Cancer Genome Interpreter (CGI)

https://www.cancergenomeinterpreter.org/home
Educating for best practices in clinical cancer genomics

Stars indicate additional information in upcoming slides
OncodriveMUT

Identifies potential driver mutations based on prevalence in disease population, computational analyses, prior knowledge of gene, and mutation-centric measurements.

Educating for best practices in clinical cancer genomics
Cancer Biomarkers Database is currently curated and maintained under the European Union’s Horizon 2020 funding and part of the collaborative effort Global Alliance for Genomics and Health (GA4GH)
Cancer Genome Interpreter is designed to support the identification of tumor alterations that drive the disease and detect those that may be therapeutically actionable. CGI relies on existing knowledge collected from several resources and on computational methods to annotate the alterations in a tumor according to distinct levels of evidence.

With a list of genomic alterations and the cancer type as input, the CGI identifies validated driver alterations and annotates and classifies the remaining variants of unknown significance. Then, alterations that are biomarkers of drug response or interact with existing chemical compounds are identified according to current knowledge.
Required Mutation Input Format - Follows HGVS Protein Change Format

- Missense:
  - NM_005157:p.T315I

- Stop:
  - TSC2:p.Q1178*

- In-frame insertion:
  - ENST00000326724:p.P1331_A1332insTP

- In-frame deletion:
  - TP53:p.I254_T256delIT
  - TP53:p.I254_T256del3
  - TP53:p.I254_T256del

- Frameshift
  - APC:p.I1557fs*30
    - The longitude of the frameshift (until the new reading frame ends in a stop codon) needs to be stated to retrieve the corresponding nucleotide change, which is used to calculate certain metrics used by the CGI; if not available, APC:p.I1557fs is also allowed.
Mutation Input Format: Nucleotide Changes

HGVS = green
Genomic tabular format = orange
Either format is acceptable

- **Point mutations:**
  - chr3:g.178936091G>A
  - chr3:g.178936091G>A

- **Block substitution:**
  - chr3:g.41266066TG>AA
  - chr3:g.41266066_41266067delinsA
  - chr3:533873 CT AC

- **Insertions:**
  - chr5:g.170837546_170837547insCTG
  - chr5:g.170837545C >CTCTG
  
- **Duplications:**
  - chr3:g.30732988_30732989dupTG
  - chr3:g.30732988_30732989delinsTG
  - chr3:30732989_30732990insTG

- **Deletions:**
  - chr2:g.234183368_234183372delA
  - chr2:g.234183368_234183372del
  - chr2:g.234183368_234183372del15
  
- **Complex indels:**
  - chr10:g.52595929_52595931delG
  - chr10:g.52595929_52595931delinsTA
  - chr7:140453155 CA TCC
Mutation Input Format:
Copy Number Alterations (CNAs) and Translocations

• Amplification:
  – ERBB2:amp

• Deletion
  – TP53:del

• Translocation
  – BCR__ABL1 (equivalent to ABL1__BCR)
    • Two underscores in this instance
Mutation Input Format:
Uploaded Files

- VCF files (in hg19) or text files can be uploaded into CGI
  - See column title specifications on the website by clicking the ‘?’ next to alterations
Selecting the Appropriate Cancer Type

- Enter cancer type manually into search bar
- Search through disease ontology tree
- Click to select cancer type and mark pink.
- If disease is not on the list, that means that no specific information for that cancer type resides in CGI
- Some classification of cancer type (even if it is generic) needs to be selected in order to search
Running a Query

• Input Mutations
  – Manually or Add File

• Select Cancer Type

• Click “Run”

• Allow analysis to Run.
  – If you see an error, check your mutation input format
  – Analysis may take several minutes
Results (Normal View)

- Tab for alterations has sub-tabs
  - Mutations
  - CNAs
  - Translocations
- Can choose to display all data or only abnormalities that are oncogenic or driver mutations
Oncogenic Classifications

Oncogenic classification

Oncogenic potential of the mutation:

- **known**: the mutation is well-demonstrated to be oncogenic in the tumor type of the sample(s) or in another cancer
- **predicted driver or predicted passenger**: according to the oncodriveMUT method (tier 1 and 2 represent higher and lower level of stringency of the driver prediction, respectively)
- **polymorphism**: mutation found at a major allele frequency higher than 1% across the population
- **no protein affecting**: the mutation does not alter the protein sequence

This pop-up displays when the ‘?’ next to “Oncogenic classification” is clicked.
Advanced View of Results: Mutations

- Adds the following information to your results tables:
  - Mutations
    - GDNA
    - Transcript
    - Exon
    - Location in relation to last exon of gene (right)
    - Tumor Driver according to CGI group publication (PMID:25759023) – hover over to specify if the gene is a driver in this tumor type or other tumor types.
    - Role – mechanism of action (OG, TSG, ambiguous)
    - In Cluster – does this mutation fall within more commonly mutated regions in that gene.
**Therapeutic Information**

- **Biomarkers**

<table>
<thead>
<tr>
<th>Sample Id</th>
<th>Observed alteration</th>
<th>Biomarker</th>
<th>Drugs</th>
<th>Effect</th>
<th>Resist.</th>
<th>Tumor type</th>
<th>Evidence level</th>
<th>Reference</th>
<th>TumorM</th>
<th>BioM</th>
</tr>
</thead>
<tbody>
<tr>
<td>lcg1</td>
<td>PTEN:del</td>
<td>PTEN deletion</td>
<td>Sirolimus (MTOR inhibitor)</td>
<td>Responsive</td>
<td></td>
<td>CANCER</td>
<td>Early trials (abstr 25...)</td>
<td>ASCO 2013</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>lcg1</td>
<td>PTEN:del</td>
<td>PTEN oncog.</td>
<td>Sirolimus (MTOR inhibitor)</td>
<td>Responsive</td>
<td></td>
<td>CANCER</td>
<td>Early trials (abstr 25...)</td>
<td>ASCO 2013</td>
<td></td>
<td>DA</td>
</tr>
<tr>
<td>default_id</td>
<td>BRAF (D594E)*18</td>
<td>BRAF mutation</td>
<td>Pan-RAF inhibitors</td>
<td>Responsive</td>
<td></td>
<td>CANCER</td>
<td>Pre-clinical</td>
<td>PMID:2673205</td>
<td></td>
<td>DA</td>
</tr>
<tr>
<td>lcg1</td>
<td>PTEN:del</td>
<td>PTEN oncog.</td>
<td>PI3K pathway inhibitors</td>
<td>Responsive</td>
<td></td>
<td>CANCER</td>
<td>Pre-clinical</td>
<td>PMID:21289267</td>
<td>PMI</td>
<td>DA</td>
</tr>
<tr>
<td>lcg1</td>
<td>PTEN:del</td>
<td>PTEN oncog.</td>
<td>PI3K pathway inhibitors</td>
<td>Responsive</td>
<td></td>
<td>CANCER</td>
<td>Pre-clinical</td>
<td>PMID:21289267</td>
<td>PMI</td>
<td>DA</td>
</tr>
<tr>
<td>lcg1</td>
<td>ERBB2:amp</td>
<td>ERBB2 (D76)</td>
<td>Trastuzumab (ERBB2 mAb inhib)</td>
<td>Responsive</td>
<td></td>
<td>CANCER</td>
<td>Pre-clinical</td>
<td>PMID:22203688</td>
<td>PMI</td>
<td>DA</td>
</tr>
<tr>
<td>default_id</td>
<td>BRAF (D594E)*18</td>
<td>BRAF (MTOR inhibitor)</td>
<td>Venlafaxine (BRAF inhibitor)</td>
<td>Resistant</td>
<td></td>
<td>CANCER</td>
<td>Pre-clinical</td>
<td>PMID:2673205</td>
<td></td>
<td>DA</td>
</tr>
</tbody>
</table>

- **Effect** – Responsive, No responsive, Resistant, Increased toxicity
- **Resist.** – Indicates any additional alterations that would confer resistance to the therapy
- **Evidence level** – Different from approval status; Early trials (phase I & II), Late trials (phase III, IV), pre-clinical, Clinical
- **Reference** – indicates reference, may include links to PubMed
- **TumorM** – checked if the tumor type matches a tumor type in which the biomarker has been observed
- **BioM** – indicates match between alteration and the observed biomarker; C = complete match, DM = different mutation (different amino acid change), DA = different alteration (biomarker is not a mutation)
### Therapeutic Information

- **Bioactivities**

<table>
<thead>
<tr>
<th>Sample Id</th>
<th>Observed alteration</th>
<th>Gene symbol</th>
<th>Compound</th>
<th>Binding potency</th>
<th>MOA</th>
<th>Match</th>
<th>Type</th>
<th>Status</th>
<th>p-Activity</th>
<th>Type</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>default.id</td>
<td>BRAF (D594E*18)</td>
<td>BRAF</td>
<td>CHEMBL3354844</td>
<td>highly potent</td>
<td>inhibitor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>default.id</td>
<td>BRAF (D594E*18)</td>
<td>BRAF</td>
<td>CHEMBL1724072</td>
<td>highly potent</td>
<td>inhibitor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>default.id</td>
<td>BRAF (D594E*18)</td>
<td>BRAF</td>
<td>CHEMBL108650</td>
<td>highly potent</td>
<td>inhibitor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>default.id</td>
<td>BRAF (D594E*18)</td>
<td>BRAF</td>
<td>CHEMBL108650</td>
<td>highly potent</td>
<td>inhibitor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>default.id</td>
<td>BRAF (D594E*18)</td>
<td>BRAF</td>
<td>CHEMBL108650</td>
<td>highly potent</td>
<td>inhibitor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Compound** – ChEMBL compound ID and link to ChEMBL Compound report card
- **Binding potency** – Highly Potent = >9 (1nM); Potent = >6 (1uM); Weak = >3 (1mM)
  - See Binding p-Activity
- **MOA** – mechanism of action
- **Match** – Checked when mechanism of action coincides with role of gene in cancer
- **Type** – molecular type of compound: oligonucleotide, oligosaccharide, protein, small molecule, unknown
- **Status** – status of clinical approval of compound: Approved, early clinical trials, late clinical trials, pre-clinical, unknown
Archiving/ Saving Results

- Can Download results in zipped folder.
  - No Log in needed
- Can Share results via email as well
- To save your analysis, you will need to log into your account
  - Can log in with Google account
Other Available Downloads

- Select Datasets from top banner
- Zipped folder available for download
- for the following
  - Cancer Genes
  - Validated Oncogenic Mutations
  - Cancer Biomarkers
  - Cancer Bioactivities
Scenario #1

• You want to make a Pan Cancer gene list from the Cancer Genome Interpreter datasets.
  – Download TSV files and sort to make Cancer list
Making a Pan Cancer List from CGI

- Click on “Datasets” on top banner.
  - Select “Cancer Genes” tab.
    - Click “Download”
    - Save files from downloaded folder
    - Open gene_MoA.tsv (mechanism of action)
      - All genes are listed alphabetically and categorized as Act (activating mutation/alteration), LOF (loss of function), or ambiguous.
Scenario #2

- You have a complex abnormal NGS panel you would like interpreted. Which abnormalities are targetable?
  - PTEN deletion
  - IDH1:R132H
  - EGFR amplification
  - CIC: R215W
  - PRKDC frameshift deletion
Converting mutations to correct input format (See slides 6-9)

- PTEN:del
- IDH1:p.R132H
- EGFR:amp
- CIC:p.R215W
- PRKDC:p.1351fs

- Select ‘Glioma’ for cancer type
Results - Go to Prescriptions

- Drugs for EGFR amplification have “No Responsive” effect.
- Target PTEN deletion with MTOR inhibitor or PI3K pathway inhibitor
- Target IDH1 with IDH1 inhibitor
- Cite references in report if necessary
Scenario #3

• You are building an interface to aid in interpretation and analysis of genomic testing results. You would like to incorporate CGI into your product.
  – Use of API
Shared API

- [https://www.cancergenomeinterpreter.org/rest_api](https://www.cancergenomeinterpreter.org/rest_api)
- API is shared online
- Some of the functionality of the API requires log in to account
  - Can sign up with Google account
Education/ Tutorials

• https://www.cancergenomeinterpreter.org/faq#q01
Contacts

• bbglab@irbbarcelona.org – Contact for comments, suggestions, bug reports