G-DOC *Plus*
Precision medicine use case

Innovation Center for Biomedical Informatics
Georgetown University

Background

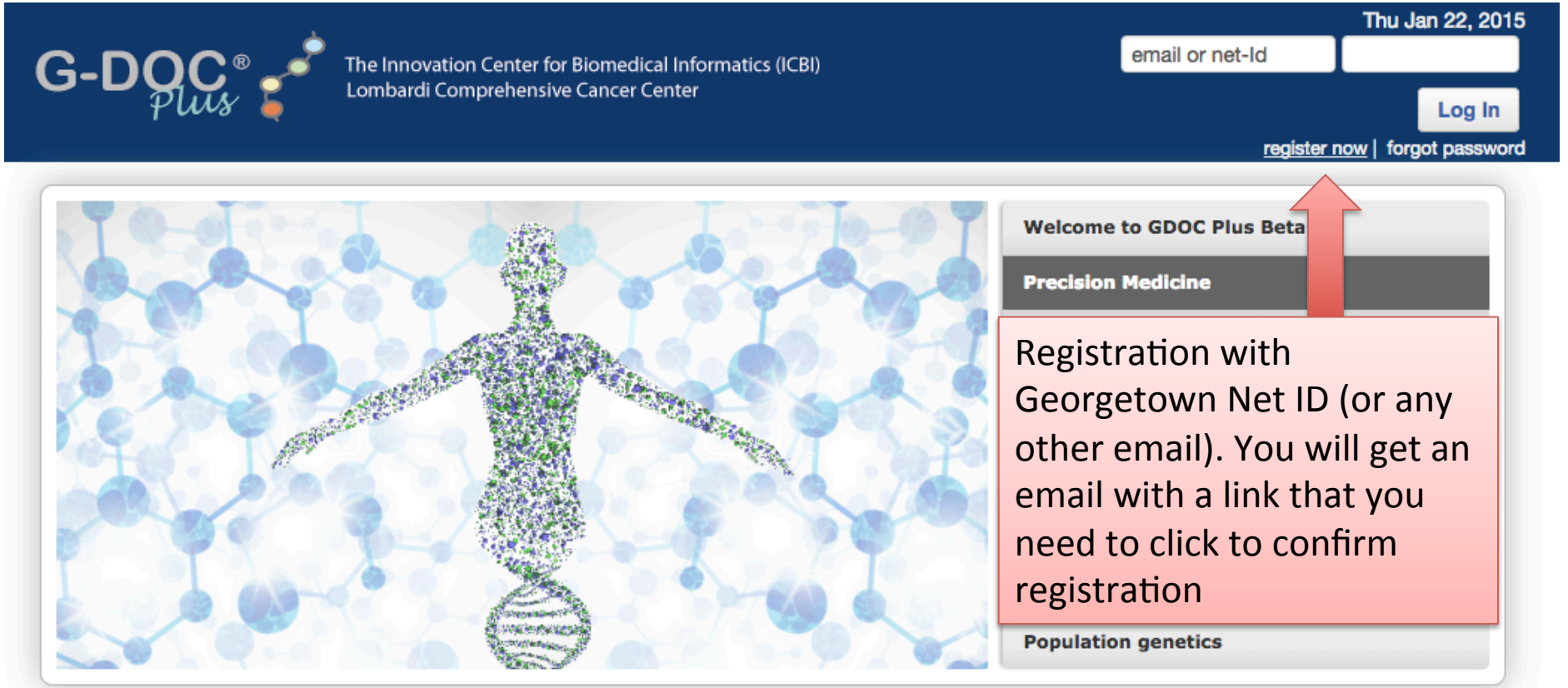
- Explore somatic and germline variants in cancer samples
- Inactivation of tumor suppressor genes could be caused by accumulation of mutations or, loss of heterozygosity (LOH), causing progression of cancer
  - one such location is Chromosome 8,
  - abnormalities in Chr 8 reported in BRC, CRC

- Example query: ***which genes are affected by novel deletions in chromosome 8 with potential impact on protein function, and what major pathways might these genes be involved in?***

Overview

- Register
- Login
- Navigation
- Precision medicine workflow
  - Variant search

First time user



Understanding Data in G-DOC Plus

It all begins with a study...

All data in G-DOC Plus derives from studies on topics such as breast cancer, wound healing, or even 1,000 Genomes. Each study may contain clinical and/or biospecimen data. Below is an overview of studies by topic.
*

* private studies, ones which are uploaded and marked private, are not counted here

News

October 02, 2014: ICBI Symposium 2014 [read]

May 02, 2014: Featured in Frontiers' Top 10 2013 Most viewed Genetics Research articles [read]

March 12, 2014: AAAS Big Data Blog [read]

Login

G-DOC® Plus

The Innovation Center for Biomedical Informatics (ICBI)
Lombardi Comprehensive Cancer Center

Thu Jan 22, 2015

kb472 ••••••••••••••

Log In

register now | forgot password

Welcome to GDOC Plus Beta!

The Georgetown Database of Cancer Plus other diseases (G-DOC Plus) is a precision medicine platform containing molecular and clinical data from thousands of patients and cell lines, along with tools for analysis and data visualization. The platform enables the integrative analysis of multiple data types to understand disease

Precision Medicine

Translational research

Population genetics

Understanding Data in G-DOC Plus

It all begins with a study...

All data in G-DOC Plus derives from studies on topics such as breast cancer, wound healing, or even 1,000 Genomes. Each study may contain clinical and/or biospecimen data. Below is an overview of studies by topic.

* private studies, ones which are uploaded and marked private, are not counted here

News

October 02, 2014: ICBI Symposium 2014 [read]

May 02, 2014: Featured in Frontiers' Top 10 2013 Most viewed Genetics Research articles [read]

March 12, 2014: AAAS Big Data Blog [read]

G-DOC Plus | Home | Studies | Lists | Analyses | Groups | Notifications | Study Options | Help | kb472

G-DOC Plus Launch Pad!

Welcome! The G-DOC Plus Launch Pad is your one-stop resource for learning more about G-DOC and getting started on the platform.

Studies

Lists

Groups

Notifications 0

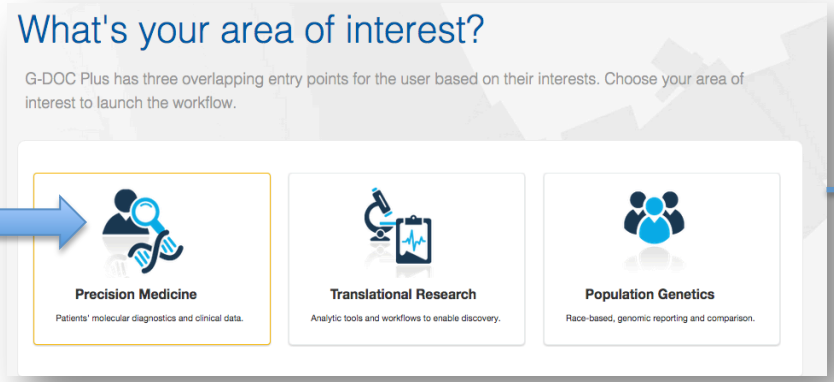
It All Starts Here!

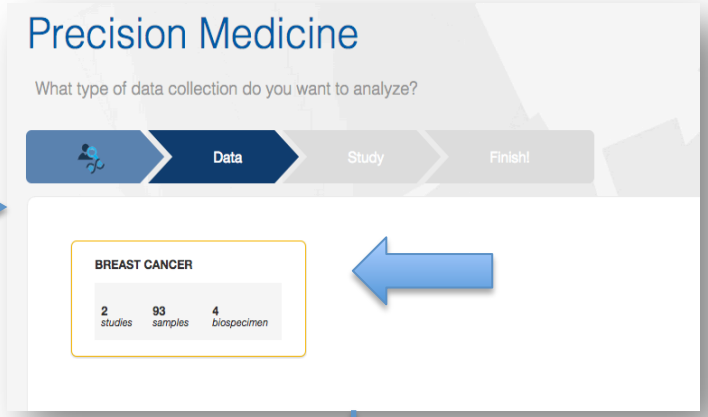
G-DOC has over seventy studies, We know this can be overwhelming! Let us guide you to choose the study that is relevant for your research.

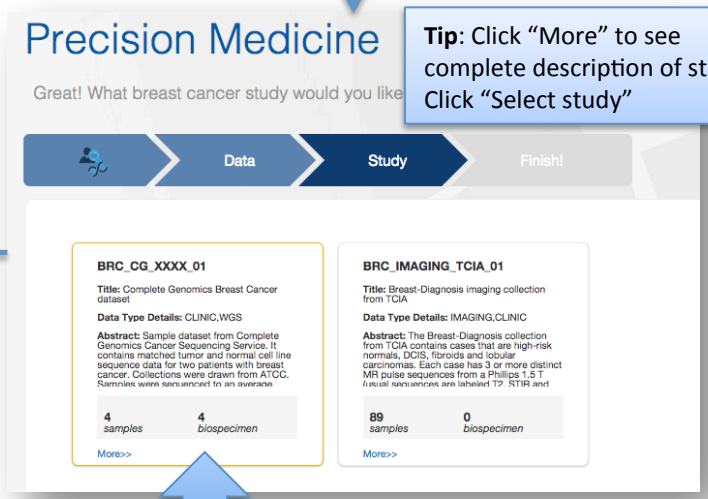
Let's Go! ›

Selections

- "What's your area of interest" : Precision medicine - it should go to the workflow page.
- Select "data collection" : BREAST CANCER
- Select a study : BRC_CG_XXXX_01 (Note: If you click on "More", you will be able to see the complete description of the study)
- Select the tool: Variant Search





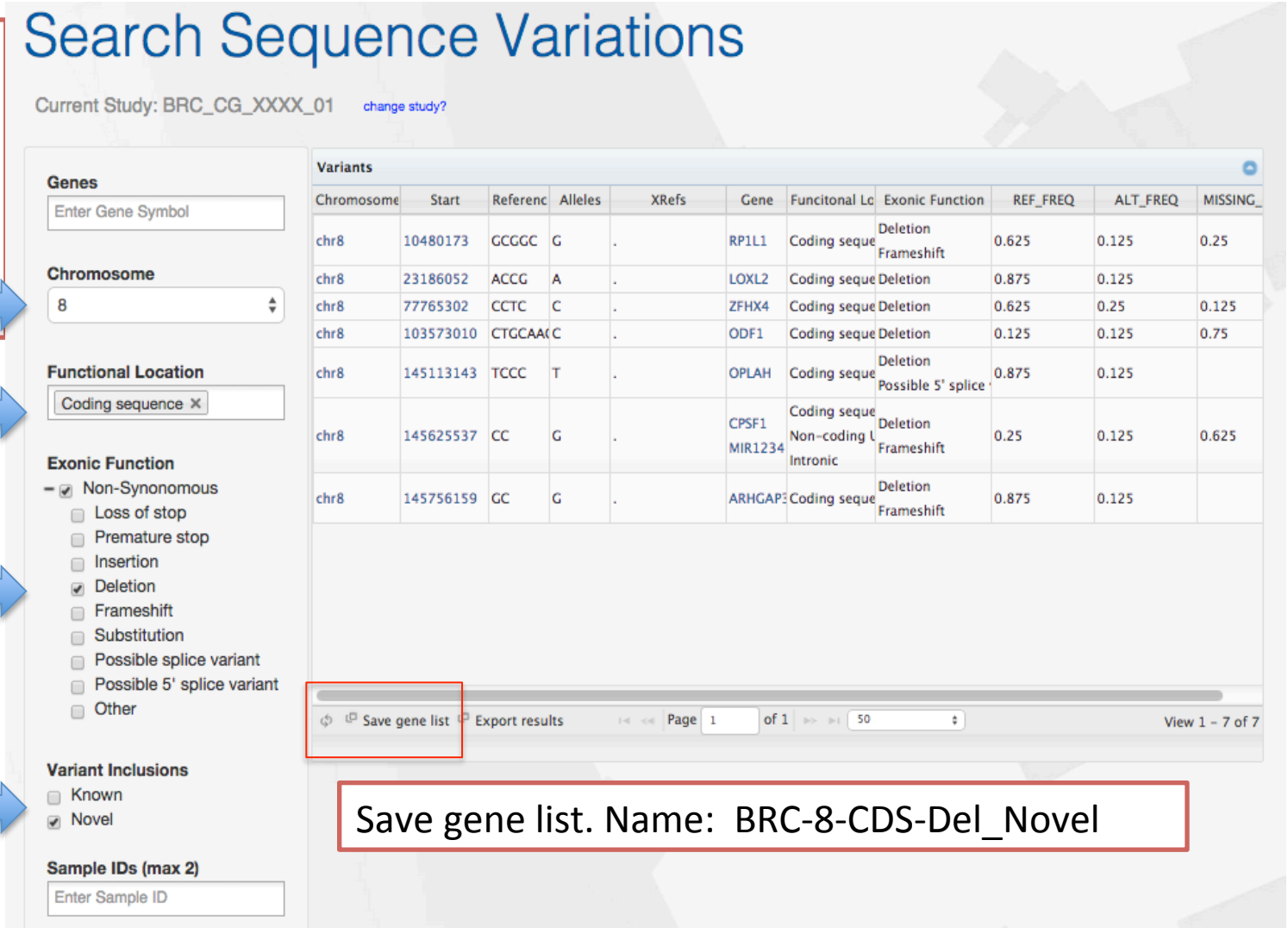


Tip: Click "More" to see complete description of study. Click "Select study"

- Make the following selections:
  - **Chromosome: c***hromosome 8*, **Functional Location:** *Coding sequence*, **Exonic Function:** *Non Synonymous -> Deletion,* ***Variant inclusion*** *: Novel*

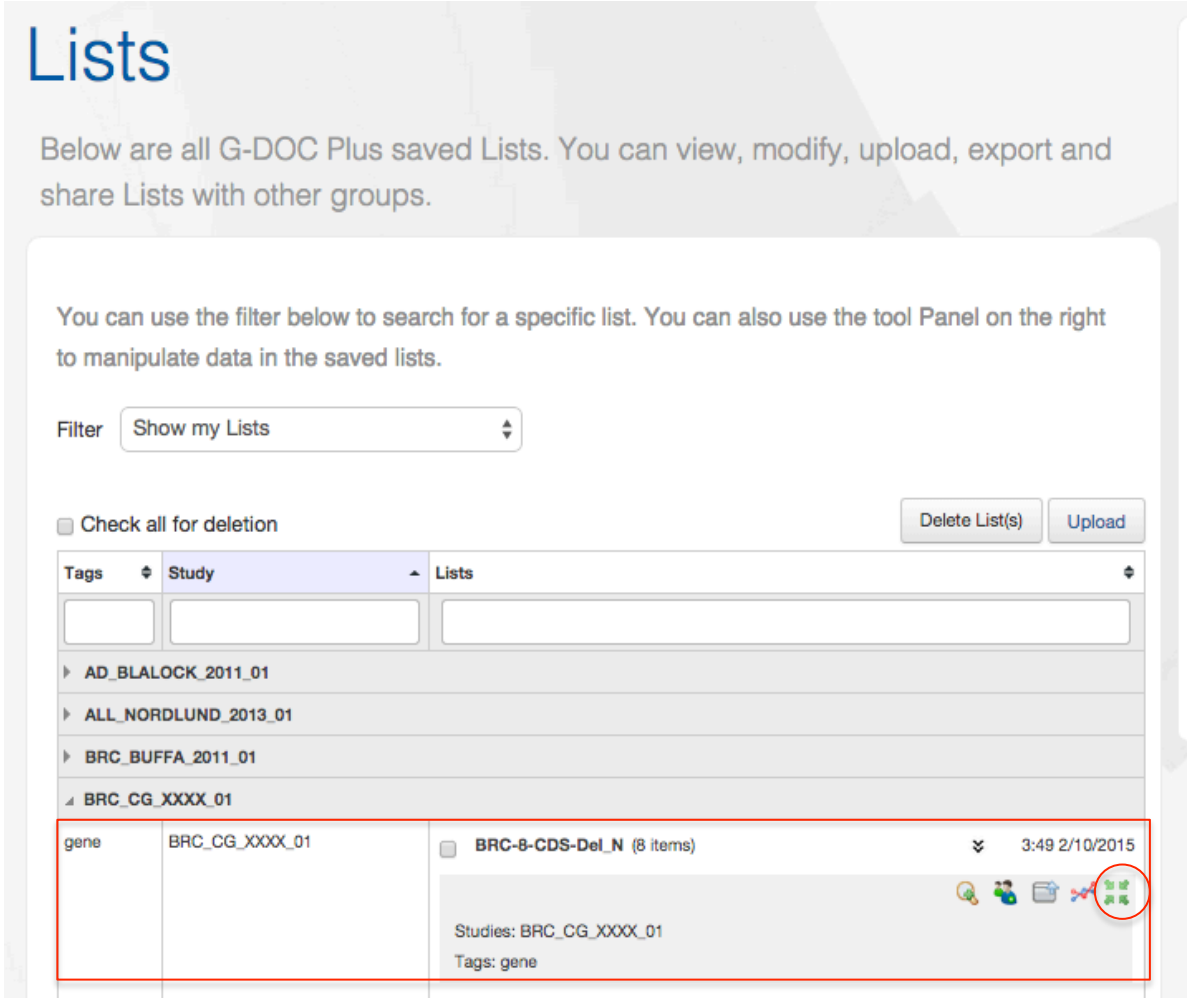
Data set: WGS from breast cancer cell line

Data stored in Amazon cloud



Save gene list. Name: BRC-8-CDS-Del_Novel

- Once the gene list is saved, click on **Lists** on the Top panel. You will see the gene list grouped under the study name (BRC_CG_XXXX_01)
- On the right side there will be several icons which indicate further downstream analysis. Click on the green icon which indicates **Enrich Gene List** analysis



Pathway enrichment using Reactome

Pathway Enrichment Results

Pathway Enrichment Results

Pathway	p-value	Overlapping Genes
Inactivation of Cdc42 and Rac	5.483×10^{-3}	ARHGAP39
Glutathione synthesis and recycling	6.699×10^{-3}	OPLAH
Processing of Intronless Pre-mRNAs	8.521×10^{-3}	CPSF1
Post-Elongation Processing of Intronless pre-mRNA	1.397×10^{-2}	CPSF1
Processing of Capped Intronless Pre-mRNA	1.397×10^{-2}	CPSF1
Glutathione conjugation	1.518×10^{-2}	OPLAH
Signaling by Robo receptor	1.941×10^{-2}	ARHGAP39
Post-Elongation Processing of Intron-Containing pre-mRNA	2.061×10^{-2}	CPSF1
mRNA 3-end processing	2.061×10^{-2}	CPSF1
Transport of Mature mRNA Derived from an Intronless Transcript	2.121×10^{-2}	CPSF1
Transport of Mature mRNAs Derived from Intronless Transcripts	2.181×10^{-2}	CPSF1
Cleavage of Growing Transcript in the Termination Region	2.602×10^{-2}	CPSF1
Post-Elongation Processing of the Transcript	2.602×10^{-2}	CPSF1
RNA Polymerase II Transcription Termination	2.602×10^{-2}	CPSF1
Rho GTPase cycle	2.662×10^{-2}	ARHGAP39
Signaling by Rho GTPases	2.662×10^{-2}	ARHGAP39
Gene Expression	3.140×10^{-2}	CPSF1

Export results Page 1 of 1 50 View 1 - 27 of 27

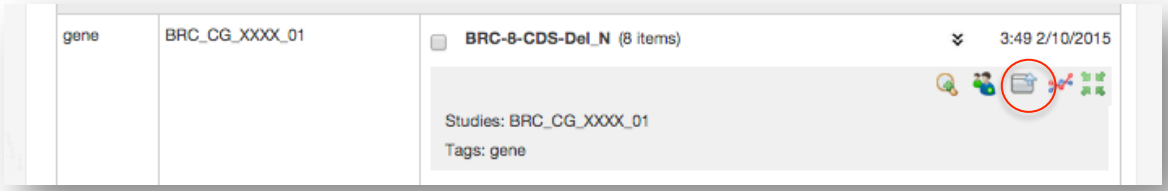
Discussion

- Pathways enriched - related to processing of pre-mRNA and mature mRNA,
  - are known to potentially alter gene expression and in-turn turn tumor suppressor genes off
- The top pathway “Inactivation of CD42 and RAC”
  - shows that when such RNA processing pathways are turned negative, it causes other tumor pathways to go out of control.

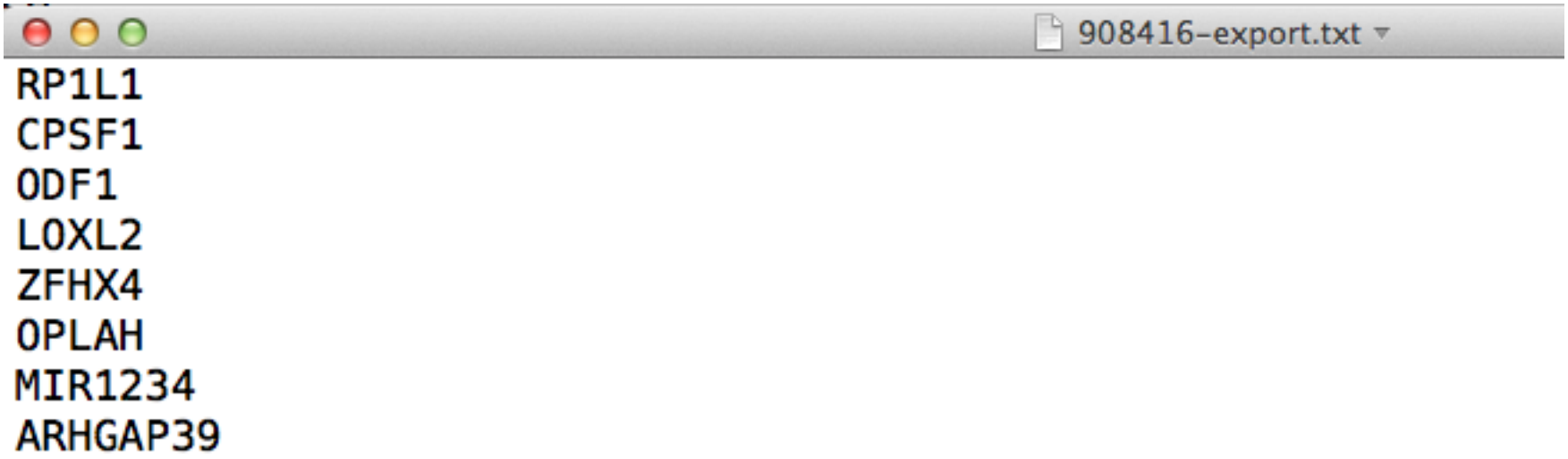
Application

- Case study shown on cell line data - represents individual tumors
  - can be applied on other patient datasets
  - allows in-depth analysis of germline mutations for individual patients based on NGS data.
- The variant search tool also allows exploration of somatic variants (variants present in tumor, but not in matched normal sample)

- To view the list of genes in the saved gene list, once again go to **Lists** on the top panel.
- Select the middle icon on the right side which indicates **Export list.** You will now see a text file downloaded to your system



- You can open the exported file to view the list of genes



Appendix

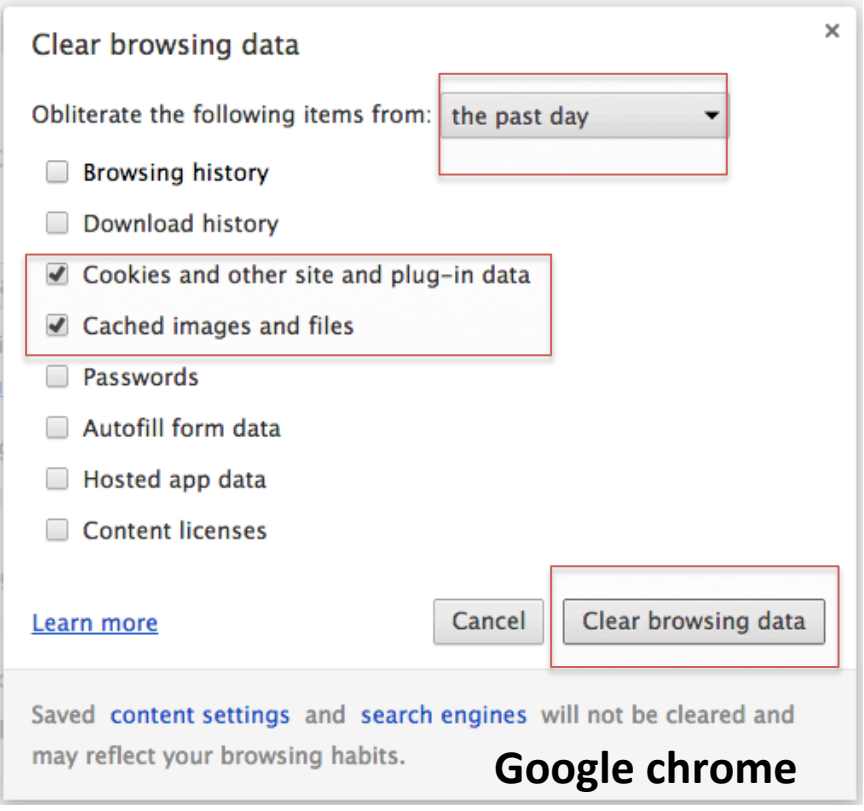
- **Functional Location**: Physical location of Variants on the chromosome.
  - Functional location filter options: coding sequence, intronic, downstream of gene, upstream of gene, 5' UTR, 3' UTR, non-coding UTR, intergenic
- **Coding gene**: Within a gene which codes for a protein. Within this region are search terms for Intron, downstream of the specified gene, upstream of the specified gene, in the 5'untranslated region, 3'untranslated region.
- **Non-coding region**: Region that does not cover coding genes, it's flanking regions and UTR.
- **Intergenic**: Excludes Coding regions and non-coding UTRs.
- **Non-Synonymous changes**
  - **Exonic Functions**: Includes regions covered only exons of a gene coding for a protein where variation has resulted in non-synonymous changes.
  - Exonic function filter options: Loss of stop (structurally impacted protein), Premature stop (truncated protein), Insertion, Deletion, Frameshift, Substitution, Possible splice variant, Possible 5' splice variant, other

General tips

- G-DOC *Plus* works best is you don't use the **back** button in the web browser repeatedly.
- Once you select a study, most tools will be easily available from the the top menu bar inside G-DOC *Plus*.
- The Pathway enrichment and the Lists tool may sometimes take a few seconds longer to execute than other tools (since they are directly connecting to the server every time). Your patience is highly appreciated.

Clearing cache

- If the G-DOC web page does not respond after several seconds, try:
  - refreshing the page.
  - Log out and log back in, and try again
  - If the above two do not work, its possible that your web browser cache may need to be cleared
    - For Google chrome, go to Settings -> Show Advanced Settings -> Under “Privacy”, select Clear Browsing data
    - For Mozilla Firefox, go to Preferences -> Advanced -> Network -> Under “Cached Web Content” -> Clear now



Google chrome

Mozilla Firefox