**eMERGE Network Proposal for Analysis**

Project/Manuscript Concept Sheet

|  |  |
| --- | --- |
| **Reference Number** | NT262 |
| **Submission Date** | November 16, 2017 |
| **Project Title** | Comparison of clinical interpretation of genetic variation in the eMERGEseq cohort using multiple methods. |
| **Tentative Lead Investigator (first author)** | David Crosslin |
| **Tentative Senior Author (last author)** | Gail Jarvik |
| **All other authors** | Ian Stanaway, Taryn Hall, David Carrell, Eric Larson, Kathy Leppig, James Ralston. |
| **Sites Involved** | Any sites interested in participating. |
| **Background / Significance** | The Electronic Medical Records and Genomics Network (eMERGE) has developed eMERGEseq, a CLIA-grade sequencing panel of 106 genes comprising the ACMG 56 and site-requested genes focused on clinical actionability and approximately 1500 SNPs. The goal is to eventually return these results through the electronic health record (EHR) for clinical decision support for 25,000 participants over a three-year period.  As the focus of clinical sequencing research moves from targeted panels to exomes and ultimately genomes, it will be advantageous to assess tools that could assist in the clinical interpretation on a larger scale. We propose assessing the correlation of clinical interpretation of genetic variation found in the current eMERGEseq cohort using three sources, with one considered the “gold standard”:   1. Both Baylor and Broad (the CSGs) are producing CLIA-grade clinical reports for return of results activities in eMERGE, and the clinical interpretation in these reports will be considered the “gold standard” for this proposal. 2. Using the ensemble called SNVs from the eMERGEseq mulisample produced by the CC, we propose to annotate using the *in silico* clinical interpretation tool Intervar1, 2. InterVar provides automatic clinical interpretation of genetic variants by the ACMG/AMP 2015 guideline1, 2. 3. Finally, we will take the same variation both provided by the CSGs and annotated by InterVar, and crowdsource participating eMERGE investigators for clinical interpretation using molecular annotation and phenotype indication via RedCap. Multiple presenters at the eMERGE IV planning meeting referred to the need for crowdsourcing clinical interpretation to scale up to larger genome targets.   1. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet* 2017; **100**(2)**:** 267-280.  2. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J*, et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in medicine : official journal of the American College of Medical Genetics* 2015; **17**(5)**:** 405-424. |
| **Outline of Project** | * Run PCA on eMERGEseq cohort, establish relationship genetically determined ancestry to self/observed reported ancestry. * Ensemble-calling of SNVs using the eMERGEseq data, producing a multisample. * Annotate multisample via ANNOVAR * Annotate ANNOVAR multisample with InterVar * Crowdsource clinical interpretation with eMERGE investigators using RedCap survey features (simple email). * Assess correlation of interpretation across three groups. * Write manuscript |
| **Desired**  **Variables (essential for analysis**  **indicated by \*)** | * De-identified positive clinical interpretation of the genetic variation from the CSGs * Phenotype indication for CSG intake * Basic demographics (self-report sex, self-report race, age, site) |
| **Desired data** | * eMERGEseq BAMs and VCFs * De-identified positive clinical interpretation of the genetic variation from the CSGs * Phenotype indication for CSG intake * Basic demographics |
| **Planned Statistical Analyses** | * Basic Table 1 demographics. * Assess correlation of interpretation across three groups. |
| **Ethical considerations** | There are no physical risks involved. |
| **Target Journal** | Bioinformatics, AJHG, etc. |
| **Milestones\*\*** |  |

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