White Paper Application

Project Title: **Role of microbiome diversity, microbial transcriptome and host gene expression in the respiratory track of humans in influenza virus infection.**

Authors: Rafael A. Medina, David E. Wentworth, Karen Nelson, Barbara Methe, Marcela Ferres, and Adolfo García-Sastre.

Primary Investigator Contact:

Name	Adolfo García-Sastre
Position	Professor
Institution	Mount Sinai School of Medicine
Address	One Gustave L. Levy Place, Annenberg Bldg. 16.90
State	New York
ZIP Code	10029
Telephone	1-212-241-7769
Fax	1-212-241-1684
E-Mail	Adolfo.Garcia-Sastre@mssm.edu

1. Executive Summary *(Please limit to 500 words.)*

Provide an executive summary of the proposal.

Infection susceptibility and disease severity is regulated by a complex balance between the genotype of the pathogen, the genotype of the host, and environmental factors. While many studies are being conducted to evaluate the contribution of these three factors in many infectious diseases, another variable that until recently has been overlooked is the metagenome (or microbial community including viruses, bacteria and microeukaryotes which inhabit the human body). Many viruses, including influenza virus, infect through mucosal surfaces that are colonized with normal flora. Years of co-evolution of a virus with its host and its metagenome have likely selected for viruses that are able to infect a host in the presence of the most prevalent metagenome, so the characteristics of the metagenome are likely to make an impact in infection susceptibility and disease outcome. Moreover, the host innate and adaptive immune factors elicited in response to viral infection are likely to make a big impact in the normal (basal) flora, which may influence the susceptibility to pathogenic bacteria colonization. We are specifically interested in: i) the effect of seasonality in the diversity and expression profile of the metagenome, ii) the role of the respiratory tract metagenome in promoting or inhibiting infection by influenza viruses, and iii) how changes in the metagenome after viral infection may promote secondary bacterial colonization and enhance disease. Currently we still lack fundamental studies that address how variable the metagenome at the URT is among humans, how this metagenome affects susceptibility to respiratory virus infection and how the metagenome responds to virus infection in the general population.

2. Justification

Gaps in knowledge:

Our present understanding of the nature and extent of the core or at least the set of common microbial members of the upper respiratory track (URT) of humans is limited. An effort to establish these basic components is ongoing as part of the Human Microbiome Project (HMP), a NIH Roadmap initiative. Furthermore, there we have little if any understanding of the influence acute respiratory infections, such as influenza virus, have on

the URT microbiome, nor do we understand the role the URT microbiome plays in host susceptibility to infection.

Information that will be obtained:

This study will fill some of these gaps by determining the baseline status of the human URT microbiome and changes that occur during the course of an influenza viral infection. In addition, we propose to obtain the full viral genome sequence as well as the host gene expression profile. In depth analysis of the changes in the viral genome and the host gene expression profiles during an acute influenza infection of the URT will complement the results obtained through the HMP. Additionally, comparison of microbiome of uninfected individuals with those experiencing mild, or sever disease will begin to elucidate the role the microbiome plays in and disease severity. These studies will also provide further assessment of the power and limitations of deep sequencing techniques to obtain in depth information and gain a global picture of the factors that affect disease status.

We propose to carry out this investigation to complement data originated through the HMP and as a companion investigation to a study in which the in the microbial diversity and microbiome expression in the upper respiratory track of ferrets (*Mustela putorius furo*) during influenza virus infection is being pursued.

Since we have such a limited understanding of the microbiome-influenza interaction(s), the proposed study has broad implications and could advance: 1) diagnostics by identifying signatures of influenza infection in the microbiome profile; 2) therapeutics or preventatives that act through strengthening the host microbiome; and/or 3) epidemiology by understanding the role the microbiome plays in disease severity and potentially in host susceptibility to human influenza virus infection.

3. Rationale

Hypothesis: We hypothesize that, during an acute viral infection the normal residing microbiome of the human URT will be affected by the antiviral state triggered in cells of the local URT resulting in the modulation in the diversity and/or quantity of the microbial population.

We aim to:

a) Establish the base line microbial communities present in the URT of humans under a normal health state (i.e. in the absence of known opportunistic viral infection or other microbial infection);

b) Determine the changes that microbial communities undergo over the course of influenza A virus infection.

- This will be assessed by identifying specific microbial populations that might be favored, altered, or reduced during influenza infection. We propose to look at these aspects through 16S sequencing of the V1-V3 and V3-V5 hypervariable regions of the rDNA gene. We will use the 16S sequencing on all samples to group individuals and time points studied, in order to guide the comprehensive metagenomics analysis on ~20% of the specimens.

c) Obtain the full viral genome to characterize the genotype composition of the virus isolated from each infected individual.

d) Determine the local URT gene expression levels of the host throughout the course of infection.

Our long terms goals are to:

a) Determine whether the increase or decrease of specific microorganisms correlates with the modulation of disease severity.

b) Determine whether specific metagenomes increase susceptibility to influenza virus infection and weather they participate in promoting influenza virus seasonality.

In order to do this, we will:

a) Use procedures developed for human microbiome projects to assess the transcriptome of the metagenome in the URT of humans by RNAseq.

4a. Approach to Data Production: Data Generation

Study design

We will conduct this study in two phases to optimize the type and amount of data that will be generated from the human samples analyzed. In the first phase we will determine the methodological parameters needed to obtain basal levels of the microbiome and the host gene expression of 10 age matched uninfected control individuals enrolled in the study. In the second phase, we will use the information obtained during the first phase to optimize and implement a standard procedure to examine changes in the microbiome diversity and transcription status, determine the changes in the host gene expression levels and obtain the full viral genome of samples obtained from individuals throughout the course of influenza A virus infection.

Phase 1

To optimize the overall experimental design, we will perform base line experiments to define specific parameters that will influence the analysis we are proposing. Thus, we will:

a) Conduct comparative analysis between nasal washes and nasal swabs to determine if both sampling procedures provide the same bacterial community and host gene profile information;

b) Determine if the nasal sampling disrupts the normal nasal microbial community of subjects that are being sampled longitudinally at specific intervals (e.g. days 0, 1, 2, 3, 5, 7 and 28);

c) Define the basal levels of the human URT microbiome diversity, as well as identify any changes and the host gene expression levels in uninfected individuals overtime.

d) Establish the level of human RNA present in the samples that will be processed for RNAseq to see if it will be sufficient for reliable in depth analysis of host gene expression.

To answer points a-d above, we will process samples from 10 control individuals (matched by age and gender) from whom nasal swabs and nasal washes will be sampled longitudinally. The samples from the 10 uninfected individuals will be used to establish the initial overall baseline levels overtime to: firstly evaluate the variation in the metagenome

diversity and in host gene expression levels from individual to individual; and to compare the level of variation of both of these parameters for each subject overtime. Additionally, we will use the data generated from these experiments to assess whether the comparison among samples and subjects allows sufficient power to find differences or changes in viral and/or other microbiome populations and in gene expression. The sampling will consist of obtaining from the same subject, firstly a nasal swab, followed by a nasal wash at each time point (e.g. days 0, 1, 2, 3, 5, 7 and 28). All samples will be subjected to 16S PCR and sequencing to survey the microbiome diversity of the URT. The data from these samples will help us better gage the level of sequence coverage we should be aiming for in order to sample, viral and host transcripts at the appropriate level.

Phase 2

Once the appropriate sampling technique is determined empirically, pediatric sequential nasal samples will be collected from 40 severe and 40 non-severe individuals with confirmed influenza infection, and from 40 uninfected controls.

Viral shedding has been shown to occur for approximately one-week post infection. Thus, nasopharyngeal swabs and/or nasal washes (if deemed necessary) will be obtained at days 0 (this is the day of recruitment, e.g. baseline sample), 1, 2, 3, 5, 7 and 28 from individuals with confirmed infection with a human influenza A virus. To rule out co-infections, all samples from day 1 will be analyzed by a commercially available RT-PCR test against 14 common respiratory viral infections. The sequential sampling proposed will allow us to evaluate the: metagenome diversity, virus infection kinetics (titer), viral genome diversity (quaciespecies) and the microbionta status over the course of a human A influenza infection. Finally, the sample collected on day 28 will be useful to evaluate whether the individuals have returned baseline levels after the infection or if a new baseline is set by infection. At this time we will also collect a blood sample to assess the strain specific seroconversion (antibody response) of each individual. Samples will be collected or resuspended in viral transport media (VTM) for further processing. The sample suspensions will be submitted for processing and sequencing at the J. Craig Venter Institute.

To analyze the reproducibility of the results, we will conduct the first part of phase 2 in pediatric samples obtained from 40 uninfected (control) children. Subsequently, pediatric samples obtained from 40 severe and 40 non-severe individuals with confirmed influenza infection will be conducted. We have defined as “severe”, as those individuals requiring hospitalization, and “non-severe” as those individuals with confirmed infection but not needing hospitalization. Moreover, for the severe individuals, most of these patients are administered antivirals and therefore we will only analyze samples from these individuals.

Metadata:

In all cases we are collecting the standard metadata and additional clinical data as established by the CEIRS human metadata excel sheet (see attached file). The data collected will included: Zip; STATE; City; TEMPERATURE_ F; Sudden_Onset; ONSET_HOURS; VACC_STATUS ; VACC_DATE; Fever; SYMPTOM_MYALGIA; SYMPTOM_COUGH; SYMPTOM_RUNNOSE; SYMPTOM_HEADACHE; SYMPTOM_FATIGUE; SYMPTOM_THROAT; PRE_VISIT_MED; PRE_VISIT_PAIN_RELIEVER; PRE_VISIT_ANTIHISTAMINE; PRE_VISIT_THROAT; PRE_VISIT_OTHER; PRE_VISIT_OTHER_Description;

MEDICAL_ASTHMA; MEDICAL_CHRONIC; MEDICAL_CONGESTIVE; MEDICAL_DIABETES; MEDICAL_IMMUNO; MEDICAL_CHRONICRENAL; MEDICAL_ATHERO; TEST_DATE; RAPID_TEST_RESULT; RAPID_TEST_RESULT2; Test for influenza virus;Influenza test result; HA subtype determined by NA subtype determined by;patient diagnosis Influenza; patient diagnosis bronchitis; patient diagnosis pneumonia; patient diagnosis strep; patient diagnosis other viral illness; patient diagnosis other bacterial illness patient diagnosis other viral illness details patient diagnosis other bacterial illness details; POST_VISIT_MED; POST_VISIT_NEURA; POST_VISIT_TAMIFLU; POST_VISIT_RELENZA; POST_VISIT_OTHER_ANTIVIRAL; POST_VISIT_AMANTADINE; POST_VISIT_rimantidine; POST_VISIT_antibiotic; POST_VISIT_antihistamine; POST_VISIT_steroids; POST_VISIT_other;POST_VISIT_other_details; TRAVEL_US;TRAVEL_OUTSIDE_US.
For this study we have a full-time nurse collecting samples and a number of Infectious Disease MD Fellows that are actively working to obtain the clinical metadata.

4b. Approach to Data Production: Data Analysis

For this phase of the project we will extract DNA for 16S sequencing and total RNA for RNAseq, and to obtain the full viral genome from each individual. The bacterial microbiome of all the specimens will be examined by sequencing a fragment of the 16S rDNA (V3-V1) hypervariable region using multiplex Roche 454 sequencing platform. For RNAseq analysis of the transcripts present we will remove ribosomal RNA from the total RNA, cDNA created and libraries constructed for sequencing on the Illumina HiSeq platform. The transcripts will be mapped to the human genome and their relative abundance determined. Viral sequencing will be done using the next generation influenza sequencing pipeline at the JCVI which uses specialized multi-segment RT-PCR (M-RTPCR) procedure, followed by multiplexed Next Generation (NextGen) sequencing strategies to enable high throughput sequencing of hundreds of bar coded samples.

5. Community Support and Collaborator Roles:

Relevance to the scientific community:

Our current understanding of the variability of the basal metagenome of the URT in humans is limited. Moreover, how this metagenome responds to virus infection and how the metagenome might affect the susceptibility to influenza virus infection has not been previously evaluated. Recent data has indicated that comorbidities and/or pre-existing conditions, such as diabetes, obesity, chronic lung disease, including asthma, among other can increase the risk of severe influenza virus infection. Establishing the role that the metagenome of the URT plays during influenza virus infection will provide the influenza virus community an in depth global understanding of additional factors that influence influenza virus pathogenesis. This study, together with the studies currently being performed in the ferret model, will also yield additional information to validate and establish the strengths and limitations of this animal model of viral pathogenesis. Moreover, data arising from this study might be useful in establishing novel approaches to inhibit viral infection of the nasopharyngeal region and will further establish the effect of the antiviral state induced upon influenza virus infection and how this affects the microenviroment and the microbial population of the humans. Collectively the data in this study has the great potential to opening novel areas of research and to provide new tools and targets for potential therapies. Therefore this study will generate highly relevant data

that will be of great interest to clinical and basic scientists in the field.

Project sites, collaborators and roles on the project:
This study will be performed in close collaboration between Mount Sinai School of Medicine, Pontificia Universidad Católica de Chile and the J. Craig Venter Institute. The project collaborators are as follows:

Adolfo García-Sastre, Ph.D. (PI), Mount Sinai School of Medicine: Dr. García-Sastre will be the project coordinator in the USA and will lead the study and serve as liaison between the investigators at the 3 different institutions involved in the study. He will also be responsible for coordinating follow up experiments during the different stages of the pilot study and to validate the results obtained.

Rafael A. Medina, Ph.D. (Co-PI), School of Medicine, Pontificia Universidad Católica de Chile: Dr. Medina will be in charge of collecting and processing the human samples in the laboratory under the approved IRB protocol. He will maintain the patient clinical metadata database and will also maintain close communication with Dr. Wentworth to coordinate the shipments of samples to JCVI for deep sequencing.

David Wentworth, Ph.D. (Co-PI), J. Craig Venter Institute: Dr. Wentworth will supervise the sample sequencing and data analysis. He will communicate directly with Drs. García-Sastre and Medina to coordinate sample shipments and interpretation of results through the different stages of the study.

Karen Nelson, PhD. (co-investigator) J. Crag Venter Institute: Dr. Nelson will participate in the data generation and analysis of the microbiome of the human upper respiratory tracts and will coordinate/communicate with Dr. Wentworth as well as Drs. García-Sastre and Medina.

Barbra Methe, Ph.D. (co-investigator), J Craig Venter Institute: Dr. Methe will help in the interpretation of the microbiome data and identify any changes needed in the study design after phase 1.

Marcela Ferres, M.D., M.P.H. (Co-PI), School of Medicine, Pontificia Universidad Católica de Chile: Dr. Ferres will supervise and oversee the patient enrolments and the collection of clinical metadata. Her laboratory will also perform the diagnostic test to confirm influenza virus infections and co-infections.

The current available funding to support this study are:

CONICYT 1121172: Virulence and antigenic determinants of influenza A virus infection in humans.
PI: Medina, Co-PI: Ferres
The purpose of this study is to use a clinical and basic research approach to gain a better understanding of the viral factors that contribute to the morbidity produced by seasonal influenza virus infection in the adult and pediatric population. U$ 355,059 (2012 – 2014)

CONICYT 79100014: Project of Insertion of Human Capital to the Academy: Strengthening the interdisciplinary research in molecular virology.

PI: Ferres, Co-PI: Medina
This grant provides funds to incorporate Dr. Medina Silva as a new investigator in the field of virology and to conduct novel collaborative and interdisciplinary research at the School of Medicine at Pontificia Universidad Catolica de Chile. U$ 104,536 (2011 – 2013)

NIH-CEIRS program HHSN266200700010C**:** Center for Research on Influenza Pathogenesis (CRIP), an NIAID funded Center for Excellence in Influenza Research and Surveillance (CEIRS).
PI: García-Sastre
This center is dedicated to determine factors affecting the pathogenicity and innate immune responses of influenza viruses. U$ 6,091,103 (this year, total costs) (2007-2014)

6. Availability & Information of Strains:

1.

We will focus this study on samples obtained from pediatric individuals with confirmed influenza virus infection, such as the H1N1 pandemic virus or the H3N2 seasonal virus (determined by qRT-PCR diagnosis). We have 16 samples during 2011 including detailed metadata that already available for this study. In addition, we have begun to recruit new individuals during the 2012 influenza season and we are currently recruiting the samples from the control individuals. Thus, we anticipate having the first group of control samples available to begin Phase I of the study (see below), in the next month. All the samples have been and will be collected by the laboratory of Dr. Medina through an approved IRB clinical protocol at Pontificia Universidad Católica de Chile School of Medicine, in Santiago, Chile. This IRB protocol aims to collect sequential nasopharyngeal samples from adults and children during the 2011 – 2013 winter seasons in Chile.
All the human samples sequenced through this pilot project will include clinical metadata required and established by the CEIRS network and will be made available as per NIAID data release policy.

Alternative samples:
If we are not able to obtain enough pediatric specimens for the proposed study we will use adult samples which we expect to be more abundantly available during the 2012 influenza season.

7. Compliance Requirements:

7a. Review NIAID's Reagent, Data & Software Release Policy:

NIAID supports rapid data and reagent release to the scientific community for all sequencing and genotyping projects funded by NIAID 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/LabsAndResources/resources/mscs/data.htm

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

Once a white paper project is approved, NIAID GSC will develop with the collaborators a detailed data and reagent release plan to be reviewed and approved by NIAID.

Accept ☒ Decline ☐

7b. Public Access to Reagents, Data, Software and Other Materials:

All sequence data generated under this project will be released to GenBank after 45 days of pre-access. The influenza virus genomes will be assembled via JCVI's assembly pipeline, annotated and deposited in GenBank, the metagenomic sequencing data will be submitted to the short-read archive (SRA) database while the attached metadata will be submitted to dbGAP. Phase 1 and phase 2 data will be released independently.

7c. Research Compliance Requirements

Upon project approval, NIAID review of relevant IRB/IACUC documentation is required prior to commencement of work. Please contact the GSC Principal Investigator(s) to ensure necessary documentation are filed for / made available for timely start of the project.

Investigator Signature:

Investigator Name: **Adolfo García-Sastre** **Date** **June 12, 2012**