|  |  |
| --- | --- |
| **External Collaborator Proposal** *for* **eMERGE Network Analysis**  Project/Manuscript Concept Sheet | |
| **Reference Number** | NT225 |
| **Submission Date** | March 20, 2017 |
| **Tentative Lead Investigator** *(first author with contact information and affiliation)* | Liuyang Wang  Duke University  liuyang.wang@duke.edu |
| **Tentative Senior Author**  *(last author)* | Dennis C. Ko  Duke University  dennis.ko@duke.edu |
| **eMERGE Site Sponsor & Contact** | David R. Crosslin, Joshua C. Denny, Gail P. Jarvik |
| **Project Title** | Consequences of human genetic variation in cellular traits on human disease |
| **All Other Authors** | Kelly J. Pittman, Kyle D. Gibbs, Jeffrey Barker, Anusha Gopalakrishnan, Raul E. Salinas, Alejandro Antonia, Luke Glover, Tom Balmat, Andrew Ingham, Mark R. Delong, Yanlu Cao, Soo-Chan Lee, Joseph Heitman, Raphael Valdivia, and |
| **Other eMERGE Sites Involved** | Any eMERGE investigator interested in the project. |
| **Background / Significance** | Our understanding of the genetic architecture underlying human diversity in complex traits has been transformed by the application of genome-wide association studies (GWAS). However, GWAS are just the first step in understanding how genetic variants contribute to disease risk by impacting genes that encode components of cellular physiology. Elucidating this chain of causality from SNP to gene to cell biology to disease can serve to not only functionally validate the genetic association with disease but also reveal potential therapeutic targets. Therefore, there is a need for approaches that can facilitate identification of the cellular pathways regulated by human genetic variants associated with disease.  The objective of this application is to systematically reveal the shared genetic architecture connecting cellular physiology and disease susceptibility. For GWAS of cellular traits, we developed and extensively validated a cellular GWAS platform called Hi-HOST (High-throughput human in vitro susceptibility testing) that uses pathogens to stimulate fundamental cellular pathways 1-4. Hi-HOST combines precise measurement of cellular phenotypes in lymphoblastoid cell lines (LCLs) from nearly a thousand people with genome-wide association using 15 million genetic variants to identify genetic differences that underlie the phenotypic variation. Building on the Hi-HOST platform, we recently completed data acquisition and genome-wide association for the Hi-HOST Phenome Project (H2P2), encompassing cellular responses of infectivity and replication, cytokine levels, and host cell death using 8 different pathogens and 79 cellular traits. Dozens of SNPs pass the commonly used genome-wide significance threshold of p=5x10-8 and are undergoing experimental validation and mechanistic studies. However, a key question is whether these same SNPs associated with cellular traits are also associated with more complex human clinical and disease traits. . The eMERGE Consortium, with EMR data and genotype data on ~80,000 people and >2,000 EMR-derived phenotypes provides an unparalleled opportunity to determine the relevance of SNPs associated with cellular traits on human health and disease. |
| **Outline of Project** | Our central hypothesis is that the pathophysiology of SNP associations with human disease can be effectively modeled using cellular traits. Identifying cross-phenotype associations with both cellular and clinical traits provides a cellular model for validation and mechanistic testing and also suggests that modulation of the cellular trait could affect the disease. Therefore, we propose to integrate eMERGE clinical GWAS data with H2P2 cellular GWAs data:   1. PheWAS of the eMERGE dataset using genome-wide significant hits from H2P2. Using the PheWAS pipeline of Joshua Denny 5, we will test each of the 17 H2P2 GWAS peaks for association with clinical traits in the eMERGE dataset. |
| **Desired Variables**  *(essential for analysis*  *indicated by* ***\*****)* | PheWAS data outlined by Denny algorithm\* |
| **Desired Data** | Site, sex, self-identified race and decade of birth (if available). |
| **Planned Statistical Analyses** | Assess association of H2P2 genome-wide significant SNPs in eMERGE PheWAS. |
| **Ethical Considerations** | There are no physical risks involved. |
| **Available Funding or Resources** | Most of the analyses has been performed at Duke University. We are simply assessing 17 SNPs in our PheWAS data set to validate previous findings. |
| **Milestones\*\*** | 1. Assess 17 H2P2 GWAS hits in eMERGE set using PheWAS codes to assesses pleiotropic immunology traits. (4/15/2017) 2. Publish result of PheWAS of H2P2 genome-wide significant hits, as part of paper introducing the H2P2 project and web database. (6/1/2017) 3. Share H2P2 tool with eMERGE Network. (8/1/2017) 4. Present H2P2 tool at the in-person eMERGE meeting in Bethesda (10/2017) |

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

1. Ko, D.C.*, et al.* Functional genetic screen of human diversity reveals that a methionine salvage enzyme regulates inflammatory cell death. *Proc Natl Acad Sci U S A* **109**, E2343-2352 (2012).

2. Ko, D.C.*, et al.* A genome-wide in vitro bacterial-infection screen reveals human variation in the host response associated with inflammatory disease. *Am J Hum Genet* **85**, 214-227 (2009).

3. Ko, D.C. & Urban, T.J. Understanding Human Variation in Infectious Disease Susceptibility through Clinical and Cellular GWAS. *PLoS Pathog* **9**, e1003424 (2013).

4. Salinas, R.E.*, et al.* A cellular genome-wide association study reveals human variation in microtubule stability and a role in inflammatory cell death. *Mol Biol Cell* **25**, 76-86 (2014).

5. Denny, J.C.*, et al.* PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* **26**, 1205-1210 (2010).

6. Wang, L.*, et al.* CPAG: software for leveraging pleiotropy in GWAS to reveal similarity between human traits links plasma fatty acids and intestinal inflammation. *Genome biology* **16**, 190 (2015).