**eMERGE Network Proposal for Analysis**

Project/Manuscript Concept Sheet

|  |  |
| --- | --- |
| **Submission Date** | **Reference number: NT113.1**  8/13/2014 (original approved version)  2nd version: 2022 |
| **Project Title** | Multiscale Analysis Of Influenza Host-Pathogen Interactions: Fluomics |
| **Tentative Lead Investigator (first author)** | Ellie Sang Sukerman |
| **Tentative Senior Author (last author)** | Steven M. Wolinsky |
| **All other authors** | Eun-Young Kim, Sudhir Penugonda, Adolfo Garcia-Sastre, Megan Shaw, Kelsey Haas, Nevan Krogan, Jennifer Allen Pacheco, Maureen E. Smith, Daniel Whorf, Caroline Keller, Tiffany Kim, Heejin Kim, Jacqueline Kirby, Joshua C. Denny, other eMERGE members from sites that provided DNA & data in eMERGE phase II |
| **Sites Involved** | Northwestern, Marshfield, Mt. Sinai, Vanderbilt, CCHMC |
| **Background / Significance** | Influenza A virus causes significant morbidity and mortality, with more than 200,000 hospitalizations and 36,000 deaths attributable to influenza in the United States annually. A differential advantage in responding to infectious disease is associated with differences in the genes relevant to the virus life cycle and immunity. |
| **Outline of Project** | *Please note this is an update to an already approved MCS, the data has already been collected and analyzed, and no new analysis is planned.*  To gain insight into the biological basis of influenza A virus replication and the host innate immune response to infection, we are conducting a cohort study to determine the contribution of identified host factors for which a plausible biological mechanism of action is found to the virus life cycle. We will use DNA samples from individuals with documented Influenza A infection as determined by PCR-based assays or viral culture performed as part of routine medical care. |

|  |  |
| --- | --- |
| **Desired**  **Variables (essential for analysis**  **indicated by \*)** | * **Demographics inc. gender, race, ethnicity & age\*** * Median BMI\* * Co-morbid conditions from standard codes & meds\* * Markers of disease severity\*   + **Treatment setting (outpatient v. inpatient v. ICU)**   + from standard codes   + Clinical outcome (i.e. death, or discharged to home or other place, if known) * Complications of influenza\* * Treatment (Tamiflu) & (flu) vaccines\* * Method of flu testing & flu subtype * Any respiratory or nasal culture results |
| **Desired data** | Phenotype data:  Influenza infection phenotype will be determined based on clinical information obtained from electronic health records. We will take everyone with a diagnosis of influenza A by PCR assay or culture, and then divide them by phenotype based on the clinical information with which we are provided. Subjects will be divided into “extreme” phenotypes based primarily on treatment setting.  Genotype data:  Genomic DNA samples from each patient will be used for whole exome sequencing, with 40-100x coverage depth, to identify rare and disruptive variants in candidate host factors and characterize the genetic and epigenetic variations in host cell modifiers of influenza A virus infection. The NU fluomics team has funding that will allow them to complete whole exome sequencing on 1,290individuals at a conservative level of multiplexing. |
| **Planned Statistical Analyses** | Genetic Association Analyses for common variants and haplotypes  Rare variants analysis (like burden analysis, or kernel association analysis)  Logistic regression for binary traits  Principal component analysis to assess population stratification  Bayesian admixture analysis to assess individual ancestral heterogeneity |
| **Ethical considerations** | None |
| **Target Journal** | PLoS Genetics or Cell Host and Microbe |
| **Milestones\*\*** | * Completed milestones 8/2014-9/2022:DNA samples sent to NU, w/ **demographic & treatment setting** phenotype data, by Oct. 1, 2014 * Remaining phenotype data sent to NU by Dec. 1, 2014 * Extreme phenotype sequencing done by June, 2015 * Remainder of sequencing done by Fall 2015 * Genotyped data for phenotypic extremes returned to sites by Fall 2015 * Genotype-phenotype analysis for phenotypic extremes to be completed by Fall 2015   Genotype-phenotype analysis for other subjects to be completed by Winter 2015-2016Proposed timeline for remainder of project:   * Submission of data to dbGaP by the end of 2022 * Genotyped data for other subjects returned to sites (from dbGaP) by spring of 2023 * 1st draft of manuscript by August, 2022 * 2nd [final] draft of manuscript by Oct 2022 * Submit by Nov 2022 |

**\*\*** This section should include: Timeline for completion of project, including approval, project duration, first and second draft of the paper and submission.