Data Release and Sharing

The primary goal of the Centers for Mendelian Genomics (CMG) Program is to discover "causal" genes and variants for human Mendelian phenotypes. Inclusion of causal gene and variant data is conventional in published reports of Mendelian gene discoveries. The data produced by this program have additional utility to the biomedical research community, including, but not limited to, identifying other disease genes, testing genotype-phenotype relationships, and exploring genetic and molecular mechanisms of disease. Therefore, it is a program priority to make public the causal genes and variants found to underlie each Mendelian phenotype and their associated phenotypes as well as well-supported candidate causal genes for conditions for which no causal gene has been definitively identified.

The CMGs are required to perform the following data release and sharing activities. The mechanisms of data release and sharing may be modified in the future in order to best serve the community.

Data Release (institutional certification not necessary)

1. Names of Mendelian conditions studied by the CMG Program

The names or brief descriptions, and, if they exist, OMIM numbers, of all Mendelian conditions that have been/are being sequenced and analyzed by the CMGs will be posted on the CMG website (www.mendelian.org). This cumulative list will be updated every 6 months.

2. Candidate genes for each Mendelian condition studied by the CMG Program

A. Each CMG, or the project investigator, will submit candidate genes to a MatchMaker Exchange (MME) node at least once each quarter. Each CMG is encouraged to submit candidate genes as soon as they are identified, but they must be submitted to a MME node at most seven months after data are released to the submitting investigator. At a minimum, all tier 2/"suggestive criteria" genes will be submitted, but tier 1/"conservative criteria" can also be submitted.

B. The names of all tier 1 genes and corresponding condition names or very brief phenotype description (e.g. HPO terms) will be posted on the public CMG website thirteen months after the data are released to the submitting investigator.

3. *Causal variants for each Mendelian phenotype for which a causal gene is found*

Each CMG will submit each published causal variant to ClinVar within three months of publication, while pre-publication submission is strongly encouraged. In addition, candidate causal variants can be submitted to ClinVar at any time.

Data Sharing (Institutional certification required)

1. Data sharing via dbGaP

For appropriately consented Mendelian conditions, each CMG will submit variant call files (VCFs) and BAMs to the NCBI database of Genotypes and Phenotypes (dbGaP). Data will be submitted to dbGaP quarterly with the intent to be made public no earlier than seven months after data are released to the submitting investigator. Phenotypic data including the name of the disorder or very brief phenotype description, sex, ancestry, etc. are linked to the sequence data and made available to qualified investigators using the NHGRI DAC system.