|  |  |
| --- | --- |
| **External Collaborator Proposal** *for* **eMERGE Network Analysis**  Project/Manuscript Concept Sheet | |
| **Reference Number** | NT292 |
| **Submission Date** | 6/5/2018 |
| **Tentative Lead Investigator** *(first author with contact information and affiliation)* | First Author of paper to use Replication data and Concept Sheet contact: Veera Manikandan Rajagopal MD veera@biomed.au.dk (Department of Biomedicine, Aarhus University, Denmark)  VUMC Lead of existing approved NT208 concept sheet, associated with the present concept sheet: Reyna L. Gordon, PhD, ([reyna.gordon@vanderbilt.edu](mailto:reyna.gordon@vanderbilt.edu)) |
| **Tentative Senior Author**  *(last author)* | Ditte Demontis, Associate Professor, Department of Biomedicine, Aarhus University, Denmark |
| **eMERGE Site Sponsor & Contact** | VGER (VGER member Lea Davis, PhD)  lea.k.davis@vanderbilt.edu |
| **Project Title** | Genomewide association studies of developmental disorders of speech, language and scholastic skills |
| **All Other Authors** | (from Vanderbilt: Reyna Gordon, Jennifer Below, Lauren Petty, Lea Davis)  (from iPSYCH: Esben Agerbo, Jakob Grove, Thomas D Als, Anders Børglum, David Hugaard, Merete Nordentoft, Preben Bo Mortensen, Thomas Werge, Ole Mors) |
| **Other eMERGE Sites Involved** | This replication request is highly related to our ongoing NT208 eMERGE dataset, which includes pediatric data from CHOP, Cincinnati, Boston Children’s, and Vanderbilt. |
| **Background / Significance** | Developmental speech and language disorders are heritable deficits that interfere with children’s communication and academic skills. GWAS approaches to the genetic basis of these disorders have thus far occurred in small sample sizes and warrant further study. |
| **Outline of Project** | A PhD student at Aarhus university (Veera Manikandan) working in the Integrative Psychiatric Research consortium (iPSYCH) has conducted a GWAS on several language phenotypes from ICD codes in the iPSYCH data. Their approach has some commonalities with our **NT208 project** (a separate eMERGE project/paper focused on developmental language disorder). The iPSYCH group requests look-ups of the association results for 22 SNPs and polygenic risk score analysis in the NT 208 dataset as a replication dataset. In addition, we request access to eMERGE genetic data associated with the following codes (a subset of which are not present in NT208):  F80\_speech\_disorders  F80.1\_expressive\_disorder  F80.2\_receptive\_disorder  F80.9\_Unspecified speech/language disorder  F81.0\_Dyslexia  F80.0\_Articulation\_disorder  F81\_Scholastic\_disorders  F81.3\_mixed\_scholastic\_disorders |
| **Desired Variables**  *(essential for analysis*  *indicated by* ***\*****)* | The iPSYCH group has requested effect sizes and significance of the SNPs listed below for each of the ICD codes mentioned above (vs controls), in addition to estimates from pooled analysis for speech/language disorders (F80.xx vs controls) and scholastic skills disorders (F81.xx vs controls). |
| **Desired Data** | GWAS data from eMERGE III, from CHOP, CCHMC, BCH, and VUMC, is requested with the ICDs listed above, along with other ICDs pulled from the eMERGE phenotype set (in order to covary for psychiatric diagnoses). From the demographics only sex is needed. The replication request using the NT208 data is also expected to include the following SNPs :  rs34485877, rs76996482, rs12222027, rs7859347  rs13004636, rs2503696, rs59498466, rs76502789  rs140929816, rs149819210, rs117527688, rs117625151, rs117924479  rs147041538, rs183805351, rs202211268, rs62460842, rs117710669  rs141987367, rs60670513, rs34823998, rs66814427  The closest SNP in LD is to be used if not available in eMERGE imputed data. We do not expect to re-access phenotype data from NT208, as cases and controls have already been defined. |
| **Planned Statistical Analyses** | Genome-wide association testing is ongoing for NT208 project (which has a different phenotyping approach and goes beyond the scope of this replication request); we propose to retrieve summary statistics from preliminary analysis of our eMERGE dataset (our NT208 project takes several further steps which are currently underway and not needed for this replication request). We will also work together across sites to conduct polygenic risk score analyses for the replications. Covariates include sex, site, psychiatric diagnosis, and genotyping batch (if applicable). |
| **Ethical Considerations** | There are no additional risks to participants. |
| **Available Funding or Resources** |  |
| **Milestones\*\*** | Spring 2018: iPsych group has completed their own analysis  May 2018: replication request submitted to eMERGE CC  June 2018: Vanderbilt collaborators will perform look-ups from a subset of our ongoing GWAS for NT208, and will access and analyze additional data  July/August 2018: lead author will draft paper, solicit feedback, and submit for publication. |

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