Background: The GREGoR Consortium (Genomics Research to Elucidate the Genetics of Rare diseases) seeks to develop and apply approaches to discover the cause of currently unexplained Mendelian genetic disorders.

GREGoR Consortium Data is available on NHGRI’s AnVIL platform. Researchers can access GREGoR data via dbGaP study phs003047. The GREGoR Consortium data model available at https://github.com/UW-GAC/gregor_data_models.

Consortium Data Summary

<table>
<thead>
<tr>
<th></th>
<th>Participants</th>
<th>WES/WGS +phenotype</th>
<th>+WES/WGS +RNAseq +phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probands</td>
<td>951</td>
<td>854</td>
<td>75</td>
</tr>
<tr>
<td>Other Affected</td>
<td>182</td>
<td>157</td>
<td>7</td>
</tr>
<tr>
<td>Unaffected</td>
<td>1239</td>
<td>86</td>
<td>59</td>
</tr>
<tr>
<td>Possibly Affected or Unknown Status</td>
<td>140</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2512</td>
<td>1107</td>
<td>141</td>
</tr>
</tbody>
</table>

Most Frequent GREGoR Phenotypes (HPO Term Counts)

Using AnVIL workspaces for GREGoR Data Sharing and Analysis

Supported Data Types

Current
- Family structure
- Phenotypes (HPO)
- Short Read
- ES/WGS
- Short Read RNaseq
- Candidate Variants
- Long Read
- WGS (ONT)
- Long Read WGS (Pa Bio)
- ATAC-Seq
- Joint Caliset

Coming Soon
- Optical Mapping
- Metabolomics
- Functional Data
- + more!

Who should stop here?

- Rare disease curators/analysts/tool developers
- Clinicians who want to query patients’ variants in a rare disease cohort
- Biomedical informaticists interested in data processing and modeling
- Anyone interested in accessing a rare disease dataset

Family sizes in the GREGoR Dataset (N = 990 families)

Most Frequent GREGoR Phenotypes

- Global developmental delay
- Intellectual disability
- Microcephaly
- Hypertension
- Seizures
- Diabetes
- Hemiparesis
- Cataracts
- Leg length discrepancy
- Intellectual disability with epilepsy
- Polycystic kidney disease
- Disproportionate dwarfism
- Microphthalmia
- Delayed speech and language development
- Abnormality of the face
- Abnormality of movement
- Congenital heart defects
- Attention deficit hyperactivity disorder
- Abnormal abdomen morphologies
- Heterotaxy
- Short stature
- Central apraxia
- Delayed gross motor development
- Neurodevelopmental delay
- Autism spectrum

Data Sharing in the GREGoR Consortium to Support Rare Genetic Disease Research


The GREGoR Consortium is funded by the National Human Genome Research Institute of the National Institutes of Health, through the following grants: U01HG011758, U01HG011755, U01HG011745, U01HG011762, U01HG011744, and U24HG011746. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.