Division of Laboratory Systems

Two New Search Tools to Find Genomic DNA Reference Materials

Scheinfeldt L[1], Kusic D[1], Kalman LV[2]
[1]Coriell Institute for Medical Research, Camden NJ, [2]Division of Laboratory Systems, Centers for Disease Control and Prevention, Atlanta GA

GeT-RM
Genetic Testing Reference Materials
PROGRAM

Introduction

The Coriell Institute for Medical Research manages several National Institutes of Health (NIH) sponsored biorepositories, including the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository and the National Human Genome Research Institute (NHGRI) Sample Repository for Human Genetic Research.

- The NIGMS Repository distributes DNA samples from patients diagnosed with heritable diseases comprising over 1,000 Online Mendelian Inheritance in Man (OMIM) disorders.
- The NHGRI Repository distributes DNA samples contributed by thousands of people living around the world to facilitate studies of genetic and genomic variation.

These collections maintain and distribute immortalized cell lines that provide renewable sources of derived DNA samples that can be used as reference materials for clinical genetic testing. We developed two web-based search tools for the identification of DNA samples containing specific variants.

The Genetic Testing Reference Material Program (GeT-RM) is a collaborative CDC-based program that helps to improve the availability of genomic DNA reference materials for genetic testing.

Methods

GeT-RM PGx Search: created using integrated pharmacogenetic (PGx) and human leukocyte antigen (HLA) consensus annotations for 34 genes/loci characterized during nine Centers for Disease Control and Prevention's Genetic Testing Reference Material Program (GeT-RM) studies.
- Each sample (N=363) was characterized with a variety of methodologies and test platforms in two or more laboratories.
- Results were assessed for quality, discordances, and determination of consensus genotype for each sample.

SNP Search: created from a dataset of 30x coverage whole genome sequencing (WGS) data collected through the 1000 Genomes Project (N=3,202 sequenced genomes).

Results

We designed and implemented two user friendly, web-based genetic variation search tools to facilitate the identification of DNA samples with genetic variation profiles needed to serve as positive controls for genetic testing.
- **GeT-RM PGx Search** allows users to find DNA samples with specific PGx and HLA annotations.
- **SNP Search** allows users to search up to 100 variants at a time and find DNA samples with user specified genotype profiles.

GeT-RM PGx Search

An Excel file containing consensus genotypes covering 363 DNA samples characterized during nine GeT-RM PGx or *HLA* studies for 34 genes/loci (**Table 1**) is available for download on the GeT-RM website. A searchable web-based database tool, GeT-RM PGx Search, is available on the Coriell Genomic Data Search tool website.

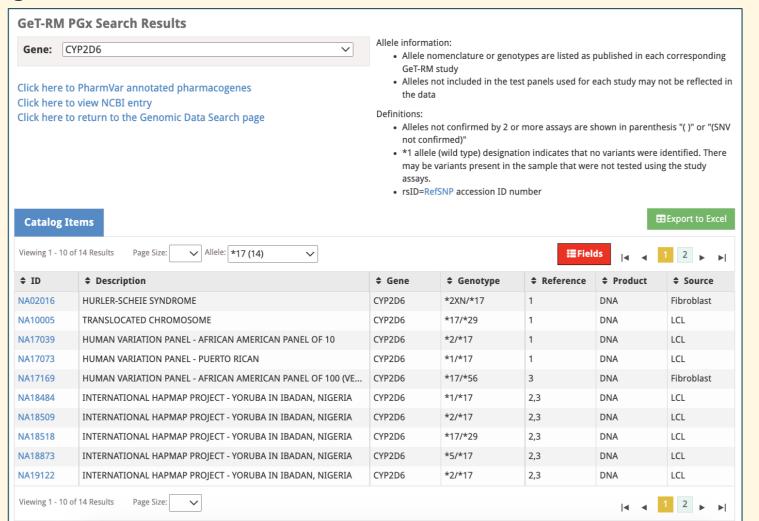**Figure 1. Screenshot of searchable web-based database, GeT-RM PGx Search.**

Figure 1. This example displays results for *CYP2D6*, chosen from the dropdown filter on the top of the table. Results for each gene can be filtered by allele. The reference for each genotype is provided in the adjacent column. Links to PharmVar and the NCBI entry for the selected gene are provided. Results can be exported to a CSV file by clicking the green button.

Table 1. Genes/loci in the GeT-RM PGx Search database.

CYP1A1	CYP2D6	GSTP1	SLCO1B1	HLA-A	HLA-DQA1
CYP1A2	CYP2E1	GSTT1	SLCO2B1	HLA-B	HLA-DQB1
CYP2A6	CYP3A4	NAT1	TPMT	HLA-C	HLA-DPA1
CYP2B6	CYP3A5	NAT2	UGT1A1	HLA-DRB1	HLA-DPB1
CYP2C8	CYP4F2	NUDT15	UGT2B7	HLA-DRB3	
CYP2C9	DPYD	SLC15A2	UGT2B15	HLA-DRB4	
CYP2C19	GSTM1	SLC22A2	UGT2B17	HLA-DRB5	

- *CYP2C* Cluster NC_000010.10: g.96405502G>A, rs12777823
- *GGCX* NM_000821.6:c.214+597G>A, rs12714145
- *GGCX* NM_000821.6:c.2084+45G>C rs11676382
- *VKORC1* NM_024006.5:c.-1639G>A, rs9923231
- *VKORC1* NM_024006.6:c.106G>A, rs61742245
- *VKORC1* NM_024006.6:c.196G>A, rs72547529

SNP Search

A searchable web-based tool of WGS data, SNP Search, is available on the Coriell Genomic Data Search tool website. Users can search up to 100 genome-wide SNPs in a single query.

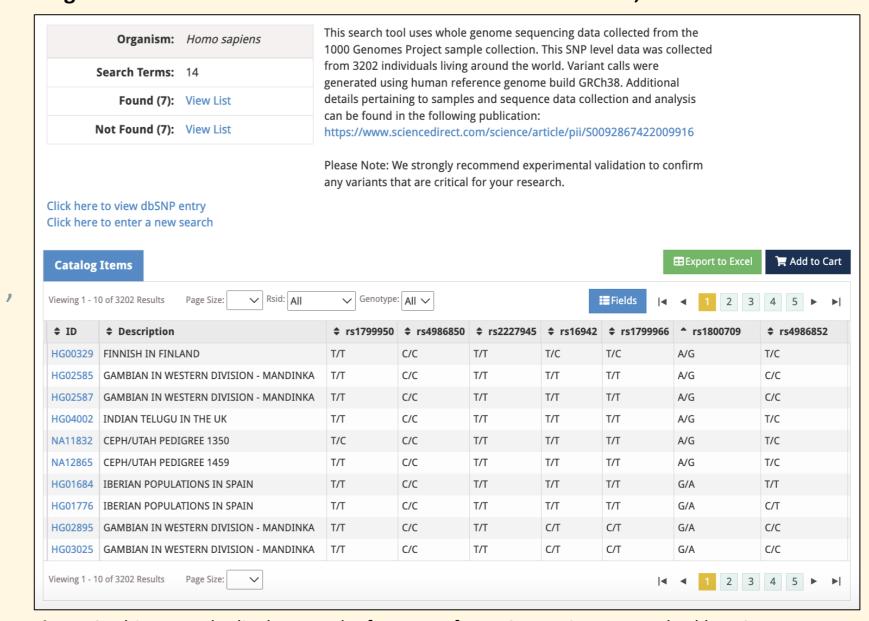**Figure 2. Screenshot of searchable web-based WGS dataset, SNP Search.**

Figure 2. This example displays results for a set of 7 *BRCA1* variants, searched by RSID. Results for each SNP can be filtered by allele. Links to NCBI entry for the searched SNP(s) are provided. Results can be exported to a CSV file by clicking the green button.

Conclusions

- These two new web-based search tools provide easily accessible ways to find information about publicly available DNA samples that can be used to support quality assurance programs of laboratories performing clinical genetic testing.
- Both tools are located on the Coriell Institute for Medical Research website.
- All reference materials developed by GeT-RM are publicly available from the NIGMS and NHGRI Repositories at the Coriell Institute for Medical Research.

GeT-RM PGx Search

https://www.coriell.org/GeTRM/PGxSearch

SNP Search

https://www.coriell.org/SNPSearch/WGS

GeT-RM Website

https://www.cdc.gov/lab-quality/php/get-rm/index.html

Available on the GeT-RM Website:
- **Info about RMs characterized by GeT-RM**
- **Links to GeT-RM publications**

CONTACT INFO

Lisa Kalman PhD
LKalman@cdc.gov