International Data Sharing Efforts: Lessons from Canadian Open Genetics Repository (COGR)

Marina Wang [1], Shana White [2], Samantha Baxter [2], Michael Oates [2], Chet Graham [2], Matthew S Lebo [2,3], Jordan Lerner-Ellis[1], and the COGR Working Groups

[1] Mt Sinai Hospital, University of Toronto, Ontario Institute for Cancer Research ; [2] Laboratory for Molecular Medicine, PPM; [3]Brigham and Woman's Hospital, Harvard Medical School

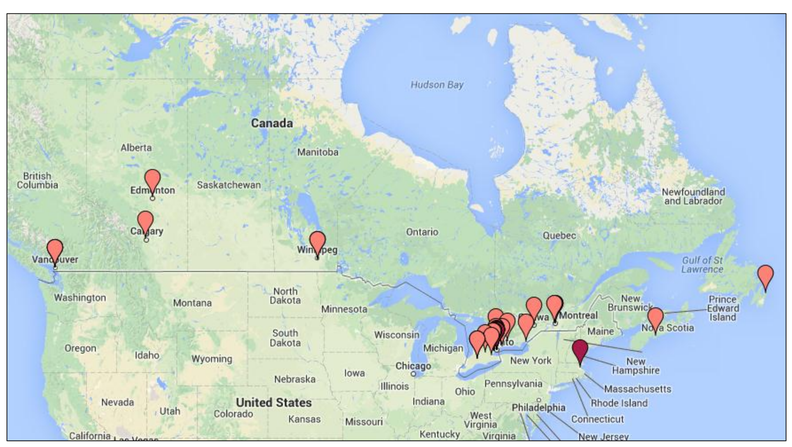
Background

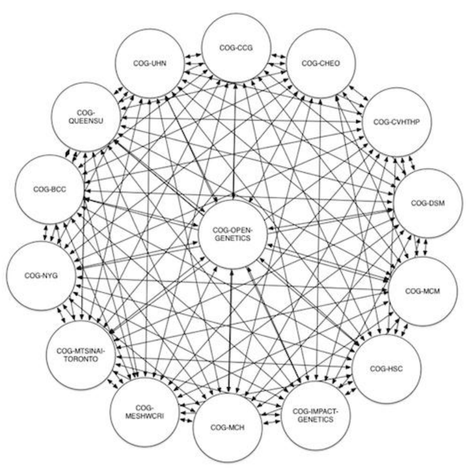
Individual laboratory knowledge and soiled data sets impede our knowledge of variants and prevents clinicians from receiving the most accurate interpretations for variants found in their patients. This in turn may prevent patients from receiving the most appropriate care.

The Canadian Open Genetics Repository is a collaborative effort to create a unified, open-access, clinical-grade database of variants reported by medical diagnostics laboratories across Canada. Utilizing GeneInsight® as a common platform, labs will collect, store and share variant information in real time.

As clinical laboratories adopt modern genomics technologies, the need for this type of collaborative framework is increasingly important.

Participating Centers and Data Sharing Model





Participating COGR Centers with a GeneInsight Instance

- Mt Sinai Hospital
- Impact Genetics
- Memorial University
- Sick Kids Hospital
- Western University
- McGill University Health
- United Health Network
- University of Manitoba
- Alberta Children's Hospital
- British Columbia Cancer Agency
- Children's Hospital of Eastern Ontario
- British Columbia Children's Hospital
- North York General Hospital
- Hamilton Health Sciences-McMaster University
- Credit Valley Hospital-Trillium Health Partners
- Canada's Michael Smith Genome Sciences Centre
- Edmonton Molecular Diagnostics-University of Alberta
- Research Molecular Genetics Lab-Women's College Hospital
- Kingston General Hospital- Queen's University

Experiences

Our progress thus far has highlighted the critical need for robust and sustainable IT infrastructure. Our goal is that continued use of a common variant assessment tool and real-time variant sharing will increase standardization and discussion.

Survey results from COGR laboratories

Use formal tracking system for variants in the literature	Yes: 46% **No: 54%**
Lab uses consistent set of terms for classification	**Yes: 65%** No: 35%
Lab has written rules for evidence-based classification of variants	Yes: 40% **No: 60%**
Variant data are linked to disease type	**Yes: 61%** No: 39%
Maintains database tracking families associated with particular variant	**Yes: 55%** No: 45%
Reassess variants every time seen in new patient	**Yes: 65%** No: 35%

Aims & Platform

Aim 1. Design freely available and consistent variant assessment procedures.

- Individual variant data will be transformed into a unified format, eliminating discrepancies, omissions, and duplications.

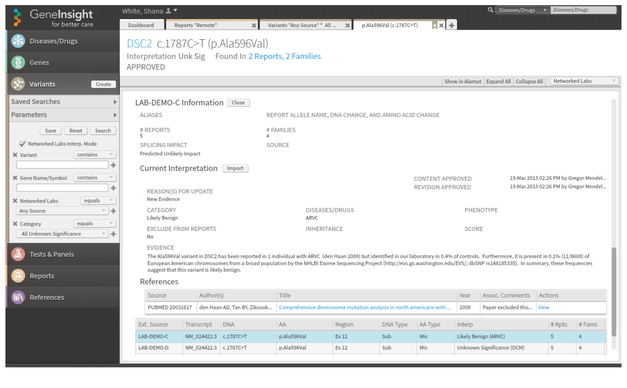
Aim 2. Data extraction and transfer.

- The project team will devise and optimize procedures to support the extraction of disease, gene and variant data currently held within participating laboratories.

Aim 3. Data access and dissemination.

- Methods will be developed such that accurate and readily accessible data will be presented to different stakeholder groups in appropriate summary levels. Individual laboratories see their variants with specific detail while the diagnostic community views a more general summary that is mindful of privacy and confidentiality.

COGR will provide an instance of the GeneInsight platform to all participating labs.



Current Variant Sharing

Participating laboratories are currently sharing:

- 7 Diseases: HBOC, HCM, Lynch syndrome, FAP, MUTYH-Associated Polyposis (MAP), Alpha-1 Antitrypsin Deficiency (A1AD), Somatic Cancer
- 52 Genes
- >3,000 variants

P- Pathogenic

LP- Likely Pathogenic

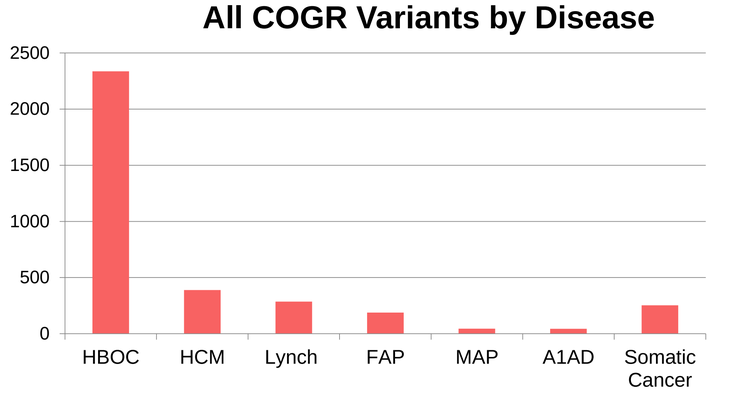
VUS- Variant of Unknown Significance

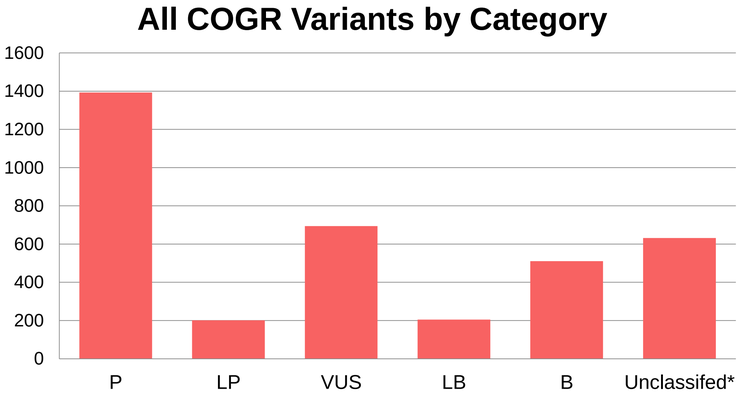
LB- Likely Benign

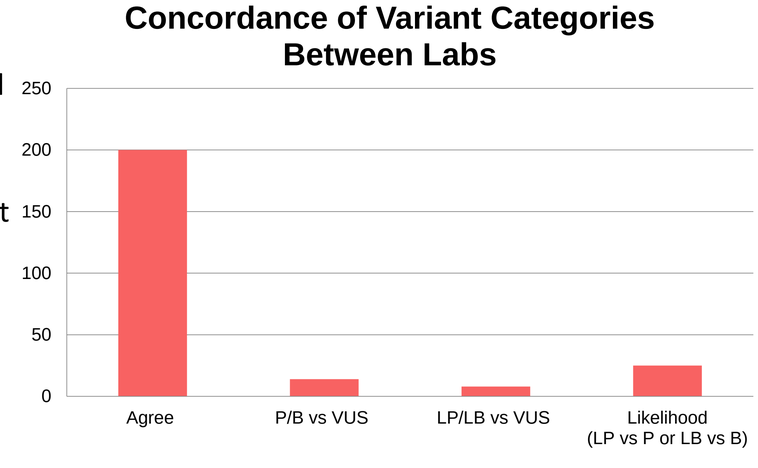
B- Benign

*Unclassified variants are under review, pending classification.

- Over 3,000 variants are being shared
  - 391 have been reported by 2 different labs
  - 50 have been seen by 3 different labs
- The majority of variants categorized by multiple labs have concordant categories.







Future Directions

The COGR project is positioned to help remove some of the chief obstacles to advancements in personalized healthcare including the lack of standardized resources and protocols for interpreting the ever-increasing volumes of patient data being generated by clinical labs.

Continuing initiatives include:

- Establishing committees of disease area experts to create consensus interpretations for variants categorized by multiple laboratories.
- Make consensus variant interpretations available to stakeholder groups with different but appropriate levels of summary.
- Continue to collaborate with other international data-sharing efforts including ClinVar.

In summary, the COGR serves as a focal point for the collaboration of Canadian laboratories with themselves and other countries in the development of tools and methods that take full advantage of laboratory data in diagnosing, managing and treating genetic diseases. Furthermore, as more laboratories share data, knowledge will improve and ultimately lead to better patient care.

Canadian Open Genetics Repository
Email: cogr@opengenetics.ca
Web: **http://opengenetics.ca**