
[Ext] Inquiry: GenBank MW434123

gb-admin@ncbi.nlm.nih.gov <gb-admin@ncbi.nlm.nih.gov>
To: amy.kistler@czbiohub.org

Fri, Feb 19, 2021 at 9:08 AM

Dear Dr. Kistler:

We are currently processing your GenBank record MW434123 and have come across a point that requires your attention.

The following coding regions lack start codons:

gb MW434123 :CDS	maturation protein	gb MW434123 :56-1216
gb MW434123 :CDS	capsid protein	gb MW434123 :1170-1652
gb MW434123 :CDS	RNA-dependent RNA polymerase	gb MW434123 :1668->3331

Please provide the correct nucleotide spans for these features with valid start codons.

Send your reply to: gb-admin@ncbi.nlm.nih.gov

Appended for your reference is the GenBank flatfile containing the information we have currently.

Thank you for your attention, and we look forward to hearing from you soon.

Please reply using the current Subject line.

Sincerely,

Linda Frisse, PhD
The GenBank Submissions Staff
Bethesda, Maryland USA

gb-admin@ncbi.nlm.nih.gov (for replies/updates to records in GenBank)
info@ncbi.nlm.nih.gov (for general questions regarding GenBank)

Working GenBank flatfile:

```
LOCUS      MW434123          3331 bp  RNA   linear  ENV 19-FEB-2021
DEFINITION  Ulae virus isolate CMS001_026_ALCO genomic sequence.
ACCESSION  MW434123
VERSION    MW434123
DBLINK     BioProject: PRJNA605178
           BioSample: SAMN14051466
           Sequence Read Archive: SRR11035298
KEYWORDS   ENV.
SOURCE     Ulae virus (mosquito metagenome)
ORGANISM   Ulae virus
           Viruses; Riboviria; unclassified Riboviria.
REFERENCE  1 (bases 1 to 3331)
AUTHORS    Batson,J., Dudas,G., Haas-Stapleton,E., Kistler,A.L., Li,L.M.,
           Logan,P., Ratnasiri,K. and Retallack,H.
TITLE      Single mosquito metatranscriptomics identifies vectors, emerging
           pathogens and reservoirs in one assay
JOURNAL    bioRxiv (2020) In press
REMARK     DOI: 10.1101/2020.02.10.942854
```

REFERENCE 2 (bases 1 to 3331)

AUTHORS Batson,J., Dudas,G., Haas-Stapleton,E., Kistler,A.L., Li,L.M.,
Logan,P., Ratnasiri,K. and Retallack,H.

TITLE Direct Submission

JOURNAL Submitted (10-DEC-2020) Infectious Disease Initiative, CZ Biohub,
499 Illinois St, San Francisco, CA 94158, USA

COMMENT ##Assembly-Data-START##

Assembly Method :: IDSeq pipeline v. 3.2

Coverage :: 4.147203x

Sequencing Technology :: Illumina

##Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..3331

/organism="Ulae virus"

/mol_type="genomic RNA"

/isolate="CMS001_026_ALCO"

/host="Culiseta incidens"

/db_xref="taxon:2800947"

/environmental_sample

/country="USA: Alameda County, California"

/lat_lon="37.50953 N 121.93113 W"

/collection_date="19-Sep-2017"

/note="female, blood fed, collected in park habitat;
metagenomic; derived from metagenome: mosquito metagenome"

CDS 56..1216

/codon_start=1

/product="maturation protein"

/translation="-

SFLSYRGRGVFQMMKPYNNSESDYLGLKKEESLSFSGSKSPRVNGQLRLRENPFSLKKI
RHVYGRFGALNNEGFFVIKTRSSYMITVSNFNYYCSFDQRLHDKALLDAFSDLTAKVA
NVADLIRTRKQTADMVVRNLSSLVQFARDLRKGNISASSRVGSSLHLRPKTRRGEPV
GARWLEYSYGWKPLVQDIYNILDKGFGPILLEGKGYASSNETTRFSKSFSGSPSNFYGG
ESYTGGLVCSRSVKVKLKATAEASISGSAWGLDNPALLAWAIPYSFVIDWFLPVG
NYLSAATNIGYKPRIYQASITEKYTVTMNGYVGKNNRYKTYGDGAFHSSFQKQIRRS
TNIPTVPLPSFKNPISVSTRIANQLGLLNVELNKRIR"

CDS 1170..1652

/codon_start=1

/product="capsid protein"

/translation="-

NSVFLTWNSIREYVRPNSEITQMGQITPIVLKDGGKATPIAHTFGVQSPQTDPTKPATW
YESNSGSPLGYYQVTASVKFVPAGISKVRFKISLPVLATTDNSNCKDSNTPVVSYTI
ADLSFSIPTNATLDNRKDLLAFATNLLQTTMANDAVVQLQPAW"

CDS 1668..>3331

/codon_start=1

/product="RNA-dependent RNA polymerase"

/translation="-

ITVNNMSKIKSATRNIECGVNIILELLEKIDTPISLSCYLHVKYKTGQCPEIDPNQYN
TPYSLKIDFQAVAIKKNPWIFPKSSRKEDLALAKFLECEDKNRVVNDTFRRLRIPFR
ERSVLETARNFLFRILGSTPNPSYPDFGPGSTFSLKGSSTLVHKLEALPECTSSCHD
QVIETILSKMPGYALSCGMVKTNPYPNSEKSFYLAQRHLPVVKGNRLFFVPKTSQISR
PICVEPMGNLLIQKSFNGAIRTCLMRFGLDIRDGKDDSKAQALHKSLARQASIDGLNA
TIDLSSASDTISYELVKFILPTQWFDLSNCSRSAFTQLPCGTWHRNEKFSSMGNGFTF
ELETAIFFALALATAECNNTQCNPKVFGDDIIVPSSLSEPLIHILTVCGFSVNKEKTF
TSGPFRESCGGDYFLGHDVRAKYLKVNENEKPLEFGIHILNFIRSMAANLCDDVYLHP
AFKRTWQRLHKNLSTKPLCYGPVGREGVVHIGYRSYKRKQTRPQFYRKKYREDFIVF
SSTPSGELGAALYGLSSDGTVPRNCPYYYYVR"

ORIGIN

1 cttatagctc tggacctagt ccagtttgaa gaggcctaaa tgagagggtt cggctataag
61 tttcctttct tactatagag gtgttttca gatgatgaag ccgtataata actctgaatc
121 ggattactta ggtctgaaga aagaggaaac tctttcattt tcagggtcca agtctccgag
181 agttaacgga caacttcgtc ttcgtgaaaa tcctttctcc ttaaagaaga taaggcacgt
241 ctatggtagg ttggcgctt taaataacga aggcgggttt gttataaaaa ctgcgtcaag
301 ttatatgatt acggtatcaa attcaatta ttattgttca ttgatcaac gtcttcatga
361 taaagcgttg ttagacgcat ttagtgatct aacggcaaag gttgcgaatg ttgcagactt
421 aattagaact agaaagcaga ccgccgatat ggtggtcaat agactttctt ctttagttca

481 gtttccccgc gatcttcgca aaggtaatat aaagtccgcc tcttctagag tcgggtcttc
541 attacacctt agaccaaaaa cacgtcgagg cgagccggta ggcgcgcgtt ggccttagta
601 ctctatgga tggaaacctc tcgttcagga tattataat atactgaca aaggtttgg
661 tccaatccta ctgaaggaa aaggttatgc gtcgtcaaat gagacaacgc gtttctctaa
721 atctttcggg agcccttcta acttttacgg tcaggagagc tatacaggag ggctagtatg
781 ttctcgtagc gttaaagtaa aacttaaagc tactgcagaa gttccgcta ttagcgggat
841 ttctgcttg ggtcttgata accccgcatt actagcctgg gaagctattc cctactctt
901 tgttatagat tggttctcc ctgttgggaa ttacctagt gctgcgacta atataggcta
961 taagcctcgt atttatcaag cctctataac agagaaatac actgtaacca tgaatggta
1021 tgcgggtat aagaataacc gctataaac atatggcgac ggtgcattc attctagctt
1081 tcagaaacag attcgccgga gtactaatat tccgactgtt ccgcttcaa gttcaaaaa
1141 ccctatctcg gtttcgacct ggatagccaa tcaactcggg ctcttaacg tggaactcaa
1201 taagagaata cgtagacca aatagtga ttaactcaat gggacaaatt acccctattg
1261 ttttaagga cggcaaagct actcctatcg cacacacttt cgggtgacaa tcaccacaga
1321 cagacctac taaactgca acatggatg aaagtaacag tggttctcca ctaggttact
1381 accaagttac agcctcagtt aagtttggc ctgctggtat tagtaaagtg cgtttaaaa
1441 tctcgcttcc agtattagcg accactgata gtaactgtg taaagacagt aatactccag
1501 tagtctctta tactcaaatc gcagattgt ccttagtat tctactaat gcaactctgg
1561 ataaccgaaa agacttacta gcttttgcta caaatcttt acagactact atggccaatg
1621 atgctgtgt acagcttcag cggcctggg aatctattaa taattaaatt attacggta
1681 acaatatgag caagattaaa tctgtacac gcaacattga atgtggcgta aatattatcc
1741 tagagttgct ggaaaagatt gataccccga ttagcctctc ttgctatctt catgtgaagt
1801 acaagacagg ccaatgtccg gaaattgat ctaaccagta taacactcct tactcattga
1861 aaattgattt tcaagccgc gcatgatca agaagaatcc ctggattttt ccaaagtctt
1921 ctagaaaaga ggatttagct ctgtctaaat ttctgcaatg cgaggataaa aatagggttg
1981 tgaatgatac ctccgaaga ttgcggattc cattccgcga gcgatctgtt ctgaaacag
2041 ctcgtaattt tctctcaga atattagga gtactcctaa cccttctac cgggattcg
2101 gtccagggtc gacttttct ttgaaaggca gtttaagtac attagtacat aaactagagg
2161 ctttaccgga atgtacaagt agttgtcatg accagggtat tgagactata ctgtcaaaaa
2221 tgctggcta tgcgcttct tgtggtatgg taaaaactaa tccatatcca aatagcgaaa
2281 aaagcttcta ctggctcag agacatctac ccgtagttaa aggtaacaga ttatttttg
2341 ttcaaaaaac tagccagatt agtaggcaa ttgctgtga acctatggga aatctgttaa
2401 tacagaaatc ctttgtaac gcaattcgta cctgcctaag gaggtttggg ttagacattc
2461 gtatgggaa agatgattct aaggcgcaag ccttacataa atctctgcc cgtcaagcga
2521 gtatcgatgg tcttaacgca accattgatt tgtctctgc ttctgacacg atttctacg
2581 agttggtaa atttatttg ccaactcaat ggttcgattc tcttaattct ttagatctg
2641 cctttacca gctccctgc gggacctggc atagaaatga gaaattctca tctatggga
2701 atggatttac tttgaacta gagactgcaa tattcttgc cttagcttta gctactgctg
2761 aatgaataa cactcagtgt aatcctaaag ttttggcga tgatattatc gtgccgtcgt
2821 cattatcgga gcctttaata catattctca ccgtatgtg attttctgtt aataaggaaa
2881 agacttttac gtctggctct ttcagagagt cctgcggcgg ggattacttc ctcgggcatg
2941 atgtccgggc gaaatattta aaggtaatg aaaatgaaa accactgaa ttggcatcc
3001 atattctaa tttattagg agtatggctg ctaactgtg cgacgatgtt tattacatc
3061 cggcttttaa gcgaacatg caacggctcc ataaaaacct ctccacaaa ccattatgtt
3121 atgtccggt tgggagagaa ggagtcgtcc acattggcta cagaagctat aaacgtaaac
3181 aaacaggaag gccacaattc tatcggaaga aatacagaga agactttatc gtcttctcca
3241 gtacaccttc aggagaatta gggcgagcgc ttacggatt aagtagtgat ggtaccgttc
3301 ctaggaaactg ccatactat tacgtccgga g

//