**Investigating Sequence Similarity**

**Exercise 4.2: Tracking a virus**

**Instructor Preface**

This document includes:

1. A student handout for Exercise 4.2
2. A postlab worksheet containing a compilation of the questions integrated throughout the student handout for Exercise 4.2 that students can complete and turn in for a post-lab assignment
3. An appendix with the FASTA files required for this project

Instructors should distribute the student handout and the question compilation worksheet to students, and make a separate FASTA file with the sequences to distribute as a separate document.

**Investigating Sequence Similarity**

EXERCISE 4.2

**Inquiry-Based Investigation – Tracking a Virus**

**Objectives**

After completing this exercise, you should be able to:

1. Apply bioinformatics tools to address a real-life biological question

Zika virus was first characterized in 1947 and spread through both Africa and Asia for decades before increasing dramatically in its virulence in South America in 2015. Investigate whether the 2015-2016 Brazilian Zika strain evolved from:

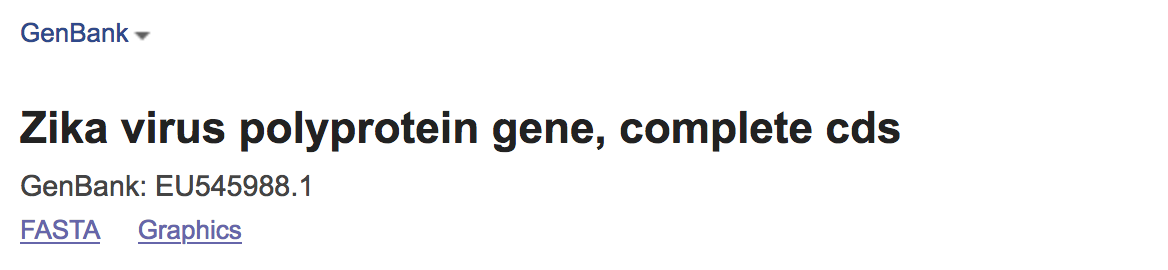
1. The Asian lineage, or
2. The African lineage..

… by aligning the full length viral genome sequences isolated from clinical samples, mosquitos, and apes. Your investigation will parallel the efforts of virologists seeking to understand the history of Zika as well as what molecular changes in the Zika genome have led to increased virulence (Wang et al., 2016).

To investigate this question, use your knowledge of bioinformatics tools and apply them to retrieving FASTA sequences, performing a multiple sequence alignment, and generating a phylogeny.

Answer Postlab question 1

1. Start by retrieving the FASTA genomic sequences using the **NCBI Genbank** accession numbers found in **Table 1** and paste them into the instructor provided **.fas** file. You will notice that a subpopulation of the sequences have already been imported into the .fas file to speed up your analysis.
   1. Use the Genbank accession numbers to look up the sequence
   2. To easily access the FASTA-formatted sequence, click on the “FASTA” link on the sequence record (see below). Then you can cut and paste the new sequences into your .fas file (we’ve done the first few for you).
      1. Remember to re-name your sequences something simple- we’ve included the Genbank #, the year, and the location at minimum.
      2. Remember to do this in a proper text editing program, i.e. TextMate (Mac) or NotePad (PC), to maintain it in .fas format.





1. Next, using MEGA, align the genomic sequences for the different isolated strains of Zika virus and create a phylogeny to reveal sequence similarity and infer phylogeny.
   1. Due to the relatively large size of the sequences (10,000+ bp), it may take a few minutes for MEGA to do the alignment and to make the phylogeny
   2. I suggest doing a standard **maximum likelihood tree** with ~100 bootstrap replicates (this may take a while)
   3. Make sure the tree has **branch lengths** (all of the tips won’t line up perfectly), **Image>Save as pdf,** and save a copy of the tree to use in your postlab.

***Answer postlab question 2***

**Table 1.** Zika viral strain genomic sequence information associated with historical host isolation descriptive data. Sequences in **bold** are already included in your .fas file; you will have to upload the remaining from Genbank

| **Year Isolated** | **Location** | **Host Isolated From** | **NCBI GenBank Accession #** | **Strain Name** |
| --- | --- | --- | --- | --- |
| **1947** | **Uganda** | **Monkey**  **(*Macaca mulatta*)** | **LC002520** | **MR766-NIID (Reference Strain)** |
| **1966** | **Malaysia** | **Mosquito**  **(*Aedes aegypti*)** | **HQ234499** | **P6-740** |
| **1976** | **Central African Republic** | **Mosquito**  **(*Aedes africanus*)** | **KF268948** | **ARB13565** |
| **2001** | **Senegal** | **Mosquito**  **(*Aedes dalzieli*)** | **KF383118** | **ArD157995** |
| 2007 | Micronesia | Human | EU545988 | FSM |
| 2015 | Brazil | Human | KU321639 | ZikaSPH2015 |
| 2016 | Brazil | Human | KU926309 | Rio-U1 |

***Answer Postlab question 3***

**References**

Wang, L., Valderramos, S. G., Wu, A., Ouyang, S., Li, C., Brasil, P., ... & Aliyari, R. (2016). From mosquitos to humans: genetic evolution of Zika virus. *Cell Host & Microbe*, 19(5), 561-565. doi: 10.1016/j.chom.2016.04.006.

**Investigating Sequence Similarity**

**Postlab**

**Exercise 4.2 – Inquiry-Based Investigation – Tracking a Virus**

1. Postlab question 1: The intro proposed 2 hypotheses (4 pts):

H1: The 2015-2016 Brazilian Zika strain evolved from the Asian lineage of Zika

H2: The 2015-2016 Brazilian Zika strain evolved from the African lineage of Zika

What would you expect a phylogeny of the Zika genomes to look like under hypothesis 1 vs hypothesis 2? Explain in words.

Postlab question 2: Upload a pdf of your tree and refer to it in your answer, **annotating on the tree** the clade with the 2015-16 Brazilian Zika genomes

P*ostlab question 3. Do the data support one of the proposed hypotheses? Explain, clearly referring to the tree in Postlab Q2 (3 pts)*

**Appendix 1. FASTA-formatted sequences for this exercise**

>LC002520\_1947\_Uganda\_Monkey\_Zika\_MR766-NIID

AGTTGTTGATCTGTGTGAGTCAGACTGCGACAGTTCGAGTCTGAAGCGAGAGCTAACAACAGTATCAACA

GGTTTAATTTGGATTTGGAAACGAGAGTTTCTGGTCATGAAAAACCCAAAGAAGAAATCCGGAGGATTCC

GGATTGTCAATATGCTAAAACGCGGAGTAGCCCGTGTAAACCCCTTGGGAGGTTTGAAGAGGTTGCCAGC

CGGACTTCTGCTGGGTCATGGACCCATCAGAATGGTTTTGGCGATACTAGCCTTTTTGAGATTTACAGCA

ATCAAGCCATCACTGGGCCTTATCAACAGATGGGGTTCCGTGGGGAAAAAAGAGGCTATGGAAATAATAA

AGAAGTTCAAGAAAGATCTTGCTGCCATGTTGAGAATAATCAATGCTAGGAAAGAGAGGAAGAGACGTGG

CGCAGACACCAGCATCGGAATCATTGGCCTCCTGCTGACTACAGCCATGGCAGCAGAGATCACTAGACGC

GGGAGTGCATACTACATGTACTTGGATAGGAGCGATGCCGGGAAGGCCATTTCGTTTGCTACCACATTGG

GAGTGAACAAGTGCCACGTACAGATCATGGACCTCGGGCACATGTGTGACGCCACCATGAGTTATGAGTG

CCCTATGCTGGATGAGGGAGTGGAACCAGATGATGTCGATTGCTGGTGCAACACGACATCAACTTGGGTT

GTGTACGGAACCTGTCATCACAAAAAAGGTGAGGCACGGCGATCTAGAAGAGCCGTGACGCTCCCTTCTC

ACTCTACAAGGAAGTTGCAAACGCGGTCGCAGACCTGGTTAGAATCAAGAGAATACACGAAGCACTTGAT

CAAGGTTGAAAACTGGATATTCAGGAACCCCGGGTTTGCGCTAGTGGCCGTTGCCATTGCCTGGCTTTTG

GGAAGCTCGACGAGCCAAAAAGTCATATACTTGGTCATGATACTGCTGATTGCCCCGGCATACAGTATCA

GGTGCATTGGAGTCAGCAATAGAGACTTCGTGGAGGGCATGTCAGGTGGGACCTGGGTTGATGTTGTCTT

GGAACATGGAGGCTGCGTTACCGTGATGGCACAGGACAAGCCAACAGTTGACATAGAGTTGGTCACGACG

ACGGTTAGTAACATGGCCGAGGTAAGATCCTATTGCTACGAGGCATCGATATCGGACATGGCTTCGGACA

GTCGTTGCCCAACACAAGGTGAAGCCTACCTTGACAAGCAATCAGACACTCAATATGTCTGCAAAAGAAC

ATTAGTGGACAGAGGTTGGGGAAACGGTTGTGGACTTTTTGGCAAAGGGAGCTTGGTGACATGTGCCAAG

TTTACGTGTTCTAAGAAGATGACCGGGAAGAGCATTCAACCGGAAAATCTGGAGTATCGGATAATGCTAT

CAGTGCATGGCTCCCAGCATAGCGGGATGACTGTCAATGATATAGGATATGAAACTGACGAAAATAGAGC

GAAAGTCGAGGTTACGCCTAATTCACCAAGAGCGGAAGCAACCTTGGGAGGCTTTGGAAGCTTAGGACTT

GACTGTGAACCAAGGACAGGCCTTGACTTTTCAGATCTGTATTACCTGACCATGAACAATAAGCATTGGT

TGGTGCACAAAGAGTGGTTTCATGACATCCCATTGCCTTGGCATGCTGGGGCAGACACTGGAACTCCACA

CTGGAACAACAAAGAGGCATTGGTAGAATTCAAGGATGCCCACGCCAAGAGGCAAACCGTCGTCGTTCTG

GGGAGCCAGGAAGGAGCCGTTCACACGGCTCTCGCTGGAGCTCTAGAGGCTGAGATGGATGGTGCAAAGG

GAAAGCTGTTCTCTGGCCATTTGAAATGCCGCCTAAAAATGGACAAGCTTAGATTGAAGGGCGTGTCATA

TTCCTTGTGCACTGCGGCATTCACATTCACCAAGGTCCCAGCTGAAACACTGCATGGAACAGTCACAGTG

GAGGTGCAGTATGCAGGGACAGATGGACCCTGCAAGATCCCAGTCCAGATGGCGGTGGACATGCAGACCC

TGACCCCAGTTGGAAGGCTGATAACCGCCAACCCCGTGATTACTGAAAGCACTGAGAACTCAAAGATGAT

GTTGGAGCTTGACCCACCATTTGGGGATTCTTACATTGTCATAGGAGTTGGGGACAAGAAAATCACCCAC

CACTGGCATAGGAGTGGTAGCACCATCGGAAAGGCATTTGAGGCCACTGTGAGAGGCGCCAAGAGAATGG

CAGTCCTGGGGGATACAGCCTGGGACTTCGGATCAGTCGGGGGTGTGTTCAACTCACTGGGTAAGGGCAT

TCACCAGATTTTTGGAGCAGCCTTCAAATCACTGTTTGGAGGAATGTCCTGGTTCTCACAGATCCTCATA

GGCACGCTGCTAGTGTGGTTAGGTTTGAACACAAAGAATGGATCTATCTCCCTCACATGCTTGGCCCTGG

GGGGAGTGATGATCTTCCTCTCCACGGCTGTTTCTGCTGACGTGGGGTGCTCAGTGGACTTCTCAAAAAA

GGAAACGAGATGTGGCACGGGGGTATTCATCTATAATGATGTTGAAGCCTGGAGGGACCGGTACAAGTAC

CATCCTGACTCCCCCCGCAGATTGGCAGCAGCAGTCAAGCAGGCCTGGGAAGAGGGGATCTGTGGGATCT

CATCCGTTTCAAGAATGGAAAACATCATGTGGAAATCAGTAGAAGGGGAGCTCAATGCTATCCTAGAGGA

GAATGGAGTTCAACTGACAGTTGTTGTGGGATCTGTAAAAAACCCCATGTGGAGAGGTCCACAAAGATTG

CCAGTGCCTGTGAATGAGCTGCCCCATGGCTGGAAAGCCTGGGGGAAATCGTATTTTGTTAGGGCGGCAA

AGACCAACAACAGTTTTGTTGTCGACGGTGACACACTGAAGGAATGTCCGCTTGAGCACAGAGCATGGAA

TAGTTTTCTTGTGGAGGATCACGGGTTTGGAGTCTTCCACACCAGTGTCTGGCTTAAGGTCAGAGAAGAT

TACTCATTAGAATGTGACCCAGCCGTCATAGGAACAGCTGTTAAGGGAAGGGAGGCCGCGCACAGTGATC

TGGGCTATTGGATTGAAAGTGAAAAGAATGACACATGGAGGCTGAAGAGGGCCCACCTGATTGAGATGAA

AACATGTGAATGGCCAAAGTCTCACACATTGTGGACAGATGGAGTAGAAGAAAGTGATCTTATCATACCC

AAGTCTTTAGCTGGTCCACTCAGCCACCACAACACCAGAGAGGGTTACAGAACCCAAGTGAAAGGGCCAT

GGCACAGTGAAGAGCTTGAAATCCGGTTTGAGGAATGTCCAGGCACCAAGGTTTACGTGGAGGAGACATG

CGGAACTAGAGGACCATCTCTGAGATCAACTACTGCAAGTGGAAGGGTCATTGAGGAATGGTGCTGTAGG

GAATGCACAATGCCCCCACTATCGTTTCGAGCAAAAGACGGCTGCTGGTATGGAATGGAGATAAGGCCCA

GGAAAGAACCAGAGAGCAACTTAGTGAGGTCAATGGTGACAGCGGGGTCAACCGATCATATGGACCACTT

CTCTCTTGGAGTGCTTGTGATTCTACTCATGGTGCAGGAGGGGTTGAAGAAGAGAATGACCACAAAGATC

ATCATGAGCACATCAATGGCAGTGCTGGTAGTCATGATCTTGGGAGGATTTTCAATGAGTGACCTGGCCA

AGCTTGTGATCCTGATGGGTGCTACTTTCGCAGAAATGAACACTGGAGGAGATGTAGCTCACTTGGCATT

GGTAGCGGCATTTAAAGTCAGACCAGCCTTGCTGGTCTCCTTCATTTTCAGAGCCAATTGGACACCCCGT

GAGAGCATGCTGCTAGCCCTGGCTTCGTGTCTTCTGCAAACTGCGATCTCTGCTCTTGAAGGTGACTTGA

TGGTCCTCATTAATGGATTTGCTTTGGCCTGGTTGGCAATTCGAGCAATGGCCGTGCCACGCACTGACAA

CATCGCTCTACCAATCTTGGCTGCTCTAACACCACTAGCTCGAGGCACACTGCTCGTGGCATGGAGAGCG

GGCCTGGCTACTTGTGGAGGGATCATGCTCCTCTCCCTGAAAGGGAAAGGTAGTGTGAAGAAGAACCTGC

CATTTGTCATGGCCCTGGGATTGACAGCTGTGAGGGTAGTAGACCCTATTAATGTGGTAGGACTACTGTT

ACTCACAAGGAGTGGGAAGCGGAGCTGGCCCCCTAGTGAAGTTCTCACAGCCGTTGGCCTGATATGTGCA

CTGGCCGGAGGGTTTGCCAAGGCAGACATTGAGATGGCTGGACCCATGGCTGCAGTAGGCTTGCTAATTG

TCAGCTATGTGGTCTCGGGAAAGAGTGTGGACATGTACATTGAAAGAGCAGGTGACATCACATGGGAAAA

GGACGCGGAAGTCACTGGAAACAGTCCTCGGCTTGACGTGGCACTGGATGAGAGTGGTGATTTCTCCTTG

GTAGAGGAAGATGGTCCACCCATGAGAGAGATCATACTTAAGGTGGTCCTGATGGCCATCTGTGGCATGA

ACCCAATAGCTATACCTTTTGCTGCAGGAGCGTGGTATGTGTATGTGAAGACTGGGAAAAGGAGTGGCGC

CCTCTGGGACGTGCCTGCTCCCAAAGAAGTGAAGAAAGGAGAGACCACAGATGGAGTGTACAGAGTGATG

ACTCGCAGACTGCTAGGTTCAACACAGGTTGGAGTGGGAGTCATGCAAGAGGGAGTCTTCCACACCATGT

GGCACGTTACAAAAGGAGCCGCACTGAGGAGCGGTGAGGGAAGACTTGATCCATACTGGGGGGATGTCAA

GCAGGACTTGGTGTCATACTGTGGGCCTTGGAAGTTGGATGCAGCTTGGGATGGACTCAGCGAGGTACAG

CTTTTGGCCGTACCTCCCGGAGAGAGGGCCAGAAACATTCAGACCCTGCCTGGAATATTCAAGACAAAGG

ACGGGGACATCGGAGCAGTTGCTCTGGACTACCCTGCAGGGACCTCAGGATCTCCGATCCTAGACAAATG

TGGAAGAGTGATAGGACTCTATGGCAATGGGGTTGTGATCAAGAATGGAAGCTATGTTAGTGCTATAACC

CAGGGAAAGAGGGAGGAGGAGACTCCGGTTGAATGTTTCGAACCCTCGATGCTGAAGAAGAAGCAGCTAA

CTGTCTTGGATCTGCATCCAGGAGCCGGAAAAACCAGGAGAGTTCTTCCTGAAATAGTCCGTGAAGCCAT

AAAAAAGAGACTCCGGACAGTGATCTTGGCACCAACTAGGGTTGTCGCTGCTGAGATGGAGGAGGCCTTG

AGAGGACTTCCGGTGCGTTACATGACAACAGCAGTCAACGTCACCCATTCTGGGACAGAAATCGTTGATT

TGATGTGCCATGCCACTTTCACTTCACGCTTACTACAACCCATCAGAGTCCCTAATTACAATCTCTACAT

CATGGATGAAGCCCACTTCACAGACCCCTCAAGTATAGCTGCAAGAGGATATATATCAACAAGGGTTGAA

ATGGGCGAGGCGGCTGCCATTTTTATGACTGCCACACCACCAGGAACCCGTGATGCGTTTCCTGACTCTA

ACTCACCAATCATGGACACAGAAGTGGAAGTCCCAGAGAGAGCCTGGAGCTCAGGCTTTGATTGGGTGAC

AGACCATTCTGGGAAAACAGTTTGGTTCGTTCCAAGCGTGAGAAACGGAAATGAAATCGCAGCCTGTCTG

ACAAAGGCTGGAAAGCGGGTCATACAGCTCAGCAGGAAGACTTTTGAGACAGAATTTCAGAAAACAAAAA

ATCAAGAGTGGGACTTTGTCATAACAACTGACATCTCAGAGATGGGCGCCAACTTCAAGGCTGACCGGGT

CATAGACTCTAGGAGATGCCTAAAACCAGTCATACTTGATGGTGAGAGAGTCATCTTGGCTGGGCCCATG

CCTGTCACGCATGCTAGTGCTGCTCAGAGGAGAGGACGTATAGGCAGGAACCCTAACAAACCTGGAGATG

AGTACATGTATGGAGGTGGGTGTGCAGAGACTGATGAAGGCCATGCACACTGGCTTGAAGCAAGAATGCT

TCTTGACAACATCTACCTCCAGGATGGCCTCATAGCCTCGCTCTATCGGCCTGAGGCCGATAAGGTAGCC

GCCATTGAGGGAGAGTTTAAGCTGAGGACAGAGCAAAGGAAGACCTTCGTGGAACTCATGAAGAGAGGAG

ACCTTCCCGTCTGGCTAGCCTATCAGGTTGCATCTGCCGGAATAACTTACACAGACAGAAGATGGTGCTT

TGATGGCACAACCAACAACACCATAATGGAAGACAGCGTACCAGCAGAGGTGTGGACAAAGTATGGAGAG

AAGAGAGTGCTCAAACCGAGATGGATGGATGCTAGGGTCTGTTCAGACCATGCGGCCCTGAAGTCGTTCA

AAGAATTCGCCGCTGGAAAAAGAGGAGCGGCTTTGGGAGTAATGGAGGCCCTGGGAACACTGCCAGGACA

CATGACAGAGAGGTTTCAGGAAGCCATTGACAACCTCGCCGTGCTCATGCGAGCAGAGACTGGAAGCAGG

CCTTATAAGGCAGCGGCAGCCCAACTGCCGGAGACCCTAGAGACCATTATGCTCTTAGGTTTGCTGGGAA

CAGTTTCACTGGGGATCTTCTTCGTCTTGATGCGGAATAAGGGCATCGGGAAGATGGGCTTTGGAATGGT

AACCCTTGGGGCCAGTGCATGGCTCATGTGGCTTTCGGAAATTGAACCAGCCAGAATTGCATGTGTCCTC

ATTGTTGTGTTTTTATTACTGGTGGTGCTCATACCCGAGCCAGAGAAGCAAAGATCTCCCCAAGATAACC

AGATGGCAATTATCATCATGGTGGCAGTGGGCCTTCTAGGTTTGATAACTGCAAACGAACTTGGATGGCT

GGAAAGAACAAAAAATGACATAGCTCATCTAATGGGAAGGAGAGAAGAAGGAGCAACCATGGGATTCTCA

ATGGACATTGATCTGCGGCCAGCCTCCGCCTGGGCTATCTATGCCGCATTGACAACTCTCATCACCCCAG

CTGTCCAACATGCGGTAACCACTTCATACAACAACTACTCCTTAATGGCGATGGCCACACAAGCTGGAGT

GCTGTTTGGCATGGGCAAAGGGATGCCATTTTATGCATGGGACCTTGGAGTCCCGCTGCTAATGATGGGT

TGCTATTCACAATTAACACCCCTGACTCTGATAGTAGCTATCATTCTGCTTGTGGCGCACTACATGTACT

TGATCCCAGGCCTACAAGCGGCAGCAGCGCGTGCTGCCCAGAAAAGGACAGCAGCTGGCATCATGAAGAA

TCCCGTTGTGGATGGAATAGTGGTAACTGACATTGACACAATGACAATAGACCCCCAGGTGGAGAAGAAG

ATGGGACAAGTGTTACTCATAGCAGTAGCCATCTCCAGTGCTGTGCTGCTGCGGACCGCCTGGGGATGGG

GGGAGGCTGGAGCTCTGATCACAGCAGCGACCTCCACCTTGTGGGAAGGCTCTCCAAACAAATACTGGAA

CTCCTCTACAGCCACCTCACTGTGCAACATCTTCAGAGGAAGCTATCTGGCAGGAGCTTCCCTTATCTAT

ACAGTGACGAGAAACGCTGGCCTGGTTAAGAGACGTGGAGGTGGGACGGGAGAGACTCTGGGAGAGAAGT

GGAAAGCTCGTCTGAATCAGATGTCGGCCCTGGAGTTCTACTCTTATAAAAAGTCAGGTATCACTGAAGT

GTGTAGAGAGGAGGCTCGCCGTGCCCTCAAGGATGGAGTGGCCACAGGAGGACATGCCGTATCCCGGGGA

AGTGCAAAGCTCAGATGGTTGGTGGAGAGAGGATATCTGCAGCCCTATGGGAAGGTTGTTGACCTCGGAT

GTGGCAGAGGGGGCTGGAGCTATTATGCCGCCACCATCCGCAAAGTGCAGGAGGTGAGAGGATACACAAA

GGGAGGTCCCGGTCATGAAGAACCCATGCTGGTGCAAAGCTATGGGTGGAACATAGTTCGTCTCAAGAGT

GGAGTGGACGTCTTCCACATGGCGGCTGAGCCGTGTGACACTCTGCTGTGTGACATAGGTGAGTCATCAT

CTAGTCCTGAAGTGGAAGAGACACGAACACTCAGAGTGCTCTCTATGGTGGGGGACTGGCTTGAAAAAAG

ACCAGGGGCCTTCTGTATAAAGGTGCTGTGCCCATACACCAGCACTATGATGGAAACCATGGAGCGACTG

CAACGTAGGCATGGGGGAGGATTAGTCAGAGTGCCATTGTCTCGCAACTCCACACATGAGATGTACTGGG

TCTCTGGGGCAAAGAGCAACATCATAAAAAGTGTGTCCACCACAAGTCAGCTCCTCCTGGGACGCATGGA

TGGCCCCAGGAGGCCAGTGAAATATGAGGAGGATGTGAACCTCGGCTCGGGTACACGAGCTGTGGCAAGC

TGTGCTGAGGCTCCTAACATGAAAATCATCGGCAGGCGCATTGAGAGAATCCGCAATGAACATGCAGAAA

CATGGTTTCTTGATGAAAACCACCCATACAGGACATGGGCCTACCATGGGAGCTACGAAGCCCCCACGCA

AGGATCAGCGTCTTCCCTCGTGAACGGGGTTGTTAGACTCCTGTCAAAGCCTTGGGACGTGGTGACTGGA

GTTACAGGAATAGCCATGACTGACACCACACCATACGGCCAACAAAGAGTCTTCAAAGAAAAAGTGGACA

CCAGGGTGCCAGATCCCCAAGAAGGCACTCGCCAGGTAATGAACATAGTCTCTTCCTGGCTGTGGAAGGA

GCTGGGGAAACGCAAGCGGCCACGCGTCTGCACCAAAGAAGAGTTTATCAACAAGGTGCGCAGCAATGCA

GCACTGGGAGCAATATTTGAAGAGGAAAAAGAATGGAAGACGGCTGTGGAAGCTGTGAATGATCCAAGGT

TTTGGGCCCTAGTGGATAGGGAGAGAGAACACCACCTGAGAGGAGAGTGTCACAGCTGTGTGTACAACAT

GATGGGAAAAAGAGAAAAGAAGCAAGGAGAGTTCGGGAAAGCAAAAGGTAGCCGCGCCATCTGGTACATG

TGGTTGGGAGCCAGATTCTTGGAGTTTGAAGCCCTTGGATTCTTGAACGAGGACCATTGGATGGGAAGAG

AAAACTCAGGAGGTGGAGTCGAAGGGTTAGGATTGCAAAGACTTGGATACATTCTAGAAGAAATGAATCG

GGCACCAGGAGGAAAGATGTACGCAGATGACACTGCTGGCTGGGACACCCGCATTAGTAAGTTTGATCTG

GAGAATGAAGCTCTGATTACCAACCAAATGGAGGAAGGGCACAGAACTCTGGCGTTGGCCGTGATTAAAT

ACACATACCAAAACAAAGTGGTGAAGGTTCTCAGACCAGCTGAAGGAGGAAAAACAGTTATGGACATCAT

TTCAAGACAAGACCAGAGAGGGAGTGGACAAGTTGTCACTTATGCTCTCAACACATTCACCAACTTGGTG

GTGCAGCTTATCCGGAACATGGAAGCTGAGGAAGTGTTAGAGATGCAAGACTTATGGTTGTTGAGGAAGC

CAGAGAAAGTGACCAGATGGTTGCAGAGCAATGGATGGGATAGACTCAAACGAATGGCGGTCAGTGGAGA

TGACTGCGTTGTGAAGCCAATCGATGATAGGTTTGCACATGCCCTCAGGTTCTTGAATGACATGGGAAAA

GTTAGGAAAGACACACAGGAGTGGAAACCCTCGACTGGATGGAGCAATTGGGAAGAAGTCCCGTTCTGCT

CCCACCACTTCAACAAGCTGTACCTCAAGGATGGGAGATCCATTGTGGTCCCTTGCCGCCACCAAGATGA

ACTGATTGGCCGAGCTCGCGTCTCACCAGGGGCAGGATGGAGCATCCGGGAGACTGCCTGTCTTGCAAAA

TCATATGCGCAGATGTGGCAGCTCCTTTATTTCCACAGAAGAGACCTTCGACTGATGGCTAATGCCATTT

GCTCGGCTGTGCCAGTTGACTGGGTACCAACTGGGAGAACCACCTGGTCAATCCATGGAAAGGGAGAATG

GATGACCACTGAGGACATGCTCATGGTGTGGAATAGAGTGTGGATTGAGGAGAACGACCATATGGAGGAC

AAGACTCCTGTAACAAAATGGACAGACATTCCCTATCTAGGAAAAAGGGAGGACTTATGGTGTGGATCCC

TTATAGGGCACAGACCCCGCACCACTTGGGCTGAAAACATCAAAGACACAGTCAACATGGTGCGCAGGAT

CATAGGTGATGAAGAAAAGTACATGGACTATCTATCCACCCAAGTCCGCTACTTGGGTGAGGAAGGGTCC

ACACCCGGAGTGTTGTAAGCACCAATTTTAGTGTTGTCAGGCCTGCTAGTCAGCCACAGTTTGGGGAAAG

CTGTGCAGCCTGTAACCCCCCCAGGAGAAGCTGGGAAACCAAGCTCATAGTCAGGCCGAGAACGCCATGG

CACGGAAGAAGCCATGCTGCCTGTGAGCCCCTCAGAGGACACTGAGTCAAAAAACCCCACGCGCTTGGAA

GCGCAGGATGGGAAAAGAAGGTGGCGACCTTCCCCACCCTTCAATCTGGGGCCTGAACTGGAGACTAGCT

GTGAATCTCCAGCAGAGGGACTAGTGGTTAGAGGAGACCCCCCGGAAAACGCAAAACAGCATATTGACGC

TGGGAAAGACCAGAGACTCCATGAGTTTCCACCACGCTGGCCGCCAGGCACAGATCGCCGAACAGCGGCG

GCCGGTGTGGGGAAATCCATGGTTTCT

>HQ234499\_1966\_Malaysia\_Mosquito\_Zika\_P6-740

ATGAAAAACCCAAAAAAGAAATCCGGAGGATTCCGGATTGTCAATATGCTAAAACGCGGAGTAGCCCGTG

TGAGCCCCTTTGGGGGCTTGAAGAGGCTACCAGCTGGACTTCTGCTGGGTCATGGACCCATCAGGATGGT

CTTGGCGATACTAGCCTTCTTGAGATTCACGGCAATCAAGCCATCACTGGGTCTCATCAATAGATGGGGT

TCCGTGGGGAAAAAAGAGGCTATGGAAATAATAAAGAAGTTCAAGAAAGATCTGGCTGCCATGCTGAGAA

TAATCAATGCTAGGAAGGAGAAGAAGAGACGTGGCGCAGACACCAGTGTCGGAATTGTTGGCCTCCTGCT

GACCACAGCCATGGCAGTGGAGGTCACCAGACGTGGGAGTGCATACTATATGTACTTAGACAGAAGCGAT

GCTGGGGAGGCCATATCTTTTCCAACCACACTGGGGGTGAATAAGTGTTACATACAGATCATGGATCTTG

GACACATGTGTGATGCCACAATGAGCTATGAATGCCCTATGTTGGATGAGGGGGTAGAACCAGATGACGT

CGATTGCTGGTGCAACACGACATCGACTTGGGTTGTGTACGGAACCTGCCATCACAAAAAAGGTGAGGCA

CGGAGATCTAGAAGAGCTGTGACGCTCCCCTCTCATTCCACTAGGAAGCTGCAAACGCGGTCGCAGACCT

GGTTGGAATCAAGAGAATACACAAAGCACTTGATCAGAGTCGAAAATTGGATATTCAGGAACCCTGGCTT

TGCGTTGGCAGCAGCTGCCATTGCTTGGCTTTTGGGAAGCTCAACGAGCCAAAAAGTCATATACTTGGTC

ATGATACTGTTGATTGCCCCGGCATACAGTATCAGGTGCATAGGAGTCAGCAATAGGGATTTTGTGGAAG

GTATGTCAGGTGGGACCTGGGTTGATGTTGTCTTGGAACATGGAGGTTGTGTTACCGTAATGGCACAGGA

CAAGCCAACTGTTGATATAGAGTTGGTCACAACAACGGTTAGCAACATGGCGGAGGTAAGATCCTACTGC

TACGAGGCATCAATATCGGACATGGCTTCGGACAGCCGCTGCCCAACACAAGGTGAAGCCTACCTTGACA

AGCAGTCAGACACTCAATATGTTTGCAAAAGAACGTTAGTGGACAGAGGTTGGGGAAATGGATGTGGACT

CTTTGGCAAAGGGAGCCTGGTGACATGCGCCAAGTTTGCATGCTCCAAGAAAATGACTGGGAAGAGCATC

CAGCCAGAGAACCTGGAGTACCGGATAATGCTGTCAGTTCATGGCTCCCAGCACAGTGGGATGATTGTTA

ATGACANAGGACATGAAACTGATGAGAATAGAGCGAAGGTTGAGATAACGCCCAATTCACCAAGAGCCGA

AGCCACCCTGGGAGGTTTTGGAAGCCTAGGACTTGATTGTGAACCGAGGACAGGCCTTGACTTTTCAGAT

TTGTATTACTTGACTATGAATAACAAGCATTGGTTGGTGCACAAGGAGTGGTTCCATGACATTCCACTAC

CTTGGCATGCTGGGGCAGACACCGGAACTCCACATTGGAACAACAAAGAAGCATTGGTAGAGTTCAAGGA

CGCACATGCCAAAAGGCAAACTGTCGTGGTTCTAGGGAGTCAAGAAGGAGCCGTTCACACGGCTCTTGCT

GGAGCCCTGGAGGCTGAGATGGATGGTGCAAAGGGAAGGCTGTCCTCTGGCCACTTGAAATGTCGCTTGA

AAATGGACAAACTTAGATTGAAGGGCGTGTCATACTCCTTATGTACCGCGGCGTTCACATTCACCAAGAT

CCCGGCTGAAACGCTGCATGGGACAGTCACAGTGGAGGTACAGTATGCAGGGACAGATGGACCCTGCAAG

GTTCCAGCTCAGATGGCGGTGGATATGCAAACTCTGACCCCAGTTGGGAGGTTGATAACCGCTAACCCTG

TGATCACTGAAAGCACTGAGAATTCAAAGATGATGTTGGAACTTGACCCACCATTTGGGGATTCTTACAT

TGTCATAGGAGTTGGGGATAAGAAGATCACCCACCACTGGNACAGGAGTGGCAGCACCATCGGAAAAGCA

TTTGAAGCCACTGTGAGAGGCGCCAAGAGAATGGCAGTCTTGGGAGACACAGCCTGGGACTTTGGATCAG

TCGGAGGTGCTCTCAACTCATTGGGCAAGGGCATCCATCAAATTTTTGGAGCAGCTTTCAAATCATTGTT

TGGAGGAATGTCCTGGTTCTCACAAATCCTCATAGGAACGTTGCTGGTGTGGTTGGGTCTGAACACAAAG

AATGGATCTATTTCCCTTACGTGCTTGGCCTTAGGGGGAGTGTTGATCTTCCTATCTACAGCCGTCTCTG

CTGATGTGGGGTGTTCGGTGGACTTCTCAAAGAAGGAAACGAGATGCGGTACGGGGGTGTTCGTCTATAA

CGACGTTGAAGCCTGGAGGGACAGGTACAAGTACCATCCTGACTCCCCTCGTAGATTGGCAGCAGCAGTC

AAGCAGGCCTGGGAAGATGGGATCTGTGGGATCTCCTCTGTTTCAAGAATGGAAAACATTATGTGGAGAT

CAGTAGAAGGGGAGCTCAACGCAATTCTGGAAGAGAATGGAGTTCAACTGACGGTCGTTGTGGGATCTGT

AAAAAACCCCATGTGGAGAGGTCCGCAGAGGTTGCCTGTGCCTGTGAATGAGCTGCCCCACGGTTGGAAG

GCCTGGGGGAAATCGTACTTTGTCAGGGCAGCAAAGACCAACAACAGCTTTGTTGTGGATGGTGACACAC

TGAAGGAATGCCCGCTCAAACACAGAGCATGGAACAGCTTTCTTGTGGAGGATCACGGGTTCGGGGTATT

TCACACTAGTGTCTGGCTTAAAGTCAGAGAGGATTACTCATTAGAGTGTGATCCAGCCGTCATAGGAACA

GCTGCTAAGGGAAAGGAGGCCGTGCACAGTGATCTAGGCTACTGGATTGAGAGTGAAAAGAACGACACAT

GGAGGCTGAAGAGGGCTCACCTGATCGAGATGAAAACATGTGAATGGCCAAAGTCCCACACACTGTGGAC

AGATGGAATAGAAGAAAGTGATCTGATCATACCTAAGTCTTTAGCTGGGCCACTCAGCCACCACAACACC

AGAGAGGGCTACAGGACTCAAGTGAAAGGGCCGTGGCATAGTGAAGAGCTTGAAATCCGGTTTGAGGAAT

GTCCAGGCACCAAGGTCCACGTGGAGGAAACATGTGGAACGAGAGGACCGTCCCTGAGATCAACCACTGC

AAGCGGAAGGGTGATCGAGGAATGGTGCTGCAGGGAATGCACAATGCCCCCATTGTCGTTCCGGGCAAAA

GATGGCTGTTGGTATGGAATGGAGATAAGGCCCAGGAAGGAACCAGAGAGTAACCTAGTAAGGTCAATGG

TGACTGCAGGATCAACTGATCACATGGATCACTTCTCCCTTGGAGTGCTTGTGATTCTGCTCATGGTGCA

GGAAGGGCTGAAGAAGAGAATGACCACAAAGATCATCATAAGCACATCAATGGCAGTGTTGGTAGCTATG

ATCCTGGGAGGATTTTCAATGAGTGACTTGGCTAAGCTTGCAATTCTGATGGGTGCCACCTTCGCGGAAA

TGAACACTGGAGGAGATGTAGCTCATCTGGCGCTGATAGCGGCATTCAAAGTCAGACCCGCGTTGCTGGT

CTCTTTCATCTTCAGAGCCAATTGGACACCCCGTGAGAGCATGCTGCTGGCCTTGGCCTCGTGCCTTCTG

CAAACTGNGATCTCCGCCCTGGAAGGCGACCTGATGGTTCTCATCAATGGTTTTGCTTTGGCCTGGTTGG

CAATACGAGCGATGGCTGTTCCACGCACTGACAACATCACCTTGGCAATCCTGGCTGCTCTGACACCACT

GGCCCGAGGCACACTGCTTGTAGCGTGGAGAGCAGGCCTTGCTACTTGTGGGGGGTTCATGCTCCTCTCT

CTGAAGGGGAAAGGTAGTGTGAAGAAGAACCTACCATTTGTCATGGCCTTGGGACTAACCGCTGTGAGGC

TGGTTGACCCCATCAACGTGGTGGGACTGCTGTTGCTCACAAGGAGTGGGAAGCGGAGCTGGCCCCCTAG

TGAAGTACTCACAGCTGTTGGCCTGATATGTGCACTGGCCGGAGGGTTCGCCAAAGCAGATATAGAGATG

GCTGGGCCCATGGCTGCAGTTGGCCTGCTAATTGTTAGTTACGTGGTCTCAGGAAAGAGTGTGGACATGT

ACATTGAAAGAGCAGGTGACATCACATGGGAAAAAGATGCGGAAGTTACTGGAAACAGCCCCCGGCTCGA

TGTGGCACTAGATGAGAGTGGTGATTTCTCCCTGGTGGAGGATGATGGTCCCCCCATGAGAGAGATCATA

CTCAAGGTGGTCCTGATGACCATCTGTGGCATGAACCCAATAGCCATACCCTTTGCAGCTGGAGCGTGGT

ATGTGTATGTGAAGACTGGAAAGAGGAGTGGTGCTCTATGGGATGTGCCTGCTCCCAAGGAAGTAAAAAA

GGGGGAGACCACAGATGGAGTGTATAGAGTGATGACTCGCAGACTGCTAGGTTCAACACAAGTTGGAGTG

GGAGTCATGCAAGAGGGGGTCTTCCACACTATGTGGCACGTCACAAAAGGATCCGCGCTGAGGAGCGGTG

AAGGGAGACTTGATCCATACTGGGGAGATGTTAAGCAGGATCTGGTGTCATACTGTGGCCCGTGGAAGCT

AGATGCCGCTTGGGACGGACACAGCGAGGTGCAGCTTTTGGCCGTGCCCCCCGGAGAGAGAGCGAGGAAC

ATCCAGACTCTGCCCGGAATATTCAAGACAAAGGATGGGGACATCGGAGCAGTTGCTCTGGACTACCCAG

CAGGAACTTCAGGATCTCCGATCCTAGACAAGTGTGGGAGAGTGATAGGACTCTATGGCAATGGGGTCGT

GATCAAAAATGGAAGTTATGTTAGTGCCATCACCCAAGGGAGGAGGGAGGAAGAGACTCCTGTTGAATGC

TTCGAACCTTCGATGCTGAAGAAGAAGCAGCTAACTGTCTTGGATCTGCATCCTGGAGCTGGGAAAACCA

GGAGAGTTCTTCCTGAAATAGTCCGTGAAGCCATAAAAACAAGACTCCGCACGGTGATCCTGGCTCCAAC

CAGGGTTGTCGCTGCTGAAATGGAGGAAGCCCTTAGAGGGCTTCCAGTGCGTTACATGACAACAGCAGTT

AATGTCACCCACTCTGGGACAGAAATCGTTGATTTAATGTGCCATGCCACCTTCACTTCACGCCTACTAC

AACCCATTAGAGTCCCCAACTACAATCTTTACATTATGGATGAGGCCCACTTCACAGATCCCTCAAGTAT

AGCAGCAAGAGGATACATATCAACAAGGGTTGAGATGGGCGAGGCGGCTGCCATCTTCATGACCGCCACA

CCACCAGGAACCCGCGACGCATTTCCGGACTCTAACTCACCAATCATGGACACAGAAGTGGAAGTCCCAG

AGAGAGCCTGGAGCTCAGGCTTTGATTGGGTGACGGATCATTCTGGAAAAACAGTTTGGTTTGTTCCAAG

CGTGAGGAACGGCAACGAGATCGCGGCTTGTCTGACAAAAGCTGGAAAACGGGTCATACAGCTCAGCAGA

AAGACTTTTGAGACAGAGTTCCAGAAAACAAAAAATCAAGAGTGGGACTTCGTCGTAACAACTGACATCT

CAGAGATGGGCGCCAACTTCAAAGCTGACCGGGTCATAGATTCCAGGAGATGCCTGAAGCCGGTCATACT

TGATGGCGAGAGAGTCATTCTGGCTGGACCCATGCCTGTCACACATGCCAGCGCTGCCCAGAGGAGGGGG

CGCATAGGCAGGAATCCCAACAAACCTGGAGATGAGTATATGTATGGAGGTGGGTGCGCAGAGACTGATG

AAGACCATGCACACTGGCTTGAAGCAAGAATGCTTCTTGATAACATTTACCTCCAAGATGGCCTCATAGC

CTCGCTCTATCGACCTGAGGCCGATAAGGTAGCAGCCATTGAGGGAGAGTTCAAGCTTAGGACGGAGCAA

AGGAAGACCTTTGTGGAACTCATGAAAAGAGGAGATCTTCCTGTTTGGCTGGCCTATCAGGTTGCATCTG

CCGGAATAACCTACACAGATAGAAGATGGTGTTTTGATGGCACGACCAACAACACCATAATGGAAGACAG

TGTGCCGGCAGAGGTGTGGACCAGATACGGAGAGAAAAGAGTGCTCAAACCGAGGTGGATGGACGCCAGA

GTTTGTTCAGATCATGCGGCCCTGAAGTCATTCAAAGAATTTGCCGCTGGGAAAAGAGGAGCGGCCTTTG

GAGTGATGGAAGCCCTGGGAACACTGCCAGGACACATGACAGAGAGGTTTCAGGAAGCCATTGACAACCT

CGCTGTGCTCATGCGGGCAGAGACTGGAAGCAGGCCCTACAAAGCCGCGGCGGCCCAATTACCGGAGACC

TTAGAGACCATCATGCTTTTGGGTTTGCTGGGAACAGTCTCGCTGGGAATCTTCTTTGTCTTGATGCGGA

ACAAGGGCATAGGGAAGATGGGCTTTGGAATGGTGACCCTTGGGGCCAGTGCATGGCTTATGTGGCTCTC

GGAAATTGAGCCAGCCAGAATTGCATGTGTCCTCATTGTCGTGTTTCTATTGCTGGTGGTGCTCATACCT

GAGCCAGAAAAGCAGAGATCTCCCCAGGACAACCAAATGGCAATTATCATCATGGTAGCAGTGGGTCTTC

TGGGCTTGATAACCGCCAATGAACTCGGATGGTTGGAGAGAACAAAAAGTGACCTAGGCCATCTAATGGG

AAGGAGAGAGGAGGGGGCAACCATGGGATTCTCAATGGACATTGACTTGCGGCCAGCCTCAGCTTGGGCT

ATCTATGCCGCTCTGACAACTCTCATCACCCCAGCCGTCCAACATGCGGTAACCACTTCATACAACAACT

ACTCCTTAATGGCGATGGCCACGCAAGCCGGAGTGTTGTTTGGCATGGGCAAAGGGATGCCATTCTATGC

GTGGGACTTCGGAGTCCCGCTGCTAATGATGGGTTGCTACTCACAATTAACACCCTTGACCTTAATAGTG

GCCATCATTCTGCTCGTGGCGCACTACATGTACTTGATCCCAGGTCTACAGGCAGCAGCGGCGCGCGCTG

CCCAGAAGAGAACGGCAGCTGGCATCATGAAGAACCCTGTTGTGGATGGAATAGTGGTGACTGACATTGA

CACAATGACAATTGACCCCCAAGTGGAGAAAAAGATGGGACAAGTGCTACTCATAGCAGTAGCCATCTCC

AGTGCCGTTCTGCTGCGCACCGCCTGGGGGTGGGGGGAGGCTGGGGCCCTGATCACAGCCGCAACTTCCA

CTTTGTGGGAAGGCTCTCCGAATAAATACTGGAACTCCTCCACAGCCACTTCACTGTGTAACATTTTTAG

GGGAAGTTACTTGGCTGGAGCTTCTCTTATTTACACAGTAACAAGAAACGCTGGCCTGGTCAAGAGACGT

GGAGGTGGAACGGGAGAGACCCTGGGGGAGAAATGGAAGGCCCGCCTGAACCAGATGTCGGCCCTGGAGT

TTTACTCCTACAAAAAGTCAGGCATCACCGAAGTGTGCAGAGAAGAAGCCCGCCGCGCCCTCAAGGACGG

AGTGGCAACAGGAGGCCATGCTGTGTCCCGAGGAAGCGCAAAGCTTAGATGGTTGGTGGAGAGAGGATAC

CTGCAGCCCTATGGAAAGGTCATTGATCTTGGATGTGGCAGAGGGGGCTGGAGTTACTACGCCGCCACCA

TCCGCAAAGTTCAAGAGGTGAAAGGATACACAAAGGGAGGCCCTGGTCATGAAGAACCCACGTTGGTGCA

AAGCTATGGATGGAACATAGTCCGTCTTAAGAGTGGGGTGGACGTCTTTCACATGGCGGCGGAGTCGTGT

GACACTTTGCTGTGTGACATAGGTGAGTCATCATCTAGTCCTGAAGTGGAAGAAGCACGGACGCTCAGAG

TACTCTCCATGGTGGGGGATTGGCTTGAAAAAAGACCAGGGGCCTTTTGTATAAAGGTGTTGTGCCCATA

CACCAGCACCATGATGGAAACCCTAGAGCGACTGCAGCGTAGGTATGGGGGAGGACTGGTCAGAGTGCCA

CTCTCCCGCAACTCTACACATGAGATGTACTGGGTCTCTGGAGCGAAAAGCAACATCATAAAAAGTGTGT

CCACCACGAGCCAGCTCCTCTTGGGACGCATGGACGGGCCCAGGAGGCCAGTGAAATATGAGGAGGATGT

GAATCTCGGCTCCGGCACGCGAGCTGTGGCAAGCTGCGCCGAAGCTCCCAACCTGAAGATCATTGGTAAC

CGCGTTGAGAGGATCCGCAGTGAGCATGCGGAAACGTGGTTCTTTGATGAGAACCACCCATACAGGACAT

GGGCTTACCATGGGAGCTACGAGGCCCCTACACAAGGGTCAGCGTCTTCTCTCATAAACGGGGTTGTCAG

GCTCCTGTCAAAGCCCTGGGATGTGGTGACTGGAGTCACAGGAATAGCCATGACCGACACCACACCGTAT

GGCCAGCAAAGAGTTTTCAAGGAAAAAGTGGACACTAGGGTGCCAGACCCCCAGGAAGGCACTCGTCAGG

TGATGAACATGGTCTCTTCCTGGCTATGGAAGGAGCTAGGTAAACACAAACGGCCACGAGTTTGCACCAA

AGAAGAGTTCATCAATAAGGTTCGCAGCAATGCAGCACTGGGGGCAATATTTGAAGAGGAGAAAGAATGG

AAGACTGCAGTGGAAGCTGTGAACGATCCAAGGTTCTGGGCCCTAGTGGACAAGGAAAGAGAGCACCACT

TGAGAGGAGAGTGTCAGAGCTGTGTGTACAACATGATGGGAAAAAGAGAAAAGAAGCAAGGGGAATTTGG

AAAGGCCAAGGGCAGCCGCGCCATTTGGTACATGTGGCTAGGGGCTAGATTTCTAGAGTTTGAAGCCCTT

GGATTCTTGAACGAGGATCACTGGATGGGGAGAGAGAATTCAGGAGGTGGTGTTGAAGGGCTGGGATTAC

AAAGACTTGGATATGTTCTAGAAGAAATGAGCCGCACACCAGGAGGAAAGATGTATGCAGATGATACCGC

TGGCTGGGACACCCGCATCAGTAGGTTTGATCTGGAGAATGAAGCTCTGATCACCAACCAAATGGAGAAA

GGGCACAGGGCCTTGGCGTTGGCCATAATCAAGTACACATACCAAAACAAAGTGGTAAAGGTCCTTAGAC

CAGCTGAAAGAGGGAAGACAGTTATGGACATCATCTCAAGACAAGACCAAAGAGGGAGCGGACAAGTTGT

TACTTACGCTCTTAATACATTCACCAACCTGGTGGTGCAGCTCATTCGGAACATGGAGGCTGAGGAAGTT

CTAGAGATGCAAGACTTGTGGCTGTTGAGGAGGCCAGAGAAGGTGACCAGCTGGTTGCAGAGCAACGGAT

GGGATAGGCTCAAACGAATGGCAGTCAGTGGAGATGATTGTGTTGTGAAACCAATTGATGATAGGTTTGC

ACATGCCCTCAGGTTTTTGAATGACATGGGGAAAGTTAGGAAGGACACACAGGAGTGGAAACCCTCAACT

GGATGGAGCAACTGGGAAGAAGTTCCGTTTTGCTCCCATCACTTCAACAAGCTTTACCTCAAGGACGGGA

GGTCCATTGTGGTCCCCTGTCGCCACCAAGATGAACTGATTGGCCGAGCCCGCGTCTCACCAGGGGCGGG

ATGGAGCATCCGGGAGACTGCTTGCCTAGCAAAATCATATGCACAAATGTGGCAGCTTCTTTATTTCCAC

AGAAGGGACCTCCGACTGATGGCCAACGCCATTTGTTCATCTGTGCCAGTTGACTGGGTTCCAACTGGGA

GAACCACCTGGTCAATCCATGGAAAGGGAGAATGGATGACCACTGAGGACATGCTTGTGGTGTGGAACAG

AGTGTGGATTGAGGAGAACGACCACATGGAGGACAAGACCCCAGTCACGAAATGGACAGACATTCCCTAT

TTGGGAAAAAGGGAAGACTTATGGTGTGGATCTCTTATAGGGCACAGACCACGCACTACTTGGGCTGAGA

ACATTAAAGACACAGTCAACATGGTGCGCAGGATCATAGGTGATGAAGAAAAGTACATGGACTACCTATC

CACTCAAGTTCGCTACTTGGGTGAAGAAGGGTCCACACCTGGAGTGTTA

>KF383118\_2001\_Senegal\_Mosquito\_Zika\_ArD157995

ATGAAAAACCCAAAGAAGAAATCCGGAAGATTCCGGATTGTCAATATGCTAAAACGCGGAGTAGCCCGTG

TAAACCCCTTGGGAGGTTTGAAGAGGTTGCCAGCCGGACTTCTGCTGGGTCATGGACCCATCAGAATGGT

TTTGGCGATACTAGCCTTTTTGAGATTTACAGCAATCAAGCCATCACTGGGCCTTATCAACAGATGGGGT

TCCGTGGGGAAAAAAGAGGCTATGGAAATAATAAAGAAGTTCAAGAAAGATCTTGCTGCCATGTTGAGAA

TAATCAATGCTAGGAAAGAGAGGAAGAGACGTGGCGCAGACACCAGCATCGGAATCATTGGCCTCCTGCT

GACTACAGCCATGGCAGCAGAGATCACTAGACGCGGGAGTGCATACTACATGTACTTGGATAGGAGCGAT

GCCGGGAAGGCCATTTCGTTTGCTACCACATTGGGAGTGAACAAGTGCCACGTACAGATCATGGACCTCG

GGCACATGTGTGACGCCACCATGAGTTATGAGTGCCCTATGCTGGATGAGGGAGTGGAACCAGATGATGT

CGATTGCTGGTGCAACACGACATCAACTTGGGTTGTGTACGGAACCTGTCATCACAAAAAAGGTGAGACA

CGGCGATCTAGAAGATCTGTGTCGCTCCGTTATCACTATACAAGGAAGTTGCAAACGCGGTCGCAGACAT

GGTTAGAATCAAGAGAATACAAGAAGCACTTGATCATGGTCGAAAACTGGATATTCAGGAACCCCGGGTT

TGCCATAGTGTCCGTTGCCATTACCTGGCTGATGGGAAGCTTGACGAGCCAAAAAGTCATATACTTGGTC

ATGATAGTGTTGATTGTCCCGGCATACAGTATCAGCTGCATTGGAGTCAGCAATAGAGACTTAGTGGAGG

GCATGTCAGGTGGGACCTGGGTTGATGTTGTCTTGGAACATGGAGGGTGCGTTACCGAGATGGCACAGGA

CAAGCCAACAGTTGACATAGAGTTGGTCACGATGACGGTTAGTAACATGGCCGAGGTAAGATCCTATTGC

TACGAGGCATCGTTATCCGACATGGCTTCGGCCAGTCGTTGCCCAACACAAGGCGAACCCTCCCTCGACA

AGCAATCAGACACTCAATCTGTATGCAAAAGAACATTAGGAGACAGAGGTTGGGGAAATGGTTGTGGGAT

TTTTGGCAAAGGGAGCTTGGTGACATGTTCCAAGTTCACGTGTTGTAAGAAGATGCCCGGGAAGAGCATT

CAACCGGAAAATCTGGAGTATCGGATAATGCTCCCAGTGCATGGCTCCCAGCATAGCGGGATGATTGTGA

ATGACATAGGACATGAAACTGACGAAAACAGAGCGAAAGTCGAGGTCACACCCAATTCACCAAGAGCAGA

AGCAACCTTGGGAGGTTTTGGAAGCTTGGGACTTGACTGTGAACCAAGGACAGGCCTTGACTTCTCAGAT

CTGTATTATCTGACCATGAACAACAAGCATTGGTTGGTGCACAAGGAGTGGTTTCATGACATCCCATTAC

CTTGGCATGCTGGTGCAGACACTGGAACTCCACACTGGAACAACAAAGAGGCATTGGTGGAGTTCAAGGA

CGCCCACGCCAAGAGGCAAACTGTTGTGGTTCTGGGGAGCCAAGAGGGAGCTGTTCACACGGCCCTCGCT

GGAGCTCTGGAGGCTGAGATGGATGGTGCAAAGGGAAGGCTATTCTCTGGCCATTTGAAATGCCGCCTAA

AAATGGACAAGCTTAGGTTGAAGGGTGTGTCATATTCCCTGTGTACTGCAGCGTTCACATTTACCAAGGT

CCCAGCTGAAACATTGCATGGAACAGTTACAGTGGAGGTGCAGTCTGCAGGGACAGATGGACCCTGCAAG

GTCCCAGCCCAGATGGCGGTGGACATGCAGACCCTGACCCCAGTTGGAAGGCTGATAACCGCCAACCCCG

TGATTACTGAAAGCACTGAGAACTCAAAGATGATGTTGGAGCTTGACCCACCATTTGGGGATTCTTACAT

TGTCATAGGAGTTGGGGACAAGAAAATCACCCACCACTGGCATAGGAGTGGTAGCACCATCGGAAAGGCA

TTTGAAGCCACTGTGAGAGGCGCCAAGAGAATGGCAGTCCTGGGGGATACAGCCTGGGACTTCGGATCAG

TCGGGGGTGTGTTCAACTCACTGGGTAAGGGCATTCACCAGATTTTTGGAGCAGCCTTCAAATCACTGTT

TGGAGGAATGTCCTGGTTCTCACAGATCCTCATAGGCACGCTGCTAGTGTGGTTAGGTTTGAACACAAAG

AATGGATCTATCTCCCTCACATGCTTGGCCCTGGGGGGAGTGATGATCTTCCTCTCCACGGCTGTTTCTG

CTGACGTGGGGTGCTCAGTGGACTTCTCAAAAAAGGAAACGAGATGTGGCACGGGGGTATTCATCTATAA

TGATGTTGAAGCCTGGAGGGACCGGTACAAGTACCATCCTGACTCCCCCCGCAGATTGGCAGCAGCAGTC

AAGCAGGCCTGGGAAGAGGGGATCTGTGGGATCTCATCCGTTTCAAGAATGGAAAACATCATGTGGAAAT

CAGTAGAAGGGGAGCTCAATGCTATCCTAGAGGAGAATGGAGTTCAACTGACAGTTGTTGTGGGATCTGT

AAAAAACCCCATGTGGAGAGGTCCACAAAGATTGCCAGTGCCTGTGAATGNGCTGCCCCATGGCTGGAAA

GCCTGGGGGAAATCGTATTTTGTTAGGGCGGCAAAGACCAACAACAGTTTTGTTGTCGACGGTGACACAC

TGAAGGAATGTCCGCTTAAGCACAGAGCATGGAATAGTTTTCTTGTGGAGGATCACGGGTTTGGAGTCTT

CCACACCAGTGTTTGGCTTAAGGTCAGAGAAGATTACTCATTAGAATGTGACCCAGCCGTCATAGGAACA

GCTGTTAAGGGAAGGGAGGCCGCGCACAGTGATCTGGGCTATTGGATTGAAAGTGAAAAGAATGACACAT

GGAGGCTGAAGAGGGCCCACCTGATTGAGATGAAAACATGTGAATGGCCAAAGTCTCACACATTGTGGAC

AGATGGAGTAGAAGAAAGTGATCTTATCATACCCAAGTCTTTAGCTGGTCCACTCAGCCACCACAACACC

AGAGAGGGTTACAGAACCCAAGTGAAAGGGCCATGGCACAGTGAAGAGCTTGAAATCCGGTTTGAGGAAT

GTCCAGGCACCAAGGTTCACGTGGAGGAGACATGCGGAACTAGAGGACCATCTCTGAGATCAACTACTGC

AAGTGGAAGGGTCATTGAGGAATGGTGCTGTAGGGAATGCACAATGCCCCCACTATCGTTTCGAGCAAAA

GACGGCTGCTGGTATGGAATGGAGATAAGGCCCAGGAAAGAACCAGAGAGCAACTTAGTGAGGTCAATGG

TGACAGCGGGGTCAACCGATCATATGGACCACTTCTCTCTTGGAGTGCTTGTGATTCTACTCATGGTGCA

GGAGGGGTTGAAGAAGAGAATGACCACAAAGATCATCATGAGCACATCAATGGCAGTGCTGGTAGTCATG

ATCTTGGGAGGATTTTCAATGAGTGACCTGGCCAAGCTTGTGATCCTGATGGGTGCTACTTTCGCAGAAA

TGAACACTGGAGGAGATGTAGCTCACTTGGCATTGGTAGCGGCATTTAAAGTCAGACCAGCCTTGCTGGT

CTCCTTCATTCTCAGAGCCAATTGGACACCCCGTGAGAGCATGCTGCTAGCCCTGGCTTCGTGTCTTCTG

CAAACTGCGATCTCTGCTCTTGAAGGTGACTTGATGGTCCTCGTTAATGGATTTGCTTTGGCCTGGTTGG

CAATTCGAGCAATGGCCGTGCCACGCACTGACAACATCGCTCTAGCAATCTTGGCTGCTCTAACACCACT

AGCTCGAGGCACACTGCTCGTGGCATGGAGAGCGGGCCTGGCTACTTGTGGAGGGTTCATGCTCCTCTCC

CTGAAAGGGAAAGGTAGTGTGAAGAAGAACCTGCCATTTGTCATGGCCCTGGGATTGACAGCTGTGAGGG

TAGTAGACCCTATTAATGTGGTAGGACTACTGTTACTCACAAGGAGTGGGAAGCGGAGCTGGCCCCCTAG

TGAAGTTCTCACAGCCGTTGGCCTGATATGTGCACTGGCCGGAGGGTTTGCCAAGGCAGACATTGAGATG

GCTGGACCCATGGCTGCAGTAGGCTTGCTAATTGTCAGCTATGTGGTCTCGGGAAAGAGTGTGGACATGT

ACATTGAAAGAGCAGGTGACATCACATGGGAAAAGGACGCGGAAGTCACTGGAAACAGTCCTCGGCTTGA

CGTGGCACTGGATGAGAGTGGTGATTTCTCCTTGGTAGAGGAAGATGGTCCACCCATGAGAGAGATCATA

CTCAAGGTGGTCCTGATGGCCATCTGTGGCATGAACCCAATAGCTATACCTTTTGCTGCAGGAGCGTGGT

ATGTGTATGTGAAGACTGGGAAAAGGAGTGGCGCCCTCTGGGACGTGCCTGCTCCCAAAGAAGTGAAGAA

AGGAGAGACCACAGATGGAGTGTACAGAGTGATGACTCGCAGACTGCTAGGTTCAACACAGGTTGGAGTG

GGAGTCATGCAAGAGGGAGTCTTCCACACCATGTGGCACGTTACAAAAGGAGCCGCACTGAGGAGCGGTG

AGGGAAGACTTGATCCATACTGGGGGGATGTCAAGCAGGACTTGGTGTCATACTGTGGGCCTTGGAAGTT

GGATGCAGCTTGGGATGGACTTAGCGAGGTACAGCTTTTGGCCGTACCTCCCGGAGAGAGGGCCAGAAAC

ATTCAGACCCTGCCTGGAATATTCAAGACAAAGGACGGGGACATCGGAGCAGTTGCTCTGGACTACCCTG

CAGGGACCTCAGGATCTCCGATCCTAGACAAATGTGGAAGAGTGATAGGACTCTATGGCAATGGGGTTGT

GATCAAGAATGGAAGCTATGTTAGTGCTATAACCCAGGGAAAGAGGGAGGAGGAGACTCCGGTTGAATGT

TTCGAACCCTCGATGCTGAAGAAGAAGCAGCTAACTGTCCTGGATCTGCATCCAGGAGCCGGAAAAACCA

GGAGAGTTCTTCCTGAAATAGTCCGTGAAGCCATAAAAAAGAGACTCCGGACAGTGATCTTGGCACCAAC

TCCAGTTGAGAGATCATACTCAAGGTGGTCCTTGTGGCCATCTGTGGCACGTACCCAGACAACAGCAGTC

AACGTCACCCATTCTGGGACAGAAATCGTTGATTTGATGTGCCATGCCACTTTCACTTCACGCTTACTAC

AACCCATCAGAGTCCCTAATTACAATCTCAACATCATGGATGAAGCCCACTTCACAGACCCCTCAAGTAT

AGCTGCAAGAGGATACATATCAACAAGGGTTGAAATGGGCGAGGCGGCTGCCATTTTTATGACTGCCACA

CCACCAGGAACCCGTGATGCGTTTCCTGACTCTAACTCACCAATCATGGACACAGAAGTGGAAGTCCCAG

AGAGAGCCTGGAGCTCAGGCTTTGATTGGGTGACAGACCATTCTGGGAAAACAGTTTGGTTCGTTCCAAG

CGTGAGATCTGGAGAAGAAAGCGCAGCCTGTCTGACAAAGGCTGGAAAGCGGGTCATACAGCTCAGCAGG

AAGACTTTTGAGACAGAATTTCAGAAAACAAAAAATCAAGAGTGGGACTTTGTCATAACAACTGACATCT

CAGAGATGGGCGCCAACTTCAAGGCTGACCGGGTCATAGACTCTAGGAGATGCCTAAAGCCAGTCATACT

TGATGGTGAGAGAGTCATCTTGGCTGGGCCCATGCCTGTCACGCATGCTAGTGCTGCTCAGAGGAGAGGA

CGTATAGGCAGGAACCCTAACAAACCTGGAGATGAGTACATGTATGGAGGTGGGTGTGCAGAGACTGATG

AAGACCATGCACACTGGCTTGAAGCAAGAATGCTTCTTGACAACATCTACCTCCAGGATGGCCTCATAGC

CTCGCTCTATCGGCCTGAGGCCGATAAGGTAGCCGCCATTGAGGGAGAGTTTAAGCTGAGGACAGAGCAA

AGGAAGACCTTCGTGGAACTCATGAAGAGAGGAGACCTTCCCGTCTGGCTAGCCTATCAGGTTGCATCTG

CCGGAATAACTTACACAGACAGAAGATGGTGCTTTGATGGCACAACCAACAACACCATAATGGAAGACAG

CGTACCAGCAGAGGTGTGGACAAAGTATGGAGAGAAGAGAGTGCTCAAACCGAGATGGATGGATGCTAGG

GTCTGTTCAGACCATGCGGCCCTGAAGTCGTTCAAAGAATTCGCCGCTGGAAAAAGAGGAGCGGCTTTGG

GAGTAATGGAGGCCCTGGGAACACTGCCAGGACACATGACAGAGAGGTTTCAGGAAGCCATTGACAACCT

CGCCGTGCTCATGCGAGCAGAGACTGGAAGCAGGCCTTATAAGGCAGCGGCAGCCCAACTGCCGGAGACC

CTAGAGACCATTATGCTCTTAGGTTTGCTGGGAACAGTTTCACTGGGGATCTTCTTCGTCTTGATGCGGA

ATAAGGGCATCGGGAAGATGGGCTTTGGAATGGTAACCCTTGGGGCCAGTGCATGGCTCATGTGGCTTTC

GGAAATTGAACCAGCCAGAATTGCATGTGTCCTCATTGTTGTGTTTTTATTACTGGTGGTGCTCATACCC

GAGCCAGAGAAGCAAAGATCTCCCCAAGATAACCAGATGGCAATTATCATCATGGTGGCAGTGGGCCTTC

TAGGTTTGATAACTGCAAACGAACTTGGATGGCTGGAAAGAACAAAAAATGACATAGCTCATCTAATGGG

AAGGAGAGAAGAAGGAGCAACCATGGGATTCTCAATGGACATTGATCTGCGGCCAGCCTCCGCCTGGGCT

ATCTATGCCGCATTGACAACTCTCATCACCCCAGCTGTCCAACATGCGGTAACCACTTCATACAACAACT

ACTCCTTAATGGCGATGGCCACACAAGCTGGAGTGCTGTTTGGCATGGGCAAAGGGATGCCATTTATGCA

TGGGGACCTTGGAGTCCCGCTGCTAATGATGGGTTGCTATTCACAATTAACACCCCTGACTCTGATAGTA

GCTATCATTCTGCTTGTGGCGCACTACATGTACTTGATCCCAGGCCTACAAGCGGCAGCAGCGCGTGCTG

CCCAGAAAAGGACAGCAGCTGGCATCATGAAGAATCCCGTTGTGGATGGAATAGTGGTAACTGACATTGA

CACAATGACAATAGACCCCCAGGTGGAGAAGAAGATGGGACAAGTGTTACTCATAGCAGTAGCCATCTCC

AGTGCTGTGCTGCTGCGGACCGCCTGGGGATGGGGGGAGGCTGGAGCTCTGATCACAGCAGCGACCTCCA

CCTTGTGGGAAGGCTCTCCAAACAAATACTGGAACTCCTCTACAGCCACCTCACTGTGCAACATCTTCAG

AGGAAGCTATCTGGCAGGAGCTTCCCTTATCTATACAGTGACGAGAAACGCTGGCCTGGTTAAGAGACGT

GGAGGTGGGACGGGAGAGACTCTGGGAGAGAAGTGGAAAGCTCGTCTGAATCAGATGTCGGCCCTGGAGT

TCTACTCTTATAAAAAGTCAGGTATCACTGAAGTGTGTAGAGAGGAGGCTCGCCGTGCCCTCAAGGATGG

AGTGGCCACAGGAGGACATGCCGTATCCCGGGGAAGTGCAAAGCTCAGATGGTTGGTGGAGAGAGGATAT

CTGCAGCCCCATGGGAAGGTTGTTGACCTCGGATGTGGCAGAGGGGGCTGGAGCTATTATGCCGCCACCA

TCCGCAAAGTGCAGGAGGTGAGAGGATACACAAAGGGAGGTCCCGGTCATGAAGAACCCATGCTGGTGCA

AAGCTATGGGTGGAACATAGTTCGTCTCAAGAGTGGAGTGGACGTCTTCCACATGGCGGCTGAGCCGTGT

GACACTCTGCTGTGTGACATAGGTGAGTCATCATCTAGTCCTGAAGTGGAAGAGACACGAACACTCAGAG

TGCTCTCTATGGTGGGGGACTGGCTTGAAAAAAGACCAGGGGCCTTCTGTATAAAGGTGCTGTGCCCATA

CACCAGCACTATGATGGAAACCATGGAGCGACTGCAACGTAGGCATGGGGGAGGATTAGTCAGAGTGCCA

TTGTCTCGCAACTCCACACATGAGATGTACTGGGTCTCTGGGGCAAAAAGCAACATCATAAAAAGTGTGT

CCACCACAAGTCAGCTCCTCCTGGGACGCATGGATGGCCCCAGGAGGCCAGTGAAATATGAGGAGGATGT

GAACCTCGGCTCGGGTACACGAGCTGTGGCAAGCTGTGCTGAGGCTCCTAACATGAAAATCATCGGCAGG

CGCATTGAGAGAATCCGCAATGAACATGCAGAAACATGGTTTCTTGATGAAAACCACCCATACAGGACAT

GGGCCTACCATGGGAGNTACGAAGCCCCCACGCAAGGATCAGCGTCTTCCCTCGTGAACGGGGTTGTTAG

ACTCCTGTCAAAGCCTTGGGACGTGGTGACTGGAGTTACAGGAATAGCCATGACTGACACCACACCATAC

GGCCAACAAAGAGTCTTCAAAGAAAAAGTGGACACCAGGGTGCCAGATCCCCAAGAAGGCACTCGCCAGG

TAATGAACATGGTCTCTTCCTGGCTGTGGAAGGAGCTGGGGAAACGCAAGCGGCCACGCGTCTGCACCAA

AGAAGAGTTTATCAACAAGGTGCGCAGCAATGCAGCACTGGGAGCAATATTTGAAGAGGAAAAAGAATGG

AAGACGGCTGTGGAAGCTGTGAATGATCCAAGGTTTTGGGCCCTAGTGGATAGGGAGAGAGAACACCACC

TGAGAGGAGAGTGTCACAGCTGTGTGTACAACATGATGGGAAAAAGAGAAAAGAAGCAAGGAGAGTTCGG

GAAAGCAAAAGGTAGCCGCGCCATCTGGTACATGTGGTTGGGAGCCAGGTTTCTGGAGTTTGAATCACTG

GGGTTTCTGAATGAAGATCACTGGATGGGAAGAGAGAACTCTGGAGGCGGAGTTGAAGGACTGGGACTGC

AGAGACTGGGCTATGTCCTTGAGGAGATGAGCCAGGCACCAGGAGGGAAGATGTACGCAGATGACACTGC

TGGCTGGGACACCCGCATTAGTAAGTTTGATCTGGAGAATGAAGCTTTGATTACCAACCAAATGGAGGAA

GGGCACAGAACTCTGGCGTTGGCCGTGATTAAATACACATACCAAAACAAAGTGGTGAAGGTCCTCAGAC

CAGCTGAAGGAGGAAAAACAGTTATGGACATCATTTCAAGACAAGACCAGAGAGGGAGTGGACAAGTTGT

CACTTATGCTCTCAACACATTCACCAACTTGGTGGTGCAGCTTATCCGGAACATGGAAGCTGAGGAAGTG

TTAGAGATGCAAGACTTATGGCTGTTGAGGAAGCCAGAGAAAGTGACCAGATGGTTGCAGAGCAATGGAT

GGGATAGACTCAAACGAATGGCGGTCAGTGGAGATGACTGCGTTGTGAAGCCAATCGATGATAGGTTTGC

ACATGCCCTCAGGTTCTTGAATGACATGGGAAAAGTTAGGAAAGACACACAGGAGTGGAAACCCTCGACT

GGATGGAGCAATTGGGAAGAAGTCCCGTTCTGCTCCCACCACTTCAACAAGCTGTACCTCAAGGATGGGA

GATCCATTGTGGTCCCTTGCCGCCACCAAGATGAACTGATTGGCCGAGCTCGCGTTTCACCAGGGGCAGG

ATGGAGCATCCGGGAGACTGCCTGTCTTGCAAAATCATATGCGCAGATGTGGCAGCTCCTTTATTTCCAC

AGAAGAGACCTTCGACTGATGGCTAATGCCATTTGCTCGGCTGTACCAGTTGACTGGGTACCAACTGGGA

GAACCACCTGGTCAATCCATGGAAAGGGAGAATGGATGACCACTGAGGACATGCTCATGGTGTGGAATAG

AGTGTGGATTGAGGAGAACGACCATATGGAGGACAAGACTCCTGTAACAAAATGGACAGACATTCCCTAT

CTAGGAAAAAGGGAGGACTTATGGTGTGGATCCCTTATAGGGCACAGACCCCGCACCACTTGGGCTGAAA

ACATCAAAGACACAGTCAACATGGTGCGCAGGATCATAGGTGATGAAGAAAAGTACATGGACTATCTATC

CACCCAAGTCCGCTACTTGGGTGAGGAAGGGTCCACACCCGGAGTGTTGTAA

>KF268948\_1976\_Central\_African\_Republic\_Mosquito\_Zika\_ARB13565

AGTTGTTGATCTGTGTGAATCAGACTGCGACAGTTCGAGTTTGAAGCGAGAGCTAACAACAGTATCAACA

GGTTTAATTTGGATTTGGAAGCGAGAGTTTCTGGTCATGAAAAACCCAAAGAAGAAATCCGGAGGATTCC

GGATTGTCAATATGCTAAAACGCGGAGTAGCCCGTGTAAACCCCTTGGGAGGTTTGAAGAGGCTGCCAGC

CGGACTTCTGCTGGGTCATGGACCCATCAGAATGGTTTTGGCAATACTAGCCTTTTTGAGATTTACAGCA

ATCAAGCCATCACTGGGCCTCATCAACAGATGGGGTTCTGTGGGGAAAAAAGAGGCTATGGAAATAATAA

AGAAGTTTAAGAAAGATCTTGCTGCCATGTTGAGAATAATCAATGCTAGGAAGGAGAGGAAGAGACGTGG

TGCAGACACCAGCATCGGAATCATTGGTCTCCTGCTGACTACAGCCATGGCAGCGGAGATTACCAGACGT

GGAAGTGCATACTACATGTACTTGGACAGGAGCGATGCTGGGAAGGCCATCTCCTTTGCTACCAACTTGG

GAGTTAACAAATGCCATGTACAGATCATGGATCTCGGGCACATGTGTGACGCCACCATGAGTTATGAGTG

CCCTATGTTGGACGAGGGAGTGGAACCAGATGATGTCGATTGCTGGTGCAACACGACATCAACTTGGGTT

GTGTACGGAACCTGTCATCACAAAAAAGGTGAAGCACGGCGATCTAGAAGAGCTGTGACGCTCCCTTCTC

ACTCTACAAGGAAGTTGCAAACGCGGTCGCAGACCTGGTTAGAATCAAGAGAATACACGAAGCACCTGAT

CAAGGTTGAAAACTGGATATTTAGGAACCCCGGGTTTGCGCTCGCAGCTGTTGCCATTGCCTGGCTTTTG

GGAAGCTCGACGAGCCAAAAAGTCATATACTTGGTTATGATACTGCTGATTGCCCCGGCATACAGCATCA

GGTGCATTGGAGTCAGCAATAGAGACTTCGTGGAAGGCATGTCAGGTGGGACCTGGGTTGATGTCGTCTT

GGAACATGGAGGTTGCGTCACCGTTATGGCACAGGACAAGCCAACAGTTGACATAGAGTTGGTCACGACA

ACGGTTAGTAACATGGCCGAGGTAAGATCCTACTGTTATGAGGCATCAATATCAGACATGGCTTCGGACA

GTCGCTGCCCAACACAAGGTGAAGCCTACCTTGATAAGCAATCAGACACCCAATATGTGTGCAAAAGAAC

ATTGGTGGACAGAGGTTGGGGAAATGGGTGTGGACTTTTTGGCAAAGGGAGCCTGGTGACATGTGCCAAG

TTCACGTGTTCCAAGAAGATGACTGGGAAGAGCATTCAGCCGGAGAATCTGGAGTATCGAATAATGCTAT

CAGTGCATGGCTCCCAGCACAGCGGGATGATTGTTAATGACATAGGACATGAAACTGACGAAAACAGAGC

GAAAGTCGAGGTCACGCCCAATTCACCAAGAGCAGAAGCAACCTTGGGAGGCTTTGGAAGCTTAGGACTT

GATTGTGAACCAAGGACAGGCCTGGACTTTTCAGATTTGTACTACTTGACCATGAACAACAAGCACTGGT

TGGTGCACAAAGAGTGGTTTCATGACATCCCATTGCCTTGGCATGCTGGGGCAGATACCGGAACTCCACA

CTGGAACAACAAGGAGGCATTGGTAGAGTTCAAAGATGCCCACGCCAAGAGGCAAACCGTCGTTGTTCTA

GGGAGCCAGGAAGGAGCCGTTCACACGGCTCTTGCTGGAGCTCTGGAGGCTGAGATGGATGGTGCAAAAG

GAAGGCTGTTCTCCGGCCATTTGAAATGCCGCTTAAAAATGGACAAGCTTAGATTGAAGGGCGTGTCATA

TTCCTTGTGCACCGCAGCGTTCACATTTACCAAGGTCCCAGCTGAAACATTGCATGGAACAGTTACAGTG

GAGGTGCAGTATGCAGGGACAGATGGACCCTGCAAGGTCCCAGCCCAGATGGCGGTGGACATGCAGACCC

TGACCCCAGTCGGGAGGCTGATAACCGCCAACCCTGTGATTACTGAAAGCACTGAGAATTCAAAGATGAT

GTTGGAGCTCGACCCACCATTTGGGGATTCTTACATTGTCATAGGAGTTGGGGACAAGAAAATCACCCAT

CACTGGCATAGGAGTGGTAGCACCATCGGAAAAGCATTTGAGGCCACTGTGAGAGGTGCCAAGAGAATGG

CAGTTCTGGGGGATACAGCCTGGGACTTCGGATCAGTCGGGGGCGTGTTCAACTCATTGGGCAAGGGTGT

TCACCAGATTTTTGGAGCAGCTTTCAAATCACTGTTTGGAGGAATGTCCTGGTTCTCACAGATCCTCATA

GGCACGCTGCTGGTGTGGTTGGGTTTGAACACAAAGAATGGATCCATCTCCCTCACATGCTTGGCCCTGG

GGGGAGTGATGATCTTCCTCTCCACGGCTGTCTCTGCTGACGTGGGGTGCTCGGTGGACTTCTCAAAAAG

AGAAACGAGATGTGGCACGGGGGTGTTCGTCTACAATGATGTTGAAGCCTGGAGGGACAGGTACAAGTAC

CATCCTGACTCCCCCCGCAGATTGGCAGCAGCAGTCAAGCAGGCTTGGGAAGAGGGGATTTGTGGGATCT

CCTCCGTTTCAAGAATGGAAAACATCATGTGGAAATCAGTGGAAGGGGAGCTCAATGCTATCTTAGAGGA

GAATGGAGTTCAACTGACAGTTGTTGTAGGATCTGTAAAAAACCCCATGTGGAGAGGCCCACAAAGATTG

CCAGTGCCTGTGAATGAGCTGCCCCATGGCTGGAAAGCCTGGGGGAAATCATATTTTGTCAGAGCGGCAA

AGACCAACAACAGTTTTGTTGTCGATGGTGACACACTGAAGGAATGTCCGCTCAAACACAGAGCATGGAA

TAGTTTCCTTGTGGAGGATCATGGGTTTGGGGTCTTCCACACCAGTGTTTGGCTCAAGGTTAGAGAAGAC

TACTCACTAGAATGTGACCCAGCCGTCATAGGAACAGCTGTTAAGGGAAAGGAGGCCGCCCACAGTGACC

TGGGCTACTGGATTGAAAGTGAAAAGAATGACACATGGAGGCTGAAGAGGGCCCATCTGATTGAGATGAA

AACATGTGAGTGGCCAAAGTCTCACACTCTGTGGACAGATGGAGTAGAAGAAAGTGATCTGATCATACCT

AAATCTTTAGCTGGTCCACTCAGCCACCACAACACTAGAGAGGGTTACAGAACTCAAGCAAAAGGGCCAT

GGCATAGTGAAGAGCTTGAAATCCGGTTTGAGGAATGTCCAGGCACCAAGGTTCACGTGGAGGAGACATG

CGGAACTAGAGGACCATCTCTGAGATCAACCACTGCAAGTGGAAGGGTCATTGAGGAATGGTGCTGTAGA

GAATGCACAATGCCCCCACTATCGTTTCGAGCAAAAGACGGCTGCTGGTATGGAATGGAGATAAGGCCCA

GGAAAGAACCAGAGAGCAACTTAGTGAGGTCAATGGTGACAGCGGGATCAACCGATCACATGGACCACTT

CTCTCTTGGAGTGCTTGTGATTCTACTCATGGTGCAGGAAGGGCTGAAGAAGAGAATGACCACAAAGATC

ATTATGAGCACATCAATGGCAGTGCTGGTGGCCATGATCCTGGGAGGATTTTTAATGAGTGACTTGGCTA

AGCTTGTGATCCTGATGGGTGCCACCTTCGCAGAAATGAACACTGGAGGAGATGTGGCTCACTTGGCATT

GGTTGCGGCATTTAAAGTCAGACCAGCCTTGCTGGTCTCCTTTATTTTCAGAGCTAATTGGACACCTCGT

GAGAGCATGCTGCTAGCCCTGGCTTCGTGTCTTTTGCAAACTGCAATCTCTGCTCTCGAAGGTGACTTGA

TGGTCCTCGTCAATGGATTTGCTTTGGCCTGGTTGGCAATACGAGCAATGGCCGTGCCACGCACTGACAA

CATTGCTCTGGCAATCCTGGCTGCTCTGACACCACTAGCCCGAGGTACACTGCTCGTGGCATGGAGAGCG

GGCCTCGCTACTTGTGGAGGGTTTATGCTCCTCTCCCTGAAAGGGAAAGGTAGTGTGAAGAAGAACCTGC

CATTTGTCATGGCCTTGGGATTGACTGCTGTGAGGATAGTAGACCCCATTAATGTGGTGGGACTACTGTT

ACTCACAAGGAGTGGGAAGCGGAGCTGGCCCCCTAGTGAAGTGCTTACAGCTGTTGGCCTGATATGTGCA

CTGGCCGGAGGCTTTGCCAAGGCAGACATTGAGATGGCTGGGCCCATGGCTGCAGTAGGCTTGCTGATTG

TCAGTTATGTGGTCTCAGGAAAGAGTGTGGATATGTACATTGAAAGAGCAGGTGACATCACATGGGAAAA

GGACGCGGAAGTTACTGGAAATAGTCCTCGGCTTGACGTGGCACTAGATGAGAGTGGTGATTTCTCTCTG

GTGGAGGAAGATGGTCCACCCATGAGAGAGATCATACTCAAGGTGGTCCTGATGGCCATCTGTGGCATGA

ACCCAATAGCCATACCCTTTGCTGCAGGAGCGTGGTATGTGTATGTGAAGACTGGAAAAAGGAGTGGTGC

CCTCTGGGATGTGCCCGCTCCCAAAGAAGTGAAAAAAGGAGAGACCACAGATGGGGTATACAGAGTGATG

ACCCGCAGACTGCTGGGTTCAACACAGGTTGGAGTGGGAGTCATGCAAGAGGGAGTCTTCCACACCATGT

GGCACGTCACAAAAGGAGCTGCACTGAGGAGCGGTGAAGGGAGACTTGATCCATACTGGGGGGATGTCAA

GCAGGACTTAGTGTCATACTGTGGGCCTTGGAAGTTGGATGCAGCTTGGGATGGACTTAGTGAGGTGCAG

CTCTTGGCAGTACCCCCCGGAGAGAGGGCCAGAAACATTCAGACTCTGCCTGGAATATTCAAGACAAAGG

ATGGGGACATCGGAGCAGTCGCTCTGGACTACCCCGCAGGAACTTCAGGATCCCCGATCCTAGACAAATG

CGGAAGAGTGATAGGACTTTATGGCAATGGGGTTGTGATCAAGAATGGAAGCTATGTTAGTGCCATAACC

CAGGGAAAGAGGGAGGAGGAGACTCCGGTTGAATGTTTCGAACCCTCGATGCTGAAGAAGAAGCAGCTAA

CTGTCCTGGATCTGCATCCAGGAGCCGGAAAGACTAGGAGAGTTCTTCCTGAAATAGTCCGTGAAGCCAT

AAAAAAGAGACTCCGCACAGTGATTTTGGCTCCAACCAGGGTTGTCGCTGCTGAGATGGAGGAAGCCTTG

AGAGGACTTCCGGTGCGTTACATGACAACAGCAGTTAACGTCACCCACTCTGGGACAGAAATCGTTGATT

TGATGTGCCATGCCACCTTCACCTCACGCCTATTACAACCCATCAGAGTCCCTAATTACAATCTCTACAT

CATGGATGAAGCCCATTTCACAGACCCCTCAAGTATAGCTGCAAGAGGATACATATCAACAAGGGTTGAA

ATGGGCGAGGCGGCTGCCATCTTCATGACAGCCACGCCACCAGGAACCCGAGATGCATTTCCAGACTCCA

ACTCACCAATCATGGACACAGAAGTGGAAGTCCCAGAGAGAGCCTGGAGCTCAGGCTTTGATTGGGTGAC

AGACCATTCTGGGAAAACAGTTTGGTTCGTTCCAAGTGTGAGAAATGGAAATGAAATCGCAGCCTGTCTG

ACAAAGGCTGGAAAGCGGGTCATACAACTCAGCAGGAAGACTTTTGAGACAGAGTTTCAGAAAACAAAAA

ATCAAGAGTGGGACTTTGTCATAACAACTGACATCTCAGAGATGGGCGCGAATTTCAAAGCTGACCGGGT

CATAGACTCTAGGAGATGCCTAAAGCCAGTCATACTTGATGGTGAGAGAGTCATCTTGGCTGGGCCCATG

CCTGTCACGCATGCTAGTGCTGCTCAAAGGAGAGGGCGTATAGGCAGGAACCCTAACAAACCTGGAGATG

AGTACATGTATGGAGGTGGATGTGCAGAGACCGATGAAGACCATGCACACTGGCTTGAAGCAAGAATGCT

TCTTGACAACATCTACCTCCAGGATGGCCTCATAGCCTCGCTCTACCGGCCTGAGGCCGATAAGGTAGCT

GCCATTGAGGGAGAGTTCAAGCTGAGGACAGAGCAAAGGAAAACCTTTGTGGAACTTATGAAGAGAGGAG

ACCTTCCCGTTTGGCTAGCCTATCAGGTTGCATCTGCCGGAATAACTTACACAGACAGAAGATGGTGCTT

TGATGGCACAACCAACAACACCATAATGGAAGATAGTGTACCAGCAGAAGTGTGGACAAAATATGGAGAG

AAGAGAGTGCTCAAACCAAGATGGATGGATGCGAGGGTCTGTTCCGATCATGCTGCCCTGAAGTCGTTCA

AAGAATTTGCTGCCGGGAAAAGAGGAGTGGCTTTAGGAGTAATGGAGGCCCTGGGAACATTGCCAGGACA

CATGACAGAGAGGTTTCAGGAGGCCATTGACAATCTCGCTGTGCTCATGCGAGCAGAGACTGGAAGCAGG

CCTTACAAGGCAGCGGCAGCCCAACTGCCGGAGACCCTAGAGACCATCATGCTCTTAGGCTTGTTGGGAA

CAGTTTCGTTGGGGATCTTCTTTGTCTTGATGCGGAACAAGGGCATCGGGAAGATGGGTTTTGGAATGGT

AACTCTTGGGGCCAGTGCATGGCTCATGTGGCTTTCGGAAATTGAACCAGCCAGAATTGCATGTGTCCTC

ATTGTTGTGTTCTTGTTACTGGTGGTACTCATACCCGAGCCAGAGAAGCAAAGATCTCCCCAGGACAACC

AGATGGCAATCATCATCATGGTGGCAGTGGGCCTTCTGGGTTTGATAACTGCAAACGAACTTGGATGGCT

GGAAAGAACAAAAAATGACATAGCTCATCTAATGGGAAGGAGAGAAGAAGGAGTAACTATGGGATTCTCA

ATGGACATCGATTTGCGGCCAGCCTCCGCCTGGGCTATCTATGCCGCACTGACAACTCTCATCACCCCAG

CCGTCCAACACGCGGTAACCACTTCATACAACAACTACTCCCTAATGGCGATGGCCACACAAGCTGGAGT

GCTATTTGGCATGGGCAAAGGGATGCCATTTTATGCATGGGACTTTGGAGTCCCGCTGCTAATGATGGGT

TGCTACTCACAATTAACACCCCTAACCTTGATAGTGGCTATCATTCTGCTTGTGGCACACTATATGTACT

TGATCCCAGGCCTACAGGCAGCAGCAGCGCGTGCTGCCCAGAAGAGAACAGCAGCTGGCATCATGAAGAA

TCCCGTTGTGGATGGAATAGTGGTAACTGACATTGACACAATGACAATTGACCCCCAAGTGGAGAAGAAG

ATGGGACAAGTGTTACTTATAGCAGTAGCCATCTCCAGTGCTGTGCTGCTGCGGACTGCTTGGGGATGGG

GGGAGGCTGGAGCTCTGATTACTGCAGCAACCTCCACCTTGTGGGAAGGTTCTCCAAACAAATACTGGAA

CTCCTCTACAGCCACCTCACTGTGCAACATCTTCAGAGGAAGTTACCTGGCAGGCGCTTCTCTTATTTAT

ACAGTGACAAGAAATGCTGGCCTAGTCAAGAGACGTGGAGGTGGAACGGGAGAGACTCTAGGAGAGAAGT

GGAAAGCTCGCCTGAATCAGATGTCGGCCCTGGAGTTCTACTCTTACAAAAAGTCAGGCATCACTGAAGT

GTGTAGAGAGGAGGCTCGCCGCGCCCTCAAGGATGGAGTGGCCACAGGAGGACATGCTGTATCCCGGGGA

AGCGCAAAGCTCAGATGGCTGGTGGAGAGAGGATACCTGCAACCCCATGGAAAGGTTGTTGACCTCGGAT

GTGGCAGAGGGGGTTGGAGCTATTATGCCGCCACCATCCGGAAAGTGCAGGAGGTGAAAGGATACACAAA

GGGAGGTCCCGGTCATGAAGAACCCATGCTGGTGCAAAGCTATGGGTGGAACATCGTTCGTCTCAAGAGT

GGAGTGGACGTCTTCCACATGGCGGCCGAGCCGTGTGATACCTTGCTGTGTGACATAGGTGAGTCATCAT

CTAGTCCTGAAGTGGAAGAGACGCGAACACTCCGAGTGCTCTCTATGGTGGGGGACTGGCTTGAGAAAAG

ACCGGGGGCCTTCTGTATAAAGGTGCTGTGTCCATACACCAGCACTATGATGGAGACCATGGAGCGCCTG

CAACGTAGGTATGGGGGAGGATTAGTCAGAGTGCCATTGTCTCGCAACTCCACACATGAGATGTATTGGG

TCTCTGGAGCAAAAAGTAACGTCATAAAAAGTGTGTCCACCACAAGTCAGCTCCTCCTGGGACGCATGGA

TGGTCCCAGGAGGCCAGTGAAATATGAGGAAGATGTGAACCTCGGCTCGGGCACACGAGCTGTGGCAAGC

TGTGCTGAAGCTCCCAACATGAAGGTCATTGGCAAGCGCATTGAGAGAATCCGCAGTGAACATGCAGAAA

CATGGTTCTTTGATGAAAACCATCCATACAGGACATGGGCCTACCATGGGAGCTACGAAGCCCCCACACA

AGGGTCAGCATCTTCCCTCGTGAATGGGGTTGTTAGACTCCTGTCAAAGCCCTGGGATGTAGTGACTGGA

GTTACAGGAATAGCTATGACTGATACAACACCATACGGCCAACAAAGAGTCTTCAAAGAAAAAGTGGACA

CCAGGGTGCCAGATCCCCAAGAAGGCACTCGCCAGGCAATGAACATGGTCTCTTCCTGGCTGTGGAAGGA

GTTGGGGAAACGCAAGCGGCCACGTGTCTGCACCAAAGAAGAGTTCATCAACAAGGTGCGCAGCAATGCA

GCACTGGGAGCAATATTTGAAGAGGAAAAAGAATGGAAGACGGCCGTGGAGGCTGTGAATGATCCAAGGT

TTTGGGCTCTAGTGGACAAGGAAAGAGAACATCACCTGAGAGGAGAGTGTCATAGCTGTGTGTACAACAT

GATGGGAAAAAGAGAGAAGAAGCAAGGAGAATTCGGAAAAGCAAAAGGCAGCCGCGCCATCTGGTACATG

TGGTTGGGAGCCAGATTCTTGGAGTTTGAAGCCCTTGGATTCTTGAATGAGGACCATTGGATGGGAAGAG

AAAACTCAGGAGGTGGTGTCGAAGGGCTAGGACTGCAAAGACTTGGATACATTCTAGAAGAAATGAATCG

GGCACCAGGAGGAAAGATGTATGCAGATGACACCGCTGGCTGGGACACCCGCATTAGTAAGTTTGATCTG

GAGAATGAAGCCTTAATTACCAACCAAATGGAGGAAGGGCACAGAGCTCTGGCGTTGGCTGTGATTAAGT

ACACAAAAAACAAAGTGGTGAAGGTCCTCAGACCAGCTGAAGGAGGGAAAACAGTTATGGACATCATTTC

AAGACAGGACCAGAGAGGGAGTGGACAAGTTGTCACCTATGCTCTCAACACATTTACCAACTTAGTGGTG

CAGCTTATCCGGAATATGGAGGCTGAGGAAGTGTTAGAGATGCAAGACTTATGGCTGTTGAGGAAGCCAG

AGAAAGTGACCAGATGGTTGCAGTGCAATGGATGGGACAGACTCAAACGAATGGCAGTCAGTGGAGATGA

CTGCGTTGTGAAGCCAATTGATGATAGGTTTGCGCATGCCCTCAGGTTCTTAAATGACATGGGAAAAGTT

AGGAAAGACACACAGGAATGGAAACCCTCGACTGGATGGAGCAATTGGGAAGAAGTCCCGTTCTGCTCCC

ACCATTTCAACAAGCTGCACCTCAAGGATGGTAGATCCATTGTGGTCCCTTGCCGCCACCAAGATGAATT

GATTGGCCGAGCCCGTGTATCACCAGGGGCAGGATGGAGCATTCGGGAGACTGCCTGTCTAGCAAAATCA

TATGCACAGATGTGGCAGCTTCTTTACTTCCACAGAAGAGACCTTCGACTGATGGCCAATGCTATTTGTT

CGGCTGTGCCAGTTGACTGGGTACCAACCGGGAGAACCACCTGGTCAATCCACGGAAAGGGAGAATGGAT

GACTACTGAGGACATGCTCATGGTGTGGAATAGAGTGTGGATTGAGGAGAACGACCATATGGAGGACAAG

ACTCCTGTGACAAAGTGGACAGACATCCCCTATCTGGGAAAAAGGGAGGACTTATGGTGTGGATCCCTTA

TAGGGCACAGGCCCCGCACCACTTGGGCTGAGAACATCAAAGACACAGTCAACATGGTGCGCAGGATCAT

AGGTGATGAAGAAAAATACATGGACTACTTATCCACCCAGGTCCGCTACTTGGGTGAGGAAGGGTCCACA

CCTGGAGTGTTGTAAGCACCAATTTCAGTGTTGTCAGGCCTGCTAGCCAGCCACAGTTTGGGGAAAGCTG

TGCAGCCTGTAGCCCCCCCAGGAGAAGCTGGGAAACCAAGCTCATAGTCAGGCCGAGAACGCCATGGCAC

GGAAGAAGCCATGCTGCCTGTGAGCCCCTCAGAGGACACTGAGTTAAAAAACCCCACGCGCTTGGAAGCG

CAGGATGGGAAAAGAAGGTGGCGACCTTCCCCACCCTTTAATCTGGGGCCTGAACTGGAGACTAGCTGTG

AATCTCCAGCAGAGGGACTAGTGGTTAGAGGAGACCCCCCGGAAAACGCAAAACAGCATATTGACGCTGG

GAAAGACCAGAGACTCCATGAGTTTCCACCACGCTGGCCGCCAGGCACAGATCGCCGAACAGCGGCGGCC

AGTGTGGG