RosaAsantewaa\_9 (BI2),  Runhaar\_22 (DJ), Salk\_13 (AW), Scap1\_9 (BI2), Schwartz33\_23 (DJ),  ScienceWizSam\_14 (AU1), Secretariat\_20 (DJ), SendItCS\_11 (BI4),  ShaggyRogers\_25 (DJ), SheRa\_9 (BI2), Shepard\_14 (AU2), Shiba\_13 (AW),  Sloopyjoe\_13 (AW), Soshi\_14 (BI1), Spectropatronm\_14 (BI1), Sporto\_14 (AW),  StarLord\_13 (AW), Stayer\_13 (AW), Synepsis\_12 (AU1), TaidaOne\_14 (BI1),  Talon44\_21 (DJ), Tanis\_19 (DJ), Tatanka\_12 (AU1), Teacup\_13 (AU1),  TenaciousP\_22 (DJ), Tenney120\_14 (AU6), Tenno\_14 (AU1), Thiqqums\_11 (BI4),  TieDye\_13 (BI1), Tokki\_14 (AU2), TonyStarch\_14 (BI1), TrixiePhattel\_13 (AU6),  Truckee\_13 (AU5), Trufflozitus\_22 (DJ), Trustiboi\_14 (AU1), Untouchable\_21 (DJ),  Uzumaki\_13 (AU6), Vardy\_22 (DJ), Zeina\_13 (AU6), Zeph\_23 (DJ), Zippen\_15  (AU3), | |
| Blast Results | |
| Function  High Match 99% Major Capsid Protein | |

|  |
| --- |
| Gene Information Stop 11208 |
| Start 10867 kept, originally called by glimmer and genemark. Coding potential captured, and 96% coverage with (Uzumaki), this was also shown by BLAST results. Starteraator agrees with the high number of manual annotations. HHpred suggested head completion protein at 58% which was insignificant. |
| Function- Hypothetical protein |
| Hypothetical protein to the left of tail terminator and other surrounding genes match |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 11208 |
| Glimmer | 10867 |
| GeneMark | 10867 |
| Length(s) | 342 |
| Overlap(s) | -18 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | Gap -18 |
| Best Z-score? | 1.896 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~10890 |
| Staterator Screenshot  <http://phages.wustl.edu/starterator/Pham716Report.pdf> | |
|  | |
|  | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**    **Hypothetical protein to the left of tail terminator and other surrounding genes match** |
| **BLASTp (Identified function)**  **96% aligned with Uzumaki for being a hypothetical protein (Function unknown)** |
| **HHpred**    **58% probability with head completion protein. Not likely** |
| **Conserved Domain: No data** |
| **DeepTMHMM: No data** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 11479 |
| Changed, originally called by glimmer at 11261 and not originally called by genemark. Moving start increased Z-Value and allowed coding potential to be captured. Moving starts also improved alignment (1:1 with Uzumaki). Starterator agrees with a high number of manual annotations for new start. HHpred shows alignment (98.38% probability) with Lambda Phage SPP1 5A21 Chain G as required by official function list. |
| Function- Tail terminator |
| -Same surrounding genes as Uzumaki and to the left of major tail protein. Also, tail terminators are usually on the “left” side of the genome |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction |  |
| Stop | 11479 |
| Glimmer | 11261 |
| GeneMark | Not Called on GeneMark |
| Length(s) | 219 |
| Overlap(s) | 35 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 36 (overlap) |
| Best Z-score? | 2.124 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~ |
| Staterator Screenshot | |
| [Pham172666Report.pdf](http://phages.wustl.edu/starterator/Pham172666Report.pdf) | |

|  |
| --- |
| BLAST results |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**    **-Same surrounding genes as Uzumaki and to the left of major tail protein** |
| **BLASTp (Identified function) 100 %aligned with Uzumaki with an identified function of tail terminator. (Usually on left side of genome)** |
| **HHpred** |
| **Conserved Domain: No data** |
| **DeepTMHMM:** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 12356 |
| Start 11490 kept, originally called by glimmer and genemark. Coding potential captured, and full coverage with another phage in cluster (Zenia). BLAST also showed 100% alignment. Starterator agrees with a high number of manual annotations. HHpred shows 99% alignment for tail tube protein, which appears to be a major tail protein in SEA-Phage naming guide. |
| Function- Major tail protein |
| -Similar surrounding genes as Zenia. (Found on left side of genome) |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 12356 |
| Glimmer | 11490 |
| GeneMark | 11490 |
| Length(s) | 867 |
| Overlap(s) | -10 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -10 (gap) |
| Best Z-score? | 2.400 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~11490 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham84731Report.pdf> | |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Similar surrounding genes as Zenia. (Found on left side of genome)** | | **BLASTp (Identified function)**  **100% aligned with Zenia with identified function of major tail protein** | | **HHpred**    **99.9% probability score with “phage tail tube protein” otherwise known as major tail protein** | | **Conserved Domain: No data** | | **DeepTMHMM** |  |  | | --- | | **Other Evidence?** | | |

|  |
| --- |
| Gene Information Stop 12951 |
| Start called at 12379 by both Glimmer and Genemark. Original call kept. Coding potential fully captured according to Gene Coding Report. Starterator shows large amount of manual annotations of gene in clustermates by other research teams. Blast Results match best with Lewando and Zeina. But no specific protein was identified. HHPred also shows no particular matches above 90% for this gene. |
| Function: Hypothetical Protein |
| Gene appears to have high occurrence across clustermates. With some clustermates like Lewando having high misalignment on the frame but similar genes. Uzumaki and Zeina have the gene in very similar locations as Trixie. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 12951 |
| Glimmer | 12379 |
| GeneMark | 12379 |
| Length(s) | 573 |
| Overlap(s) | -22 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -22 |
| Best Z-score? | 3, Z=3.286, 12 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~12370 |
| Staterator Screenshot | |
| [Starterator](http://phages.wustl.edu/starterator/Pham85043Report.pdf)  Start 32:  • Found in 216 of 216 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 166 of 167  • Called 99.1% of time when present  • Phage (with cluster) where this start called: Acai\_17 (AU1), AddiRose\_21 (JA),  Aflac\_26 (DJ), AlainaMarie\_26 (DJ), Annihilus\_13 (BI2), Arataki\_24 (DJ), Arcadia\_16  (AM), AreFloNak\_17 (AU1), Argan\_16 (AU6), Artorias\_23 (DJ), Arzan\_14 (FI),  Avazak\_23 (DJ), BarbieDoll\_20 (AU6), BenitoAntonio\_16 (AM), Benllo\_15 (AM),  BetterYeti\_28 (DJ), Bilo\_13 (BI2), Bing\_18 (BI5), Biscute\_21 (AU6), Boog\_16 (AW),  Bouchard\_18 (AU2), Bowling\_15 (AM), Breylor17\_18 (AU1), BronxBay\_17 (AW),  Brunswick\_17 (AU1), Burley\_26 (DJ), Camille\_18 (EL), CapnMurica\_17 (AU1),  Captrips\_23 (DJ), Capybara\_23 (DJ), CastorTray\_19 (AU1), Caterpillar\_17 (AU4),  Cheesy\_16 (AM), CherryBlossom\_18 (BI1), Chikenjars\_26 (DJ), Chlochlo\_18 (AU2),  Circum\_16 (AM), Citrus\_25 (DJ), ClubPenguin\_16 (BI7), Conley\_23 (DJ), Correa\_16  (AM), Count\_23 (EL), CricKo\_15 (BI4), Crocheter\_25 (DJ), Damascus\_18 (EL),  Dancer\_25 (DJ), Darby\_17 (AU1), Demure\_25 (DJ), DevitoJr\_17 (AU1),  DizzyRudy\_18 (EL), Djungelskog\_17 (AW), DoctorPepper\_17 (AW), Doxi13\_13 (BI2),  DrGrey\_18 (BI1), Duffington\_26 (DJ), Dynamite\_16 (AM), Eddiemania\_27 (DJ),  Eevee\_22 (JA), Egad\_17 (AW), ElephantMan\_17 (AU1), Elsa\_16 (AM), Elver\_31  (FK), EndAve\_26 (DJ), Esketit\_18 (BI1), Fazel\_24 (DJ), FidgetOrca\_18 (BI1),  Figliar\_26 (DJ), FortCran\_23 (DJ), FrodoSwaggins\_18 (BI1), FromIndy\_16 (BI4),  Gandionco\_33 (FK), GantcherGoblin\_16 (AU6), Giantsbane\_19 (AU2), Gill\_23 (DJ),  GirlPower\_17 (BI1), GoCrazy\_15 (AM), GoblinVoyage\_13 (BI2), Gordon\_17 (AU1),  Gravy\_23 (DJ), Hankly\_15 (AM), HazuAndZazu\_18 (BI1), Heisenberger\_16 (AM),  HerbBucket\_16 (AW), Hortihabitatio\_27 (DJ), Hoshi\_18 (BI1), HotFries\_12 (BI2),  Hydrus\_27 (DJ), IceWarrior\_18 (BI1), Indigenous\_18 (BI1), Ingrid\_18 (AU3),  Inked\_19 (AU), Issa\_17 (AU1), JEGGS\_16 (AM), Jaylociraptor\_18 (BI1),  Jazzy4900\_16 (FI), Jodelie19\_26 (DJ), JoyLin\_22 (JA), JulesRay\_25 (DJ),  Kardashian\_16 (BI6), Kardesai\_15 (AM), KeaneyLin\_15 (AM), Kenosha\_25 (DJ),  Kerry\_23 (DJ), KevinMinion\_16 (AU6), Kinny\_20 (AU6), Leathea\_17 (AU6),  Lewando\_20 (AU6), LibertyBell\_16 (BI3), LilHuddy\_18 (AU2), Linda\_17 (AW),  Loretta\_18 (AU3), LucySwiss\_17 (AU1), Madamato\_18 (BI1), Madvan\_26 (DJ),  Magel\_23 (DJ), MajinBuu\_27 (DJ), Makai\_18 (AU5), Maya\_18 (BI1), MediumFry\_17  (AU4), Meibysrarus\_18 (BI1), Michelle\_17 (AW), Miek\_15 (BI4), MiniBagel\_17 (AU1),  Mooshroom\_16 (AM), Moozy\_13 (BI2), Mossy\_27 (DJ), MrAaronian\_17 (AW),  Mudcat\_16 (AM), Mysticmoon\_26 (DJ), Nadmeg\_26 (DJ), Namo\_18 (BI1),  NapoleonB\_16 (AM), Nason\_16 (AM), Natasha\_17 (AW), Navi1117\_18 (AU6),  Nightmare\_18 (AU1), Niktson\_17 (AU1), Nithya\_26 (DJ), Nivinsha\_17 (AU1),  NoPickles\_23 (DJ), Odesza\_23 (DJ), OhMyWard\_26 (DJ), OlgasClover\_26 (DJ),  OlympicHelado\_18 (BI1), Paella\_32 (FK), Pepy6\_029 (CC), Perkunas\_25 (DJ),  Petito\_24 (DJ), Phaby\_18 (AU2), Phaila\_16 (AU6), Phepper\_26 (DJ), Pherobrine\_25  (DJ), PherryCruz\_12 (BI2), Pippin15\_21 (JA), Poco6\_030 (CC), Popy\_18 (BI1),  Powelldog\_19 (AW), ProfFrink\_17 (AW), Qui\_32 (FK), Rainydai\_15 (BI4), Raunak\_17  (AW), RavenPuff\_12 (BI2), Regigigas\_16 (BI4), Renaldo\_22 (AU6), Rickmore\_24  (DJ), Rima\_18 (BI1), Roney\_23 (DJ), RosaAsantewaa\_13 (BI2), Runhaar\_26 (DJ),  RustyBoy\_17 (AW), Salk\_17 (AW), Scap1\_13 (BI2), Schwartz33\_27 (DJ),  ScienceWizSam\_18 (AU1), Secretariat\_24 (DJ), SendItCS\_14 (BI4), Serenabean\_21  (JA), SheRa\_13 (BI2), Shepard\_18 (AU2), Shiba\_17 (AW), Sloopyjoe\_17 (AW),  Soshi\_18 (BI1), Spectropatronm\_18 (BI1), Sporto\_18 (AW), StarLord\_17 (AW),  Stayer\_17 (AW), Sunny4976\_15 (FI), Sweetclover\_20 (JA), Synepsis\_16 (AU1),  TaidaOne\_18 (BI1), Talon44\_25 (DJ), Tanis\_23 (DJ), Tatanka\_16 (AU1), Teacup\_17  (AU1), TenaciousP\_26 (DJ), Tenney120\_18 (AU6), Tenno\_18 (AU1), Thiqqums\_15  (BI4), TieDye\_17 (BI1), Tokki\_18 (AU2), TonyStarch\_18 (BI1), Tribby\_16 (AM),  TrixiePhattel\_17 (AU6), Truckee\_17 (AU5), Trufflozitus\_26 (DJ), Trustiboi\_18 (AU1),  Untouchable\_25 (DJ), Uzumaki\_17 (AU6), Vardy\_26 (DJ), Xenomorph\_15 (AM),  Yotsuba\_22 (JA), Zeina\_17 (AU6), Zeph\_27 (DJ), Zippen\_19 (AU3)  Gene: TrixiePhattel\_17 Start: 12379, Stop: 12951, Start Num: 32  Candidate Starts for TrixiePhattel\_17:  (15, 12217), (25, 12280), (Start: 32 @12379 has 166 MA's), (52, 12520), (55, 12562), (58, 12571), (60,  12577), (70, 12661), (77, 12751), (81, 12772) | |
| Blast Result | |
| Function  No results over 90%. Best Match | |

|  |
| --- |
| Gene Information Stop 14735 |
| Start 12969 kept, originally called by glimmer and genemark. Coding potential captured, and 100% alignment with GancherGoblin. BLAST shows 100% alignment. Starterator agrees with a high number of manual annotations. HHpred shows 98% probability of minor tail protein. |
| Function- Minor tail protein |
| **Similar genes surround GancherGoblin, and have other minor tail proteins surrounding.** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 14735 |
| Glimmer | 12969 |
| GeneMark | 12969 |
| Length(s) | 1767 |
| Overlap(s) | 17 (Gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 17 (Gap) |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~14760 |
| Staterator Screenshot  <http://phages.wustl.edu/starterator/Pham85662Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Similar genes surrounding GancherGoblin, and surrounded by other minor tail proteins.** |
| **BLASTp (Identified function)**  **100% aligned with Gantcher goblin** |
| **HHpred**    **98% probability with minor tail protein for bacteriophages** |
| **Conserved Domain** |
| **DeepTMHMM: No data** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 15741 |
| Start of 14746 called by both Genemark and Glimmer. Coding potential fully captured acrossing to Gene Coding Report. Starterator shows moderate amount of calls for this location. Blast Results show high match for minor tail protein in Uzumaki, Zeina, Lewando. HHPred shows high match for structural protein. |
| Function: Minor Tail Protein |
| Synteny suggests based on cluster mates and neighboring genes. That this is a Minor Tail Protein |
|  |

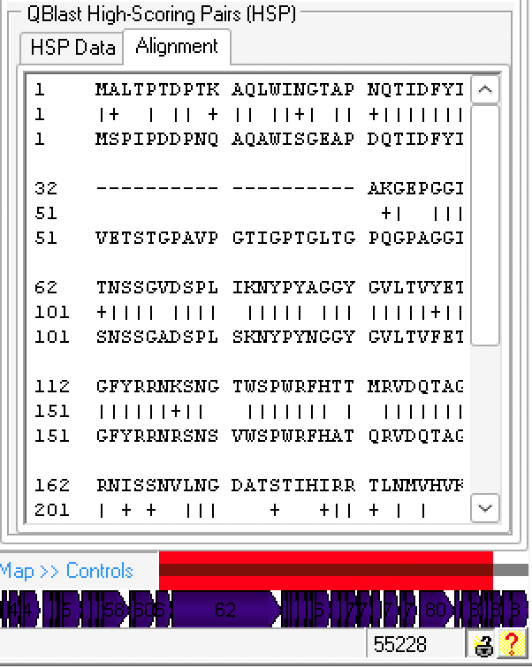
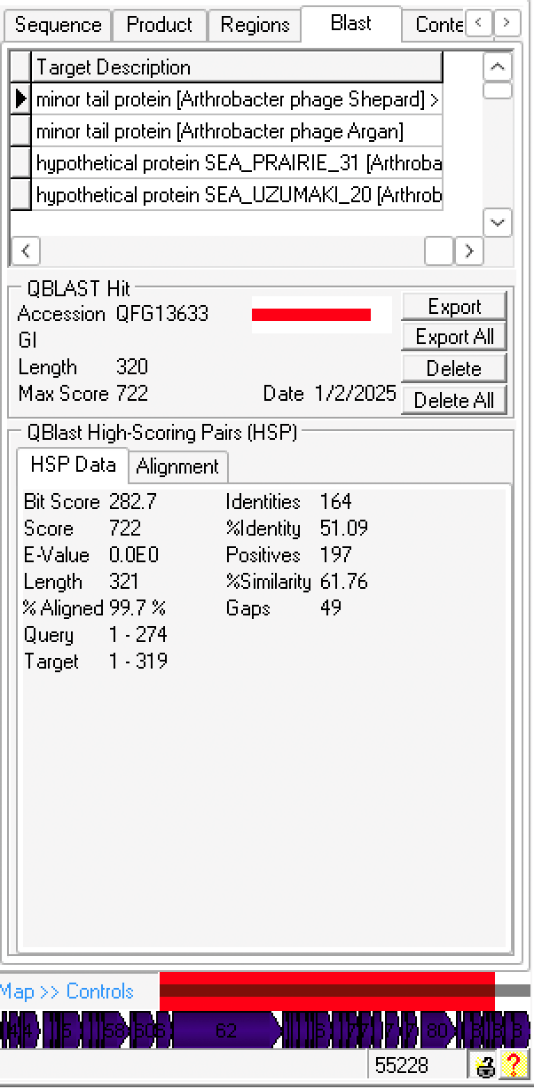
|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 15741 |
| Glimmer | 14746 |
| GeneMark | 14746 |
| Length(s) | 996 |
| Overlap(s) | -10 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -10 |
| Best Z-score? | 2.156 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~14750 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham85240Report.pdf>  Gene: TrixiePhattel\_19 Start: 14746, Stop: 15741, Start Num: 2  Candidate Starts for TrixiePhattel\_19:  (Start: 2 @14746 has 21 MA's), (6, 14815), (24, 14962), (159, 15403), (204, 15472), (230, 15529), (293, 15655),  Start 2:  • Found in 36 of 136 ( 26.5% ) of genes in pham  • Manual Annotations of this start: 21 of 99  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Acai\_22 (AU1), AreFloNak\_22 (AU1),  Argan\_18 (AU6), BarbieDoll\_22 (AU6), Biscute\_23 (AU6), Breylor17\_23 (AU1),  Brunswick\_22 (AU1), CapnMurica\_22 (AU1), Darby\_22 (AU1), DevitoJr\_22 (AU1),  ElephantMan\_22 (AU1), GantcherGoblin\_18 (AU6), Gordon\_22 (AU1), Issa\_22  (AU1), KevinMinion\_19 (AU6), Kinny\_22 (AU6), Leathea\_19 (AU6), Lewando\_22  (AU6), LucySwiss\_22 (AU1), MiniBagel\_22 (AU1), Navi1117\_20 (AU6),  Nightmare\_23 (AU1), Niktson\_22 (AU1), Nivinsha\_22 (AU1), Phaila\_18 (AU6),  Renaldo\_24 (AU6), ScienceWizSam\_23 (AU1), Synepsis\_21 (AU1), Tatanka\_21  (AU1), Teacup\_22 (AU1), Tenney120\_20 (AU6), Tenno\_23 (AU1), TrixiePhattel\_19  (AU6), Trustiboi\_23 (AU1), Uzumaki\_19 (AU6), Zeina\_19 (AU6), | |
| Blast Results | |
| Function:  Strong Match above 90% | |

|  |
| --- |
| Gene Information Stop 16582 |
| Start 15755 kept, called by glimmer and genemark.coding potential is captured, and full coverage is shown with Argan. High Z-Value is shown for this start. HHpred suggested a receptor binding protein, which could fit as a tail protein but will not be called. High glycine also suggests a role in tail function. Default call for minor tail protein. |
| Function- Minor tail protein |
| **Same gene as Argan, also to the right of another minor tail protein. Also before tape measure protein.** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 16582 |
| Glimmer | 15755 |
| GeneMark | 15755 |
| Length(s) | 828 |
| Overlap(s) | -13 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | #1 15755 Z= 3.297 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~15760 |
| Staterator Screenshot | |
| [Pham85240Report.pdf](http://phages.wustl.edu/starterator/Pham85240Report.pdf) | |



|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Same gene as argan, also to the right of another minor tail protein. Also before tape measure protein.** |
| **BLASTp (Identified function)**  **100% aligned with Argan for minor tail protein** |
| **HHpred**    **97% probability score with siphoviridae receptor binding protein** |
| **Conserved Domain** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 16723 |
| Start 1592 kept, called by glimmer and genemark. Coding potential captured, and full coverage with Zenia. BLAST also shows 100% alignment. Starterator agrees with a high number of manual annotations. HHpred shows low probabilities for other functions. |
| Function- Hypothetical protein |
| **Similar genes surrounding as Zenia, and to the left of tape measure protein** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 16723 |
| Glimmer | 16592 |
| GeneMark | 16592 |
| Length(s) | 132 |
| Overlap(s) | -9 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -9 (gap) |
| Best Z-score? | 2.156 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~16620 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham218020Report.pdf> | |

|  |
| --- |
|  |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **Similar genes with Zenia, and to the left of tape measure protein** |
| **BLASTp (Identified function)**  **100% aligned with Zenia Uzumaki, and Lewando for hypothetical protein** |
| **HHpred** |
| **Conserved Domain: No data** |
| **DeepTMHMM: No data** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 17555 |
| Start of 16758 called by both Genemark and Glimmer. Coding potential fully captured and original calls kept. Starterator large amount of manual annotations of gene in clustermates by other research teams. Blast Results shows high match for Major Tail Protein. HHPred shows a match for Tail Tube Protein, which Trixie\_Phattel is a siphoviridae. Suggesting this is a Major Tail Protein |
| Function: Major Tail Protein |
| Synteny shows gene presence in other Cluster Mates with offset along the reading frame. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 17555 |
| Glimmer | 16758 |
| GeneMark | 16758 |
| Length(s) | 789 |
| Overlap(s) | -34 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -34 |
| Best Z-score? | 1, Z=3.286, 9 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~16760 |
| Staterator Screenshot | |
| [Starterator](http://phages.wustl.edu/starterator/Pham84731Report.pdf)  Gene: TrixiePhattel\_22 Start: 16758, Stop: 17555, Start Num: 15  Candidate Starts for TrixiePhattel\_22:  (Start: 15 @16758 has 305 MA's), (40, 16914), (66, 17076), (83, 17187), (101, 17274), (102, 17277),(124, 17382),  Start 15:  • Found in 393 of 437 ( 89.9% ) of genes in pham  • Manual Annotations of this start: 305 of 337  • Called 99.5% of time when present  • Phage (with cluster) where this start called: Acai\_16 (AU1), AddiRose\_20 (JA),  AddiRose\_22 (JA), Aflac\_25 (DJ), Aflac\_27 (DJ), AlainaMarie\_25 (DJ),  AlainaMarie\_27 (DJ), Annihilus\_12 (BI2), Annihilus\_14 (BI2), Arataki\_23 (DJ),  Arataki\_25 (DJ), Arcadia\_15 (AM), Arcadia\_17 (AM), AreFloNak\_16 (AU1), Argan\_15  (AU6), Argan\_21 (AU6), Artorias\_22 (DJ), Artorias\_24 (DJ), Arzan\_13 (FI),  Avazak\_22 (DJ), Avazak\_24 (DJ), BarbieDoll\_19 (AU6), BarbieDoll\_25 (AU6),  BenitoAntonio\_15 (AM), BenitoAntonio\_17 (AM), Benllo\_14 (AM), Benllo\_16 (AM),  BetterYeti\_27 (DJ), BetterYeti\_29 (DJ), Bilo\_12 (BI2), Bilo\_14 (BI2), Bing\_17 (BI5),  Bing\_19 (BI5), Biscute\_20 (AU6), Biscute\_26 (AU6), BluerMoon\_25 (DJ),  BluerMoon\_27 (DJ), Boog\_15 (AW), Boog\_18 (AW), Bouchard\_17 (AU2), Bowling\_14  (AM), Bowling\_16 (AM), Breylor17\_17 (AU1), BronxBay\_16 (AW), BronxBay\_19  (AW), Brunswick\_16 (AU1), Burley\_25 (DJ), Burley\_27 (DJ), Camille\_17 (EL),  Camille\_20 (EL), CapnMurica\_16 (AU1), Captrips\_22 (DJ), Captrips\_24 (DJ),  Capybara\_22 (DJ), Capybara\_24 (DJ), CastorTray\_18 (AU1), Caterpillar\_16 (AU4),  Cheesy\_15 (AM), Cheesy\_17 (AM), CherryBlossom\_17 (BI1), CherryBlossom\_19  (BI1), Chikenjars\_25 (DJ), Chikenjars\_27 (DJ), Chlochlo\_17 (AU2), Circum\_15 (AM),  Circum\_17 (AM), Citrus\_24 (DJ), Citrus\_26 (DJ), ClubPenguin\_15 (BI7),  ClubPenguin\_17 (BI7), Conley\_22 (DJ), Conley\_24 (DJ), Correa\_15 (AM), Correa\_17  (AM), Count\_24 (EL), Count\_25 (EL), CricKo\_14 (BI4), CricKo\_16 (BI4),  Crocheter\_24 (DJ), Crocheter\_26 (DJ), Daddyjeff\_20 (JA), Daddyjeff\_22 (JA),  Damascus\_17 (EL), Damascus\_20 (EL), Dancer\_24 (DJ), Dancer\_26 (DJ), Darby\_16  (AU1), Demure\_24 (DJ), Demure\_26 (DJ), DevitoJr\_16 (AU1), DizzyRudy\_17 (EL),  DizzyRudy\_20 (EL), Djungelskog\_16 (AW), Djungelskog\_19 (AW), DoctorPepper\_16  (AW), DoctorPepper\_19 (AW), Doxi13\_12 (BI2), Doxi13\_14 (BI2), DrGrey\_17 (BI1),  DrGrey\_19 (BI1), Duffington\_25 (DJ), Duffington\_27 (DJ), Dynamite\_15 (AM),  Dynamite\_17 (AM), Eddiemania\_26 (DJ), Eddiemania\_28 (DJ), Eevee\_21 (JA),  Eevee\_23 (JA), Egad\_16 (AW), Egad\_19 (AW), ElephantMan\_16 (AU1), Elsa\_15  (AM), Elsa\_17 (AM), Elver\_30 (FK), Elver\_32 (FK), EndAve\_25 (DJ), EndAve\_27  (DJ), Esketit\_17 (BI1), Esketit\_19 (BI1), Fazel\_23 (DJ), Fazel\_25 (DJ), FidgetOrca\_17  (BI1), FidgetOrca\_19 (BI1), Figliar\_25 (DJ), Figliar\_27 (DJ), FortCran\_22 (DJ),  FortCran\_24 (DJ), FrodoSwaggins\_17 (BI1), FrodoSwaggins\_19 (BI1), FromIndy\_15  (BI4), FromIndy\_17 (BI4), Gandionco\_32 (FK), Gandionco\_34 (FK),  GantcherGoblin\_15 (AU6), GantcherGoblin\_21 (AU6), Giantsbane\_18 (AU2), Gill\_22  (DJ), GirlPower\_16 (BI1), GirlPower\_18 (BI1), GoCrazy\_14 (AM), GoCrazy\_16 (AM),  GoblinVoyage\_12 (BI2), GoblinVoyage\_14 (BI2), Gordon\_16 (AU1), Gravy\_22 (DJ),  Gravy\_24 (DJ), Gusicorn\_27 (DJ), Gusicorn\_30 (DJ), Hankly\_14 (AM), Hankly\_16  (AM), HazuAndZazu\_17 (BI1), HazuAndZazu\_19 (BI1), Heisenberger\_15 (AM),  Heisenberger\_17 (AM), HerbBucket\_15 (AW), HerbBucket\_18 (AW),  Hortihabitatio\_26 (DJ), Hortihabitatio\_28 (DJ), Hoshi\_17 (BI1), Hoshi\_19 (BI1),  HotFries\_11 (BI2), HotFries\_13 (BI2), Hydrus\_26 (DJ), Hydrus\_28 (DJ),  IceWarrior\_17 (BI1), IceWarrior\_19 (BI1), Indigenous\_17 (BI1), Indigenous\_19 (BI1),  Ingrid\_17 (AU3), Inked\_18 (AU), Issa\_16 (AU1), JEGGS\_15 (AM), JEGGS\_17 (AM),  Jaylociraptor\_17 (BI1), Jaylociraptor\_19 (BI1), Jazzy4900\_15 (FI), Jodelie19\_25 (DJ),  Jodelie19\_27 (DJ), JoyLin\_21 (JA), JoyLin\_23 (JA), JulesRay\_24 (DJ), JulesRay\_26  (DJ), Kardashian\_15 (BI6), Kardashian\_17 (BI6), Kardesai\_14 (AM), Kardesai\_16  (AM), KeaneyLin\_14 (AM), KeaneyLin\_16 (AM), Kenosha\_24 (DJ), Kenosha\_26 (DJ),  Kerry\_22 (DJ), Kerry\_24 (DJ), KevinMinion\_15 (AU6), KevinMinion\_22 (AU6),  Kinny\_19 (AU6), Kinny\_25 (AU6), Leathea\_16 (AU6), Leathea\_22 (AU6),  Lewando\_19 (AU6), Lewando\_25 (AU6), LibertyBell\_15 (BI3), LibertyBell\_17 (BI3),  LilHuddy\_17 (AU2), Linda\_16 (AW), Linda\_19 (AW), Loretta\_17 (AU3), LucySwiss\_16  (AU1), Madamato\_17 (BI1), Madamato\_19 (BI1), Madvan\_25 (DJ), Madvan\_27 (DJ),  Magel\_22 (DJ), MajinBuu\_26 (DJ), MajinBuu\_28 (DJ), Makai\_17 (AU5), Maya\_17  (BI1), Maya\_19 (BI1), MediumFry\_16 (AU4), Meibysrarus\_17 (BI1), Meibysrarus\_19  (BI1), Michelle\_16 (AW), Michelle\_19 (AW), Miek\_14 (BI4), Miek\_16 (BI4),  MiniBagel\_16 (AU1), Mooshroom\_15 (AM), Mooshroom\_17 (AM), Moozy\_12 (BI2),  Moozy\_14 (BI2), Mossy\_26 (DJ), Mossy\_28 (DJ), MrAaronian\_16 (AW),  MrAaronian\_19 (AW), Mudcat\_15 (AM), Mudcat\_17 (AM), Mysticmoon\_25 (DJ),  Mysticmoon\_27 (DJ), Nadmeg\_25 (DJ), Nadmeg\_27 (DJ), Namo\_17 (BI1), Namo\_19  (BI1), NapoleonB\_15 (AM), NapoleonB\_17 (AM), Nason\_15 (AM), Nason\_17 (AM),  Natasha\_16 (AW), Navi1117\_17 (AU6), Navi1117\_23 (AU6), Nightmare\_17 (AU1),  Niktson\_16 (AU1), Nithya\_25 (DJ), Nithya\_27 (DJ), Nivinsha\_16 (AU1), NoPickles\_22  (DJ), NoPickles\_24 (DJ), Odesza\_22 (DJ), Odesza\_24 (DJ), OhMyWard\_25 (DJ),  OhMyWard\_27 (DJ), OlgasClover\_25 (DJ), OlgasClover\_27 (DJ), OlympicHelado\_17  (BI1), OlympicHelado\_19 (BI1), Paella\_31 (FK), Paella\_33 (FK), Pepy6\_028 (CC),  Pepy6\_030 (CC), Perkunas\_24 (DJ), Perkunas\_26 (DJ), Petito\_23 (DJ), Petito\_25  (DJ), Phaby\_17 (AU2), Phaila\_15 (AU6), Phaila\_21 (AU6), Phepper\_25 (DJ),  Phepper\_27 (DJ), Pherobrine\_24 (DJ), Pherobrine\_26 (DJ), PherryCruz\_11 (BI2),  PherryCruz\_13 (BI2), Pippin15\_20 (JA), Pippin15\_22 (JA), Poco6\_029 (CC),  Poco6\_031 (CC), Popy\_17 (BI1), Popy\_19 (BI1), Powelldog\_18 (AW), ProfFrink\_16  (AW), ProfFrink\_19 (AW), Qui\_31 (FK), Qui\_33 (FK), Rainydai\_14 (BI4), Rainydai\_16  (BI4), Raunak\_16 (AW), Raunak\_19 (AW), RavenPuff\_11 (BI2), RavenPuff\_13 (BI2),  Regigigas\_15 (BI4), Regigigas\_17 (BI4), Renaldo\_21 (AU6), Renaldo\_27 (AU6),  Rickmore\_23 (DJ), Rickmore\_25 (DJ), Rima\_17 (BI1), Rima\_19 (BI1), Roney\_22  (DJ), Roney\_24 (DJ), RosaAsantewaa\_12 (BI2), RosaAsantewaa\_14 (BI2),  Runhaar\_25 (DJ), Runhaar\_27 (DJ), RustyBoy\_16 (AW), Salk\_16 (AW), Salk\_19  (AW), Scap1\_12 (BI2), Scap1\_14 (BI2), Schwartz33\_26 (DJ), Schwartz33\_28 (DJ),  ScienceWizSam\_17 (AU1), Secretariat\_23 (DJ), Secretariat\_25 (DJ), SendItCS\_13  (BI4), SendItCS\_15 (BI4), Serenabean\_20 (JA), Serenabean\_22 (JA),  ShaggyRogers\_28 (DJ), ShaggyRogers\_30 (DJ), SheRa\_12 (BI2), SheRa\_14 (BI2),  Shepard\_17 (AU2), Shiba\_16 (AW), Shiba\_19 (AW), Sloopyjoe\_16 (AW),  Sloopyjoe\_19 (AW), Soshi\_17 (BI1), Soshi\_19 (BI1), Spectropatronm\_17 (BI1),  Spectropatronm\_19 (BI1), Sporto\_17 (AW), Sporto\_20 (AW), StarLord\_16 (AW),  StarLord\_19 (AW), Stayer\_16 (AW), Stayer\_19 (AW), Sunny4976\_14 (FI),  Sweetclover\_19 (JA), Sweetclover\_21 (JA), Synepsis\_15 (AU1), TaidaOne\_17 (BI1),  TaidaOne\_19 (BI1), Talon44\_24 (DJ), Talon44\_26 (DJ), Tanis\_22 (DJ), Tanis\_24  (DJ), Tatanka\_15 (AU1), Teacup\_16 (AU1), TenaciousP\_25 (DJ), TenaciousP\_27  (DJ), Tenney120\_17 (AU6), Tenney120\_23 (AU6), Tenno\_17 (AU1), Thiqqums\_14  (BI4), Thiqqums\_16 (BI4), TieDye\_16 (BI1), TieDye\_18 (BI1), Tokki\_17 (AU2),  TonyStarch\_17 (BI1), TonyStarch\_19 (BI1), Tribby\_15 (AM), Tribby\_17 (AM),  TrixiePhattel\_16 (AU6), TrixiePhattel\_22 (AU6), Truckee\_16 (AU5), Trufflozitus\_25  (DJ), Trufflozitus\_27 (DJ), Trustiboi\_17 (AU1), Untouchable\_24 (DJ), Untouchable\_26  (DJ), Uzumaki\_16 (AU6), Uzumaki\_22 (AU6), Vardy\_25 (DJ), Vardy\_27 (DJ),  Xenomorph\_14 (AM), Xenomorph\_16 (AM), Yotsuba\_21 (JA), Yotsuba\_23 (JA),  Zeina\_16 (AU6), Zeina\_22 (AU6), Zeph\_26 (DJ), Zeph\_28 (DJ), Zippen\_18 (AU3), | |
| Blast Results | |
| Function  High Match Above 90% | |
| Gene Information Stop 18491 | |
| Start 17568 kept, called by glimmer and genemark coding potential captured, and full coverage with GancherGoblin. BLAST also shows 100% alignment. High Z-value and agreement from starterator with a good number of manual annotations. HHpred showed results for protein with unknown function. | |
| Function- Hypothetical protein | |
| Next to similar genes as Ganchergoblin and near tape measure protein | |
|  | |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 18491 |
| Glimmer | 17568 |
| GeneMark | 17568 |
| Length(s) | 924 |
| Overlap(s) | 12 (Gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 12 (gap) |
| Best Z-score? | 3.201 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~17590 |
| Staterator Screenshot        <http://phages.wustl.edu/starterator/Pham212650Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **Next to similar genes as Ganchergoblin and near tape measure protein** |
| **BLASTp (Identified function)**  **100% aligned with Ganchergoblin for hypothetical protein** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM: No data** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 18802 |
| Start 18491 kept, called by genemark and glimmer. Coding potential captured, and full coverage with phage of the same cluster (Zeina). BLAST also shows 100% alignment. High Z-value and high number of manual annotations from Starterator agrees. HHpred shows Tail fiber chaperone with 97% but may not meet requirements. |
| Function-Hypothetical protein |
| **Similar surrounding genes as Zeina, to the right of minor tail proteins, and to the left of tape measure protein** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 18802 |
| Glimmer | 18491 |
| GeneMark | 18491 |
| Length(s) | 312 |
| Overlap(s) | 1 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 2.722 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~18495 |
| Staterator Screenshot | |
|  | |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Similar surrounding genes as Zinea, to the right of minor tail proteins, and to the left of tape measure protein** |
| **BLASTp (Identified function)**  **100% aligned with Zeina** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 19207 |
| Gene begins at start of 18824. Uncalled by Genemark or Glimmer. Coding potential captured according to gene coding report. Starterator possess manual annotations matching our gene from clustermates with a high percentage of annotations. Blast Results shows match with Lewando. HHPred also suggests high probability of gene presence. |
| Function: Hypothetical Protein |
| Clustermates possess this gene however alignment varies between clustermates but the gene remains similar throughout. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 19207 |
| Glimmer | NA |
| GeneMark | NA |
| Length(s) | 384 |
| Overlap(s) | -9 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 1, Z=2.357, 10 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~18850 |
| Staterator Screenshot | |
| [Starterator](http://phages.wustl.edu/starterator/Pham85042Report.pdf)  Gene: TrixiePhattel\_25 Start: 18812, Stop: 19207, Start Num: 39  Candidate Starts for TrixiePhattel\_25:  (39, 18812), (Start: 42 @18824 has 116 MA's), (47, 18866), (62, 19004), (72, 19082), (80, 19190), (81,  19193),  Start 42:  • Found in 161 of 218 ( 73.9% ) of genes in pham  • Manual Annotations of this start: 118 of 169  • Called 97.5% of time when present  • Phage (with cluster) where this start called: Acai\_25 (AU1), Aflac\_31 (DJ),  AlainaMarie\_31 (DJ), Arataki\_29 (DJ), Arcadia\_21 (AM), AreFloNak\_25 (AU1),  Argan\_24 (AU6), Artorias\_28 (DJ), Arzan\_21 (FI), Avazak\_28 (DJ), BenitoAntonio\_21  (AM), Benllo\_19 (AM), BetterYeti\_33 (DJ), Biscute\_29 (AU6), BluerMoon\_31 (DJ),  Boog\_21 (AW), Bouchard\_26 (AU2), Bowling\_20 (AM), Breylor17\_26 (AU1),  BronxBay\_22 (AW), Brunswick\_25 (AU1), Burley\_31 (DJ), Camille\_23 (EL),  CapnMurica\_25 (AU1), Captrips\_28 (DJ), Capybara\_28 (DJ), CastorTray\_27 (AU1),  Caterpillar\_23 (AU4), Cheesy\_21 (AM), Chikenjars\_31 (DJ), Chlochlo\_26 (AU2),  Circum\_21 (AM), Citrus\_30 (DJ), Conley\_29 (DJ), Correa\_20 (AM), Crocheter\_31  (DJ), Damascus\_23 (EL), Dancer\_30 (DJ), Darby\_25 (AU1), Demure\_30 (DJ),  DevitoJr\_25 (AU1), DizzyRudy\_23 (EL), Djungelskog\_22 (AW), DoctorPepper\_22  (AW), Duffington\_31 (DJ), Dynamite\_20 (AM), Eddiemania\_32 (DJ), Egad\_22 (AW),  ElephantMan\_25 (AU1), Elsa\_21 (AM), Elver\_38 (FK), EndAve\_31 (DJ), Fazel\_29  (DJ), Figliar\_31 (DJ), FortCran\_28 (DJ), Gandionco\_40 (FK), GantcherGoblin\_24  (AU6), Giantsbane\_26 (AU2), Gill\_28 (DJ), GoCrazy\_19 (AM), Gordon\_25 (AU1),  Gravy\_28 (DJ), Gusicorn\_34 (DJ), Hankly\_19 (AM), Heisenberger\_20 (AM),  HerbBucket\_21 (AW), Hortihabitatio\_32 (DJ), Hydrus\_32 (DJ), Ingrid\_25 (AU3),  Inked\_27 (AU), Issa\_25 (AU1), JEGGS\_20 (AM), Jazzy4900\_23 (FI), Jodelie19\_32  (DJ), JulesRay\_30 (DJ), Kardesai\_19 (AM), KeaneyLin\_19 (AM), Kenosha\_31 (DJ),  Kerry\_28 (DJ), Kinny\_28 (AU6), Leathea\_25 (AU6), Lewando\_28 (AU6), LilHuddy\_26  (AU2), Linda\_22 (AW), Loretta\_25 (AU3), LucySwiss\_25 (AU1), Madvan\_31 (DJ),  Magel\_28 (DJ), MajinBuu\_32 (DJ), Makai\_25 (AU5), MediumFry\_23 (AU4),  Michelle\_22 (AW), MiniBagel\_25 (AU1), Mooshroom\_20 (AM), Mossy\_32 (DJ),  MrAaronian\_22 (AW), Mudcat\_20 (AM), Nadmeg\_32 (DJ), NapoleonB\_20 (AM),  Nason\_21 (AM), Natasha\_21 (AW), Navi1117\_26 (AU6), Nightmare\_26 (AU1),  Niktson\_25 (AU1), Nithya\_31 (DJ), Nivinsha\_25 (AU1), NoPickles\_28 (DJ),  Odesza\_28 (DJ), OhMyWard\_31 (DJ), OlgasClover\_31 (DJ), Paella\_39 (FK),  Perkunas\_30 (DJ), Petito\_30 (DJ), Phaby\_26 (AU2), Phaila\_24 (AU6), Phepper\_31  (DJ), Pherobrine\_30 (DJ), ProfFrink\_22 (AW), Qui\_39 (FK), Raunak\_22 (AW),  Renaldo\_30 (AU6), Rickmore\_30 (DJ), Roney\_28 (DJ), Runhaar\_31 (DJ),  RustyBoy\_21 (AW), Salk\_22 (AW), Schwartz33\_32 (DJ), ScienceWizSam\_27 (AU1),  Secretariat\_29 (DJ), ShaggyRogers\_34 (DJ), Shepard\_26 (AU2), Shiba\_22 (AW),  Sloopyjoe\_22 (AW), Sporto\_23 (AW), StarLord\_22 (AW), Stayer\_22 (AW),  Sunny4976\_22 (FI), Synepsis\_24 (AU1), Talon44\_30 (DJ), Tanis\_28 (DJ),  Tatanka\_25 (AU1), Teacup\_25 (AU1), TenaciousP\_32 (DJ), Tenney120\_26 (AU6),  Tenno\_26 (AU1), Tokki\_26 (AU2), Tribby\_21 (AM), Truckee\_24 (AU5),  Trufflozitus\_31 (DJ), Trustiboi\_26 (AU1), Untouchable\_31 (DJ), Uzumaki\_25 (AU6),  Vardy\_31 (DJ), Xenomorph\_19 (AM), Zeina\_25 (AU6), Zeph\_32 (DJ), Zippen\_25  (AU3), | |
| BLAST DATA | |
| Function  High Match above 90% | |

|  |
| --- |
| Gene Information Stop 23763 |
| Start 19204 changed, Original glimmer call was 19423, and genemark call was at 19423. Coding potential was captured with a new start, Created better overlap.High amount of manual annotations in starterator HHpred shows 99.9% probability score with *Staphylococcus* tape measure protein |
|  |
| Function-Tape measure protein |
| **Same arrangement of genes with Zeina, and similarity for tape measure protein** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 23763 |
| Glimmer | 19423 |
| GeneMark | 19183 |
| Length(s) | 4341 |
| Overlap(s) | 4 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | #2 Z=2.622, 19204 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~19215 |
|  |  |
| Staterator Screenshot | |
| Gene: TrixiePhattel\_27 Start: 19423, Stop: 23763, Start Num: 47 Candidate Starts for TrixiePhattel\_27: (Start: 16 @19183 has 2 MA's), (Start: 17 @19204 has 129 MA's), (Start: 20 @19234 has 3 MA's), (28, 19288), (32, 19309), (45, 19402), (47, 19423), (58, 19489), (59, 19492), (67, 19549), (68, 19570), (74, 19630), (75, 19672), (80, 19732), (89, 19858), (92, 19873), (95, 19888), (101, 19921), (112, 19999), (114, 20014), (120, 20062), (135, 20215), (140, 20260), (145, 20281), (148, 20314), (158, 20389), (161, 20398), (166, 20440), (169, 20458), (180, 20500), (189, 20560), (190, 20566), (193, 20575), (198, 20590), (199, 20599), (220, 20698), (221, 20710), (222, 20719), (224, 20731), (294, 21016), (308, 21076), (312, 21085), (333, 21205), (336, 21226), (347, 21298), (349, 21304), (358, 21349), (361, 21358), (367, 21382), (371, 21400), (382, 21448), (385, 21457), (398, 21538), (406, 21565), (410, 21577), (413, 21586), (417, 21607), (433, 21688), (457, 21799), (460, 21811), (461, 21814), (477, 21898), (480, 21910), (498, 21979), (507, 22018), (511, 22039), (519, 22075), (528, 22108), (530, 22117), (536, 22129), (537, 22132), (557, 22201), (558, 22210), (562, 22240), (574, 22279), (577, 22288), (586, 22318), (592, 22339), (604, 22378), (627, 22465), (631, 22480), (635, 22498), (638, 22513), (641, 22522), (643, 22528), (654, 22567), (660, 22600), (661, 22603), (667, 22624), (668, 22627), (687, 22714), (696, 22744), (697, 22747), (702, 22765), (708, 22789), (714, 22810), (724, 22864), (725, 22867), (740, 22936), (747, 22990), (772, 23119), (778, 23164), (806, 23344), (810, 23368), (816, 23392), (827, 23437), (830, 23443), (837, 23482), (840, 23512), (845, 23530), (877, 23650),    Start 17: • Found in 180 of 218 ( 82.6% ) of genes in pham • Manual Annotations of this start: 129 of 167 • Called 92.8% of time when present • Phage (with cluster) where this start called: Acai\_26 (AU1), AddiRose\_25 (JA), Aflac\_32 (DJ), AlainaMarie\_32 (DJ), Annihilus\_19 (BI2), Arataki\_30 (DJ), Arcadia\_22 (AM), AreFloNak\_26 (AU1), Argan\_25 (AU6), Artorias\_29 (DJ), Avazak\_29 (DJ), BenitoAntonio\_22 (AM), Benllo\_20 (AM), BetterYeti\_34 (DJ), Bilo\_19 (BI2), Biscute\_30 (AU6), BluerMoon\_35 (DJ), Bouchard\_27 (AU2), Bowling\_21 (AM), Breylor17\_27 (AU1), BronxBay\_23 (AW), Brunswick\_26 (AU1), Burley\_32 (DJ), Camille\_24 (EL), CapnMurica\_26 (AU1), Captrips\_29 (DJ), Capybara\_29 (DJ), CastorTray\_28 (AU1), Caterpillar\_24 (AU4), Cheesy\_22 (AM), Chikenjars\_32 (DJ), Chlochlo\_27 (AU2), Circum\_22 (AM), Citrus\_31 (DJ), Conley\_30 (DJ), Correa\_21 (AM), Count\_30 (EL), Crocheter\_32 (DJ), Daddyjeff\_25 (JA), Damascus\_24 (EL), Dancer\_31 (DJ), Darby\_26 (AU1), Demure\_31 (DJ), DevitoJr\_26 (AU1), DizzyRudy\_24 (EL), Djungelskog\_23 (AW), DoctorPepper\_23 (AW), Doxi13\_19 (BI2), Duffington\_32 (DJ), Dynamite\_21 (AM), Eddiemania\_33 (DJ), Eevee\_26 (JA), ElephantMan\_26 (AU1), Elsa\_22 (AM), Elver\_39 (FK), EndAve\_32 (DJ), Fazel\_30 (DJ), Figliar\_32 (DJ), FortCran\_29 (DJ), GantcherGoblin\_25 (AU6), Giantsbane\_27 (AU2), Gill\_29 (DJ), GoCrazy\_20 (AM), GoblinVoyage\_19 (BI2), Gordon\_26 (AU1), Gravy\_29 (DJ), Gusicorn\_35 (DJ), Hankly\_20 (AM), Heisenberger\_21 (AM), HerbBucket\_22 (AW), Hortihabitatio\_33 (DJ), HotFries\_18 (BI2), Hydrus\_33 (DJ), Ingrid\_26 (AU3), Inked\_28 (AU), Issa\_26 (AU1), JEGGS\_21 (AM), Jodelie19\_33 (DJ), JoyLin\_26 (JA), JulesRay\_31 (DJ), Kardesai\_20 (AM), KeaneyLin\_20 (AM), Kenosha\_32 (DJ), Kerry\_29 (DJ), Kinny\_29 (AU6), Leathea\_26 (AU6), Lewando\_29 (AU6), LilHuddy\_27 (AU2), Linda\_23 (AW), Loretta\_26 (AU3), LucySwiss\_26 (AU1), Madvan\_32 (DJ), MajinBuu\_33 (DJ), Makai\_26 (AU5), MediumFry\_24 (AU4), Michelle\_23 (AW), MiniBagel\_26 (AU1), Mooshroom\_21 (AM), Mossy\_33 (DJ), MrAaronian\_23 (AW), Mudcat\_21 (AM), Mysticmoon\_32 (DJ), Nadmeg\_33 (DJ), NapoleonB\_21 (AM), Nason\_22 (AM), Navi1117\_27 (AU6), Nightmare\_27 (AU1), Niktson\_26 (AU1), Nithya\_32 (DJ), Nivinsha\_26 (AU1), NoPickles\_29 (DJ), Odesza\_29 (DJ), OhMyWard\_32 (DJ), OlgasClover\_32 (DJ), Paella\_40 (FK), Pepy6\_035 (CC), Perkunas\_31 (DJ), Petito\_31 (DJ), Phaby\_27 (AU2), Phaila\_25 (AU6), Phepper\_32 (DJ), Pherobrine\_31 (DJ), PherryCruz\_18 (BI2), Pippin15\_25 (JA), Poco6\_036 (CC), Powelldog\_24 (AW), ProfFrink\_23 (AW), Qui\_40 (FK), RavenPuff\_18 (BI2), Rickmore\_31 (DJ), Roney\_29 (DJ), RosaAsantewaa\_19 (BI2), Runhaar\_32 (DJ), RustyBoy\_22 (AW), Salk\_23 (AW), Schwartz33\_33 (DJ), ScienceWizSam\_28 (AU1), Secretariat\_30 (DJ), Serenabean\_25 (JA), ShaggyRogers\_35 (DJ), SheRa\_19 (BI2), Shepard\_27 (AU2), Sloopyjoe\_23 (AW), Sporto\_24 (AW), StarLord\_23 (AW), Stayer\_23 (AW), Sweetclover\_24 (JA), Synepsis\_25 (AU1), Talon44\_31 (DJ), Tanis\_29 (DJ), Tatanka\_26 (AU1), Teacup\_26 (AU1), TenaciousP\_33 (DJ), Tenney120\_27 (AU6), Tenno\_27 (AU1), Tokki\_27 (AU2), Tribby\_22 (AM), Truckee\_25 (AU5), Trufflozitus\_32 (DJ), Trustiboi\_27 (AU1), Untouchable\_32 (DJ), Uzumaki\_26 (AU6), Vardy\_32 (DJ), Xenomorph\_20 (AM), Zeina\_26 (AU6), Zeph\_33 (DJ), Zippen\_26 (AU3)  [Pham210816Report.pdf](http://phages.wustl.edu/starterator/Pham210816Report.pdf) | |

|  |
| --- |
| **Function** |
| **Context & Synteny**    **Same arrangement of genes with Zeina, and similarity for tape measure protein** |
| **BLASTp (Identified function)** |
| **HHpred**    **99.9% probability score with *Staphylococcus* tape measure protein** |
| **Conserved Domain**    **Conserved domain with tape measure protein** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 24583 |
| Start called at 23747 by both Glimmer and Genemark. Original call kept due to full capture of coding potential. Starterator also has a large amount of calls on a similar position for this gene. Blast results show match for minor tail protein on Zeina and Uzumaki. HHPred suggests Tail Protein at high probability |
| Function: Minor Tail Protein |
| Synteny with cluster mates suggests based on positioning and pham that this gene is a Minor Tail Protein. With virus being siphovirirdae it suggests Minor Tail Protein |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 24583 |
| Glimmer | 23747 |
| GeneMark | 23747 |
| Length(s) | 837 |
| Overlap(s) | 17 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 17 |
| Best Z-score? | 1, Z=2.545, 14 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~23790 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham219403Report.pdf>  Gene: TrixiePhattel\_28 Start: 23747, Stop: 24583, Start Num: 2  Candidate Starts for TrixiePhattel\_28:  (Start: 2 @23747 has 122 MA's), (8, 23786), (17, 23816), (18, 23825), (34, 23954), (41, 23990), (45,24017), (50, 24038), (58, 24068), (64, 24086), (68, 24107), (123, 24338), (124, 24341), (139, 24428),(154, 24512), (156, 24518),  Start 2:  • Found in 157 of 218 ( 72.0% ) of genes in pham  • Manual Annotations of this start: 122 of 168  • Called 95.5% of time when present  • Phage (with cluster) where this start called: Acai\_27 (AU1), AddiRose\_26 (JA),  Annihilus\_20 (BI2), Arcadia\_23 (AM), AreFloNak\_27 (AU1), Argan\_26 (AU6),  BenitoAntonio\_23 (AM), Benllo\_21 (AM), Bilo\_20 (BI2), Bing\_25 (BI5), Biscute\_31  (AU6), Boog\_23 (AW), Bouchard\_28 (AU2), Bowling\_22 (AM), Breylor17\_28 (AU1),  BronxBay\_24 (AW), Brunswick\_27 (AU1), Camille\_25 (EL), CapnMurica\_27 (AU1),  CastorTray\_29 (AU1), Caterpillar\_25 (AU4), Cheesy\_23 (AM), CherryBlossom\_24  (BI1), Chlochlo\_28 (AU2), Circum\_23 (AM), ClubPenguin\_23 (BI7), Correa\_22 (AM),  Count\_31 (EL), CricKo\_22 (BI4), Daddyjeff\_26 (JA), Damascus\_25 (EL), Darby\_27  (AU1), DevitoJr\_27 (AU1), DizzyRudy\_25 (EL), Djungelskog\_24 (AW),  DoctorPepper\_24 (AW), Doxi13\_20 (BI2), DrGrey\_24 (BI1), Dynamite\_22 (AM),  Eevee\_27 (JA), Egad\_24 (AW), ElephantMan\_27 (AU1), Elsa\_23 (AM), Elver\_40  (FK), Esketit\_24 (BI1), FidgetOrca\_24 (BI1), FrodoSwaggins\_24 (BI1), FromIndy\_23  (BI4), GantcherGoblin\_26 (AU6), Giantsbane\_28 (AU2), GirlPower\_23 (BI1),  GoCrazy\_21 (AM), GoblinVoyage\_20 (BI2), Gordon\_27 (AU1), Hankly\_21 (AM),  HazuAndZazu\_24 (BI1), Heisenberger\_22 (AM), HerbBucket\_23 (AW), Hoshi\_24  (BI1), HotFries\_19 (BI2), IceWarrior\_24 (BI1), Indigenous\_24 (BI1), Ingrid\_27 (AU3),  Inked\_29 (AU), Issa\_27 (AU1), JEGGS\_22 (AM), Jaylociraptor\_24 (BI1),  Jazzy4900\_25 (FI), JoyLin\_27 (JA), Kardashian\_23 (BI6), Kardesai\_21 (AM),  KeaneyLin\_21 (AM), KevinMinion\_27 (AU6), Leathea\_27 (AU6), Lewando\_30 (AU6),  LibertyBell\_22 (BI3), LilHuddy\_28 (AU2), Linda\_24 (AW), Loretta\_27 (AU3),  LucySwiss\_27 (AU1), Madamato\_24 (BI1), Makai\_27 (AU5), Maya\_24 (BI1),  MediumFry\_25 (AU4), Meibysrarus\_24 (BI1), Michelle\_24 (AW), Miek\_22 (BI4),  MiniBagel\_27 (AU1), Mooshroom\_22 (AM), Moozy\_20 (BI2), MrAaronian\_24 (AW),  Mudcat\_22 (AM), Namo\_24 (BI1), NapoleonB\_22 (AM), Nason\_23 (AM), Natasha\_23  (AW), Navi1117\_28 (AU6), Nightmare\_28 (AU1), Niktson\_27 (AU1), Nivinsha\_27  (AU1), OlympicHelado\_24 (BI1), Paella\_41 (FK), Phaby\_28 (AU2), Phaila\_26 (AU6),  PherryCruz\_19 (BI2), Pippin15\_26 (JA), ProfFrink\_24 (AW), Qui\_41 (FK),  Rainydai\_22 (BI4), Raunak\_24 (AW), RavenPuff\_19 (BI2), Regigigas\_23 (BI4),  Rima\_24 (BI1), RosaAsantewaa\_20 (BI2), RustyBoy\_23 (AW), Salk\_24 (AW),  Scap1\_20 (BI2), ScienceWizSam\_29 (AU1), SendItCS\_21 (BI4), Serenabean\_26  (JA), SheRa\_20 (BI2), Shepard\_28 (AU2), Shiba\_24 (AW), Sloopyjoe\_24 (AW),  Soshi\_24 (BI1), Spectropatronm\_24 (BI1), Sporto\_25 (AW), StarLord\_24 (AW),  Stayer\_24 (AW), Sunny4976\_24 (FI), Sweetclover\_25 (JA), Synepsis\_26 (AU1),  TaidaOne\_24 (BI1), Tatanka\_27 (AU1), Teacup\_27 (AU1), Tenney120\_28 (AU6),  Tenno\_28 (AU1), Thiqqums\_22 (BI4), TieDye\_23 (BI1), Tokki\_28 (AU2),  TonyStarch\_24 (BI1), Tribby\_23 (AM), TrixiePhattel\_28 (AU6), Truckee\_26 (AU5),  Trustiboi\_28 (AU1), Uzumaki\_27 (AU6), Xenomorph\_21 (AM), Yotsuba\_27 (JA),  Zeina\_27 (AU6), Zippen\_27 (AU3), | |
| Blast Results | |
| Function  High Match Above 90% | |

|  |
| --- |
| Gene Information Stop 25747 |
| Gene called at 25747. Start called at 24770 by Glimmer and Genemark. Original called changed to Start 24584. Coding potential fully captured according to gene coding report. Starterator shows large amounts of manual annotations for this gene and start. Blast results show match with uzumaki for Minor Tail Protein. HHPred suggests high match with Minor Tail Protein for this gene. |
| Function: minor tail protein |
| Synteny shows region for tail proteins. Many Clustermates also have a similar gene but offset on the reading frame from the others. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 25747 |
| Glimmer | 24770 |
| GeneMark | 24770 |
| Length(s) | 978 |
| Overlap(s) | -186 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 0 |
| Best Z-score? | 1, Z=2.478, 12 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~24620 |
| Staterator Screenshot | |
| [Starterator](http://phages.wustl.edu/starterator/Pham212289Report.pdf)  Start 12: • Found in 136 of 218 ( 62.4% ) of genes in pham • Manual Annotations of this start: 95 of 167 • Called 92.6% of time when present • Phage (with cluster) where this start called: Acai\_28 (AU1), AddiRose\_27 (JA), Annihilus\_21 (BI2), AreFloNak\_28 (AU1), Argan\_27 (AU6), Arzan\_24 (FI), BarbieDoll\_31 (AU6), Bilo\_21 (BI2), Bing\_26 (BI5), Biscute\_32 (AU6), Boog\_24 (AW), Bouchard\_29 (AU2), Breylor17\_29 (AU1), BronxBay\_25 (AW), Brunswick\_28 (AU1), CapnMurica\_28 (AU1), CastorTray\_30 (AU1), Caterpillar\_26 (AU4), CherryBlossom\_25 (BI1), Chlochlo\_29 (AU2), ClubPenguin\_24 (BI7), Count\_32 (EL), CricKo\_23 (BI4), Daddyjeff\_27 (JA), Damascus\_26 (EL), Darby\_28 (AU1), DevitoJr\_28 (AU1), DizzyRudy\_26 (EL), Djungelskog\_25 (AW), DoctorPepper\_25 (AW), Doxi13\_21 (BI2), DrGrey\_25 (BI1), Eevee\_28 (JA), Egad\_25 (AW), Elver\_41 (FK), FidgetOrca\_25 (BI1), FrodoSwaggins\_25 (BI1), FromIndy\_24 (BI4), Gandionco\_43 (FK), GantcherGoblin\_27 (AU6), Giantsbane\_29 (AU2), GirlPower\_24 (BI1), GoblinVoyage\_21 (BI2), Gordon\_28 (AU1), HazuAndZazu\_25 (BI1), HerbBucket\_24 (AW), Hoshi\_25 (BI1), HotFries\_20 (BI2), IceWarrior\_25 (BI1), Indigenous\_25 (BI1), Ingrid\_28 (AU3), Inked\_30 (AU), Issa\_28 (AU1), Jaylociraptor\_25 (BI1), Jazzy4900\_26 (FI), JoyLin\_28 (JA), Kardashian\_24 (BI6), KevinMinion\_28 (AU6), Kinny\_31 (AU6), Leathea\_28 (AU6), Lewando\_31 (AU6), LibertyBell\_23 (BI3), LilHuddy\_29 (AU2), Linda\_25 (AW), Loretta\_28 (AU3), LucySwiss\_28 (AU1), Madamato\_25 (BI1), Makai\_28 (AU5), Maya\_25 (BI1), MediumFry\_26 (AU4), Meibysrarus\_25 (BI1), Michelle\_25 (AW), Miek\_23 (BI4), MiniBagel\_28 (AU1), Moozy\_21 (BI2), MrAaronian\_25 (AW), Namo\_25 (BI1), Natasha\_24 (AW), Navi1117\_29 (AU6), Nightmare\_29 (AU1), Nivinsha\_28 (AU1), Paella\_42 (FK), Phaby\_29 (AU2), Phaila\_27 (AU6), Pippin15\_27 (JA), Popy\_25 (BI1), Powelldog\_26 (AW), ProfFrink\_25 (AW), Qui\_42 (FK), Rainydai\_23 (BI4), Raunak\_25 (AW), RavenPuff\_20 (BI2), Regigigas\_24 (BI4), Renaldo\_33 (AU6), RosaAsantewaa\_21 (BI2), RustyBoy\_24 (AW), Salk\_25 (AW), Scap1\_21 (BI2), ScienceWizSam\_30 (AU1), SendItCS\_22 (BI4), Serenabean\_27 (JA), SheRa\_21 (BI2), Shepard\_29 (AU2), Shiba\_25 (AW), Sloopyjoe\_25 (AW), Soshi\_25 (BI1), Spectropatronm\_25 (BI1), Sporto\_26 (AW), StarLord\_25 (AW), Stayer\_25 (AW), Sunny4976\_25 (FI), Sweetclover\_26 (JA), TaidaOne\_25 (BI1), Tatanka\_28 (AU1), Teacup\_28 (AU1), Tenney120\_29 (AU6), Tenno\_29 (AU1), Thiqqums\_23 (BI4), TieDye\_24 (BI1), Tokki\_29 (AU2), TonyStarch\_25 (BI1), Truckee\_27 (AU5), Uzumaki\_28 (AU6), Yotsuba\_28 (JA), Zeina\_28 (AU6), Zippen\_28 (AU3)  Gene: TrixiePhattel\_29 Start: 24770, Stop: 25747, Start Num: 42 Candidate Starts for TrixiePhattel\_29: (Start: 12 @24584 has 95 MA's), (25, 24689), (27, 24695), (39, 24746), (42, 24770), (44, 24776), (50, 24815), (85, 24980), (88, 25001), (112, 25121), (133, 25220), (137, 25238), (151, 25319), (167, 25355), (171, 25376), (173, 25379), (174, 25382), (207, 25553), (209, 25559), (213, 25589), | |
| BLAST DATA | |
| Function  High Match with two possible proteins | |

|  |
| --- |
| Gene Information Stop 55163 |
| Not called by glimmer or genemark. Found via coding potential report. Starterator produced low manual annotations for start 54906. Blast results suggest 54906 is a hypothetical protein related to Zeina. Nothing significant from HHPred. |
| Function: Hypothetical Protein |
| Gene appears to match location well with Renaldo. But all cluster mates seem to end with this gene. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 55163 |
| Glimmer | NA |
| GeneMark | NA |
| Length(s) | 267 |
| Overlap(s) | -1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -1 |
| Best Z-score? | 2.957 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~54920 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham707Report.pdf>  Gene: TrixiePhattel\_93 Start: 54897, Stop: 55163, Start Num: 10  Candidate Starts for TrixiePhattel\_93:  (10, 54897), (Start: 13 @54906 has 24 MA's), (18, 54918), (30, 55023), (37, 55077),  Start 10:  • Found in 8 of 218 ( 3.7% ) of genes in pham  • No Manual Annotations of this start.  • Called 12.5% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_93 (AU6),  Start 13: • Found in 43 of 218 ( 19.7% ) of genes in pham • Manual Annotations of this start: 24 of 168 • Called 97.7% of time when present • Phage (with cluster) where this start called: Argan\_92 (AU6), Arzan\_97 (FI), BarbieDoll\_92 (AU6), Biscute\_91 (AU6), Boog\_91 (AW), BronxBay\_87 (AW), Camille\_77 (EL), Count\_133 (EL), Damascus\_82 (EL), DizzyRudy\_86 (EL), Djungelskog\_86 (AW), DoctorPepper\_86 (AW), Egad\_88 (AW), GantcherGoblin\_91 (AU6), HerbBucket\_87 (AW), Jazzy4900\_92 (FI), KevinMinion\_91 (AU6), Kinny\_93 (AU6), Leathea\_98 (AU6), Lewando\_89 (AU6), Linda\_87 (AW), Michelle\_87 (AW), MrAaronian\_87 (AW), Natasha\_90 (AW), Navi1117\_97 (AU6), Phaila\_96 (AU6), Powelldog\_93 (AW), ProfFrink\_87 (AW), Raunak\_88 (AW), Renaldo\_93 (AU6), RustyBoy\_90 (AW), Salk\_87 (AW), Schwartz33\_88 (DJ), Shiba\_86 (AW), Sloopyjoe\_87 (AW), Sporto\_84 (AW), StarLord\_87 (AW), Stayer\_87 (AW), | |
| Blast Results      Start 54906 | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 26407 |
| Gene Stop at 26407 with gene start originally called by Genemark at 26105. However start 26000 captures full coding potential. Blast Results showed alignment with Hypothetical protein of Zeina. Starterator shows researchers calling gene in similar location of other cluster mates. HHPred shows no significant matches. |
| Function: Hypothetical Protein |
| Gene location appears to be similar location and after similar genes of other cluster mates. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 26407 |
| Glimmer | Glimmer did not call |
| GeneMark | 26105 |
| Length(s) | 303 |
| Overlap(s) | -104 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | #2 26105 Z= 2.433 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~26090 |
| Staterator Screenshot    [Pham88143Report.pdf](http://phages.wustl.edu/starterator/Pham88143Report.pdf)  Gene: TrixiePhattel\_31 Start: 26105, Stop: 26407, Start Num: 5 Candidate Starts for TrixiePhattel\_31: (Start: 2 @26000 has 5 MA's), (5, 26105), (7, 26189), (8, 26237), (9, 26264), (10, 26294), (11, 26324),  Start 2: • Found in 10 of 15 ( 66.7% ) of genes in pham • Manual Annotations of this start: 5 of 7 • Called 80.0% of time when present • Phage (with cluster) where this start called: Argan\_29 (AU6), GantcherGoblin\_29 (AU6), KevinMinion\_30 (AU6), Leathea\_30 (AU6), Lewando\_33 (AU6), Navi1117\_31 (AU6), Uzumaki\_30 (AU6), Zeina\_30 (AU6), | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **To the right of minor tail protein** |
| **BLASTp (Identified function)**  **73% aligned with hypothetical protein with Zenia** |
| **HHpred** |
| **Conserved Domain:** |
| **DeepTMHMM: N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information |
| Gene begins at 26861. Glimmer called at 26358. Genemark called at 26382. Start 26358 was chosen. Genemark call failed to capture full coding potential. Start 26358 also produced a blast result in alignment with Zeina with a hypothetical protein. HHPred produced no significant results. |
| Function: Hypothetical Protein |
| This gene is present in similar locations across clustermates. Around the same genes as well. However it moves about the frame slightly. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 26861 |
| Glimmer | 26358 |
| GeneMark | 26382 |
| Length(s) | 504 |
| Overlap(s) | 50 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 50 |
| Best Z-score? | 2.180 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~26370 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham1442Report.pdf>  Track 38 : Phaila\_30, TrixiePhattel\_32  Gene: TrixiePhattel\_32 Start: 26358, Stop: 26861, Start Num: 6  Candidate Starts for TrixiePhattel\_32:  (Start: 6 @26358 has 1 MA's), (10, 26382), (Start: 20 @26415 has 3 MA's), (Start: 24 @26430 has 16 MA's), (40, 26535), (50, 26604), (51, 26610), (56, 26640), (68, 26712), (69, 26724),  Start 6:  • Found in 8 of 98 ( 8.2% ) of genes in pham  • Manual Annotations of this start: 1 of 70  • Called 50.0% of time when present  • Phage (with cluster) where this start called: KevinMinion\_31 (AU6), Phaila\_30 (AU6),  TrixiePhattel\_32 (AU6), Zeina\_31 (AU6) | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 27131 |
| Start 26868 kept, originally called by Genemark but not glimmer. Full coverage with Lewando. Coding potential captured. BLAST also shows 100% alignment. High Z-Value agrees with this start. HHpred shows 94% probability for outer membrane protein. This is not on the naming guide so Membrane protein will be called. This was changed to membrane protein due to high matches with other phages in cluster calling membrane protein. DeepTMHMM predicts two transmembrane domains. |
| -Function Membrane protein |
| **-Near similar genes as Lewando, also has 100% alignment in BLASTp. To the left of tape measure protein.** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 27131 |
| Glimmer | Not called |
| GeneMark | 26868 |
| Length(s) | 264 |
| Overlap(s) | -6 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~26905 |
| Staterator Screenshot | |
| Gene: TrixiePhattel\_33 Start: 26868, Stop: 27131, Start Num: 33 Candidate Starts for TrixiePhattel\_33: (Start: 33 @26868 has 5 MA's), (46, 26913), (49, 26928), (51, 26934),  <http://phages.wustl.edu/starterator/Pham176016Report.pdf> | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Near similar genes as Lewando, also has 100% alignment in BLASTp** |
| **BLASTp (Identified function)**  **100% aligned with Lewando for membrane protein** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 27533 |
| Gene begins at stop 27533. Genemark called at 27141 with no call from Glimmer. Full Coding potential captured by start. Gene called in large amounts by other researchers on Starterator. High Match on Blast Results and HHPred for Holin. With cluster mate matches being Argan and Gantchergoblin. |
| Function: Holin |
| Location and similarity matches Holin in other cluster mates. Though the exact location on the frame differs between cluster mates. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 27533 |
| Glimmer | N/A |
| GeneMark | 27141 |
| Length(s) | 393 |
| Overlap(s) | -9 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -9 |
| Best Z-score? | 1, Z=3.297, 9 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~27150 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham228945Report.pdf>  Gene: TrixiePhattel\_35 Start: 27141, Stop: 27533, Start Num: 3  Candidate Starts for TrixiePhattel\_35:  (Start: 3 @27141 has 46 MA's), (6, 27165), (14, 27228), (32, 27360), (45, 27501),  Start 3:  • Found in 74 of 78 ( 94.9% ) of genes in pham  • Manual Annotations of this start: 46 of 50  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Acai\_33 (AU1), AreFloNak\_33 (AU1),  Argan\_32 (AU6), Arzan\_29 (FI), Biscute\_37 (AU6), Boog\_29 (AW), Bouchard\_34  (AU2), Breylor17\_34 (AU1), BronxBay\_30 (AW), Brunswick\_33 (AU1),  CapnMurica\_33 (AU1), CastorTray\_35 (AU1), Caterpillar\_31 (AU4), Chlochlo\_34  (AU2), Darby\_33 (AU1), DevitoJr\_33 (AU1), Djungelskog\_30 (AW), DoctorPepper\_30  (AW), Egad\_30 (AW), ElephantMan\_33 (AU1), Elver\_46 (FK), Gandionco\_48 (FK),  GantcherGoblin\_32 (AU6), Giantsbane\_34 (AU2), Gordon\_33 (AU1), HerbBucket\_29  (AW), Inked\_35 (AU), Issa\_33 (AU1), Jazzy4900\_31 (FI), KevinMinion\_33 (AU6),  Kinny\_36 (AU6), Leathea\_33 (AU6), Lewando\_36 (AU6), Linda\_30 (AW),  LucySwiss\_33 (AU1), Makai\_33 (AU5), MediumFry\_31 (AU4), Michelle\_30 (AW),  MiniBagel\_33 (AU1), MrAaronian\_30 (AW), Natasha\_29 (AW), Navi1117\_34 (AU6),  Nightmare\_34 (AU1), Niktson\_33 (AU1), Nivinsha\_33 (AU1), Paella\_47 (FK),  Phaby\_34 (AU2), Phaila\_32 (AU6), Powelldog\_32 (AW), ProfFrink\_30 (AW), Qui\_47  (FK), Raunak\_31 (AW), Renaldo\_38 (AU6), RustyBoy\_29 (AW), Salk\_30 (AW),  ScienceWizSam\_35 (AU1), Shiba\_30 (AW), Sloopyjoe\_30 (AW), Sporto\_31 (AW),  StarLord\_30 (AW), Stayer\_30 (AW), Sunny4976\_30 (FI), Synepsis\_32 (AU1),  Tatanka\_33 (AU1), Teacup\_33 (AU1), Tenney120\_34 (AU6), Tenno\_34 (AU1),  Tokki\_34 (AU2), TrixiePhattel\_35 (AU6), Truckee\_32 (AU5), Trustiboi\_34 (AU1),  Uzumaki\_33 (AU6), Zeina\_33 (AU6), Zippen\_34 (AU3) | |
| Blast Results | |
| Function  High Match above 90% | |

|  |
| --- |
| Gene Information Stop 27885 |
| Start 27592 kept, originally called by glimmer and genemark. Coding potential is captured, and full coverage with GantcherGoblin. BLASTalso shows a 100% alignment. High Z-Value and High number of manual annotations on starterator agrees. HHpred shows insufficient probability for other protein functions. |
| Function-Hypothetical protein |
| **-Synteny mixed, however there is a match with GantcherGoblin by being to the right of Tape measure protein and minor tail protein.** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 27885 |
| Glimmer | 27592 |
| GeneMark | 27592 |
| Length(s) | 294 |
| Overlap(s) | 58 (Gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 58 (Gap) |
| Best Z-score? | 3.286 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~27610 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham85027Report.pdf> | |

|  |
| --- |
|  |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Synteny mixed, however there is a match with GantcherGoblin by being to the right of Tape measure protein and minor tail protein.** | | **BLASTp (Identified function)**  **100% aligned with GancherGoblin for hypothetical protein** | | **HHpred** | | **Conserved Domain N/A** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | |

|  |
| --- |
| Gene Information Stop 28306 |
| Both Glimmer and Genemark calling start at 28070. Coding potential fully captured via that start. Starterator shows minor annotation of similar gene in cluster mate at this start. Blast Results are in alignment with Ziena for a membrane protein. HHPred doesn’t cross the 90% on a match. No conclusive match on this gene. Deep TMHMM does not predict a transmembrane domain. |
| Function: Hypothetical Protein |
| Synteny is held across the clustermates with this gene. However the location isn’t in perfect alignment and there are different genes nearby. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 28306 |
| Glimmer | 28070 |
| GeneMark | 28070 |
| Length(s) | 237 |
| Overlap(s) | -184 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -184 |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~28080 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham29896Report.pdf>  Gene: TrixiePhattel\_37 Start: 28070, Stop: 28306, Start Num: 1  Candidate Starts for TrixiePhattel\_37:  (Start: 1 @28070 has 1 MA's), (4, 28142), (6, 28181),  Start 1:  • Found in 7 of 7 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 1 of 1  • Called 85.7% of time when present  • Phage (with cluster) where this start called: KevinMinion\_36 (AU6), Navi1117\_36  (AU6), Phaila\_35 (AU6), Tenney120\_37 (AU6), TrixiePhattel\_37 (AU6), Zeina\_36 (AU6), | |
| Blast Results | |
| Function  No match above 90%    DeepTMHMM | |

|  |
| --- |
| Gene Information |
|  |
| Start 28585 kept, Originally called by both glimmer and genemark. Coding potential captured, and full coverage of phage in the same cluster (Lewando). High Z-Value, and high number of manual annotations in starterator agree. HHpred shows insignificant probability for proteins of other functions. |
| Function-Hypothetical protein |
| **-Gene number is the same on both TrixiePhattel and Lewando. It is to the right of tape measure protein and minor tail protein** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 28818 |
| Glimmer | 28585 |
| GeneMark | 28585 |
| Length(s) | 234 |
| Overlap(s) | 278 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 278 (gap) |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~28590 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham66729Report.pdf> | |

|  |
| --- |
|  |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Gene number 38 is the same on both TrixiePhattel and Lewando. It is to the right of tape measure protein and minor tail protein** | | **BLASTp (Identified function)**  **100% aligned with Lewando for hypothetical protein** | | **HHpred** | | **Conserved Domain N/A** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | |

|  |
| --- |
| Gene Information Stop 29126 |
| Glimmer call at 28923. Genemark call at 28908. Both calls fully capture coding potential. Starterator has no manual annotations supporting 28908 compared to 28923 having near 100%. Blast results are high match with Zeina for a hypothetical protein with no significant match on the HHpred. |
| Function: Hypothetical Protein |
| Clustermates show similar location with some offset. Surrounding genes are different between clustermates. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 29162 |
| Glimmer | 28923 |
| GeneMark | 28908 |
| Length(s) | 240 |
| Overlap(s) | -104 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -104 |
| Best Z-score? | 3.053 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~28930 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham225987Report.pdf>  Gene: TrixiePhattel\_39 Start: 28923, Stop: 29162, Start Num: 17  Candidate Starts for TrixiePhattel\_39:  (15, 28908), (Start: 17 @28923 has 65 MA's), (37, 29112),  Start 17:  • Found in 91 of 129 ( 70.5% ) of genes in pham  • Manual Annotations of this start: 65 of 99  • Called 98.9% of time when present  • Phage (with cluster) where this start called: Acai\_39 (AU1), Arcadia\_37 (AM),  AreFloNak\_38 (AU1), Argan\_36 (AU6), BarbieDoll\_41 (AU6), BenitoAntonio\_38 (AM),  Benllo\_37 (AM), Biscute\_40 (AU6), Boog\_35 (AW), Bouchard\_39 (AU2), Bowling\_37  (AM), Breylor17\_39 (AU1), BronxBay\_36 (AW), Brunswick\_38 (AU1), CapnMurica\_37  (AU1), CastorTray\_40 (AU1), Caterpillar\_36 (AU4), Cheesy\_37 (AM), Chlochlo\_39  (AU2), Circum\_39 (AM), Correa\_35 (AM), Darby\_37 (AU1), DevitoJr\_38 (AU1),  Djungelskog\_36 (AW), DoctorPepper\_36 (AW), Dynamite\_38 (AM), Egad\_36 (AW),  ElephantMan\_37 (AU1), Elsa\_37 (AM), GantcherGoblin\_36 (AU6), Giantsbane\_40  (AU2), GoCrazy\_36 (AM), Gordon\_37 (AU1), Hankly\_37 (AM), Heisenberger\_37  (AM), HerbBucket\_35 (AW), Inked\_41 (AU), Issa\_39 (AU1), JEGGS\_37 (AM),  Kardesai\_39 (AM), KeaneyLin\_36 (AM), Kinny\_41 (AU6), Leathea\_36 (AU6),  Lewando\_40 (AU6), LilHuddy\_40 (AU2), Linda\_36 (AW), LucySwiss\_39 (AU1),  Makai\_37 (AU5), MediumFry\_37 (AU4), Michelle\_36 (AW), MiniBagel\_37 (AU1),  Mooshroom\_40 (AM), MrAaronian\_36 (AW), Mudcat\_36 (AM), NapoleonB\_38 (AM),  Nason\_37 (AM), Natasha\_36 (AW), Navi1117\_39 (AU6), Nightmare\_39 (AU1),  Niktson\_37 (AU1), Nivinsha\_39 (AU1), Phaby\_40 (AU2), Phaila\_38 (AU6),  Powelldog\_39 (AW), ProfFrink\_36 (AW), Raunak\_37 (AW), Renaldo\_42 (AU6),  RustyBoy\_36 (AW), Salk\_36 (AW), ScienceWizSam\_40 (AU1), Shepard\_41 (AU2),  Shiba\_36 (AW), Sloopyjoe\_36 (AW), Sporto\_38 (AW), StarLord\_36 (AW), Stayer\_36  (AW), Synepsis\_37 (AU1), Tatanka\_37 (AU1), Teacup\_39 (AU1), Tenney120\_39  (AU6), Tenno\_38 (AU1), Tokki\_39 (AU2), Tribby\_37 (AM), TrixiePhattel\_39 (AU6),  Truckee\_36 (AU5), Trustiboi\_39 (AU1), Uzumaki\_36 (AU6), Xenomorph\_34 (AM),  Zeina\_38 (AU6), Zippen\_41 (AU3), | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 29442 |
| Start of 29215 selected based on gene coding potential report.  Gene possesses no result on blast or starterator. However results occurred on a HHPred result run through the MPI Bioinformatics Tool. Suggesting gene and protein activity with unknown output. |
| Function: Hypothetical Protein |
| Unable to determine due to lack of associated pham number. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction |  |
| Stop | 29442 |
| Glimmer | NA |
| GeneMark | NA |
| Length(s) | 228 |
| Overlap(s) | -52 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -52 |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~29225 |
| Staterator Screenshot | |
|  | |
| <https://toolkit.tuebingen.mpg.de/jobs/Stop29442> | |

|  |
| --- |
| Gene Information Stop 29622 |
| Start 29485 kept, called originally by glimmer and genemark. Coding potential was captured. BLAST shows 74% alignment but 100% similarity with Zeina. Best Z-value (with reasonable amount of base pairs) available. All of the manual starterator annotations agree. HHpred shows insignificant hits for other protein function possibilities. Deep TMHMM predicts a transmembrane domain. |
| Function-Membrane protein |
| **Near similar genes as Zeina (after holin protein)** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 29622 |
| Glimmer | 29485 |
| GeneMark | 29485 |
| Length(s) | 138 |
| Overlap(s) | 322 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 322 (gap) |
| Best Z-score? | 1.896 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~29490 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham11721Report.pdf> | |

|  |
| --- |
|  |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Near similar genes as Zenia (after holin protein)** | | **BLASTp (Identified function)**  **-74.5% aligned with Zenia for membrane protein** | | **HHpred** | | **Conserved Domain** | | **DeepTMHMM** |  |  | | --- | | **Other Evidence?** | |

|  |
| --- |
| Gene Information Stop 29871 |
| Glimmer call at 29758. No call from genemark. Coding potential fully captured. Starterator report has no manual annotations of this area. It produces no significant blast results. HHPred also produces no significant results. |
| Function: Hypothetical Protein |
| Synteny is only found with KevinMinion. Genes which surround stop 29781 are similar between the two. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 29871 |
| Glimmer | 29758 |
| GeneMark | NA |
| Length(s) | 114 |
| Overlap(s) | -135 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? |  |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~29770 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham195198Report.pdf>  Gene: TrixiePhattel\_41 Start: 29758, Stop: 29871, Start Num: 4  Candidate Starts for TrixiePhattel\_41:  (1, 29704), (4, 29758), (5, 29782), (6, 29815), (7, 29845),  Start 4:  • Found in 2 of 2 ( 100.0% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: KevinMinion\_40 (AU6), TrixiePhattel\_41  (AU6), | |
|  | |
|  | |

|  |
| --- |
| Gene Information |
| Start 29861 kept, originally called by genemark but not glimmer. Coding potential captured. 94% alignment with Uzumaki. Highest Z-value available. HHpred shows results for SKG6 protein, but no high hits. |
| Function- Hypothetical protein |
| **-Next to similar genes as Uzumaki, right of minor tail protein** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 30022 |
| Glimmer | Not called |
| GeneMark | 29861 |
| Length(s) | 162 |
| Overlap(s) | 114 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 114 (overlap) |
| Best Z-score? | 3.286 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~29890 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham4767Report.pdf> | |

|  |
| --- |
|  |
| |  | | --- | | **Function** | | **Context & Synteny**    **-Next to similar genes, to right of minor tail protein** | | **BLASTp (Identified function)**  **94% aligned with Uzumaki for hypothetical protein** | | **HHpred-** | | **Conserved Domain** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | |

|  |
| --- |
| Gene Information Stop 30770 |
| Glimmer calling at 30009. Genemark calling at 30027. Start 30009 was selected due to producing a functional blast in alignment with Uzumaki for a hypothetical protein. HHPred shows a high match with a Hypothetical Protein. |
| Function: Hypothetical Protein |
| High Synteny with clustermates for this gene. With multiple places throughout the genome. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 30770 |
| Glimmer | 30009 |
| GeneMark | 30027 |
| Length(s) | 762 |
| Overlap(s) | 14 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 14 |
| Best Z-score? | 1.977 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~30040 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham183Report.pdf>  Gene: TrixiePhattel\_43 Start: 30009, Stop: 30770, Start Num: 38  Candidate Starts for TrixiePhattel\_43:  (Start: 31 @29982 has 45 MA's), (Start: 38 @30009 has 15 MA's), (44, 30027), (56, 30075), (60,30090), (63, 30102), (100, 30201), (115, 30252), (130, 30318), (138, 30375), (146, 30423), (169, 30489), (174, 30498), (177, 30525),  Start 38:  • Found in 25 of 436 ( 5.7% ) of genes in pham  • Manual Annotations of this start: 15 of 336  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Annihilus\_34 (BI2), Argan\_39 (AU6),  BarbieDoll\_46 (AU6), Bilo\_32 (BI2), Biscute\_45 (AU6), Doxi13\_34 (BI2),  GantcherGoblin\_39 (AU6), GoblinVoyage\_34 (BI2), HotFries\_33 (BI2),  KevinMinion\_42 (AU6), Kinny\_46 (AU6), Leathea\_41 (AU6), Lewando\_44 (AU6),  Moozy\_33 (BI2), Navi1117\_42 (AU6), Phaila\_42 (AU6), PherryCruz\_33 (BI2),  RavenPuff\_33 (BI2), Renaldo\_45 (AU6), Scap1\_33 (BI2), SheRa\_34 (BI2),  Tenney120\_43 (AU6), TrixiePhattel\_43 (AU6), Uzumaki\_39 (AU6), Zeina\_42 (AU6), | |
| Blast Results | |
| Function | |

|  |
| --- |
| Gene Information |
| Start 30777 kept, called by glimmer and genemark. Coding potential captured. Full coverage with Uzumaki. Good Z-Value and high number of manual starterator annotations agree. Moved as far up as possible. HHpred shows high probability for Nucleotide modification associated domain, however cannot be called due to SEA-Phage naming guide. Defaults to Hypothetical protein |
|  |
| Function-Hypothetical protein |
| **-Similar genes near each other, (after holin)** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 31151 |
| Glimmer | 30777 |
| GeneMark | 30777 |
| Length(s) | 375 |
| Overlap(s) | 6 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? |  |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~30805 |
| Staterator Screenshot      http://phages.wustl.edu/starterator/Pham218520Report.pdf | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**    **-Similar genes near each other, (after holin)** |
| **BLASTp (Identified function)**  **100% aligned with Uzumaki for hypothetical protein** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 31359 |
| Both glimmer and genemark called at 31144. With the original call preserved. Coding potential fully captured by start. Starterator shows gene is always called at that start when presented. Scores high blast with Zeina for Hypothetical Protein. No significant HHPred results. |
| Function: Hypothetical Protein |
| High Synteny for cluster mates with the gene. Location and context seems largely similar for Uzumakin. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction |  |
| Stop | 31359 |
| Glimmer | 31144 |
| GeneMark | 31144 |
| Length(s) | 216 |
| Overlap(s) | 8 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 8 |
| Best Z-score? | 1, Z=3.053,14 |
| Gene Mark Screenshot | |
| Estimate Start of Coding potential | ~31150 |
| Staterator Screenshot    [Starterator](http://phages.wustl.edu/starterator/Pham12150Report.pdf)  Gene: TrixiePhattel\_45 Start: 31144, Stop: 31359, Start Num: 9  Candidate Starts for TrixiePhattel\_45:  (Start: 9 @31144 has 4 MA's), (18, 31297), (20, 31336),  Start 9:  • Found in 11 of 11 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 4 of 4  • Called 90.9% of time when present  • Phage (with cluster) where this start called: Argan\_41 (AU6), GantcherGoblin\_42  (AU6), KevinMinion\_44 (AU6), Leathea\_43 (AU6), Navi1117\_45 (AU6), Renaldo\_47  (AU6), Tenney120\_45 (AU6), TrixiePhattel\_45 (AU6), Uzumaki\_41 (AU6), Zeina\_44  (AU6), | |
| Blast Results | |
| Function  No Match higher than 90% | |

|  |
| --- |
| Gene Information |
| Start 31420 kept, originally called by glimmer and genemark. Coding potential captured. Full coverage with clustermate Lewando. BLAST also shows 100% alignment. Highest Z-value available, and high number of manual annotations agrees with this start. HHpred shows no significant hits for another protein function. |
| Function- Hypothetical protein |
| **Gene is similar to that of Lewando & to the right of tape measure protein** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 31620 |
| Glimmer | 31423 |
| GeneMark | 31423 |
| Length(s) | 198 |
| Overlap(s) | -63 gap |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? |  |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~31430 |
| Staterator Screenshot  <http://phages.wustl.edu/starterator/Pham214772Report.pdf> | |
|  | |

|  |
| --- |
|  |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **Gene is similar to that of Lewando & to the right of tape measure protein (olive green)** |
| **BLASTp (Identified function)**  **100% aligned with Lewando for hypothetical protein** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 32023 |
| Glimmer and Genemark calling start at 31685. Fully captures coding potential from gene coding potential record. Starterator produces no significant manual annotations. Blast result shows alignment with gantchergoblin and argan for a hypothetical protein. HHPred shows high match over 90% for hypothetical protein |
| Function; Hypothetical Protein |
| Synteny with other clustermates for the location and genes near the called gene. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 32023 |
| Glimmer | 31685 |
| GeneMark | 31685 |
| Length(s) | 339 |
| Overlap(s) | -64 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -64 |
| Best Z-score? | 2.806 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~31690 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham182449Report.pdf>  Gene: TrixiePhattel\_47 Start: 31685, Stop: 32023, Start Num: 11  Candidate Starts for TrixiePhattel\_47:  (11, 31685), (18, 31757), (25, 31802), (33, 31907),  Start 11:  • Found in 1 of 46 ( 2.2% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_47 (AU6), | |
| Blast Results | |
| Function  High Match over 90% | |

|  |
| --- |
| Gene Information Stop 32164 |
| Start at 32042 called (Only option). However, coding potential was captured. There is full coverage with GantcherGoblin. BLAST also shows a 100% alignment. The gap is not significant. HHpred is calling protein function with no significant hits. |
| Function- Hypothetical protein |
| **-Similar genes surrounding GantcherGoblin. To the left of DNA primase/polymerase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 32164 |
| Glimmer | N/A |
| GeneMark | 32042 |
| Length(s) | 123 |
| Overlap(s) | 18 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 18 (gap) |
| Best Z-score? | 2.789 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~32050 |
| Staterator Screenshot  Starterator missing from gene list  <http://phages.wustl.edu/starterator/Pham11751Report.pdf>  The start number called the most often in the published annotations is 8, it was called in 5 of the 5 non-draft genes in the pham. Genes that call this "Most Annotated" start: • Argan\_45, BarbieDoll\_50, Biscute\_49, GantcherGoblin\_45, KevinMinion\_47, Kinny\_50, Leathea\_46, Lewando\_48, Navi1117\_49, Phaila\_48, Renaldo\_51, Tenney120\_48, Uzumaki\_44, Zeina\_47, Genes that have the "Most Annotated" start but do not call it: • Genes that do not have the "Most Annotated" start: • Summary by start number: Start 8: • Found in 14 of 14 ( 100.0% ) of genes in pham • Manual Annotations of this start: 5 of 5 • Called 100.0% of time when present | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Similar genes surrounding GantcherGoblin. To the left of DNA primase/polymerase** |
| **BLASTp (Identified function)** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 32704 |
| Glimmer and Genemark both called at 32174. Kept because it captures the fully coding potential. Starterator shows large amount of manual annotations for selected start. Blast results show high match for hypothetical protein in Zeina. HHPred has no match available. |
| Function: Hypothetical Protein |
| Synteny surrounding found gene with similar location and genes around it across clustermates |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 32704 |
| Glimmer | 32174 |
| GeneMark | 32174 |
| Length(s) | 531 |
| Overlap(s) | -9 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -9 |
| Best Z-score? | 1, Z=3.201, 9 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~32180 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham182449Report.pdf>  Gene: TrixiePhattel\_48 Start: 32174, Stop: 32704, Start Num: 15  Candidate Starts for TrixiePhattel\_48:  (Start: 15 @32174 has 76 MA's), (22, 32261), (28, 32288), (31, 32315), (36, 32345), (52, 32507), (55,32531), (62, 32600),  Start 15:  • Found in 102 of 137 ( 74.5% ) of genes in pham  • Manual Annotations of this start: 76 of 109  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Acai\_46 (AU1), Annihilus\_36 (BI2),  Arcadia\_48 (AM), AreFloNak\_45 (AU1), Argan\_46 (AU6), BarbieDoll\_51 (AU6),  BenitoAntonio\_48 (AM), Benllo\_49 (AM), Bilo\_34 (BI2), Biscute\_50 (AU6), Boog\_44  (AW), Bouchard\_44 (AU2), Bowling\_47 (AM), Breylor17\_45 (AU1), BronxBay\_43  (AW), Brunswick\_45 (AU1), CapnMurica\_42 (AU1), CastorTray\_47 (AU1),  Caterpillar\_41 (AU4), Cheesy\_48 (AM), Chlochlo\_44 (AU2), Circum\_50 (AM),  Correa\_46 (AM), Darby\_43 (AU1), DevitoJr\_43 (AU1), Djungelskog\_43 (AW),  DoctorPepper\_42 (AW), Doxi13\_36 (BI2), Dynamite\_48 (AM), Egad\_43 (AW),  ElephantMan\_44 (AU1), Elsa\_48 (AM), GantcherGoblin\_46 (AU6), Giantsbane\_46  (AU2), GoCrazy\_47 (AM), GoblinVoyage\_36 (BI2), Gordon\_42 (AU1), Hankly\_47  (AM), Heisenberger\_47 (AM), HerbBucket\_43 (AW), HotFries\_35 (BI2), Ingrid\_46  (AU3), Inked\_48 (AU), Issa\_46 (AU1), JEGGS\_47 (AM), Kardesai\_49 (AM),  KeaneyLin\_47 (AM), KevinMinion\_48 (AU6), Kinny\_51 (AU6), Leathea\_47 (AU6),  Lewando\_49 (AU6), LilHuddy\_46 (AU2), Linda\_43 (AW), Loretta\_46 (AU3),  LucySwiss\_46 (AU1), Makai\_44 (AU5), Michelle\_43 (AW), MiniBagel\_44 (AU1),  Mooshroom\_50 (AM), Moozy\_35 (BI2), MrAaronian\_43 (AW), Mudcat\_45 (AM),  NapoleonB\_48 (AM), Nason\_48 (AM), Natasha\_43 (AW), Navi1117\_50 (AU6),  Nightmare\_46 (AU1), Niktson\_44 (AU1), Nivinsha\_45 (AU1), Phaby\_46 (AU2),  Phaila\_49 (AU6), PherryCruz\_35 (BI2), Powelldog\_48 (AW), ProfFrink\_43 (AW),  Raunak\_46 (AW), RavenPuff\_35 (BI2), Renaldo\_52 (AU6), RustyBoy\_43 (AW),  Salk\_43 (AW), Scap1\_35 (BI2), ScienceWizSam\_45 (AU1), SheRa\_36 (BI2),  Shepard\_48 (AU2), Shiba\_42 (AW), Sloopyjoe\_43 (AW), Sporto\_44 (AW),  StarLord\_43 (AW), Stayer\_43 (AW), Synepsis\_43 (AU1), Tatanka\_43 (AU1),  Teacup\_45 (AU1), Tenney120\_49 (AU6), Tenno\_45 (AU1), Tokki\_48 (AU2),  Tribby\_48 (AM), TrixiePhattel\_48 (AU6), Truckee\_42 (AU5), Trustiboi\_44 (AU1),  Uzumaki\_45 (AU6), Xenomorph\_44 (AM), Zeina\_48 (AU6), Zippen\_46 (AU3) | |
| Blast Results | |
| Function | |

|  |
| --- |
| Gene Information Stop 33141 |
| Start 32839 kept, originally called by glimmer and genemark. Has the highest Z-value available. Coding potential captured, and is 81% aligned with Uzumaki. HHpred shows 90% probability for protein of an unknown function. |
| Function-Hypothetical protein |
| **-Similar genes near each other. To the left of DNA primase/polymerase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 33141 |
| Glimmer | 32839 |
| GeneMark | 32839 |
| Length(s) | 303 |
| Overlap(s) | 134 (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 134 (gap) |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~32850 |
| Staterator Screenshot | |
|  | |

|  |
| --- |
|  |
| |  | | --- | | **Function** | | **Context & Synteny**    **-Similar genes near each other. To the left of DNA primase/polymerase** | | **BLASTp (Identified function)**  **81% aligned with Uzumaki** | | **HHpred** | | **Conserved Domain N/A** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | |

|  |
| --- |
| Gene Information Stop 33347 |
| Called at 33138 by both Glimmer and Genemark. Captures full coding potential. Starterator shows manual annotation in location of call by all other clustermates. Blast Results and HHPred shows high alignment with Hypothetical Protein. Uzumaki has a strong match in blast. |
| Function: Hypothetical Protein |
| Location and surrounding genes appear to be fairly differing between custermates |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 33347 |
| Glimmer | 33138 |
| GeneMark | 33138 |
| Length(s) | 210 |
| Overlap(s) | 4 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 |
| Best Z-score? | 2.754 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~33130 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham230272Report.pdf>  Gene: TrixiePhattel\_50 Start: 33138, Stop: 33347, Start Num: 3  Candidate Starts for TrixiePhattel\_50:  (Start: 3 @33138 has 6 MA's), (5, 33180), (15, 33321),  Start 3:  • Found in 13 of 13 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 6 of 6  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Argan\_48 (AU6), Elver\_160 (FK),  Gandionco\_159 (FK), KevinMinion\_50 (AU6), Leathea\_50 (AU6), Lewando\_52 (AU6),  Paella\_163 (FK), Qui\_163 (FK), Renaldo\_55 (AU6), Tenney120\_52 (AU6),  TrixiePhattel\_50 (AU6), Uzumaki\_47 (AU6), Zeina\_51 (AU6), | |
| Blast Results | |
| Function  Match of 90% | |

|  |
| --- |
| Gene Information Stop 33550 |
| Glimmer calling 33437. Genemark calls 33452. 33437 selected as start. Captures full coding potential. Starterator leans more towards 33452 with higher annotations. Blast for 33437 shows better results than 33452. HHPred for the 33437 shows a result of higher than 90% match. Evidence points to 33437 being a better start location. |
| Function: Hypothetical Protein |
| Gene is located across the frames of the various clustermates. The location and genes surrounding it are differing between clustermates |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 33550 |
| Glimmer | 33437 |
| GeneMark | 33452 |
| Length(s) | 99 |
| Overlap(s) | 23 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 3.052 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~33450 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham230055Report.pdf>  Gene: TrixiePhattel\_52 Start: 33437, Stop: 33550, Start Num: 16  Candidate Starts for TrixiePhattel\_52:  (11, 33404), (13, 33422), (Start: 16 @33437 has 1 MA's), (Start: 18 @33452 has 3 MA's), (21, 33479), (28, 33530),  Start 16:  • Found in 15 of 97 ( 15.5% ) of genes in pham  • Manual Annotations of this start: 1 of 71  • Called 33.3% of time when present  • Phage (with cluster) where this start called: KevinMinion\_51 (AU6), Phaila\_53 (AU6),  Tenney120\_53 (AU6), TrixiePhattel\_52 (AU6), Uzumaki\_48 (AU6), | |
| Blast Results | |
|  | |

|  |
| --- |
| Gene Information Stop 34212 |
| Start 33550 kept, originally called by genemark and glimmer. Coding potential is captured. Full coverage with Lewando. BLAST also shows full alignment. Good Z-Value with very low overlap. HHpred shows 90% probability for protein with unknown function. |
| Function- Hypothetical protein |
| **-Gene matches in color and size on both TrixiePhattel and Lewando**  **-Also directly to the left of another similar gene** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 34212 |
| Glimmer | 33550 |
| GeneMark | 33550 |
| Length(s) | 663 |
| Overlap(s) | 1 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 2.157 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~33580 |
| Staterator Screenshot <http://phages.wustl.edu/starterator/Pham711Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Gene 52 matches in color and size on both**  **-Also directly to the left of another similar  gene** |
| **BLASTp (Identified function)**  **100% aligned with Zenia for hypothetical protein** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information |
| Glimmer calls at 34212. GeneMark made no call. Kept called start. Captures coding potential. Starterator calls the gene when present 100% of the time by manual annotators. Blast Results show high alignment with Hypothetical Protein for Lewando. HHPred shows no significant results. |
| Function: Hypothetical Protein |
| Genes tend to be surrounded by similar genes across cluster mates sharing this gene. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 34517 |
| Glimmer | 34212 |
| GeneMark | NA |
| Length(s) | 306 |
| Overlap(s) | 1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 1, Z=2.000, 10 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~34315 |
| Staterator Screenshot    [Starterator](http://phages.wustl.edu/starterator/Pham92563Report.pdf)  Start 1:  • Found in 4 of 4 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 3 of 3  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Argan\_51 (AU6), GantcherGoblin\_51  (AU6), Lewando\_56 (AU6), TrixiePhattel\_54 (AU6),  Gene: TrixiePhattel\_54 Start: 34212, Stop: 34517, Start Num: 1  Candidate Starts for TrixiePhattel\_54:  (Start: 1 @34212 has 3 MA's), (2, 34236), (5, 34350), (6, 34353), (7, 34359), (8, 34365), (10, 34431 | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information |
| Glimmer calls at 34783. Genemark has not made a call. Coding potential is fully captured. Starterator supports 34783 through its 100% calls through manual annotations. Blast Results are a high match with Gantchergoblin for hypothetical protein. HHPred has no conclusive results |
| Function: Hypothetical Protein |
| Synteny with cluster mates sharing this gene. Area of gene is similar in Uzamaki and nearby genes |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 34980 |
| Glimmer | 34783 |
| GeneMark | NA |
| Length(s) | 198 |
| Overlap(s) | 8 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? |  |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~34820 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham94169Report.pdf>  • Track 2 : Uzumaki\_52, KevinMinion\_55, Argan\_53, Zeina\_56, GantcherGoblin\_53,  TrixiePhattel\_56, Phaila\_58, Tenney120\_58  Gene: TrixiePhattel\_56 Start: 34783, Stop: 34980, Start Num: 2  Candidate Starts for TrixiePhattel\_56:  (Start: 2 @34783 has 6 MA's), (7, 34900), (8, 34924),  Start 2:  • Found in 14 of 19 ( 73.7% ) of genes in pham  • Manual Annotations of this start: 6 of 7  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Argan\_53 (AU6), Elver\_102 (FK),  Gandionco\_100 (FK), GantcherGoblin\_53 (AU6), KevinMinion\_55 (AU6), Leathea\_55  (AU6), Navi1117\_57 (AU6), Paella\_103 (FK), Phaila\_58 (AU6), Qui\_103 (FK),  Tenney120\_58 (AU6), TrixiePhattel\_56 (AU6), Uzumaki\_52 (AU6), Zeina\_56 (AU6), | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 35885 |
| Start 34977 kept, originally called by glimmer and genemark 98% aligned with clustermate Lewando.Captured coding potential. High Z-value agrees with this start. Moving forward would create large overlap. HHpred shows 100% probability for protein of unknown function. |
| Function- Hypothetical protein |
| **-Gene of same pham are similar**  **-To the left of DNA primase/Polymerase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 35885 |
| Glimmer | 34977 |
| GeneMark | 34977 |
| Length(s) | 909 |
| Overlap(s) | 4 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 (overlap) |
| Best Z-score? | 2.957 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~34990 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham183Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Gene of same pham are similar**  **-To the left of DNA primase/Polymerase** |
| **BLASTp (Identified function)**  **98% aligned with Lewando for hypothetical protein** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 36102 |
| Call made at 35920 by Glimmer and Genemark. Fully captures coding potential. Starterator holds no manual annotations for this gene. Solid match on a blast result with hypothetical protein for Zeina. HHPred has no significant results. |
| Function: Hypothetical Protein |
| Gene seems to appear between Pham 183 and Pham 230032 in clustermates. In Zeina and Uzumaki |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 36102 |
| Glimmer | 35920 |
| GeneMark | 35920 |
| Length(s) | 183 |
| Overlap(s) | -34 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -34 |
| Best Z-score? | 2.156 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~35950 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham193097Report.pdf>  Gene: TrixiePhattel\_58 Start: 35920, Stop: 36102, Start Num: 5  Candidate Starts for TrixiePhattel\_58:  (5, 35920), (10, 35977), (12, 36076)  Found in 1 of 15 ( 6.7% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_58 (AU6) | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 36839 |
| Start 36099 kept, originally called by genemark and glimmer. Coding potential captured with full coverage with Lewando. BLAST shows 100% alignment. Good Z-value agrees with the start. HHpred shows 90% probability for an “accessory protein” but is not specific so will not be called since there is no label for it on SEA-Phage naming guide |
| Function, Hypothetical protein |
| **-Synteny similar on right and left sides**  **-To the left of DNA primase/Polymerase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 36839 |
| Glimmer | 36099 |
| GeneMark | 36099 |
| Length(s) | 741 |
| Overlap(s) | 4(overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 (overlap) |
| Best Z-score? | 2.342 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~36110 |
| Staterator Screenshot      <http://phages.wustl.edu/starterator/Pham224412Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny similar on right and left sides**  **-To the left of DNA primase/Polymerase** |
| **BLASTp (Identified function)**  **100% aligned with Lewando for hypothetical protein** |
| **HHpred**    **HHpred website** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 37396 |
| Called at 36836 by both Glimmer and GeneMark. This spot fully captures coding potential. Starterator shows some annotations at this spot. Blast results are matching with Gantchergoblin for hypothetical protein. HHPred shows no matching results. |
| Function: Hypothetical Protein |
| Gene seems to be between Pham 230032 and Pham 699 for many clustermates. Tenny, Uzumaki, Lewando. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 37396 |
| Glimmer | 36836 |
| GeneMark | 36836 |
| Length(s) | 561 |
| Overlap(s) | 4 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 |
| Best Z-score? | 3.047 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~36840 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham700Report.pdf>  Gene: TrixiePhattel\_60 Start: 36836, Stop: 37396, Start Num: 3  Candidate Starts for TrixiePhattel\_60:  (Start: 3 @36836 has 13 MA's), (25, 36941), (29, 36959), (43, 37061), (48, 37097), (58, 37166), (66, 37217), (68, 37232), (70, 37244), (73, 37256), (78, 37349),  Start 3:  • Found in 53 of 218 ( 24.3% ) of genes in pham  • Manual Annotations of this start: 13 of 168  • Called 39.6% of time when present  • Phage (with cluster) where this start called: Arcadia\_60 (AM), Argan\_57 (AU6),  BarbieDoll\_63 (AU6), Benllo\_61 (AM), Biscute\_62 (AU6), Egad\_55 (AW), Elsa\_60  (AM), GantcherGoblin\_57 (AU6), Kinny\_63 (AU6), Leathea\_59 (AU6), Lewando\_62  (AU6), Linda\_55 (AW), Michelle\_55 (AW), NapoleonB\_60 (AM), Nason\_60 (AM),  Phaila\_62 (AU6), Raunak\_58 (AW), Sloopyjoe\_55 (AW), TrixiePhattel\_60 (AU6),  Uzumaki\_56 (AU6), Zeina\_60 (AU6), | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 41279 |
| Start 37419 kept, originally called by genemark and glimmer. The Earliest possible start is called with a good Z-value, and a high amount of manual annotations agree with this start.HHpred shows 100% probability for DNA primase/polymerase with Saccharomyces cerevisiae. |
| -Function: DNA primase/Polymerase |
| **-Synteny matched with Lewando on right and left sides**  **- similar gene also called as DNA primase/polymerase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 41279 |
| Glimmer | 37419 |
| GeneMark | 37419 |
| Length(s) | 3861 |
| Overlap(s) | 22 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 22 (Overlap) |
| Best Z-score? | 2.392 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~37440 |
| Staterator Screenshot        <http://phages.wustl.edu/starterator/Pham699Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny matched with Lewando on right and left sides**  **-Large purple gene also called as DNA primase/polymerase** |
| **BLASTp (Identified function)**  **100% aligned with Lewando for DNA primase/polymerase** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 41458 |
| 41279 called by both Glimmer and GeneMark. This call captures the coding potential of the report. Starterator shows this gene is called 100% of the time when found. Blast result shows alignment with Lewando for hypothetical protein. HHPred provided no significant result |
| Function: Hypothetical Protein |
| Gene is positioned between Pham 699 and Pham 10633 for many clustermates. Uzumaki, Tenny, Zeina. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 41458 |
| Glimmer | 41279 |
| GeneMark | 41279 |
| Length(s) | 180 |
| Overlap(s) | 1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 2.381 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~41290 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham10530Report.pdf>  Gene: TrixiePhattel\_62 Start: 41279, Stop: 41458, Start Num: 1  Candidate Starts for TrixiePhattel\_62:  (Start: 1 @41279 has 5 MA's), (3, 41387), (5, 41396), (6, 41408), (8, 41429), (9, 41441)  Start 1:  • Found in 15 of 15 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 5 of 5  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Argan\_59 (AU6), BarbieDoll\_65 (AU6),  Biscute\_64 (AU6), GantcherGoblin\_59 (AU6), KevinMinion\_61 (AU6), Kinny\_65  (AU6), Leathea\_61 (AU6), Lewando\_64 (AU6), Navi1117\_63 (AU6), Phaila\_64 (AU6),  Renaldo\_67 (AU6), Tenney120\_64 (AU6), TrixiePhattel\_62 (AU6), Uzumaki\_58  (AU6), Zeina\_62 (AU6), | |
| Blast Result | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 41574 |
| Start 41458 called, originally called by glimmer but not genemark. Best Z-value of the 2 options, and high amount of annotations agrees with this start. Overlap is small. BLAST shows 100% alignment with Zeina. HHpred shows no significant hits for protein of a different function. |
| Function: Hypothetical protein |
| **-Synteny similar on left and right of gene**  **-To the right of DNA primase/Polymerase**  **-To the left of DNA helicase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 41574 |
| Glimmer | 41458 |
| GeneMark | Not called |
| Length(s) | 117 |
| Overlap(s) | 1 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1(overlap) |
| Best Z-score? | 2.625 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~41450 |
| Staterator Screenshot        <http://phages.wustl.edu/starterator/Pham10633Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny similar on left and right of gene**  **-To the right of DNA primase/Polymerase**  **-To the left of DNA helicase** |
| **BLASTp (Identified function)**  **100% aligned with zenia for hypothetical protein** |
| **HHpred**    **-not 90%** |
| **Conserved Domain** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 41838 |
| Glimmer calls at 41584. GeneMark calls at 41662. 41584 selected to cover coding potential. Starterator has no manual annotations. Blast Results shows Zeina as a match for a hypothetical protein. HHPred shows no significant results |
| Function: Hypothetical Protein |
| Gene is commonly found between Pham 10633 and Pham 101661 in clustermates. Uzumaki, Zeina, Tenny. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 41838 |
| Glimmer | 41584 |
| GeneMark | 41662 |
| Length(s) | 255 |
| Overlap(s) | -9 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -9 |
| Best Z-score? | 2.400 |
| Gene Mark Screenshot | |
| Estimate Start of Coding potential | ~41590 |
| Staterator Screenshot    [Starterator](http://phages.wustl.edu/starterator/Pham89096Report.pdf)  Start 3:  • Found in 1 of 13 ( 7.7% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_64 (AU6),  Gene: TrixiePhattel\_64 Start: 41584, Stop: 41838, Start Num: 3  Candidate Starts for TrixiePhattel\_64:  (3, 41584), (5, 41662), (7, 41710) | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 42127 |
| Start 41831 kept, originally called by glimmer and genemark. Coding potential captured, and full coverage with Lewando, Overlap is small. Roughly 50% of manual annotations called this start in the starterator. HHpred shows no significant results for a protein of another function. |
| Function-Hypothetical protein |
| **-Synteny from left side matches**  **-To the right of DNA primase/Polymerase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 42127 |
| Glimmer | 41831 |
| GeneMark | 41831 |
| Length(s) | 297 |
| Overlap(s) | 8 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | Overlap (8) |
| Best Z-score? | 2.803 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~41810 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham101661Report.pdf> | |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Synteny from left side matches**  **-To the right of DNA primase/Polymerase** | | **BLASTp (Identified function)**  **100% aligned with Lewando for Hypothetical protein** | | **HHpred**  **-Not above 90%** | | **Conserved Domain N/A** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | | |

|  |
| --- |
| Gene Information Stop 42451 |
|  |
| Start kept at 42308, originally called by glimmer and not by genemark. Coding potential captured, and BLAST showed an 89% alignment with 94% similarity with Zenia. HHpred showed no significant hits for any other protein function. |
| Function: Hypothetical protein |
| **-Synteny mixed**  **-to the right of DNA primase/Polymerase, similar to Zenia** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 42451 |
| Glimmer | 42308 |
| GeneMark | Not called |
| Length(s) | 144 |
| Overlap(s) | 180 gap |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 50 (overlap) |
| Best Z-score? | 3.361 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~42300 |
| Staterator Screenshot        <http://phages.wustl.edu/starterator/Pham227931Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny mixed**  **-to the right of DNA primase/Polymerase** |
| **BLASTp (Identified function)**  **72% aligned with Zenia for hypothetical protein** |
| **HHpred**  **--No results above 90** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 43036 |
| 42452 was called by both genemark and glimmer. Fully captures coding potential. Starterator shows some manual annotations. Blast Results show alignment with Lewando for SSB Protein. HHPred is significant for SSB Protein. |
| Function: SSB Protein |
| Has pham 227252 in front of it on trixiephattel. Most other clustermates have 198593 behind it. Trixie is unique in not having the feature. But other clustermembers have this gene as SSB Protein |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 43036 |
| Glimmer | 42452 |
| GeneMark | 42452 |
| Length(s) | 585 |
| Overlap(s) | 0 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 0 |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~42460 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham204821Report.pdf>  Gene: TrixiePhattel\_68 Start: 42452, Stop: 43036, Start Num: 22  Candidate Starts for TrixiePhattel\_68:  (Start: 22 @42452 has 8 MA's), (40, 42578), (45, 42605), (50, 42632), (59, 42725), (69, 42764), (96, 42905), (99, 42935), (105, 42989),  Start 22:  • Found in 20 of 186 ( 10.8% ) of genes in pham  • Manual Annotations of this start: 8 of 138  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Argan\_63 (AU6), BarbieDoll\_70 (AU6),  Biscute\_69 (AU6), Elver\_117 (FK), Gandionco\_115 (FK), GantcherGoblin\_64 (AU6),  Giantsbane\_64 (AU2), KevinMinion\_65 (AU6), Kinny\_70 (AU6), Leathea\_66 (AU6),  Lewando\_69 (AU6), Navi1117\_69 (AU6), Paella\_118 (FK), Phaila\_68 (AU6), Qui\_118  (FK), Renaldo\_72 (AU6), Tenney120\_69 (AU6), TrixiePhattel\_68 (AU6), Uzumaki\_63  (AU6), Zeina\_68 (AU6), | |
| Blast Results | |
| Function  Above 90% match and synteny with clustermates for SSB Protein | |

|  |
| --- |
| Gene Information Stop 43309 |
|  |
| Changed start to 43043, originally called by glimmer at 43109, and genemark at 43043. Shortened gap from 74 to 7, slightly better Z score, captures more coding potential, 100% coverage alignment with Hypothetical protein from Zenia |
| Function: Hypothetical protein |
| **Similar genes near each other. To the right of DNA primase/polymerase** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 43309 |
| Glimmer | 43109 |
| GeneMark | 43043 |
| Length(s) | 201 |
| Overlap(s) | 72(gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 72 |
| Best Z-score? | 1.676 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~43080 |
| Staterator Screenshot    Start 4: • Found in 11 of 199 ( 5.5% ) of genes in pham • Manual Annotations of this start: 5 of 169 • Called 63.6% of time when present • Phage (with cluster) where this start called: Argan\_64 (AU6), GantcherGoblin\_65 (AU6), Leathea\_67 (AU6), Lewando\_70 (AU6), Navi1117\_70 (AU6), Uzumaki\_64 (AU6), Zeina\_69 (AU6),  <http://phages.wustl.edu/starterator/Pham224304Report.pdf> | |
| New Start 43043    [**UQS94741.1**](https://www.ncbi.nlm.nih.gov/protein/UQS94741.1?report=genbank&log$=protalign&blast_rank=1&RID=YY3FGG7F013) | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Similar genes near each other**  **-To the right of DNA primase/polymerase** |
| **BLASTp (Identified function)**  **84% aligned with Zenia, GancherGoblin, Uzumaki, Argan for hypothetical protein** |
| **HHpred**    **HHpred website**    **-Shows no real description for a potential protein** |
| **Conserved Domain**  **N/A** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 43595 |
| Called at 43311 by both Glimmer and Genemark. Captures full coding potential. Starterator has a gene called 100% of the time by annotators when present with this stop. Blast Results and HHPred show hypothetical protein. HHPred Protein is uncharacterized. |
| Function: Hypothetical Protein |
| Gene found between pham 227252 and phsm 230037. Similar on Uzumaki and Zeina |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 43595 |
| Glimmer | 43311 |
| GeneMark | 43311 |
| Length(s) | 285 |
| Overlap(s) | -1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -1 |
| Best Z-score? | 2.622 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~43315 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham228522Report.pdf>  Gene: TrixiePhattel\_70 Start: 43311, Stop: 43595, Start Num: 36  Candidate Starts for TrixiePhattel\_70:  (20, 43206), (21, 43209), (25, 43230), (28, 43242), (31, 43260), (33, 43281), (35, 43296), (Start: 36 @43311 has 35 MA's), (38, 43338), (39, 43356), (52, 43476),  Start 36:  • Found in 54 of 54 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 35 of 35  • Called 100.0% of time when present  • Phage (with cluster) where this start called: Acai\_66 (AU1), AreFloNak\_69 (AU1),  Argan\_65 (AU6), Arzan\_71 (FI), Bouchard\_71 (AU2), Breylor17\_65 (AU1),  Brunswick\_67 (AU1), CapnMurica\_63 (AU1), CastorTray\_69 (AU1), Caterpillar\_65  (AU4), Chlochlo\_68 (AU2), Darby\_64 (AU1), DevitoJr\_65 (AU1), ElephantMan\_66  (AU1), Elver\_161 (FK), Gandionco\_160 (FK), GantcherGoblin\_66 (AU6),  Giantsbane\_69 (AU2), Gordon\_63 (AU1), Ingrid\_72 (AU3), Inked\_71 (AU), Issa\_67  (AU1), Jazzy4900\_69 (FI), KevinMinion\_67 (AU6), Leathea\_68 (AU6), LilHuddy\_70  (AU2), Loretta\_69 (AU3), LucySwiss\_67 (AU1), Makai\_70 (AU5), MediumFry\_66  (AU4), MiniBagel\_65 (AU1), Navi1117\_71 (AU6), Nightmare\_67 (AU1), Niktson\_66  (AU1), Nivinsha\_69 (AU1), Paella\_164 (FK), Phaby\_71 (AU2), Phaila\_70 (AU6),  Qui\_164 (FK), ScienceWizSam\_67 (AU1), Shepard\_71 (AU2), Sunny4976\_68 (FI),  Synepsis\_64 (AU1), Tatanka\_61 (AU1), Teacup\_67 (AU1), Tenney120\_71 (AU6),  Tenno\_66 (AU1), Tokki\_75 (AU2), TrixiePhattel\_70 (AU6), Truckee\_68 (AU5),  Trustiboi\_65 (AU1), Uzumaki\_65 (AU6), Zeina\_70 (AU6), Zippen\_71 (AU3), | |
| Blast Results | |
| Function  Above 90%. Not characterized | |

|  |
| --- |
| Gene Information Stop 43930 |
| Start kept at 43595, originally called by both glimmer and genemark. Full coverage with Zeina,The highest Z-Value and high amount of manual annotation in starterator agrees with this start. HHpred shows 91% probability for DmmA-like-C, which is not shown on the SEA-Phage guide. This will remain hypothetical due to no strong evidence. |
| Function- Hypothetical protein |
| **Synteny similar**  **-Near PnuC-like Nicotinamide riboside transporter, like Zeina** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 43930 |
| Glimmer | 43595 |
| GeneMark | 43595 |
| Length(s) | 336 |
| Overlap(s) | 1 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 (overlap) |
| Best Z-score? | 2.753 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~43610 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham198469Report.pdf> | |

|  |
| --- |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny similar**  **-Near PnuC-like Nicotinamide riboside transporter, like Zeina** |
| **BLASTp (Identified function)**  **100% aligned with Zenia for hypothetical protein** |
| **HHpred**    **HHpred (website)** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 44043 |
| Called at 43930 by both Glimmer and GeneMark. Captures full coding potential. Starterator report shows some annotation of this gene and start. Blast results match with Zeina for a hypothetical protein. HHPred has produced no significant results |
| Function: Hypothetical Protein |
| Gene appears shifted around in other cluster mates. With dissimilar surrounding genes compared to Trixie |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 44043 |
| Glimmer | 43930 |
| GeneMark | 43930 |
| Length(s) | 114 |
| Overlap(s) | 1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 2.957 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~43940 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham10257Report.pdf>  Gene: TrixiePhattel\_72 Start: 43930, Stop: 44043, Start Num: 7  Candidate Starts for TrixiePhattel\_72:  (3, 43900), (Start: 7 @43930 has 3 MA's), (8, 43963),  Start 7:  • Found in 14 of 18 ( 77.8% ) of genes in pham  • Manual Annotations of this start: 3 of 5  • Called 92.9% of time when present  • Phage (with cluster) where this start called: Argan\_67 (AU6), Arzan\_62 (FI),  BarbieDoll\_75 (AU6), Biscute\_74 (AU6), Jazzy4900\_61 (FI), KevinMinion\_69 (AU6),  Kinny\_75 (AU6), Lewando\_75 (AU6), Phaila\_72 (AU6), Renaldo\_79 (AU6),  Sunny4976\_60 (FI), TrixiePhattel\_72 (AU6), Zeina\_72 (AU6), | |
| Blast | |
|  | |

|  |
| --- |
| Gene Information Stop 44473 |
|  |
| Start kept at 44033, both called by glimmer and genemark. BLAST shows 100% alignment with Argan. Also, the highest Z-value was chosen due to this start. HHpred shows a 99% probability for Ryanodine, but no significant evidence for calling it. |
| Function: Hypothetical protein |
| **-To the left of DNA helicase like Argan.**  **-Synteny mixed** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 44473 |
| Glimmer | 44033 |
| GeneMark | 44033 |
| Length(s) | 441 |
| Overlap(s) | 11 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 11 (overlap) |
| Best Z-score? | 3.053 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~44040 |
| Staterator Screenshot | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-To the left of DNA helicase**  **-Synteny mixed** |
| **BLASTp (Identified function)**  **100% aligned with Argan for hypothetical protein** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence? -High hits but no other significant data for calling Ryanodine** |

|  |
| --- |
| Gene Information Stop 44473 |
|  |
| Start kept at 44033, both called by glimmer and genemark. BLAST shows 100% alignment with Argan. Also, the highest Z-value was chosen due to this start. HHpred shows a 99% probability for Ryanodine, but no significant evidence for calling it. |
| Function: Hypothetical protein |
| **-To the left of DNA helicase like Argan.**  **-Synteny mixed** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 44473 |
| Glimmer | 44033 |
| GeneMark | 44033 |
| Length(s) | 441 |
| Overlap(s) | 11 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 11 (overlap) |
| Best Z-score? | 3.053 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~44040 |
| Staterator Screenshot | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-To the left of DNA helicase**  **-Synteny mixed** |
| **BLASTp (Identified function)**  **100% aligned with Argan for hypothetical protein** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence? -High hits but no other significant data for calling Ryanodine** |

|  |
| --- |
| Gene Information Stop 44473 |
|  |
| Start kept at 44033, both called by glimmer and genemark. BLAST shows 100% alignment with Argan. Also, the highest Z-value was chosen due to this start. HHpred shows a 99% probability for Ryanodine, but no significant evidence for calling it. |
| Function: Hypothetical protein |
| **-To the left of DNA helicase like Argan.**  **-Synteny mixed** |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 44473 |
| Glimmer | 44033 |
| GeneMark | 44033 |
| Length(s) | 441 |
| Overlap(s) | 11 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 11 (overlap) |
| Best Z-score? | 3.053 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~44040 |
| Staterator Screenshot | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-To the left of DNA helicase**  **-Synteny mixed** |
| **BLASTp (Identified function)**  **100% aligned with Argan for hypothetical protein** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence? -High hits but no other significant data for calling Ryanodine** |

|  |
| --- |
| Gene Information Stop 45361 |
| Start changed to 44870, originally called at 44885 by glimmer, and no genemark call. New BLAST showed a 1:1 match. New start has significantly more manual annotations in the starterator. HHpred shows 100% probability for HNH endonuclease. H-N-H within 30 A.A |
| Function: HNH endonuclease |
| **HNH endonuclease in pham are similar** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 45361 |
| Glimmer | 44885 |
| GeneMark | N/A |
| Length(s) | 477 |
| Overlap(s) |  |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 1.663 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~44875 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham212288Report.pdf> | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-On further end of genome**  **-76 seems to be the same gene on TrixiePhattel and Lewando (mint green)** |
| **BLASTp (Identified function)**  **87% aligned with Lewando for HNH endonuclease** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 45613 |
| Called by glimmer at 45473. No genemark call. Only one available start. Captures coding potential. Starterator possess some annotations for that start and gene. Blast results produce similarity to Uzumakin with a hypothetical protein. No significant HHPred results |
| Function: Hypothetical Protein |
| Trixe\_Phattel seems to have two unknown genes which don’t match to the right of the pham. Locations of this gene on other clustermates seems more spread out. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 45613 |
| Glimmer | 45473 |
| GeneMark | NA |
| Length(s) | 141 |
| Overlap(s) | -111 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -111 |
| Best Z-score? | 1.921 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~45510 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham194581Report.pdf>  Gene: TrixiePhattel\_76 Start: 45473, Stop: 45613, Start Num: 3  Candidate Starts for TrixiePhattel\_76:  (Start: 3 @45473 has 2 MA's),  Start 3:  • Found in 14 of 14 ( 100.0% ) of genes in pham  • Manual Annotations of this start: 2 of 5  • Called 71.4% of time when present  • Phage (with cluster) where this start called: BarbieDoll\_78 (AU6), Biscute\_77 (AU6),  Kinny\_78 (AU6), Lewando\_78 (AU6), Navi1117\_78 (AU6), Phaila\_76 (AU6),  Renaldo\_82 (AU6), Tenney120\_77 (AU6), TrixiePhattel\_76 (AU6), Zeina\_77 (AU6) | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 45984 |
| Start changed to 45610, originally called at 45838 by glimmer and genemark. Small overlap. Coding potential captured. Best Z-value was picked. BLAST showed 95% alignment with GantcherGoblin. HHpred showed no significant hits for protein being of another function. |
| Function: Hypothetical protein |
| -After DNA helicase in match with GantcherGoblin. synteny a bit mixed |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction |  |
| Stop | 45984 |
| Glimmer | 45838 |
| GeneMark | 45838 |
| Length(s) | 375 |
| Overlap(s) | 4 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 2.998 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | 45780 |
| Staterator Screenshot | |
| Blast | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny misaligned (Gene changed)** |
| **BLASTp (Identified function)**  **95% aligned with GancherGoblin** |
| **HHpred** |
| **Conserved Domain**  **N/A** |
| **DeepTMHMM** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 46130 |
| Uncalled by glimmer or genemark. Start found via coding potential. No available starterator. Blast results show strong match with Uzumaki, Ziena for hypothetical protein. No significant results from HHPred |
| Function: Hypothetical Protein |
| Gene not listed on Pham for synteny analysis |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 46130 |
| Glimmer | NA |
| GeneMark | NA |
| Length(s) | 180 |
| Overlap(s) | 13 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 2, Z=2.622, 12 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~45980 |
| Staterator Screenshot | |
| Blast Screenshot | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 47374 |
| Called at 46130 by both Glimmer and Genemark. Fully captures coding potential. Starterator shows this start and gene are the most annotated. Blast and HHPred show high alignment for DNA Helicase. Blast matches Lewando, Uzumaki. |
| Function: DNA Helicase |
| Gene found in other clustermates in relatively the same spots. Leathea appears to have similar genes next to its Helicase. With 218395 on the right of it. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 47374 |
| Glimmer | 46130 |
| GeneMark | 46130 |
| Length(s) | 1245 |
| Overlap(s) | 1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 2.005 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~46160 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham704Report.pdf>  Gene: TrixiePhattel\_79 Start: 46130, Stop: 47374, Start Num: 10  Candidate Starts for TrixiePhattel\_79:  (Start: 10 @46130 has 29 MA's), (33, 46304), (39, 46358), (49, 46436), (56, 46511), (60, 46571), (63, 46601), (64, 46610), (79, 46757), (80, 46763), (103, 46901), (105, 46907), (155, 47282), (158, 47300),  Start 10:  • Found in 59 of 218 ( 27.1% ) of genes in pham  • Manual Annotations of this start: 29 of 168  • Called 84.7% of time when present  • Phage (with cluster) where this start called: Acai\_70 (AU1), AreFloNak\_77 (AU1),  Argan\_75 (AU6), Arzan\_75 (FI), BarbieDoll\_80 (AU6), Biscute\_79 (AU6),  Bouchard\_77 (AU2), Breylor17\_68 (AU1), Brunswick\_73 (AU1), CapnMurica\_69  (AU1), CastorTray\_75 (AU1), Caterpillar\_71 (AU4), Darby\_69 (AU1), DevitoJr\_71  (AU1), ElephantMan\_72 (AU1), Elver\_138 (FK), Gandionco\_136 (FK),  GantcherGoblin\_74 (AU6), Gordon\_69 (AU1), Issa\_70 (AU1), Jazzy4900\_73 (FI),  KevinMinion\_74 (AU6), Kinny\_80 (AU6), Leathea\_79 (AU6), Lewando\_79 (AU6),  LilHuddy\_77 (AU2), LucySwiss\_72 (AU1), Makai\_75 (AU5), MediumFry\_71 (AU4),  MiniBagel\_73 (AU1), Navi1117\_81 (AU6), Nightmare\_73 (AU1), Niktson\_72 (AU1),  Nivinsha\_74 (AU1), Paella\_140 (FK), Phaby\_79 (AU2), Phaila\_78 (AU6), Qui\_140  (FK), Renaldo\_84 (AU6), ScienceWizSam\_73 (AU1), Sunny4976\_72 (FI),  Synepsis\_67 (AU1), Teacup\_73 (AU1), Tenney120\_79 (AU6), Tenno\_72 (AU1),  TrixiePhattel\_79 (AU6), Truckee\_73 (AU5), Trustiboi\_70 (AU1), Uzumaki\_75 (AU6),  Zeina\_80 (AU6), | |
| Blast | |
| Function  High Match and synteny with clustermates on DNA Helicase | |

|  |
| --- |
| Gene Information Stop 47589 |
| Start changed to 47374, which closed the gap entirely. Best Z-value available in RBS table. Slightly more manual annotations with new start. HHpred showed insignificant evidence for protein of another function. |
| Function: Hypothetical protein |
| **To the left of helix-turn-helix DNA binding domain protein. Synteny mixed** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 47589 |
| Glimmer | 47398 |
| GeneMark | 47371 |
| Length(s) | 192 (Glimmer) & 219 (Genemark) |
| Overlap(s) | 23 (Glimmer) & 4 (Genemark) (gap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1(overlap) |
| Best Z-score? | 2.060 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~47380 |
| Staterator Screenshot      <http://phages.wustl.edu/starterator/Pham218395Report.pdf> | |
| Phagesdb BLAST       |  | | --- | | **Function** | | **Context & Synteny**      **-To the left of helix-turn-helix DNA binding domain protein**  **-Synteny mixed** | | **BLASTp (Identified function)**  **71% alignment with Argan for hypothetical protein** | | **HHpred** | | **Conserved Domain N/A** | | **DeepTMHMM** |  |  | | --- | | **Other Evidence?** | | |

|  |
| --- |
| Gene Information Stop 47965 |
| Called at 47618 by Glimmer and Genemark. Fully captures coding potential. Starterator has no manual annotations for this gene. Blast shows alignment with Argan for a hypothetical protein. HHPred has produced no useable figures. |
| Function: Hypothetical Protein |
| Frame alignment is off for this gene and there doesnt seem to be any shared genes nearby it. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 47965 |
| Glimmer | 47618 |
| GeneMark | 47618 |
| Length(s) | 348 |
| Overlap(s) | -28 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -28 |
| Best Z-score? | 3.190 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~47630 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham198452Report.pdf>  Gene: TrixiePhattel\_81 Start: 47618, Stop: 47965, Start Num: 2  Candidate Starts for TrixiePhattel\_81:  (2, 47618), (23, 47879), (27, 47909), (29, 47927), (32, 47954),  Start 2:  • Found in 1 of 18 ( 5.6% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_81 (AU6), | |
| Blast Results | |
| Function  No match above 90% | |

|  |
| --- |
| Gene Information Stop 48318 |
| Start 47956 kept, originally called by glimmer and genemark. Coding potential captured. BLAST shows 97% alignment with GantcherGoblin. Overlap is small. HHpred suggested helix-turn-helix, however further investigation found **no evidence for helix-turn-helix motif.** |
| Function- Hypothetical protein |
| **-Synteny mixed**  **-Left to HNH endonuclease** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 48318 |
| Glimmer | 47965 |
| GeneMark | 47965 |
| Length(s) | 354 |
| Overlap(s) | 1 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 (overlap) |
| Best Z-score? | 2.582 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~47970 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham194359Report.pdf> | |

|  |
| --- |
|  |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Synteny mixed**  **-Left to HNH endonuclease** |
| **BLASTp (Identified function)**  **97% aligned with GancherGoblin for hypothetical protein** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?**  **-No evidence for Helix-turn-helix suggested by HHpred** |
| Gene Information Stop 48527 |
| Glimmer called at 48318. No call from genemark. Captures full coding potential. Starterator reports no manual annotations. Blast results say HNH endonuclease from gantchergoblin. HHPred reports HNH endonuclease. Examining the Amino Acid sequence shows HNH over a 30 span |
| HNH endonuclease |
| Gene appears in other clustermates but at different locations in the frame |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 48527 |
| Glimmer | 48318 |
| GeneMark | NA |
| Length(s) | 210 |
| Overlap(s) | 1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 1 |
| Best Z-score? | 2.803 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~48320 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham194131Report.pdf>  Gene: TrixiePhattel\_83 Start: 48318, Stop: 48527, Start Num: 9  Candidate Starts for TrixiePhattel\_83:  (9, 48318), (10, 48324), (15, 48378), (16, 48396), (19, 48423), (26, 48474), (28, 48480), (34, 48516),  Start 9:  • Found in 1 of 48 ( 2.1% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_83 (AU6) | |
| Blast Results | |
| Function  Match above 90% with synteny for NHN Endonuclease | |

|  |
| --- |
| Gene Information Stop 48703 |
| Gene begins at 48703. Called by glimmer at 48533. Genemark did not make a call. Starterator reports no manual annotations. Blast results shows match with gancthergobling for hypothetical protein. HHPred shows a high match with something. But not a hypothetical protein. |
| Function: Hypothetical Protein |
| There appears some close matching of the gene location on Trixie in cluster mates. And the adjucient pham 194359 on other clustermates, Tenny. |
|  |

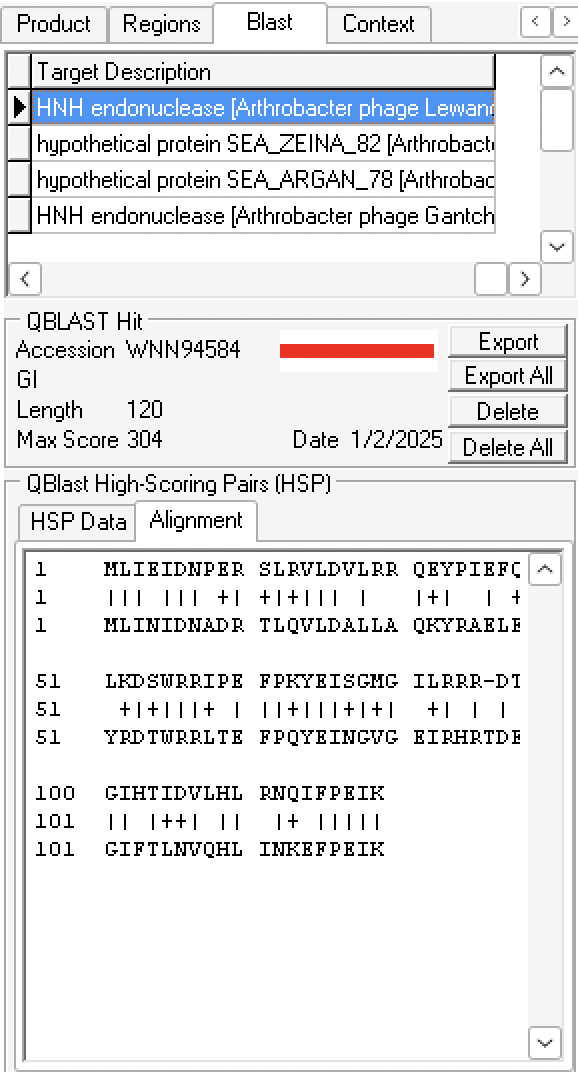
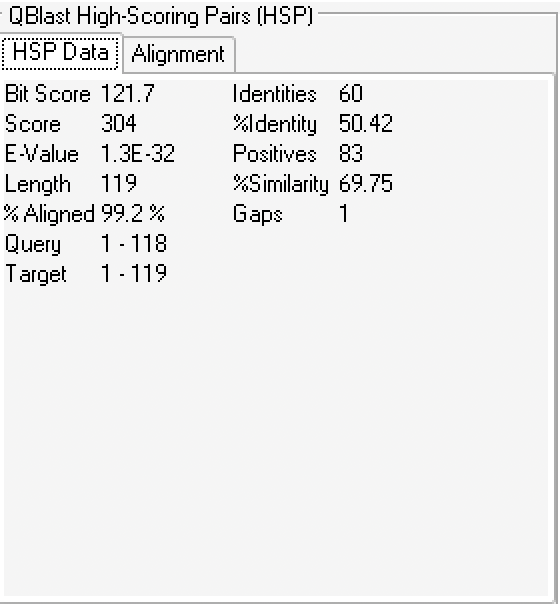
|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 48703 |
| Glimmer | 48533 |
| GeneMark | NA |
| Length(s) | 171 |
| Overlap(s) | -5 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -5 |
| Best Z-score? | 1.978 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~48540 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham221885Report.pdf>  Gene: TrixiePhattel\_84 Start: 48533, Stop: 48703, Start Num: 4  Candidate Starts for TrixiePhattel\_84:  (3, 48524), (4, 48533), (15, 48644),  Start 4:  • Found in 7 of 15 ( 46.7% ) of genes in pham  • No Manual Annotations of this start.  • Called 14.3% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_84 (AU6) | |
| Blast Results | |
| Function  High Match but no pham call | |

|  |
| --- |
| Gene Information Stop 49052 |
| Start 49052 kept, originally called by glimmer and genemark. Coding potential captured. Good amount of manually annotated starts on the starterator agree with this start. 99% coverage with Lewando. HHpred had high hits for an H-N-H endonuclease, however the H-N-H motif is missing. Upon further investigation it was found that H-N-N was a possible motif that could be used as well. |
| Function-HNH endonuclease |
| **Gene aligned with lewando, and to the right of  DNA helicase** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 49052 |
| Glimmer | 48693 |
| GeneMark | 48693 |
| Length(s) | 360 |
| Overlap(s) | 11 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 11 (overlap) |
| Best Z-score? | 2.277 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~48690 |
| Staterator Screenshot  <http://phages.wustl.edu/starterator/Pham194359Report.pdf> | |
|  | |



|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Gene aligned with lewando (light green), and near possible DNA helicase** |
| **BLASTp (Identified function)**  **99% aligned with Lewando for endonuclease** |
| **HHpred** |
| **Conserved Domain N/A** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence? Forum post about H-N-N motif** |

|  |
| --- |
| Gene Information Stop 49381 |
| Called at 49049 by Glimmer and Genemark. Captures full coding potential. Starterator. Minimal manual annotation. Blast result shows gantcher goblin match with membrane protein. HHPred shows no significant results |
| Function: Hypothetical Protein |
| Gene appears to be present in other clustermates. But no marks as to what it is. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 49381 |
| Glimmer | 49049 |
| GeneMark | 49049 |
| Length(s) | 333 |
| Overlap(s) | 4 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 |
| Best Z-score? | 2.427 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~49060 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham10743Report.pdf>  Gene: TrixiePhattel\_86 Start: 49049, Stop: 49381, Start Num: 4  Candidate Starts for TrixiePhattel\_86:  (Start: 4 @49049 has 2 MA's), (5, 49070), (14, 49169), (17, 49190), (23, 49370)  Start 4:  • Found in 13 of 14 ( 92.9% ) of genes in pham  • Manual Annotations of this start: 2 of 5  • Called 69.2% of time when present  • Phage (with cluster) where this start called: Biscute\_84 (AU6), KevinMinion\_78  (AU6), Kinny\_85 (AU6), Leathea\_83 (AU6), Lewando\_82 (AU6), Navi1117\_85 (AU6),  Renaldo\_87 (AU6), TrixiePhattel\_86 (AU6), Zeina\_83 (AU6), | |
| Blast Result | |
| Function  No match above 90% and no marks on Pham. | |

|  |
| --- |
| Gene Information Stop 50142 |
| Start 49378 kept, originally called by glimmer and genemark. Coding potential is captured and BLAST shows 99% alignment with Lewando. Good Z-value and Good amount of manually annotated starts agree with this start. HHpred suggests protein associated with H-N-H endonuclease. H-N-H motif within 30 A.A |
| Function-Endonuclease |
| **Synteny is a bit mixed but BLAST results show alignment. Both also show to be after DNA helicase and before Helix-turn-Helix DNA binding domain protein.** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 50142 |
| Glimmer | 49378 |
| GeneMark | 49378 |
| Length(s) | 765 |
| Overlap(s) | 4 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 (Overlap) |
| Best Z-score? | 2.049 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~49390 |
| Staterator Screenshot        <http://phages.wustl.edu/starterator/Pham5355Report.pdf> | |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Synteny is a bit mixed but BLAST results show alignment. Both also show to be after DNA helicase and before Helix-turn-Helix DNA binding domain protein.** | | **BLASTp (Identified function)**  **99% aligned with Lewando for endonuclease** | | **HHpred** | | **Conserved Domain** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | | |

|  |
| --- |
| Gene Information Stop 50315 |
| Start called at 50139 by Glimmer and GeneMark. Fully captures coding potential. Starterator shows little manual annotations. Blast results match uzumaki for a hypothetical protein.No significant HHPred Results. |
| Function: Hypothetical Protein |
| Gene appears at similar location as uzumaki. But it appears the gene in Trixie is smaller |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 50315 |
| Glimmer | 50139 |
| GeneMark | 50139 |
| Length(s) | 177 |
| Overlap(s) | 4 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 |
| Best Z-score? | 2.957 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~51045 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham198601Report.pdf>  Gene: TrixiePhattel\_88 Start: 50139, Stop: 50315, Start Num: 20  Candidate Starts for TrixiePhattel\_88:  (20, 50139), (22, 50166), (24, 50193), (27, 50295),  Start 20:  • Found in 1 of 10 ( 10.0% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_88 (AU6), | |
| Blast Results | |
| Function  No match above 90% nothing conclusive on Pham | |

|  |
| --- |
| Gene Information Stop 50532 |
| Call made at 50308 by Glimmer and Genemark. Fully captures coding potential.Starterator produces minimal annotation for this gene. Blast results match Zeina membrane proteins. HHPred confirms membrane protein |
| Function: Membrane Protein |
| Synteny between Uzumaki gene location and Trixie location for this membrane protein |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | forward |
| Stop | 50532 |
| Glimmer | 50308 |
| GeneMark | 50308 |
| Length(s) | 225 |
| Overlap(s) | 8 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 8 |
| Best Z-score? | 4, Z=2.722, 11 |
| Gene Mark Screenshot | |
| Estimate Start of Coding potential | ~50310 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham10540Report.pdf>  Gene: TrixiePhattel\_89 Start: 50308, Stop: 50532, Start Num: 8  Candidate Starts for TrixiePhattel\_89:  (5, 50278), (6, 50290), (7, 50293), (Start: 8 @50308 has 2 MA's), (18, 50398), (20, 50410), (21, 50413),(23, 50419), (24, 50428), (25, 50434), (27, 50452), (28, 50464), (29, 50467), (30, 50488), (32, 50503),(33, 50512),  Start 8:  • Found in 9 of 11 ( 81.8% ) of genes in pham  • Manual Annotations of this start: 2 of 4  • Called 55.6% of time when present  • Phage (with cluster) where this start called: Kinny\_89 (AU6), Tenney120\_89 (AU6),  TrixiePhattel\_89 (AU6), Uzumaki\_86 (AU6), Zeina\_87 (AU6), | |
| Blast Results | |
| Function  Match above 90% but no results on Pham    Deep TMHMM | |

|  |
| --- |
| Gene Information Stop 50926 |
| Start 50519 kept, originally called by glimmer and genemark. Coding potential is captured. BLAST shows 97% alignment with Zenia. High Z-value agrees with this start. HHpred suggests a protein function with low probability. |
| Function- Hypothetical protein |
| **-Both are surrounded by similar genes. To the left of VRR-Nuc Domain Protein.** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 50926 |
| Glimmer | 50519 |
| GeneMark | 50519 |
| Length(s) | 408 |
| Overlap(s) | 14 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? |  |
| Best Z-score? | 2.899 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~50550 |
| Staterator Screenshot    <http://phages.wustl.edu/starterator/Pham208824Report.pdf> | |
| |  | | --- | | **Function** | | **Context & Synteny**      **-Both are surrounded by similar genes. To the left of VRR-Nuc Domain Protein.** | | **BLASTp (Identified function)**  **97% alignment with Zenia for hypothetical protein** | | **HHpred** | | **Conserved Domain N/A** | | **DeepTMHMM N/A** |  |  | | --- | | **Other Evidence?** | | |

|  |
| --- |
| Gene Information Stop 51646 |
| Not called by Glimmer or GeneMark. Fully covers coding potential.  Blast results align with Uzumaki and Argan for VRR Nuc Domain Protein. HHPred also suggests that. Alpha fold confirms structure |
| Function: VRR Nuc Domain Protein |
| Comes before the Cas4 Exonclease. Missing from trixies Pham frame. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 51646 |
| Glimmer | NA |
| GeneMark | NA |
| Length(s) | 321 |
| Overlap(s) | -402 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -402 |
| Best Z-score? | 2, 2.413 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~51360 |
| Staterator Screenshot | |
|  | |
| Blast | |
| [**UVK62911.1**](https://www.ncbi.nlm.nih.gov/protein/UVK62911.1?report=genbank&log$=protalign&blast_rank=1&RID=YBP2UHA2016) | |
| Function  Above 90% match for VRR Nuc Domain Protein. Alpha Fold structure lookalike for VRR Nuc Domain | |

|  |
| --- |
| Gene Information Stop 52184 |
| Start kept at 51612, originally called by genemark and glimmer. Highest Z-Value available, and high number of manual annotations agrees with this start. BLAST shows 100% coverage with Uzumaki. HHpred also suggests Cas4 family exonuclease |
| Function- Cas4 family exonuclease |
| **Cas4 family exonuclease at “end” of genome. Also, just to the left of  helix-turn-helix DNA binding domain** |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 52184 |
| Glimmer | 51612 |
| GeneMark | 51612 |
| Length(s) | 573 |
| Overlap(s) | 35 (overlap) |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 35 (overlap) |
| Best Z-score? | 2.026 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~51640 |
| Staterator Screenshot | |
|  | |

|  |
| --- |
| **Function** |
| **Context & Synteny**      **-Cas4 family exonuclease at “end” of genone** |
| **BLASTp (Identified function)** |
| **HHpred** |
| **Conserved Domain** |
| **DeepTMHMM N/A** |

|  |
| --- |
| **Other Evidence?** |

|  |
| --- |
| Gene Information Stop 54904 |
| Called at 52181 by Glimmer and GeneMark. Fully captures coding potential. Starterator no manual annotations of this gene. Blast matching Argan and GYM 2.0 confirm helix-turn-helix DNA binding domain |
| Function: helix-turn-helix DNA binding domain |
| Clustermates seem to all possess this as there second to last gene. After cas4. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 54904 |
| Glimmer | 52181 |
| GeneMark | 52181 |
| Length(s) | 2724 |
| Overlap(s) | 4 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | 4 |
| Best Z-score? | 3.133 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~52190 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham712Report.pdf>  Gene: TrixiePhattel\_92 Start: 52181, Stop: 54904, Start Num: 20  Candidate Starts for TrixiePhattel\_92:  (20, 52181), (22, 52190), (31, 52295), (Start: 33 @52319 has 1 MA's), (36, 52346), (50, 52448), (53,  52475), (65, 52559), (68, 52574), (69, 52589), (80, 52661), (86, 52709), (94, 52772), (97, 52799), (110, 52913), (118, 52991), (120, 53021), (121, 53027), (125, 53054), (131, 53108), (132, 53117), (151, 53204), (155, 53264), (160, 53288), (170, 53348), (178, 53420), (183, 53483), (223, 53879), (225, 53909), (230, 53963), (250, 54176), (257, 54230), (266, 54275), (273, 54326), (279, 54392), (287, 54461), (295, 54533), (306, 54626), (316, 54728), (325, 54779), (332, 54812), (341, 54887),  Start 20:  • Found in 1 of 218 ( 0.5% ) of genes in pham  • No Manual Annotations of this start.  • Called 100.0% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_92 (AU6), | |
| Blast Results | |
| Function  High Match. But Synteny and Function list Suggests helix-turn-helix DNA binding domain | |
| GYM 2.0 Results      Helix Turn Helix identified by GYM 2.0 | |

|  |
| --- |
| Gene Information Stop 55163 |
| Not called by glimmer or genemark. Found via coding potential report. Starterator produced low manual annotations for start 54906. Blast results suggest 54906 is a hypothetical protein related to Zeina. Nothing significant from HHPred. |
| Function: Hypothetical Protein |
| Gene appears to match location well with Renaldo. But all cluster mates seem to end with this gene. |
|  |

|  |  |
| --- | --- |
| Gene Calling: DNA Master |  |
| Direction | Forward |
| Stop | 55163 |
| Glimmer | NA |
| GeneMark | NA |
| Length(s) | 267 |
| Overlap(s) | -1 |
|  |  |
|  |  |

|  |  |
| --- | --- |
|  |  |
| RBS Table Screenshot | |
|  | |
| Best overlap/gap? | -1 |
| Best Z-score? | 2.957 |
| Gene Mark Screenshot | |
|  | |
| Estimate Start of Coding potential | ~54920 |
| Staterator Screenshot | |
| <http://phages.wustl.edu/starterator/Pham707Report.pdf>  Gene: TrixiePhattel\_93 Start: 54897, Stop: 55163, Start Num: 10  Candidate Starts for TrixiePhattel\_93:  (10, 54897), (Start: 13 @54906 has 24 MA's), (18, 54918), (30, 55023), (37, 55077),  Start 10:  • Found in 8 of 218 ( 3.7% ) of genes in pham  • No Manual Annotations of this start.  • Called 12.5% of time when present  • Phage (with cluster) where this start called: TrixiePhattel\_93 (AU6),  Start 13: • Found in 43 of 218 ( 19.7% ) of genes in pham • Manual Annotations of this start: 24 of 168 • Called 97.7% of time when present • Phage (with cluster) where this start called: Argan\_92 (AU6), Arzan\_97 (FI), BarbieDoll\_92 (AU6), Biscute\_91 (AU6), Boog\_91 (AW), BronxBay\_87 (AW), Camille\_77 (EL), Count\_133 (EL), Damascus\_82 (EL), DizzyRudy\_86 (EL), Djungelskog\_86 (AW), DoctorPepper\_86 (AW), Egad\_88 (AW), GantcherGoblin\_91 (AU6), HerbBucket\_87 (AW), Jazzy4900\_92 (FI), KevinMinion\_91 (AU6), Kinny\_93 (AU6), Leathea\_98 (AU6), Lewando\_89 (AU6), Linda\_87 (AW), Michelle\_87 (AW), MrAaronian\_87 (AW), Natasha\_90 (AW), Navi1117\_97 (AU6), Phaila\_96 (AU6), Powelldog\_93 (AW), ProfFrink\_87 (AW), Raunak\_88 (AW), Renaldo\_93 (AU6), RustyBoy\_90 (AW), Salk\_87 (AW), Schwartz33\_88 (DJ), Shiba\_86 (AW), Sloopyjoe\_87 (AW), Sporto\_84 (AW), StarLord\_87 (AW), Stayer\_87 (AW), | |
| Blast Results      Start 54906 | |
| Function  No match above 90% | |