

Advanced cancer genomic data visualization: The Onco Query Language (OQL)

You can use the Onco Query Language (OQL) to select and define genetic alterations for all output on the cBioPortal for Cancer Genomics, including the OncoPrint, heat map, and data downloads.

Genetic Alterations

Users can define genetic alterations for three data types:

| Data Type | Keyword | Categories and Levels | Default |
|-------------------------------------|---------|--|--|
| Copy Number Alterations | CNA | AMP Amplified HOMDEL Deep Deletion GAIN Gained HETLOSS Shallow Deletion | AMP and HOMDEL |
| Mutations | MUT | MUT Show mutated cases MUT = x Specific mutations or mutation types. | All somatic, non-synonymous mutations |
| Fusions | FUSION | FUSION | Show cases with fusions |
| mRNA Expression | EXP | EXP < -x Under-expression is less than x SDs below the mean. EXP > x Over-expression is greater than x SDs above the mean. The comparison operators <= and >= also work. | At least 2 standard deviations (SD) from the mean. |
| Protein/phosphoprotein level (RPPA) | PROT | PROT < -x Protein-level under-expression is less than x SDs below the mean. PROT > x Protein-level over-expression is greater than x SDs above the mean. The comparison operators <= and >= also work. | At least 2 standard deviations (SD) from the mean. |

Basic Usage

Assuming you have selected mutations, copy number data, and mRNA expression data in step 2 of your query, you can use OQL to view only amplified cases in CCNE1:

```
CCNE1: AMP
```

or amplified and gained cases:

```
CCNE1: CNA >= GAIN
```

which can also be written:

```
CCNE1: GAIN AMP
```

To view cases with specific mutations:

```
BRAF: MUT = V600E
```

or mutations on specific position only:

BRAF: MUT = V600

or mutations of a specific type:

TP53: MUT = <TYPE>

<TYPE> could be

- MISSENSE
- NONSENSE
- NONSTART
- NONSTOP
- FRAMESHIFT
- INFRAME
- SPLICE
- TRUNC

e.g., to view TP53 truncating mutations and in-frame insertions/deletions:

TP53: MUT = TRUNC MUT = INFRAME

To view amplified and mutated cases:

CCNE1: AMP MUT

but to define over-expressed cases as those with mRNA expression greater than 3 standard deviations above the mean:

CCNE1: EXP > 3

To query cases that are over expressed in RPPA protein/phosphoprotein level:

EGFR: PROT > 2

or

EGFR_PY992: PROT > 2

Hint: inputting RPPA-PROTEIN or RPPA-PHOSPHO in the query will allow you to select from all proteins or phosphoproteins that have RPPA levels.

In general, any combination of OQL keywords and/or categories can annotate any gene.

Example: RB Pathway

Using the Defaults

Assuming these data types are selected in Step 2 of your query:

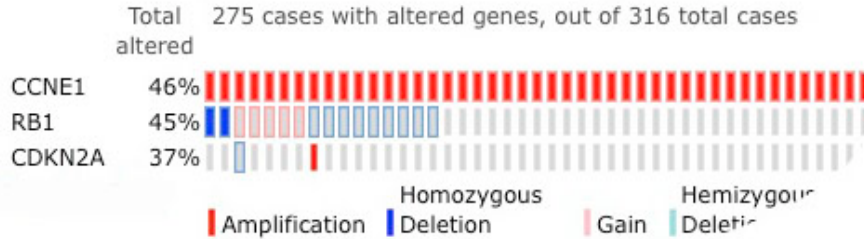
- Mutations
- Copy-number alterations

- mRNA expression

Selecting ovarian cancer and inputting the following three genes in the RB1 pathway

CCNE1 RB1 CDKN2A

displays the default visualization:



Greater Insight with the OQL Language

Given what is known about the RB pathway, the events that are most likely selected for in the tumors are CCNE1 amplification, RB1 deletions or mutations, and loss of expression of CDKN2A. To investigate this hypothesis, we use OQL to display only these events:

```

CCNE1: AMP MUT
RB1: HOMDEL MUT
CDKN2A: HOMDEL EXP < -1

```



This shows that alterations in these genes are almost entirely mutually-exclusive -- no cases are altered in all three genes, and only 8 are altered in two genes. This supports the theory that the tumor has selected for these events.

The DATATYPES Command

To save copying and pasting, the DATATYPES command sets the genetic annotation for all subsequent genes. Thus,

```
DATATYPES: AMP GAIN HOMDEL EXP > 1.5 EXP<=-1.5; CDKN2A MDM2 TP53
```

is equivalent to

```

CDKN2A : AMP GAIN HOMDEL EXP<=-1.5 EXP>1.5;
MDM2   : AMP GAIN HOMDEL EXP<=-1.5 EXP>1.5;
TP53   : AMP GAIN HOMDEL EXP<=-1.5 EXP>1.5;

```

Note that the order of datatype specifications is immaterial, and that a ' : sequence of data specifications ' command can be terminated by an end-of-line, a semicolon or both.

Please share any questions or feedback on this language with us.

cBioPortal Version 1.6.2-SNAPSHOT | MSKCC | TCGA
Questions and feedback: [cbioportal at googlegroups dot com](mailto:cbioportal@googlegroups.com) | [User discussion group](#) | [BioStars](#)
