

Gene 2

- Amino Acid Sequence(s):
 - MSYPDEPLSPIIEIPSPDAPDPDVWGEDGAYDERKAAVILEQAKSWSTEQK
HAMLASLRAAKARASMKVKYAHPAELAAALDPNYVVTPAIELISTSIERV
LTSPKQINLEISMPPQEGKSTLAAVATPLRALQHNPHRKIILATYALDLAETH
SRTMREWIETYGTDVVDPLTGLPVEDKIGLKLARGANKVTAWSVAGGRG
GLVAAGIGSRLTGMPADLMIIDDPFKNMMEADSALYRKRVDWVFSSVART
RLAPDASIIMIQTTRWHPEDLAGKVIAAERSLPRNERTWRVINIPAI AEKGIPD
ALKREPGTPMVSARDTPEAKRNFPKIRREVGERTWYALYQGSPRNPEGGI
FQQKWFDATRLPEAPLNPYTAVVGIDPADSGEGDEAGIIGGMAATIDKRLT
AVLTHDRSGQYTSQWAKVGVTLALEIGARVLAVEGYTTAKTYTRVVRE
AYNAIHREARKKQLAGVPLTPVEHRALPDLPPFTIKPWRGANKADAVARA
GGMSQSFETGRARTIDYALATFEQQAVDWQAGQHQPDRVAAGIIVHDTIF
DLMGGSMSIAAPPTGTNPSGGQGREIPPPAWMRRSIGKZ
- **Name/Function:** Terminase, large subunit
- **Start position:** 564
- **Coding Potential:** yes (couldnt find on pham report?)
- **Related/Similar Phages:** [Mycobacterium phage PG1](#)
- **HHPred:** [8DGC_E](#)
- **NCBI Blast:** [NP_943780.1](#)

Gene 4

- Amino Acid Sequence(s):
 - MSRRPARAALRRWLHDRV CVGATSPDPYCTGSGADEHARTQDANARRL
MACSTREEVARELHHQACTVWIIQQA HGGVYVELARNTCGPDGAHHVER
LIGATLISELC EMIGIVDKTDLGGZ
- **Name/Function:** Invasin IpaA
- **Start position:** 2,647
- **Coding Potential:** Yes
- **Related/Similar Phages:** [Mycobacterium phage Oline](#)

- **HHPred:** [5NL1_H](#)
- **NCBI Blast:** [YP_009014266.1](#)

Gene 26

- Amino Acid Sequence(s):
 - MATFNADGKTVTDEQLSPPAEYLPDLTDSEREADYQAEQARIEAAIADAA
AADDDELPEVEVLTPPAGLATTEQPGTEVATFERFNVEEVESWTYDKL
EFKGDLLGIRLPTKAALAGFSLASSKYVSLGVKNDLTGLFIARHLSPESYG
RVFSRLMDPDDVDYDVDAVGELFNAIVTAAVESDDEZ

Name/Function: [tail assembly chaperone](#)

- **Start position:** 18,858
- **Coding Potential:** Yes
- **Related/Similar Phages:** Mycobacterium phage PG1
- **HHPred:** [3FGX_B](#) (unknown protein)
- **NCBI Blast:** [NP_943804.1](#)

Gene 32

- Amino Acid Sequence(s):
 - MGVIYIGDTPISKIMAGTSPFAGKVYIGSTQVWPEVEFPLVYEDVNLTDALV
PAGATALVIEQLVGGGASGLAGQNQSGAGTYTNRGGTGGGGGAVIGQHII
PISELGPTFSLQIGAGGEQSTSDTTRNNGGPTIFSSGSIVLTAGGGSASGGVA
SQSGQWYVPMANGTAGGVNNTMGGAAGGAEGVSYSHSGNPQNDNEK
WDALSADPAGTAMQVLPGGLKPGNGGSQSGTGQRPGGGGGGGKGRSSS
SRGVGGRGGDGYARIYFINTSYKRDITRVYTTAGAWTWTPPPWAVAGTKI
DIVMFPGGRAGKNGATFGGAGRGGLAGTPVTATLTVGTDIPVSGALTGVV
GAGGASNGAAGGVSTCTTVGLTAPVNNADASGQDGGNAPNVTNLNGITYP

GGRGGSSGTSSSAGEDGQDPGGGGEGGGSLIAGLAGGAGGKGRVYVRVY
EVZ

- **Name/Function:** Minor tail protein
- **Start position:** 30,727
- **Coding Potential:** yes
- **Related/Similar Phages:** Mycobacterium phage PG1
- **HHPred:** [9D93_Pc](#)
- **NCBI Blast:** [NP_943810.1](#) (e value 0)

Gene 44

- Amino Acid Sequence(s):
 - MSVEPDPLTGAALVGDSLAAAGTDGVVLCGRLWRPVGGPVELPEVSQPDG
NADAHSSDEGGRHRVWVVRQIADEQHEQSHRPGGGHLVVVTTVRGPLD
GHRPPHQPRPGRVACHASMVSGSHPLCLPTSTSGGGSSMPDSERTDFMPR
RLATSSRLRYRGTGGASWVZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 38,394
- **Coding Potential:** yes
- **Related/Similar Phages:** Caudoviricetes sp.
- **HHPred:** [5MP3_C](#) (190 e value)
- **NCBI Blast:** [DAY89068.1](#)

Gene 65 (Reverse)

- Amino Acid Sequence(s):
 - MRLRDHVVTRDPDNQQPQEGTTTMTAKPTQDHPCITEARDAZ
- **Name/Function:** Hypothetical protein

- **Start position:** 55,073
- **Coding Potential:** Yes (low)
- **Related/Similar Phages:** Mycobacterium phage JakeO
- **HHPred:** [2P1L_F](#)
- **NCBI Blast:** [AZS08357.1](#)

Gene 76

- Amino Acid Sequence(s):
 - Glimmer:

MAMTPEPVTAPHMIEHRVRLTRDEAAAVAAGTYVPGTVPVHGAPTVIDKR
AQNAARAQAERDAQAAWVRQAAGTLWWSAQRVNSPARREKILAKRSEL
ITGALAKGLDLGGLTVZ
 - GeneMark:

MTTSTIASAATGMPSTGTSTAARTGARGIEFPALGGQRAYAVAMTPEPVTA
PHMIEHRVRLTRDEAAAVAAGTYVPGTVPVHGAPTVIDKRAQNAARAQA
ERDAQAAWVRQAAGTLWWSAQRVNSPARREKILAKRSELITGALAKGLD
LGGLTVZ
- **Name/Function:**
- **Start position:** Glimmer (59,690); GeneMark (59,567)
- **Coding Potential:** Strong at Glimmer position
- **Related/Similar Phages:** synthetic construct
- **HHPred:** [9FKD_B](#) , E value 63)
- **NCBI Blast:** (no hits)

Gene 82 (Reverse)

- Amino Acid Sequence(s):

- MKLCAACQAASDKWLDYRITPSIKIASGAARDDTVEGV RDARAARFQEW
RDTINRQQALIRKQCVSAGHVTIVDVELPGLDNTDIGGZ
- **Name/Function:** Unknown Function
- **Start position:** 62,216
- **Coding Potential:** Yes (low)
- **Related/Similar Phages:** Rabbit
- **HHPred:** [6CF2_E](#)
- **NCBI Blast:** No hits

Gene 91 (Reverse)

- Amino Acid Sequence(s):
 - Glimmer:

MVIATAPGVALRGTPVDCAQCGHEVTPVVGTFPDNTAVGTLSGRWREFWI
VPAHRAYGHGVCTGSGRQLGPVVRZ
 - GeneMark:

MIATAPGVALRGTPVDCAQCGHEVTPVVGTFPDNTAVGTLSGRWREFWIV
PAHRAYGHGVCTGSGRQLGPVVRZ
- **Name/Function:** Hypothetical protein
- **Start position:** Glimmer (64,300); GeneMark (64,297)
- **Coding Potential:** strong at glimmer position
- **Related/Similar Phages:** Glimmer (Mycobacterium phage Squid) Gene Mark (Mycobacterium phage PG1)
- **HHPred:** [2M6O_A](#)
- **NCBI Blast:** Glimmer ([ALH46192.1](#)); Gene Mark ([NP_943865.1](#))

Gene 101 (Reverse)

- Amino Acid Sequence(s):
 - MFDPADIVEPLTCYDCNGYITDHEDGCKTGLGITS DHENVLHDPEVELYAV
VSAMVD FEIERNGAYWAARRAAREARAZ
- **Name/Function:** hypothetical protein
- **Start position:** 67,631
- **Coding Potential:** yes
- **Related/Similar Phages:** Mycobacterium phage KlimbOn
- **HHPred:** [2APO_B](#) (e value 1.3)
- **NCBI Blast:** [AXQ64052.1](#)

Gene 17

- Amino Acid Sequence(s):
 - MSSVLEV VYSLPFAVGLLCGILGQRAYCYGRAWYKDRNDPLPNGRHRTV
AGISKVWVGGLIAVGSLGYVLYQAEATRLDTVSLAEHTQECTSDLIASVSR
GRQISTENDRLSIAHRDKLTELAQVQSVWLGRILDPPPHIAAMPADDPRRD
GYFKTITQFYKERTDEL RADIDKIREEQAKLIGDRERNPLPDPRCWPDPGPE
VKZ
- **Name/Function:** Membrane protein
- **Start position:** 13,706
- **Coding Potential:** 13706
- **Related/Similar Phages:**
- **HHPred:** Membrane Protein
- **NCBI Blast:** Membrane protein

Gene 21 (Reverse)

- Amino Acid Sequence(s):
 - MTARIITDADLDRMHERQAAALRAVAVEDRDAVLRHAYFLMNVSGGNLS
GNLRTAVRHLAEAGSAEALMQQDRADNARLMGTAFRARZ
- **Name/Function:** Hypotheical protein

- **Start position:** 16,670
- **Coding Potential:** 16670
- **Related/Similar Phages:**
- **HHPred:** Flagellar motor switch protein FliG; flagellar motor, switch complex, MOTOR PROTEIN; HET: MSE; 2.105A {Thermotoga mariti
- **NCBI Blast:** [hypothetical protein SEA_USAVI_21 \[Mycobacterium phage Usavi\]](#)

Gene 25

- Amino Acid Sequence(s):
 - MATIRARARIEIDEAALERESGEHLRAFHRSLTRRIANQSRVAVPVRTGNLG
RTIGELPQVYTPFRVRGGVEATADYAAPVHEGSRPHAIRARNAQYLHFWW
HGREMFRKSVWHPGTRARPFMRNSAQRVVTNDPRVRMTZ
- **Name/Function:** [tail assembly chaperone \[Mycobacterium phage PG1\]](#)
- **Start position:** 18,385
- **Coding Potential:** 18 050
- **Related/Similar Phages:**
- **HHPred:**

Antitoxin; Toxin-Antitoxin complex ribosome-dependent ribonuclease ribosome-independent

ribonuclease Toxin inhibitor, TO

-
- **NCBI Blast:**

[tail assembly chaperone \[Mycobacterium phage PG1\]](#)

Gene 45 (Reverse)

- Amino Acid Sequence(s):

- MGRAVTVQRTPHGGHHDKVTTAGAVGLFMLFISYLPDNPHPMSTALVGA
VCIGIAVRLRYLGQLDRTAYRPPQPPAQHYVVGARRERITHERRTRERIWFD
Z

- **Name/Function:** (Most to percentage isn't high) Hypothetical protein PBI
- **Start position:** 38,702
- **Coding Potential:** 38702
- **Related/Similar Phages:**
- **HHPred:** Complex I-MNLL; Electron transport chain, supercomplex, membrane protein, Tetrahymena thermophila, ELECTRON TRANSPORT; H
- **NCBI Blast:**

[hypothetical protein PBI PG1_44 \[Mycobacterium phage PG1\]](#)

•

Gene 73

- Amino Acid Sequence(s):
 - MPETGPLIKHTAVLSERHLVVVELPGRHHDVLPSEFPEIVSACERRARAWA
DAHGLELGGLGVYSSASCEGKPACVGVSFEVVGKGEIDADQSPAGHLD
RWRWSGGRLDWWRNVFNRSDQREGRKLVAREEGDNIVAFPRLASVDWV
RAAPGGAGGSGASGHTGQKSSQVSYHAVVDDRNIGELDRPEDGAQZ
- **Name/Function:**

[hypothetical protein PBI PG1_71 \[Mycobacterium phage PG1\]](#)

•

- **Start position:** 58,486
- **Coding Potential:** 58380
- **Related/Similar Phages:**
- **HHPred:** Uncharacterized protein; DUF487, PF16175 family, Structural Genomics, Joint Center for Structural Genomics, JCSG, Protei
- **NCBI Blast:**

-

Gene 78

- Amino Acid Sequence(s):
 - MSAATGGVGQFAKRGGRRGGGYRGFKSKKQWRWAWATKQPWARKKAH
ETKGGKVVRYYRRLPESKHSGHRGRRGPSRKGZ
- **Name/Function:**
- **Start position:** 60,697
- **Coding Potential:** 60 700
- **Related/Similar Phages:**
- **HHPred:** apical protein (not likely because it had a probaility of 7 perecnt)
- **NCBI Blast:** hypothetical protein sea klooper

Gene 83 (Reverse)

- Amino Acid Sequence(s):
 - MWPDLADVGTVALVGSVAGALYPLTRVLLAHIDRRRRARDNRHPEDL
RGERZ
- **Name/Function:** Hypothetical protein (HHPRED is low percentage)
- **Start position:** 62,374
- **Coding Potential:** 62374
- **Related/Similar Phages:**
- **HHPred:** Peroxisomal membrane protein PEX14; Peroxisomal transport, Import, PEX,
TRANSPORT PROTEIN, PROTEIN TRANSPORT;{Trypanosom
- **NCBI Blast:**[hypothetical protein PDRPv_86 \[Mycobacterium phage PDRPv\]](#)

Gene 85 (Reverse)

- Amino Acid Sequence(s):
 - MKFNNHTARDQNVRTLVASVLHEHINDHAVDHVEVEVLRTVDRKGERYE
RDFCGPTTNYSQGPRTLIRVEIELDERMFSDTDELVCAVIAAEQDALNAEIA
EAEAAATRALQRVEELRKRRRTVEVTNGZ
- **Name/Function:** [hypothetical protein SEA_STRUGGLE_83 \[Mycobacterium phage Struggle\]](#)
(highest percentage and cluster)
- **Start position:** 62,983
- **Coding Potential:** 62 983
- **Related/Similar Phages:**
- **HHPred:** CELL CYCLE PROTEIN GPSB; CELL CYCLE, BACTERIAL GROWTH
REGULATION, CELL WALL SYNTHESIS, CELL DIVISION; 2.8A {BACILLUS SUB
- **NCBI Blast:** [hypothetical protein SEA_STRUGGLE_83 \[Mycobacterium phage Struggle\]](#)

Gene 97 (Reverse)

- Amino Acid Sequence(s):
 - MGTKAVVQSRVINVNQTITSPTGVAKASKARETLATAREILYASGNMRAA
DAITPALHDLTRQIADYYRRRAEADKRAAQRZ
- **Name/Function:** hypothetical protein
- **Start position:** 65,933
- **Coding Potential:** 65933
- **Related/Similar Phages:**
- **HHPred:** L-19437; heterochiral, miniprotein, alpha helix, D-protein, D-peptide, mirror-image,
de novo, designer, DE NOVO PROTEIN;
- **NCBI Blast:** [hypothetical protein PBI_PG1_93 \[Mycobacterium phage PG1\]](#) (99%)

Gene 102 (Reverse)

- Amino Acid Sequence(s):
 - MRWHRENHTGGAYAAGAYRVHRAEAAALGWVAKGPGVETSHHETKVA
AQAECERALNARIDDPNTEVVPGDHLLLRDSGKVGQLAQRLNGDETPAD
RVVYAIRMPYGKTLCRVRSEFTLIDPYDGVDKPTLVGZ
- **Name/Function:** hyperthetical protein (highest percentage)
- **Start position:** 68,085
- **Coding Potential:** 68100
- **Related/Similar Phages:**
- **HHPred:** Tsr0524 protein; PetP, cytochrome b6f complex, SH3 domain, cyanobacteria, Thermosynechococcus elongatus, photosynthesis;
- **NCBI Blast:** hyperthetical protein

Gene 8

- Amino Acid Sequence(s):
 - MKTSLMG TARNEWQSEAWDFSESIGELSYVSWRANSCSRTTLIPSAIDP
DTGLPTGEVDIEEDPDAQTVADYVKGIADGPLGQAALIKRAVECMTVVGE
VWIAVLIRQEKDPVTGLAAPRARWYAVTREEIKSKAGETAELSPDGKTHE
FNRDLDSLVRINPRPRKASQATSPVRACLETLREIERTTRKIKNAAKSRV
MNNGVLFVPAEMSLPAAQAPIAGQAQIPGAPVPEVSGVPASEQLATMIYQ
ASVAAMEDENSQAAYIPLVASVAAEHLEKVQHIFGNEVTEVEIKTRIDAIT
RLAMGLDVSPERLLGMSKGNHWSAWAIGDEDVQLHIKPVMDLICQAIYN
DILTPLLAREGIDPTKYILWYDASGLTSDPDLSDAEVAHDRGAITS AALRR
LLNVGEDSGYDLTTLDGCREFAADVVTKNPELIAMYAPLLSSQLAGIEFPQ
PANAIESTREDEEDED SGARQQREPQTEDERSTEEAASLNDRAAYLVAER
LLVNRALDLAGKRRFKVNDAALKTKLRDVP AHEYH RVLPPVRSSEIPRLI
AGWDTALEDEVVASLGLDNEKLRNAV LATVRRQLTQPLIEGEVVZ
- **Name/Function:** Portal Protein PG1

- **Start position:** 4,548
- **Coding Potential:** Strong coding potential @bp: 4,548
- **Related/Similar Phages:** Portal Protein Bird nest, Saguaro, RCB2
- **HHPred:** No significant hits outside the portal protein
- **NCBI Blast:** No significant hits outside the portal protein

Gene 18

- Amino Acid Sequence(s):
 - MPGIQPVKGRRLRATKINSCGMPIAGPRNRLVTSGYVSLTLPVMRDATDL
TQDNAEGKECFADRTAPERRWWTPALELCNVNTGLLTMFTGWETLVDAN
DLPVGFRDQKEIESDFGVALELWSAGKMEEDCDEIPTSDDLTDTSSGRSY
GYFLFGGTEWTPGDITISGAVTTFTLTGRTIAMPAGLGPYNVQEAANGTA
RRLLTPTSKKEHLTVFRTRIAPPEPTPGTEPVPLATSTLFTDPNFYYGGPDDE
PAAVVAPAQSAZ
- **Name/Function:** Major Tail Protein PG1
- **Start position:** 14,408
- **Coding Potential:** No significant coding potential
- **Related/Similar Phages:** Major tail protein Birdnest, JAMaL
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside major tail protein

Gene 47 (Reverse)

- Amino Acid Sequence(s):
 - Glimmer:
MIDRPAQAGRYRHHRYRLAALFHHHNNHRETHHMTTETAVEAPVEVYP
PTYTEGLGELVRAHRMYLGLSQRDLAARLGMDRRDYQHIETGRDKCPPG
LISKLDALSDEFAEQVDKVIAEAIASGGITIAVNSNGEPQDEWDRLVAGRA
AIETDAEAPITLTIVGKQTTGRFARZ

- GeneMark:

MTTETAVEAPVEVYPPTYTEGLGELVRAHRMYLGLSQRDLAARLGMDR
RDYQHIETGRDKCPPGLISKLDALSDEFAEQVDKVIAEAIASGGITIAVSNS
GEPQDEWDRLVAGRAAIETDAEAPITLTIVGKQTTGRFARZ

- **Name/Function:** HDH DNA Binding Protein PG1
- **Start position:** Glimmer (39,423); GeneMark (39,324)
- **Coding Potential:** Strong coding potential @bp 39,423
- **Related/Similar Phages:** HTH DNA Binding Protein Kimbrough, HTH DNA binding protein Quesadilla
- **HHPred:** significant hits on Transcription Regulator (Revise)
- **NCBI Blast:** significant hits for HTH DNA binding protein, Helix-turn-helix DNA protein

Gene 48 (Reverse)

- Amino Acid Sequence(s):
 - MATPFDTIAPASDAATALLTKLMMAKAKATGADTAEAAVKIMTWLETHS
DSATVSANIDRLKAEGYAPERREIVAGADEVPAGRYAIDTTIHAINGTAFYRV
DRPTEGRWAGYVVFVKQIVGDEERRLSQKQGGTILRRIAEAGAEAAASARYG
HEIGECGVCGRTLTNDESRERGIGPICAEMGWZ
- **Name/Function:** Hypothetical Protein Banjo
- **Start position:** 39,978
- **Coding Potential:** Strong coding potential @bp 39,978
- **Related/Similar Phages:** Hypothetical Protein PG1, Vincenzo
- **HHPred:** No critical hits
- **NCBI Blast:** Critical hits on Hypothetical protein

Gene 53 (Reverse)

- Amino Acid Sequence(s):

- Glimmer:

MKPKPAQVSRSPHQHQHTSNTPTPHRKHPNMTSPAPRELRPYQREAVEA
 IESHWSQGVTRVGVVLPTGTGKSTVIGRTAVNGYQNREPVLMVAHRGELI
 DQMAGTIFEVDPSIPRSHVGIVRAEMDDHSAPIVVATLQTLATAHRREAVG
 FRRRILWDEVHHAGAEGFHTTFTTELGGYTDALFAGFTATMRRDDKGKSPV
 GLGDVIEKVVYEKDILWAIDSGYLVRPRGLTVRIKLNALDDVRTVAGDF
 QQSDLAEVMEAAATEYVVDAIKLHAADRRPIIFAASVDAAHHIADALTAAD
 FPAVAVTGSMSEAERQPVYEAYRNGTAKALVTVQVLTEGADFPMDCCV
 LARPTRSRNLYSQMIGRALRLYDGKQDALVLDLAGSSRSMKLVNLTQLVP
 GVDAAEVTEDGGSVIEIEPDDELPGEGSDPTPKLVRQGPVEMVVIDLLSGSD
 VTWCETTLGVPIPLMDDEIVFVWPKDGYPRLDANATSWAVATMSTKTGR
 GGWVSGSGAVGVAEPDYIGLEAAACEAAEAWIVNSGKRLPSKKDSWRRKQ
 PATAKQIAFARGLGIVGADAMTKAELSDEISTVKISRRLDRNITRIEAS

- GeneMark:

MRPYQREAVEAIESHWSQGVTRVGVVLPTGTGKSTVIGRTAVNGYQNREP
 VLMVAHRGELIDQMAGTIFEVDPSIPRSHVGIVRAEMDDHSAPIVVATLQT
 LATAHRREAVGFRRRILWDEVHHAGAEGFHTTFTTELGGYTDALFAGFTAT
 MRRDDKGKSPVGLGDVIEKVVYEKDILWAIDSGYLVRPRGLTVRIKLNAL
 LDDVRTVAGDFQQSDLAEVMEAAATEYVVDAIKLHAADRRPIIFAASVDA
 AHHIADALTAADFPVAVTGSMSEAERQPVYEAYRNGTAKALVTVQVLTEG
 ADFPMDCCVLARPTRSRNLYSQMIGRALRLYDGKQDALVLDLAGSSRS
 MKLVNLTQLVPGVDAAEVTEDGGSVIEIEPDDELPGEGSDPTPKLVRQGPVE
 MVVIDLLSGSDVTWCETTLGVPIPLMDDEIVFVWPKDGYPRLDANATSW
 AVATMSTKTGRGGWVSGSGAVGVAEPDYIGLEAAACEAAEAWIVNSGKRL
 PSKKDSWRRKQPATAKQIAFARGLGIVGADAMTKAELSDEISTVKISRRLD
 RNITRIEAS

- **Name/Function:** DNA Helicase PG1
- **Start position:** Glimmer (46,740); GeneMark (46623)
- **Coding Potential:** Strong coding potential on BP 46,632
- **Related/Similar Phages:** Birdnest, Thora

- **HHPred:** Critical hits on DNA Helicase
- **NCBI Blast:** Critical hits on DNA Helicase

Gene 84 (Reverse)

- Amino Acid Sequence(s):
 - MGDLSVSGVSRRADTVTVGEWVDIPHRHRAYLVVEWCGTDVDGMTYL
AGHTHDGVVPYVRTLSPGAKVNVIFEAVPZ
- **Name/Function:** Hypothetical Protein PG1
- **Start position:** 62,604
- **Coding Potential:** No critical hits
- **Related/Similar Phages:** Carthage
- **HHPred:** No critical hits
- **NCBI Blast:** Critical hits on Hypothetical Protein

Gene 86 (Reverse)

- Amino Acid Sequence(s):
 - MSSTNGRCNSAERGSSYARRARKLWLLSPEAGWGGDGETVPCWDCGVL
CEFEDLIADRIVPGEDGGRYTRDNIAPHCLLCSCRQGQRRTAAILAAKRRA
AZ
- **Name/Function:** HNH Endonuclease protein PG1 or Hypothetical Protein
TALLGRASS_83
- **Start position:** 63,329
- **Coding Potential:** No critical hits
- **Related/Similar Phages:** Endonuclease Protein Thonko
- **HHPred:** Critical hits on endonuclease protein
- **NCBI Blast:** Critical hits on endonuclease protein and hypothetical protein

Gene 92 (Reverse)

- Amino Acid Sequence(s):
 - MSATNLVPDLVAVMSAGRHARALLIAKQTPNYEGRHRAPEVEPPVRVGS
VWAKAIGGGPWZ
- **Name/Function:** Hypothetical Protein PG1
- **Start position:** 64,479
- **Coding Potential:** strong coding potential @bp 64,479
- **Related/Similar Phages:** TallGrassMM, Aelin
- **HHPred:** No critical hits
- **NCBI Blast:** Critical hits on Hypothetical Protein

Gene 95 (Reverse)

- Amino Acid Sequence(s):
 - MSTSPACDCLTSEDHTLPHATMCAAFAFNGHDTDHSQIRELAELAAYKDID
TALSILRAYRNGSLSZ
- **Name/Function:** Hypothetical Protein PG1
- **Start position:** 65,348
- **Coding Potential:** No strong coding potential
- **Related/Similar Phages:** Hypothetical protein Vista, Phipps, JecAttac
- **HHPred:** No significant hits
- **NCBI Blast:** Significant hits for Hypothetical protein

Gene 99 (Reverse)

- Amino Acid Sequence(s):

- MSAPHTEVEVRAQLALVSSLLAMPYYGWTNISTVRSDLRARKRELKRRL
AEVERNNNQREKESTAMEDPNGADEAPKVVMADASN PADLLAAAAEAEK
AAAAFRKQAAEIARAAEEARRPKMPTVHTGVSVVVTFRRYTSGREYNIA
AIGWRTGTSVRWAVTGTETRRFNWAGLLDFIGEANWHTLA EVTQTRTVL
ASGVEPPVAEVMGRYGKVLRTERPASPFGDSRYGDEFDITAPGYPGFZ
- **Name/Function:** Hypothetical Protein PG1
- **Start position:** 67,222
- **Coding Potential:** Strong coding potential @bp 67,222
- **Related/Similar Phages:** Gyarad, Malk 17, JakeO
- **HHPred:** No critical hits
- **NCBI Blast:** Critical hits for hypothetical protein

Gene 7

- Amino Acid Sequence(s):
 - MTIDNPATLALALVILMLAVARVTRLINSDKISDPIRVAIASRVRHHTLIAAE
AEAHSQSQRAAEAHRREARWAGVYEFVQCPWCIGMWVSFGGAAVMVW
VIKYSWTLDDWWALLPVALAVSHLVGLARFADTEEIEIEDAAEDELDDZ
- **Name/Function:** Hypothetical protein
- **Start position:** 3,961
- **Coding Potential:** at bp 3961
- **Related/Similar Phages:** Mycobacterium phage PG1
- **HHPred:** <https://www.rcsb.org/structure/2O4D>
- **NCBI Blast:** 97.97% hit with hypothetical protein (mycobacterium phage PG1). 97% hit with hypothetical protein (mycobacterium phage PDRPv)

Gene 24

- Amino Acid Sequence(s):

- MSTVVIEGSILPCGFLARGERVEVQRTAHIDRLIEQGFVNVISETGDPEPAPL
PEPVKAPARSASRDTWAEFLAEHTDIVTEDKGRDELVAEYDEWQSVHDVP
ATDDSEZ

- **Name/Function:** Northeast Structure Protein
- **Start position:** 18,051
- **Coding Potential:** Strong coding potential at bp 18051
- **Related/Similar Phages:**
- **HHPred:** <https://www.rcsb.org/structure/2HJQ>
- **NCBI Blast:** 100% hit with hypothetical protein (mycobacterium phage PG1)

Gene 37

- Amino Acid Sequence(s):
 - MSLKLCTGEYMLSTVAGLGPSRGWFPRVLTEQMLESTKDGEIKLSPDPVT
MIDGDLTWFNNSDDRVRVFLVHTAPRDIIAQNPSTVVIHDAWSHRVGRA
PSADFPSAIRNSFGGRLQLDRASVARDAVKFGRFFLTGDDQQAWVDLGIV
RPQQSFHFRYRAAVQTPGTWVTPSELEPRWEAYARWTRLVALGGPVGESZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 34,695
- **Coding Potential:** at bp 34695
- **Related/Similar Phages:**
- **HHPred:** no significant hits
- **NCBI Blast:** 100% match with hypothetical protein (macrobacterium phage PG1)

Gene 60 (Reverse)

- Amino Acid Sequence(s):

- MLGSTPITAILGAGIDNRDPEAVRAFVRQAADIGLSIMFIAPDSKAPVDLRT
PQKKKADDKAAQEAAREAGRDRWARVKSPAGLTLATSDKTVLDRYLKR
YLTLYAERYPEGVPVNVAVEVGGSSALVVVDCDTAAQLQRFYEAALPEDIGI
DDVPPPTVVTPGHAGKDGDLDADPSSWAHSDGGHFYFTVPDEIIPTLPRNL
GAMTWGGEDGFAVLWDKRYVLIPPSTRPEGSYEVVGRDYPITDAPWLLD
AIREAAARRAERTERAHRDSEAGELSEAIDGWAEATSWSAILEPLGWTPAP
RADACGCPVWTAPGVHASPKSATAHDTGCALGRYTEVNAPLHIWTDNPG
EPFEAWISDKGTKTSLKLQAAVALIDYGGNIGKAMDEIGLTPDLAVEVDPDI
APKQVDPDHAMTGDGDFELPAAPEPETTSTPDDPPFCPVCDTEAGAFVVD
GDGDMWHMPEGGSAGDEHLAEPAPATYAHDASCNTAHTAGPEPCPTPTS
YADDLSDLPADFGKSEPADEPSPYAAADPEQTDPDVYDSPHSGVPRIAPFS
HWRDMPPEYVIEGLLEHGGLTSIIGPPGIGKSSVVLDMACHIAVGRWQ
GRRTLKTKVLYLPGEGLAGAVQRLRAWEANHDVDLGNDLMLGDSIILKA
SNEAWGDLAAYIARHEIGLVVFDTFARQSSGLEENSATDVGVAVRRYDKL
RELTGAGVCVVHHTSKAHTDQGRGSSALNGALDSELVVRAATWDTESA
MIDGRLPGRPMEVMTTKQKNAEALAEALMMVSYEDEDAGIKAPLITGP
NGSVDPMQGEIVLARPRPEAVIETAVRIRQFVNDLPTQGATRAEILTAVQAD
AYARSRRDSAAYWKLHIQRAVDMGLRFNLIETLTGTASGSRYIPGQVGIEQ
ARTLAAAEVIDGTSZ

- **Name/Function:** DNA polymerase/primase
- **Start position:** 51,388
- **Coding Potential:** at bp 51388
- **Related/Similar Phages:**
- **HHPred:** <https://www.rcsb.org/structure/1NLF>
- **NCBI Blast:** 100% match with DNA polymerase/primase (Mycobacterium phage PG1).
99% match with DNA polymerase/primase (Caudovirivetes)

Gene 57 (Reverse)

- Amino Acid Sequence(s):
 - Glimmer: MTAVDLDTARQLIGDGVSTGALPAALEEKAERLKTGTEEZ
 - GeneMark:
MARMTAVDLDTARQLIGDGVSTGALPAALEEKAERLKTGTEEZ
- **Name/Function:**
- **Start position:** Glimmer (47,914); GeneMark (47,923)
- **Coding Potential:** at bp 47924
- **Related/Similar Phages:**
- **HHPred:** No sig hits... closest one is <https://www.rcsb.org/structure/7JSX> (60% match)
- **NCBI Blast:** 100% match with hypothetical protein (SEA Orfeu)

Gene 66 (Reverse)

- Amino Acid Sequence(s):
 - MAREICQDLADDVFDGDEQWCDITGEHHVCVMTDGHIARQAHECDCGIT
WZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 55,201
- **Coding Potential:** at bp 55201
- **Related/Similar Phages:**
- **HHPred:** No sig hits... closest one is <https://www.rcsb.org/structure/7NZ5> (60% match)
- **NCBI Blast:** 94% match with hypothetical protein (Macrobacterium phage PG1). 94% match with hypothetical protein (mycobacterium phage Solosis)

Gene 72

- Amino Acid Sequence(s):
 - MTDTPQTDTEVDTEFAELVAASIEERRGRVLLMRLAGSTFKTIAQAVGVS
IGTVRKDYDIALAAHLDDPPDMMIARQRAVITDIMRANYPSMLNGDKEA

AQTILKALDREAKLFGLDAPTRVLAQVNNDDFAVEAARLIEHIQKIDPNEL
KELARAGQPQQPRPQDRPAAQDTAVEPLDVELAEDGAQGPERDPAGSIDT
DTSEPDVEPAAAEPPVAEPDDDDYDDWSNLDDDDIAZ

- **Name/Function:** Unknown functional protein
- **Start position:** 57,786
- **Coding Potential:** at bp 57786
- **Related/Similar Phages:**
- **HHPred:** <https://www.rcsb.org/structure/2AO9>
- **NCBI Blast:** 100% match with Rnase E (Mycobacterium phage PG1)

Gene 81 (Reverse)

- Amino Acid Sequence(s):
 - MAHLNDFVARNARAAAIENRRRLEAELAADGIHTLDDLNAEIDAYFAZ
- **Name/Function:** Structural protein, unknown function.
- **Start position:** 61,898
- **Coding Potential:** at bp 61898
- **Related/Similar Phages:**
- **HHPred:** <https://www.rcsb.org/structure/3MAB>
- **NCBI Blast:** 100% match with hypothetical protein (mycobacterium phage PG1)

Gene 89 (Reverse)

- Amino Acid Sequence(s):
 - Glimmer:
MIVRPENGYDDPVSAEFGPMAGHYGKVVEIDPAPHDSATLITVEIVGRVD
GPVSPAICVDRVDPQWVFYPHELDHIDZ

- GeneMark:

MSCFAVGHRVIVRPENGYDDPVSAEFGPMAGHYGKVVEIDPAPHDSATLI
TVEIVGRVDGPVSPAICVDRVDPQWVFYPHELDHIDZ

- **Name/Function:** Transcription Regulator
- **Start position:** Glimmer (63,858); GeneMark (63,885)
- **Coding Potential:** at bp 63858
- **Related/Similar Phages:**
- **HHPred:** <https://www.rcsb.org/structure/2KSS>
- **NCBI Blast:** 100% match with hypothetical protein (mycobacterium phage PG1). 100% match with hypothetical protein (mycobacterium phage Badfish).

Gene 96 (Reverse)

- Amino Acid Sequence(s):
 - MSGRPLRGTVGSGRHRPWAYTVMPGDVKVVGARTTMWGLKHFVLRYYE
RHDEVGILLGWFPVAASPVPRRALRRRERRAALRARHDSTLRCZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 65,691
- **Coding Potential:** at bp 65691
- **Related/Similar Phages:**
- **HHPred:** No sig hits
- **NCBI Blast:** %98 match with hypothetical protein (mycobacterium phage PG1)

Gene 3

- Amino Acid Sequence(s):
 - MIETTATSVTEFDGTFESVITWRFDGLWDVAVAHMGTLCHAQVAASFED
ARFLAGYICGQHMRAESMVAAZ

- **Name/Function:** Hypothetical Protein
- **Start position:** 2,435
- **Coding Potential:** Strong coding potential @bp2435
- **Related/Similar Phages:** PG1, Hosp, KayaCho
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Hypothetical Protein

Gene 5

- Amino Acid Sequence(s):
 - MADQGFYFEVDGQRFDQRVEAMAHCDLYGYHYSIAIQWVDPNAPWAZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 3,077
- **Coding Potential:** Strong coding potential @bp3077
- **Related/Similar Phages:** KLUckY39, PG1, Vincenzo
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Hypothetical Protein

Gene 10

- Amino Acid Sequence(s):
 - Glimmer:
 MRDERVEIAHPEVLRPEVVETITSVTDYERDEIVEAVERVYADPYLRERAE
 DYVDTQRERVHRTVPMVQSKVEAAAREAAVAPAPARETDLPTTTGDLTPD
 IEVVITRQREAVAEVLTPGSNMLLEIAQLQGYQAAGVQNAAVIAAAQISPD
 AHLDDKVWIATMDGKTRDSHFAADGQRVELTGEFTVGGAGLLFPADPTG
 PPHEILNCRCRVGILARDEEIPDEVDRHTERLNGRDSVQVNRQGSQQDEIH
 RRARKGVVRARED TDGIGRVAAGGWTAPSEQDY SMAKTELFRTFTDQPL
 AFVGIETSDGRMLAPEIEFSVRTPPLPAMWCKQTGEGHTEAYTVGVLESA
 RIENGTVLGSYWLNKPEAD DAYADARHKVSRPSVDLAATEWKLTVDGR

ELTEEEWFDLPPDAHVVQTITKAELIGFTMVAKPAFGDTVIEFNATRESRD
 LAMVASIADDFRPRVYAADLFAQPNLAEPTALQMDPDTGRIFGHIACFGAC
 HRSVQNACVMAPKSPSNYAQFHTSPPVLLDDGTRLAVGRLTVGTGHADE
 RLRPGPAMAHYDNTGACWALVRAYETSVGIEVSGVAAPWATPEQIEMGL
 ASPLSGDWRDFGQGLDLIAALSVNTPGFAVRGRRTDDDGRPAALVASLAPA
 RKRASELSADEIGEIVARSVEQAFAAREAAEAAAKAAQDEAADIESLVELA
 IEKVGEPEPEPEPEKTPNDIVAELIEKAGLAZ

○ GeneMark:

MSGDRVRDERVEIAHPEVLRPEVVETITSVTDYERDEIVEAVERVYADPYL
 RERAEDYVDTQRERVHRTVPMVQSKVEAAAREAAVAPAPARETDLPTTT
 GDLTPDIEVVITRQREAVAELTPGSNMLLEIAQLQGYQAAGVQNAAVIAA
 AQISPD AHL LDKVWIATMDGKTRDSHFAADGQRVELTGEFTVGGAGLLFP
 ADPTGPPHEILNCRCRVGILARDEEIPDEVDRHTERLNGRDSVQVNRQGSQ
 QDEIHRRARKGVVRARED TDGIGRVAAGGWTAPSEQDY SMAKTELFRTF
 TDQPLAFVGIETSDGRMLAPEIEFSVRTPPLPAMWCKQTGEGHTEAYTVG
 VLESARIENGTVLGSGYWLNKPEAD DAYADARHKVSRPSVDLAATEWKL
 TVDGRELTEEEWFDLPPDAHVVQTITKAELIGFTMVAKPAFGDTVIEFNAT
 RESRDLAMVASIADDFRPRVYAADLFAQPNLAEPTALQMDPDTGRIFGHIA
 CFGACHRSVQNACVMAPKSPSNYAQFHTSPPVLLDDGTRLAVGRLTVGT
 GHADERLRPGPAMAHYDNTGACWALVRAYETSVGIEVSGVAAPWATPEQI
 EMGLASPLSGDWRDFGQGLDLIAALSVNTPGFAVRGRRTDDDGRPAALVAS
 LAPARKRASELSADEIGEIVARSVEQAFAAREAAEAAAKAAQDEAADIESL
 VELAIEKVGEPEPEPEPEKTPNDIVAELIEKAGLAZ

- **Name/Function:** Portal Protein
- **Start position:** Glimmer (6,734); GeneMark (6,719)
- **Coding Potential:** strong coding potential @bp6719
- **Related/Similar Phages:** PG1, PDRP_{xv}
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits other than Portal Protein

Gene 15

- Amino Acid Sequence(s):
 - MTGNDLSPRARLQALIPAKARETWYRVASAAVMFLAFGILDANEALW
AQFAVGLVTLVLFALIYAYTPARVALYAFVGVVGSVLLYYGITTEETWALITA
AVAQAFGIATAAAKTTPVVVARGQTSLNZ
- **Name/Function:** Holin
- **Start position:** 12,962
- **Coding Potential:** No significant coding potential
- **Related/Similar Phages:** JAMaL, Quesadilla
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Holin

Gene 23

- Amino Acid Sequence(s):
 - MACRTDPALVIVGHATEALLDWFKDGACPPLVGSTTNVRFFAGDSTPLAA
WDSHASQGCNEPFVWVRLMRRYRTQRFDPDTIGIDCNLPRVAPVEIGVGW
CALTEQEPRWSDYQREAEVSSDTAWRLEEALCQASTRLKRDDEQRLVGT
DALVPYGPEGGVIAWTAVLYSTYZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 17,517
- **Coding Potential:** Significant hit @bp17517
- **Related/Similar Phages:** PG1, Saguaro, Vincenzo
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Hypothetical Protein

Gene 28

- Amino Acid Sequence(s):

- MTDVKGISLGVEINGDDLAARLGEAVRTAVAPALDKLNRKLQETQRAYNK
TGDSAIEKSAVRQVAALKRVEEQADATRRaveraaaaSRLGGGGRGDSTA
GVGGNDNRRVDNRRIDNRRYDYRTYNDNRKIDQRRYDNRKVDNRKYDY
STKNDQSITYDYRSYHYDGMPPGTPGGGGGGFGRGGGRGGGGRGGVLG
FLTGPVGLNTIAQGAAALPAVATGVVNIVGAVQQLTQAGLALPGIFAGGAA
AIGTAVIGFQGLGDAIKALNEADADPAKLEEANKALEKLAPSAAAVAREVS
RLTVGPLRDFQKAIAQPLFEGLDKDIGTFADKALPRIQPMQKVAKAINST
FKEGLAALGDDKNLSLADQIFGNTAEGQQRANNAIRPLISAFGTLAKEGS
VFLPRLGDAITSVADRFNRFIQTNTNNGNIFRWIDEGLNGARAFGNTILNIL
KTITGLTKAAGQAKGSLSGDGLLGALERVTERWSTFTNSVEGQQKLGKF
FADGTADLHRYSDLLRTLWPTIKEIIAGFQTWGEIVLPIVSALASLTNTLAQI
PGLLEAVVVGFLAWRTLGPPIGGVANKLNAVGAAGKLGARGLLGGS LI
AGGGAIQASTDSGLGQLAGALTQVAGGALAGSVFGPAGAAVGAGVGAAI
AGITYVLNENARASRDAAAQAELAENLNRSRVAMELTSQATKTANDAL
LASGGAVDAATIGAVGDQISAIPDRLAGSYDEDTIKGVATALKDLSLTEQQ
ATIVTGAQPGFDALTARLREMGPGGAIQAQNLQAVRDATLGAANNAATA
APLLQNLAQTLGVDVPQAAAELTNAFSAVPKEVPINIDAPGGQAVFDM LV
AIGEKVSVDNNKNIVVEAPLAPAVLEQLKAIGYEVTQNNDKTITVRVDPQ
MYADTLAKLGNVGNVLRDLHGQSLGLPAVPAGPNNTAGGLFPGMPGPTG
PGGADGMVIPGYAPKQDTVNAV LAPGEGVLIPEAVRGLGGAQAIYAINSR
FRKGLSKRYYADGGVHGGTGALPGPAGIETELSVLVQIRDILAGKGGPASN
PLAATAANTGAVATATKATSGQKLGPFGTPIKDRDPAYEAAAAAISALGGD
PEKFLGQNPVDYAAAQAAQATQFGGGLGVTGAVNMSAYIEALTA FAMTG
DLSKVQGIGLNANSPVITALTSARNKAKGGLSDQQIADLVGTTLGPTGYT
GTLDSTNSSLVKSLTRYREQLMKQGGAAPAVGGASAVAPALGIPMTALPAG
AMDPI SAYAAHSGGKYSWGASDLAAGLSDCSGAVSDLVELLTKGQATG
ERLFSTADAGSVLKSLGAVEGALPGALQIGWDAGHMRATLPNGVAFESGG
GTGQGATYGGNAKGAAGMPNIMSLPVGALPAGLGMSGLPGGGVGGQGT
PVYVTNWPQGQGGGIPGLKNIFGAGIGAAGQAGTDVLGDIMGGVASLGNE
PWNKPGSYAALNTLVKEGNPLALAKAFGLDVADFSRAGGAAGELEKNEG

RGYDASGRLFSDTGALLDRTFTSLNAQLQAMREQMVDVIEQVSQKLNDS
 ALEPVVKAGVQSALETLKDSVVSASIGTALGQAAAPPIADAVSSAVSSLPID
 QSGAGNVGGNAAGVVTGALGMASGGPVSGGTPGKDSVPALLMPGEYVL
 NTSDVARLGGMAAIDSMRSRGFKRFATGGGVIGNDTVGADFFGVSVQVPPI
 GAIVNLLVRVLLKVIGVDIEVRDTLMEMTDDFRQFRGDAFKAFDAQGRLL
 NDTSGLIERSQSSEETA AEERIRILKIVIQUALIKYIIEKVIVPIAKAVANA AIQA
 GASAAGAAVNSQAPGAGGIVSSLISSAGQAGVDIAAEVGTDFALAISETLI
 DVVGEGLKSLLPDLTTGLFGGNGLAFLTDPVGSLLGGLLGGIAGLFSSSLFG
 GAATMIPGIPFDNGGLAHGEGLMLKATNDPELVLNPTETDLFTRFVKALE
 NGGFGGGSNTTVHAPITVIGGGPETAEQIENRLLKHMPZ

- **Name/Function:** Tape Measure Protein
- **Start position:** 19,994
- **Coding Potential:** No significant hits
- **Related/Similar Phages:** PG1, KayaCho, 39HC
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Tape Measure Protein

Gene 31

- Amino Acid Sequence(s):
 - MTVELPGRAPTTDGELARQFHERIRRLLENPRSVRVGPWVIATDPITGDLKA
 SRPGQTLMLGGDQPVEVVEQSLSLTKFVTKDDLDAALDGIEAGGGDNDS
 VWAQLYEKL TGILSPTNALNALANFFRLELGAPITSSRLPLLPLSHIRPISPN
 LLTDGSFDYEETLSGFPDWDWDDTVGRNKPGSAYTVADGTTHTIHSNSIE
 VGPDDRLNVEVYARWAGLVASGAAIKLDISCYRDDGSPVFSGPITLDSLTV
 SGTASGWTKLSVTNWDVPDEARHVVELTVTSGATAGVVHFDDAGCFKI
 GTMPQSYISGLVEGLNALWQGIQARIDDFLDLLDVFGGFNVGDHIGQLTD
 VVTRLQHLNPLNGLFDASKLGNIANIPMIAQERISGLVTALNEAGQGIRDAI
 VQALGGSGTGHSNTDVLNALMNIPASVVNTAIAGASNIENALQQAIDSVIA
 GAGDLIGSGFGFADMIAQLTGLRRATAGTAAAVVNLQSQVANLDP AANSE

VVDFGEFANSASPPSMFTKVSDTGAGSIILTNGVLAWNSADAVGREFYLY
 NGGPLQTDLFEVQVVLPSVPSHGIFGLDSTNYVYLIGRSDATGNNMVVAR
 LGWDEV RVYSRNSGTMTQIGPTISADDILTSGSSVSFKGGTIADPRYFTVSI
 NGDKVYEYSDEAPITVIGASNRFCGLGLEKGNNYATGRISTWAMLDGGSS
 AGSGVVAGYTNAGLTNLILWKGTQAQYDALPTKSPNGIYVVEGZ

- **Name/Function:** Minor Tail Protein
- **Start position:** 28,468
- **Coding Potential:** Strong coding potential @bp28468
- **Related/Similar Phages:** PG1, Bae, Cooper
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Minor Tail Protein

Gene 42

- Amino Acid Sequence(s):
 - MATIRRGAAAPYKAIYHGTTPIKQVRRGETLLWSQGKIRDDFTDILDRWL
 NELLSGDLGALCSDVTGVLRDGLGNIVGHTVAFVEDNINGLGKLVAQTGQ
 DLATAYCGVWGGTAAPNGLIGLINGIPIFGGILADWLKGELDIVSIIGSIPIIG
 DIGRLIGLIPDALGNLADPINYVVDAFGEVIGTITCGQFRPHGGDNEGICYT
 IGVLGDIKMLIPDGLLTNRHTSRLRHETLLDGGDDGYLEVQVADLGAPD
 YPTQVFRRYANNGSGATGVGMQFMNSQVSLVRRVGGNNTLVLPQVASYS
 PGDVLRLLEQFGDIHSITRNGETVGVWEDNSATAAKGVNNRSVAMIMQGA
 KELNTRKVSPLAYLEAAZ
- **Name/Function:** Minor Tail Protein
- **Start position:** 37,074
- **Coding Potential:** Significant hit @bp37074
- **Related/Similar Phages:** PG1, LilMcDreamy, Serendipitous
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside Minor Tail Protein

Gene 68 (Reverse)

- Amino Acid Sequence(s):
 - MTNNLPAPTGPRLRDAASTPLSQAPTSVLAFALGYAAARPQHANTFGPFC
AMVEELMTRPSGATDPGPLGGYDQRTVFGSIMATIPPETARQLAMLYQAQ
RGQRWASTGHREGRAHGGRZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 68,029
- **Coding Potential:** No significant hits
- **Related/Similar Phages:** Phamished, PG1, Apizium
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits outside of Hypothetical Protein

Gene 94 (Reverse)

- Amino Acid Sequence(s):
 - MSQIVTAAQLETGDLFRIRTLTGRTTYQVDSARKVHYQDHYIEIAATKVTT
HGATTMSLAPSTPVERLEPTPQRDIVAMLPYLNPGETVSVTYWDGRFWTZ
- **Name/Function:** Hypothetical Protein
- **Start position:** 65,119
- **Coding Potential:** Significant coding potential @bp65119
- **Related/Similar Phages:** PG1, Vista, Lopsy
- **HHPred:** Significant hit for Elongation Factor P tRNA 97.35%
- **NCBI Blast:** No significant hit outside Hypothetical Protein

Gene 1

- Amino Acid Sequence(s):
 - MLVMTTNHRLIYLIGAPGAGKSTMMAHLTDCFDREHVPGTDTAPAYDLL
KYTDGEVAGVELGKRRPKFSGTDALPASIEKAIPWLNTKPFRLVFAEGAR
LANKRFLDAAVAAGYEVHLGWLNHPDVDTWRAARSAEVGKVQSESWV
KGRVTAVTKLVENYINRPERGVKVYVGGPEQLTSDMEHLLGZ
- **Name/Function:** Adenylate kinase
- **Start position:** 1
- **Coding Potential:**
- **Related/Similar Phages:** Aelin, Horchata, AngelCake
- **HHPred:** <https://www.rcsb.org/structure/3GMT> , <https://www.rcsb.org/structure/6LN3>
 - High probability (99%) on both the above and several other adenylate kinase results
- **NCBI Blast:** https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_URC18108 ,
https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_XII39753
 - High query cover on both (95%+)

Gene 9 (Reverse)

- Amino Acid Sequence(s):
 - MRDLDTFVTHTVTAHTRCLDLGLTINLRSTPPTNIDINVRQLRYRYRNTGH
CFYQGQRKQYZ
- **Name/Function:** Hypothetical protein
- **Start position:** 6,763
- **Coding Potential:**
- **Related/Similar Phages:** Khan1, Plata
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits besides hypothetical protein
[\[https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_AKF12489\]](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_AKF12489)

Gene 16

- Amino Acid Sequence(s):
 - MLTFLLLVVCIIATVVTDYFGEPNNYLVGLLGTAAGAFFAAIGSDQQKKAA
DLAQTATEASTSAARAETTAKRAEHKIDEVLHERLDDDDASTKDGAGTZ
- **Name/Function:** Hypothetical protein
- **Start position:** 13,410
- **Coding Potential:**
- **Related/Similar Phages:** PG1, Newman, Olive
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits besides hypothetical protein
[https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943794]

Gene 35

- Amino Acid Sequence(s):
 - Glimmer:
MTAPTPTPDSQWAVQWEAFAPEDSSVMPEPILPEAPRSGTDEDTPENWA
AYQEAYGIYVEAMASYESLLAVLLADDANWRTITNGFPDEALARQMVG
VIRSANAGNPQTRDFRLVWSAPIVWHIADEZ
 - GeneMark:
MQWEAFAPEDSSVMPEPILPEAPRSGTDEDTPENWAAAYQEAYGIYVEAM
ASYESLLAVLLADDANWRTITNGFPDEALARQMVG
VIRSANAGNPQTRDFRLVWSAPIVWHIADEZ
- **Name/Function:** Hypothetical protein
- **Start position:** Glimmer (33,665); GeneMark (33,704)
- **Coding Potential:**
- **Related/Similar Phages:** AngelCake, Olive, PG1
- **HHPred:** No significant hits

- **NCBI Blast:** https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943813 ,
https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_817796
 - 99% query cover on the first, 100% query cover on the second

Gene 40

- Amino Acid Sequence(s):
 - MAAHNWRALRLKALGQGITNIMACPSMHVVLDLIEQLGAEASVSEAKSQ
TEARSELTRYDLYKPDITAKVINGESYFPPPPGFSDEEVEASFDAFLSNP
MRZ
- **Name/Function:** Hypothetical protein
- **Start position:** 36,316
- **Coding Potential:**
- **Related/Similar Phages:** PG1, MelsMeow, Lulwa
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits besides hypothetical protein
[\[https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943818\]](https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943818)

Gene 46 (Reverse)

- Amino Acid Sequence(s):
 - ATGACTGATCAGACCCACGAGGCACCCCCGGTGCCTCGGTACCTCAGT
CGAGACGAGGTCGCTAAGCGCCTCGGCATGAAATCCGTTTCGCTCGCTG
TCGGGCATCGAACTGCCGCCGCCGATGTTCGAGGTCGGCAAGCACAAG
GGGTGGCTGCCCCGAGACCATCGACGCGTGGCACGCCACGCGGCCCGG
TCGCGGCTGGTGGGGCGGGCGGTGA
- **Name/Function:** Helix-turn-helix DNA binding domain protein
- **Start position:** 38,902
- **Coding Potential:**

- **Related/Similar Phages:** PG1
- **HHPred:** <https://www.rcsb.org/structure/8DGL>
 - “winged helix-turn-helix”
- **NCBI Blast:** https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943823 , https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_QPX62157
 - Although for PG1 on NCBI Blast this protein is only labeled “DNA binding protein”, on the actinobacteriophage database it’s labeled “Helix-turn-helix DNA binding domain protein”

Gene 52 (Reverse)

- Amino Acid Sequence(s):
 - MVSVARQFPAWRQPDGSVWFRPADGDGWTADPEQAHADYRVVIAPDGS
VLTWDGDPDAPQTGPTAPQTHETQQPRQETATMTAIDENGDFYPRHTEW
MRYPLPPEPPRPPSKYNGRHQYMLPHPETGRLTSYPRATTVAKTLEDTTGL
AKWNRRNVVGAVLNLVHREITGKANEMGPGGDYTAAALLAELRDVYA
NHTKVGEIDRVIETIDNVTGGAEAREFGECVHAWLEALDLGMVLLRDLPD
MVRPHALAYRRVLAHRGLVALPEYVERTVLNDLGEEATAGKIDRIFQIVST
GELVLGDVKTTKVDSMQYSWLAFVQIGGVYSWAPHVLNATGDGWERM
PKLIGIPHPDDETDDERDPFAILVHVPSDHPEKASAITIDAVWGAEALIESST
TRRMRSRAKTEVPRHAVPAPTKQSVRYAEARLALTAIESAEDGERVYTEY
QDVWDDTLTDFATTVAKLLZ
- **Name/Function:** Exonuclease
- **Start position:** 44,941
- **Coding Potential:**
- **Related/Similar Phages:** PG1
- **HHPred:** <https://www.rcsb.org/structure/5ZYT> , <https://www.rcsb.org/structure/5ZYU>
 - Both specify DNA exonuclease but I wasn’t sure since the 2nd highest hit on NCBI was a hypothetical protein so I left it as just exonuclease
- **NCBI Blast:** https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943830

Gene 54

- Amino Acid Sequence(s):
 - MSDELRAVLTEALGRSYYRIVGSSSDCRVDPGEILADAVLSLPGVAVIQLPE
PTEAVTMQDCAGSGKPFKPGTLSRDGEVAKCSACRTNRYVRDDGSIEPHQ
VPVAAVVATGEEAZ
- **Name/Function:** Hypothetical protein
- **Start position:** 47,306
- **Coding Potential:**
- **Related/Similar Phages:** Pollywog, Grand2040
- **HHPred:** No significant hits
 - The single hit I got was actually pathetic
- **NCBI Blast:** https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_AYD81122

Gene 74

- Amino Acid Sequence(s):
 - MSDDLDRQRMWNNGHTDDEVIEGTGEVFDATVLYGQTDDVDEHLADAV
KTGQLPGDPAAVAARDERTFLASMRDSTDVTLRTLGLGVWLLDNGVTH
PHRGARALLAELA EYDLQTVIGADLDTFIGEIHRLRDERGGTAZ
- **Name/Function:** Hypothetical protein
- **Start position:** 59,064
- **Coding Potential:**
- **Related/Similar Phages:** Suffolk, PG1
- **HHPred:** No significant hits
- **NCBI Blast:** No significant hits besides hypothetical proteins
 - https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_YP_009005717,
https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943850

Gene 80

- Amino Acid Sequence(s):
 - MAHLNDFVARNARAAAIENRRRLEAELAADGIHTLDDLNAEIDAYFAZ
- *Name/Function:*
- *Start position:* 61,898
- *Coding Potential:*
- *Related/Similar Phages:* PG1, Newman
- *HHPred:* No significant hits
- *NCBI Blast:* No significant hits besides hypothetical protein
[https://blast.ncbi.nlm.nih.gov/Blast.cgi#alnHdr_NP_943857]

Gene 14

- Amino Acid Sequence(s):
 - MPEGIATVEDGFATIDFVNPELRGPALAAELLEGGPGTIETISRRGPRRLYR
VPEGNAEEVGLLDGDEGPATWSAGADTGRAAALKAADPNVVGDDWH
TPILEHSSRNAYVGTVANEDVLDRTQVYTGSAHSFGGHAVAPTSTDLLDG
LDAIKYPQPPEEVPPGEGDGEDGEGDGEIEGMSAQRVNLGLAEQTSGL
ASDPGATPEVGGEALGEFTTVQSTRVEDAAAYPEGEPSDWKRAELDAYA
AAHGLDTTKLANKGEVLAAINEAADGEDDAZ
- **Name/Function:** unnamed protein product / hypothetical protein
- **Start position:** 12,120
- **Coding Potential:**
- **Related/Similar Phages:** PG1, Chah
- **HHPred:** 93.86% hit hypothetical protein yqbF, e-value: 0.095, 2HJQ_A
- **NCBI Blast:** 99.29% identical to PG1,
[https://www.ncbi.nlm.nih.gov/protein/NP_943792.1?report=genbank&log\\$=protop&blast_rank=1&RID=JC022YJD014](https://www.ncbi.nlm.nih.gov/protein/NP_943792.1?report=genbank&log$=protop&blast_rank=1&RID=JC022YJD014)

Gene 11

- Amino Acid Sequence(s):
 - MAADAMREVTLARGGTVRRIRKSDPTDDYAVAKREAELAAAATTAAZ
- **Name/Function:** hypothetical protein, possible probable small nuclear ribonucleoprotein F – splicing ; caenorhabditis elegans
- **Start position:** 9,209
- **Coding Potential:**
- **Related/Similar Phages:** PG1, Manad
- **HHPred:** e-value: 79; <https://www.rcsb.org/structure/8RO0>
- **NCBI Blast:** 100% hit on PG1 & Manad, e-value: 1E-20, reference sequence stated as “identical to PBI_MANAD_10”

Gene 30

- Amino Acid Sequence(s):
 - MSAPTRVLDRGGEAHRIVTAEQGVTHITDNGITIDQFTPRQYTSCTWGLRL
RDAGTADIVIPPTADYDRLRDIEPWAHSVTIWDVDSGTTLWTGPIHKARAS
RKGMTISAKDHSAYLSRTRNPITKRWDAADPATVAGELWAAMVEAQGIN
RPIVRVDPEGDRYDFQVIADEQMLDQTLSDMNSQGLRWTVVAGTPHIGPV
STKALALLGEHDFLGDGIEFVRDGSQTYNDVLVRAADSNTARARVDYHG
KNLQTIVNLDSMFGVSNVNRAARQYVNHTGRPHTRLELTGGTELHPNAP
VSMEDLMPSARFIIIEARGVRQLFELTSVDVERRQGAVSVRVMTMGTVEDDI
ELLDATNGKQQNMTLGGQRLZ
- **Name/Function:** Tail protein
- **Start position:** 27,359
- **Coding Potential:**
- **Related/Similar Phages:** PG1, Hosp
- **HHPred:** e-value: 2.3E-17, 99.83 possibility of tail protein ;
<https://www.rcsb.org/structure/3D37>
- **NCBI Blast:** e-value: 0.0, minor tail protein, 370/370 identities w/ PG1,

Gene 33

- Amino Acid Sequence(s):
 - MPFGWNPTPQITPRVTPRGWFKDTPELPVDREIGWWAVLTLDVLDGDA
QTMTLEALKSLALANLAASQTSLLELQQIAGISFANMIPPAQALALKKIAL
NLSGVALPTQALALAKVLGIALSSMGISTQALDLGKVVTVMATDSPA
ALLTLTKKGVLSLNSTATSTQQVVLARVASLIMSTSLATSVQSGLA
VGFPLTATTVTFTNTTNIQSWTFPRNAEWLAVVCLGGGCGGRGGGPV
LAGGGGYAGSYGTVLLRRGVDPWSTTMFYMNVGPGSAGSSGGYLVS
DPSPGTASLCTTTALATLASGSGGGPNYSTQRGQDAPDLSYNGISAV
GGKASSGSATPGGVPGGGGNGANPGSLGLPGGAGGAGGRGQIWARAFQZ

- **Name/Function:** minor tail protein
- **Start position:** 32,076
- **Coding Potential:**
- **Related/Similar Phages:** Rosebush, Chah
- **HHPred:** 99.7% probability, e-value: 8.1E-16, tail spike :
<https://www.rcsb.org/structure/9D93>
- **NCBI Blast:** 92.21% hit on Rosebush, e-value: 2E-164, “glycine-rich domain”

Gene 34

- Amino Acid Sequence(s):
 - MAVGPTAYLVNKLLDHALRGVAYTPPSVIYFKAHTGDPGAAGANNASQA
TARVAVTFLAASGGTVLLTGTPEITLNATETITHGSLWDAAGPTGGNCLWT
AQATVSKGGVSGDIIRLSGFQVGFTGLAAZ
- **Name/Function:** head protein,
- **Start position:** 33,267
- **Coding Potential:**
- **Related/Similar Phages:** Rosebush, Ralstonia
- **HHPred:** 99.96%, e-value: 5.7E-28, virion-associated phage protein :
<https://www.rcsb.org/structure/8JOU>
- **NCBI Blast:** 99% hit Rosebush, e-value: 5E-84

Gene 41

- Amino Acid Sequence(s):
 - MAVTAVLFETSGTRGNRLDPDVRDEIEALAPGLETGEVTTDKIANDAVTRE
KIDAGAVGSVELDDDAVKAQHIDDGAVGTPALAAGAVTSEKTGIGVVTAE
DASGNPVEVKQVYMTVAQYNALATKDPNTDYYLSZ
- **Name/Function:** tail fiber protein

- **Start position:** 36,667
- **Coding Potential:**
- **Related/Similar Phages:** PG1
- **HHPred:** 34.54 probability, e-value: 29, “tail collar fibers, gp4”
- **NCBI Blast:** 100% hit PG1, e-value: 1E-89

Gene 51 (Reverse)

- Amino Acid Sequence(s):
 - MAGSPFSGKGGGAATAAAPAPSKGKVDPADLPDVTNVGEDKPVGSDPYAA
ADPTGVSGYKPAFFLGQLVLMHPTEYGTMATVHDKDGKESEFVRFDIPLT
VPEPGTPNAQTAVRLDNGNFGFLNRDGEVEECEPYEVGERLDDIMVFNKA
LAREGKKALDRGIAWVLGRIVKGVKKPNQSAPFILVAGSEDDKALYQQW
RQSLAKAZ
- **Name/Function:** hypothetical protein
- **Start position:** 43,480
- **Coding Potential:**
- **Related/Similar Phages:** PG1
- **HHPred:** 28.31 probability, e-value: 140, “beta sheet heterodimer LHD29”, expression system: E.Coli
- **NCBI Blast:** 100% hit PG1, e-value: 5E-148

Gene 87 (Reverse)

- Amino Acid Sequence(s):
 - MSRPIAPGKCPHITTVVITRAGRHVEICARCRRQISZ
- **Name/Function:** hypothetical protein
- **Start position:** 63,436

- **Coding Potential:**
- **Related/Similar Phages:** Kikipoo
- **HHPred:** 77.69 probability, e-value: 1.5, transposition protein, zinc-finger protein:
<https://www.rcsb.org/structure/7OXD>
- **NCBI Blast:** 94.44% hit Kikipoo, e-value: 6E-15

Gene 100 (Reverse)

- Amino Acid Sequence(s):
 - MSLSYDPSTAYVVLGLALLACFALHFSPRGSLREAVFATLTLTLTITGIIGAA
FGAVTLZ
- **Name/Function:** hypothetical protein
- **Start position:** 67,398
- **Coding Potential:**
- **Related/Similar Phages:** PG1, Vivaldi
- **HHPred:** 75.06% probability “cellulose biosynthesis protein BcsF”, e-value: 6.5,
expression system: E.coli
- **NCBI Blast:** 100% identity PG1, e-value: 1E-31

Gene 103 (Reverse)

- Amino Acid Sequence(s):
 - MAIHIIITKLGDWNTRTTSTDGGPDHYVTVAGESNVDILNSDETMFDGMLA
AQNTLGQILWGKLSAHVACGTHYGSEHGTIDRIAVIDMHGATREEVLETV
LWVNGITQGVIZ
- **Name/Function:** hypothetical protein
- **Start position:** 68,420
- **Coding Potential:**

- ***Related/Similar Phages***: Brilliant
- ***HHPred***: 32.35% probability “cell division protein SepF”, expression system: E.Coli
- ***NCBI Blast***: 97.3% identity Brilliant; 9E-73, 96.4% identity Oline; 2E-71 : mentions from both of Bacterial, Archael and Plant Plastid.

Gene 13

- Amino Acid Sequence(s):

- Glimmer:

MTTPTPPVLDVVHYEVPPVNPSPNGLFPATTWISDPDNRRFFNGVEVRGPN
YGGEDAFGTWEGHYCSVPPMGEDDQRKDGTRPDILDAFDPITVWAYDEC
DPTAPSRAEVLARAAQILRLEEQVAIEREFANRMLTDAGTPAAAATLRLAV
GYIEGVLA LTNTLGFIHIGAQWVATDPDLFIGNGGVKRSPMGHTWVIGGG
YVDALGDTIVATSQPYGWRNEPVTRGAIDAFENNYAAIAERTVLIGYEALI
AAATITEPETDEPETPZ

- GeneMark:

MLDVVHYEVPPVNPSPNGLFPATTWISDPDNRRFFNGVEVRGPNYGGEDAF
GTWEGHYCSVPPMGEDDQRKDGTRPDILDAFDPITVWAYDECDPTAPSRA
EVLARAAQILRLEEQVAIEREFANRMLTDAGTPAAAATLRLAVGYIEGVLA
LTNTLGFIHIGAQWVATDPDLFIGNGGVKRSPMGHTWVIGGGYVDALGDT
IVATSQPYGWRNEPVTRGAIDAFENNYAAIAERTVLIGYEALIAAATITEPE
TDEPETPZ

- **Name/Function:** **hypothetical** Metabolic enzyme – dehydrogenase / oxidoreductase
- **Start position:** Glimmer (11,305); GeneMark (11,326)
- **Coding Potential:** High coding potential across the ORF according to GeneMark.

Related/Similar Phages: [Mycobacterium phage LeeLot](#) ,

[Mycobacterium phage Zonia](#)

- **HHPred:** Predicted enzymatic catalytic protein (no structural domains)
- **NCBI Blast:** Hits to **hypothetical metabolic enzymes in bacteriophages**

Gene 27 (Reverse)

- Amino Acid Sequence(s):

- MTTITRTPEEIARNYEMHMGFLKDRVEELNELLGLTDATHKATVGYIGNA
RMGRDGWDYSDCRWYVFLPHPGRVGTDDDVIGGWKRKGSLEGVMAARR
DLKTLITGVKLAKQIGZ

- **Name/Function:** Putative DNA-binding regulatory protein (transcriptional regulator)
- **Start position:** 19,824
- **Coding Potential:** **High coding potential** — clear continuous ORF with no major stop disruptions.
- **Related/Similar Phages:** [Mycobacterium phage Harvey](#) [Mycobacterium phage Zonia](#)
- **HHPred:** Predicted **helix-turn-helix DNA-binding domain**
- **NCBI Blast:** Hits to **hypothetical transcriptional regulators**

Gene 38

- Amino Acid Sequence(s):
 - MTSPDCIDPDHFRVTADGGIEPQPWMQWRHVRSIGAAGRSPTYAVTGGG
NKNVLIHGLQVSYTNDTPVPQAVYGKITRGGCRVALQARSRAYLQVASGY
QKHASDPGKLEVSSRMGCGADIGRGGTLGIGTQFCVIEQRQGMVTIPLAP
ERAGWLTGPGESITARVEVKFVSEFWENTNIDGGASGSTSGYDSGATQL
DLFAVPVLZ
- **Name/Function:** Putative enzyme / hypothetical protein
- **Start position:** 35,291
- **Coding Potential:** High coding potential
- **Related/Similar Phages:** [Mycobacterium phage Nacho](#) [Mycobacterium phage Potter](#)
- **HHPred:** Predicted **enzymatic domain**
- **NCBI Blast:** Matches to **hypothetical enzymatic proteins**

Gene 49

- Amino Acid Sequence(s):
 - Glimmer:
MRAGTYTLSSGFGPRWGSQHRGLDFAAKDGTPIYAAQGGTVAYIGRADG

FGQWIVIDHPAADGGGTTVYGHMWDAFATGLRQGQRVEAGQLIAYVGTN
 GQSTGPHLHFEVHPTVWRQGSQIDPKPWLANARNPGDPVPAPAPPKGGTL
 AKLTDPFTGELWSPNRYHPRGLGDPRWIVVHTQEGGRTARDLAAYLAQKS
 SQVSYHVVDREVLKVVAEGDAPWAAAGANKYAFHICMAGSYASWSR
 NKWLDVDTSDGKNEDLQLTKTAHVIAWWCDKYGIPPVWIGGRNIPPWGL
 DGVC GHVDLGAWGGGHTDPGNFPRDELMRRVNQFLAGTELPLPTPPP
 VTPGTPKPDQYGDWMLYRGNPRNDADRVRVQRRLKAAYRSYAGHLEI
 DGDFGPLTELAVREFQRRSLLIADGIVGPNTAAALKPZ

○ GeneMark:

MPMRAGTYTLSSGFGPRWGSQHRGLDFAAKDGTPIYAAQGGTVAYIGRA
 DGFGQWIVIDHPAADGGGTTVYGHMWDAFATGLRQGQRVEAGQLIAYV
 GTNGQSTGPHLHFEVHPTVWRQGSQIDPKPWLANARNPGDPVPAPAPPK
 GGTAKLTDPFTGELWSPNRYHPRGLGDPRWIVVHTQEGGRTARDLAAYL
 AQKSSQVSYHVVDREVLKVVAEGDAPWAAAGANKYAFHICMAGSYA
 SWSRNKWLDVDTSDGKNEDLQLTKTAHVIAWWCDKYGIPPVWIGGRNIP
 PWGLDGVC GHVDLGAWGGGHTDPGNFPRDELMRRVNQFLAGTELPL
 PTPPPVTPGTPKPDQYGDWMLYRGNPRNDADRVRVQRRLKAAYRSYAG
 HLEIDGDFGPLTELAVREFQRRSLLIADGIVGPNTAAALKPZ

- **Name/Function:** Hypothetical protein.
- **Start position:** Glimmer (40,104); GeneMark (40,098)
- **Coding Potential:** Very high coding potential
- **Related/Similar Phages:**

[Mycobacterium phage Nacho](#) [Mycobacterium phage Thora](#)

-
- **HHPred:** Predicted **structural protein domains**
- **NCBI Blast:** Hits to **phage head morphogenesis / structural proteins**

Gene 55 (Reverse)

- Amino Acid Sequence(s):
 - MENAAREALKPIRELHRPRWSNCINACCSGEDCRLRDRVCEHCEVDWPC
DTAKLIYTTEELERZ
- **Name/Function:** Putative **zinc-binding DNA-binding regulatory protein**
- **Start position:** 47,494
- **Coding Potential:** High coding potential
- **Related/Similar Phages:** [Mycobacterium phage Lego3393](#) [Mycobacterium phage Orion](#)
- **HHPred:** Predicted **zinc-finger DNA-binding protein**
- **NCBI Blast:** Hits to **hypothetical zinc-binding regulatory proteins**

Gene 62 (Reverse)

- Amino Acid Sequence(s):
 - MSENWRPVTEFEGYYEVSNQGRVRSVPRVTVRSNGVPLTVRGRIIVQRPG
NKHGHLKVSLQRDGIWSTRWVHRLVALEWCERAPGQDYVLHGPEGATV
NKATNLRWGTA AENSADRKAFGAPYPPPREMCEAGHAMTPDNIYRPPKR
PNDRHCKACQRDRVRAYRERKRS AZ
- **Name/Function:** Putative **DNA-binding transcriptional regulator**
- **Start position:** 52,626
- **Coding Potential:** High coding potential
- **Related/Similar Phages:** [Mycobacterium phage KingTut](#)

[Mycobacterium phage FugateOSU](#)

-
- **HHPred:** Predicted **DNA-binding regulatory domain**
- **NCBI Blast:** Hits to **hypothetical phage regulatory proteins**

Gene 63 (Reverse)

- Amino Acid Sequence(s):
 - MTDIIDAPVRVPSRSWSDPLLGARMHTGEHAVRFGAEFVRQAERGPANTI
ALATDIETPGLDNSFTINCVTMAWAQPDSEGRVHSILLDPRRDEQHAALVAE
MYALAGPVILHNAPFDIPPLFHAGLMTGEVIKKRIVDTVVLARFALPEPPP
YGQAKTLTALAIRYLGLSEFKGGMERAFAAAGYKTQQAGYEGMDIDSPIY
RQGAMADTVATLALPEVLRRMCIAWTLDHPFEVYGATSVAAEAAILDTQE
IVHRMMLWRSALGINVDREYLDRYLNQVADEQAIYTADLAAHGLDGGTG
KGAKLVQYLAERGELPQPWPRTPKGALRATKADLEALATVNPLAKAQLS
LANIEKTLGYLNKVSTQAAVTGRCHPQVATLGASATGRMSYGSPELQQFP
KDARPILCDDGDGLTSIDWSQIEPVTMANMARDIDFLAAYEAGEDLYGPI
MLAAGIDRPTSKVVLLSSMYGSGIAKLAALIGHTEESTAQIKRQMFAAMK
RSAKWMVKVEEVAAQFGRVITAGGRILPVDEGGIFKAVNYTVQGSAYDV
LANTIVEMYRRGMTDELYLAMHDEVVSTPVAAEVQEIMSTPPEFLIRWA
ERTPVLRTDRADMQNSWKKVZ

Name/Function: Putative **large replication/assembly enzyme (ATPase/helicase-type)**

- **Start position:** 54,494
- **Coding Potential:** Very high coding potential
- **Related/Similar Phages:** [Mycobacterium phage KLUckY39](#) [Mycobacterium phage Harvey](#)
- **HHPred:** Predicted **multi-domain enzymatic protein**
- **NCBI Blast:** Hits to **phage replication or assembly proteins**

Gene 67 (Reverse)

- Amino Acid Sequence(s):

- MFTDPRIRSAFLAGVAAGAALGVTVGAAALVLTALPDDVSVPRPGLSVTET
AATKQFAMWPADPAPTPEDDPTFDCRYHGNHICGPLNVDGYAPGLYIDGV
LVDPWPVDPASTPTIEVPZ

- **Name/Function:** Putative **membrane protein / lysis-associated protein**
- **Start position:** 55,557
- **Coding Potential:** High coding potential

Related/Similar Phages:

[Mycobacterium phage PhatCats2014](#) [Mycobacterium phage BatteryCK](#)

-
- **HHPred:** Predicted **transmembrane protein**
- **NCBI Blast:** Hits to **hypothetical membrane proteins**

Gene 70

- Amino Acid Sequence(s):
 - MGTKKTDDRRSETLHEFDAALADYLGREAGFEAAVRELA AVRTDIARIEA
RRREIHDAMKKKFAGGVTIARPAGGAPVHLYTTT VRQVRPRAVPSAVIKK
ADPVLWEKAKTWKPFVTTKVPISVSAAELGLGEYRLPPAPYKATDSIDVL
VQRYQSPVFDVLAGLREREKAAVFELNRIRETVGDWDGAAVTFADGWQV
GVRRWQYDSDELKRVAPDVWDSLAVESVITYDRLRTIDLSLAVAQGMLEI
SDLEZ
- **Name/Function:** Putative **DNA-binding transcriptional regulator**
- **Start position:** 56,495
- **Coding Potential:** High coding potential

Related/Similar Phages:

[Mycobacterium phage PG1](#) [Mycobacterium phage Chunky](#)

- **HHPred:** Predicted **helix-turn-helix regulatory protein**
- **NCBI Blast:** Hits to **phage transcriptional regulators**

Gene 77

- Amino Acid Sequence(s):
 - MTDTLDRDDLDSDIEAAPLPAWQPDHSLVPAILAGSPVTMKQLRPPDRAW
AVAVLRRMGYTAELIADWLSCSLRLVHTISADSAVLVQLYLDHTETTGTNEH
RMLASEVARLAEALADAEAAAERYRQQRDRLIavgssdkatfpcgcpRTR
YNTYVAPKTGKAGCRHHRTLAVARHRARKRQATGVASGHGQSNEGAGHP
Z
- **Name/Function:** Putative **zinc-binding DNA-binding regulatory protein**
- **Start position:** 60,036
- **Coding Potential:** High coding potential

Related/Similar Phages: [Mycobacterium phage Yoshand](#)
[Mycobacterium phage Pinkman](#)

- **HHPred:** Predicted **zinc-finger DNA-binding domain**
- **NCBI Blast:** Hits to **hypothetical zinc-binding regulatory proteins**