Influenza Genome Sequencing Project Proposal

Date of Proposal (MM/DD/YY): __06/13/12_____

TITLE: Role of the Untranslated Regions of the Influenza A Virus Replication and Vaccines

Purpose/Objective-please provide a brief one to two paragraph description of the proposed project. Include justification and relevant background.

Background on influenza A virus untranslated regions.

This project represents a collaborative effort between NIH/NIAID (Dr. David Spiro), the Biomedical Advanced Research and Development Authority (BARDA) (Dr. Armen Donabedian and Dr. Mario Barro), and the J. Craig Venter Institute (Dr. David Wentworth) and provide critical insight into the untranslated regions (UTRs) play in influenza A virus replication and reassortment.

Each of the 8 viral genomic RNA segments contains UTRs at the 5' and 3' termini that play critical roles in transcription, replication, and segment specific packaging of the eight vRNAs into a virion during morphogenesis. The termini of each vRNA interact via base pairing between the conserved residues at the 5' and 3' termini of the vRNA segments. The UTRs range from ~20-40 nucleotides and all contain 13 nucleotides at the 5' end and 12 nucleotides at the 3' end are believed to be highly conserved among all vRNA segments (Fig. 1). The nucleotides that juxtapose these 12-13 termini show segment, subtype, and strain specific differences. Even within the highly conserved 3' terminus there is known variation at the 4th position (U/C). The vRNA promoter is defined as a double-stranded element formed by the conserved 5' and 3' terminal ends of the vRNA molecule. Multiple models (e.g., panhandle vs corkscrew) of this interaction exist and experimental results sometimes contradict each model (Fig 1).

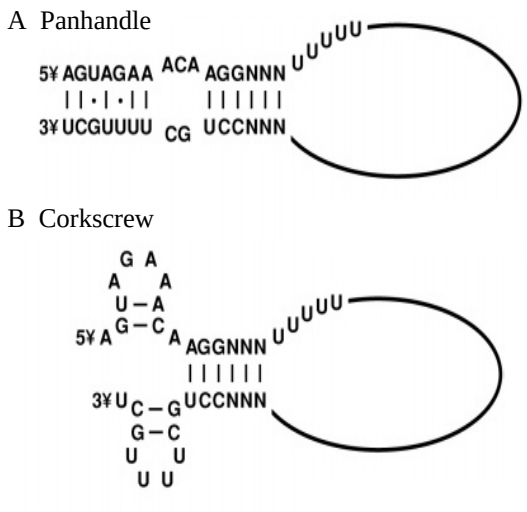

Fig 1. Proposed models of vRNA termini base pairing. (A) In the Panhandle model the 5' and 3' termini of each segment pair to form the promoter. (B) In the Corkscrew model, the 3' and 5' extreme termini form stem-loop structures within themselves before the two strands base-pair with each other. N represents positions that are segment and strain specific. (1).

Gaps in knowledge.

We don't have a complete understanding of the structure(s) formed by 5' and 3' UTRs and it is likely that different conformations exist at different stages of the virus life cycle and different models may explain different viral activities. Despite the central role the UTR's of influenza vRNAs play in replication, genome packaging, and evolution there is very limited sequence information available in sequence databases. Analysis of Genbank as of 05/21/2012 shows that the percentage of each RNA segment ranges from 7-12% (Table 1) and many of these "complete sequences" are incorrect because the primer sequences used to amplify the virus were inappropriately deposited as part of the sequence that was amplified.

Table 1. Percentage of each vRNA segment containing a UTR

RNA Segment	PB2	PB1	PA	HA	NP	NA	M	NS
Percent UTR +	8.0%	8.0%	7.7%	7.6%	10.4%	6.9%	11.9%	10.3%

The true terminal sequences for the 8 segments represent a minority of the sequences in the database because the methods used to sequence influenza viruses typically use RT-PCR amplification with oligonucleotide primers predicted to bind the termini or primers designed only to amplify the coding

region. While the current strategy used by the JCVI influenza GSC captures more of the UTR than most approaches, it is not complete. Therefore, we don't have a strong idea of the level of UTR variation between virus strains, or between specific segments within a strain of virus.

Public health significance and community resource.

A study designed to sequence the actual UTRs of influenza viruses (particularly reassortants) will elucidate how the UTR variations influence the ability of two viruses to reassort or replicate efficiently in cells from different species, which is critical to evolution of viruses with pandemic potential and to producing high yield vaccine seed stocks. Influenza virus vaccines are the result of a genetic combination of a high growth master donor virus that is adapted for growth in eggs (e.g., A/PR/8/34) with the HA and NA from a highly prevalent seasonal wild type virus. These vaccine seed stocks are created by reassortment between the donor and wild type and they are can be difficult to generate or have poor yield which delays vaccination campaigns and reduces the number of doses available. Additionally, before the reassortant can be used in a manufacturing setting, further adaptation in chicken embryo's or cell lines is required. An extensive collaborative effort is ongoing to understand the genetic rules driving the optimization of a vaccine candidate. Many changes can occur along the genome of the new vaccine and sequence information that will help us understand the impact of the HA and NA UTR regions on reassortant rescue and high yielding virus production is critical. The sequence data will be made publically available, and a paper detailing the influence the UTR's have on reassortment, which will serve as a resource for vaccine research and industry communities to improve future vaccine creation.

Goals and Study Design.

The primary goal of this project is to determine how sequence variation in the UTR's of the HA and NA impact the replication/fitness of influenza A viruses by focusing our analysis on the UTR's high yield reassortants used as vaccine seed stocks and selected naturally circulating strains. The following objectives will be used to accomplish this goal:

1. The sequence of the complete genomes including the HA and NA UTRs of low and high yield reassortant virus stocks and their parental viruses. This will be done by combining the M-RTPCR process currently used in the JCVI influenza sequencing pipeline with RNA ligation RT-PCR that we have developed.
    a. RNA ligation is used to link the 5' and 3' termini, which circularizes the genomic RNAs, then we use RT-PCR to amplify the region across the junction that contains the UTR.
    b. These will then be sequenced using our SISSPA and next generation pipline that employs Roche 454, and Illumina HiSeq technologies. As an alternative we will use Ion torrent to sequence these small RNA-ligation RT-PCR amplicons.
    c. Some of the viruses to be studied have been sequenced or are in process via the M-RTPCR based pipeline. For these samples we will only need to add RNA ligation RT-PCR data to complete the UTRs.
2. We will also investigate changes in the HA and NA UTRs of a subset of ~6-10 reassortants vaccine seeds which have been serially passaged in mammalian cell culture or eggs to determine if the substrate/species used to propagate the viruses selects for changes in the UTRs that enhance growth under specific conditions.

Completion of these objectives will provide complete genomic sequence information, including accurate UTR sequences for approximately 60 viruses and this information can be used by to ensure virus/antigen yield at the outset of creating vaccine seed stocks. Additionally, techniques developed in this project could be incorporated into the influenza genome sequencing pipeline to generate complete UTR data for all new sequencing projects, which would dramatically improve the database.

References.

1. Palese P, Shaw ML. Orthomyxoviridae: The Viruses and Their Replication. In: Knipe DM, Howley PM, Griffin DE, et al., editors. Fields Virology. 5 ed. Philadelpha, PA: Lippincott Williams and Wilkins; 2007. p. 1647-90.

Description of Collection: Please briefly describe the collection of viral strains to be sequenced and also include an Excel Spreadsheet detailing: These are parental viruses and vaccine seeds that were generated from them. All will be from allantoic fluid or tissue culture supernatants.

1. A/Vietnam/1203/04 (H5N1)
2. A/Indonesia/5/05/2005 (H5N1)
3. A/India/NIV/2006 (H5N1)
4. A/Egypt/321-Namru3/2007 (H5N1)
5. A/Egypt/3300-NAMRU3/2008 (H5N1)
6. A/Egypt/N03072/2010 (H5N1)
7. A/Hubei/1/2010 (H5N1)
8. A/Anhui/01/2005 (H5N1)
9. A/chicken/VietNam/NCVD-016/2008 (H5N1)
10. A/chicken/Vietnam/NCVD-03/2008 (H5N1)
11. A/HI/7/09(H3N2)-PR8-IDCDC-RG23 (H3N2)
12. A/Wisconsin/15/2009(H3N2)-PR8-IDCDC-RG24 (H3N2)
13. A/Texas/05/2009(H1N1)-PR8-IDCDC-RG15 (HN1)pdm09
14. A/Texas/05/2009(H1N1)-PR8-IDCDC-RG18 (HN1)pdm09
15. A/Texas/05/2009(H1N1)-PR8-IDCDC-RG20 (HN1)pdm09
16. A/New York/18/2009(H1N1)-PR8-IDCDC-RG22 (HN1)pdm09
17. A/Hong Kong/33982/2009(H9N2)
18. A/New Jersey/11/1976 (H1N1)
19. A/St. Petersburg/8/2006 (H1N1)
20. A/South Dakota/06/2007 (H1N1)
21. A/California/07/2009 (H1N1)
22. A/Beijing/32/1992 (H3N2)
23. A/Johannesburg/33/1994 (H3N2)
24. A/Panana/2007/1999 (H3N2)
25. A/Ulan Ude/01/2000 (H3N2)
26. A/New York/55/2004 (H3N2)
27. A/New York/55/2004 (H3N2)
28. A/Nepal/921/2006 (H3N2)
29. A/Uruguay/76/2007 (H3N2)
30. X-53A
31. NYMC X-163
32. NYMC X-173C
33. NYMC X-179
34. X-117
35. X-123
36. X-141
37. X-143
38. NYMC X-157A
39. NYMC X-157D
40. NYMC X-165

41. NYMC X-175C

NIAID review of relevant IRB/IACUC documentation is required prior to commencement of work.

NIAID supports rapid genomic and meta data and reagent release to the scientific community for all sequencing and genotyping projects funded by NIAID Genomc Sequencing Center (GSC). It is expected that projects will adhere to the data and reagent release policy described in the following web sites:

http://www3.niaid.nih.gov/research/resources/mscs/data.htm

http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-013.html

Agree to NIAID's Genome Sequencing Centers' Data Release Policy?

Yes _____X____ No _____

Influenza Genome Sequencing Project Proposal

Once a proposal is approved, NIAID GSC will establish an agreement to be reviewed and approved by NIAID.

Principal Investigator:

Name	David E. Wentworth
Position:	Associate Prof./Director of Viral Programs
Address:	J. Craig Venter Institute, 9704 Medical Center Drive, Rockville
State	Maryland
ZIP Code	20850
Telephone	301-795-7625
Fax	
E-Mail	dwentwor@jcvi.org

Organization/Company:

Name:	J. Craig Venter Institute
Address:	9704 Medical Center Drive, Rockville
State	MD
ZIP Code	20850

Principal Investigator Signature:

Institutional Official Signature:

NIAID Project Officer

Maria Y. Giovanni, Ph.D.
Assistant Director for Microbial Genomics & Advanced Technology
Division of Microbiology and Infectious Diseases
NIAID/NIH/DHHS
6610 Rockledge Dr. MSC 6603; Room 4807
Bethesda, MD 20892-6603
301-496-1884 Phone
301-480-4528 FAX
mgiovanni@niaid.nih.gov

NIAID Influenza Genome Sequencing Project Manager

Punam Mathur
Health Specialist for Microbial Genomics
Division of Microbiology and Infectious Diseases
NIAID/NIH/DHHS
6610 Rockledge Dr. MSC 6603; Room 4805
Bethesda, MD 20892-6603
301-496-1884 Phone
301-480-4528 FAX
mathurpu@niaid.nih.gov