

Hands-on exercises on FireBrowse
TCGA analysis Workshop
January 7, 2016

Prepared by: Parthav Jailwala, CCBR

Exercise 1: Exploring iCoMut features

[Comut plot](#) is often seen in cancer research publications as a visual summary of genetic variations/mutations in cancer cohorts. The plot itself is essentially an ensemble of multiple smaller and simpler plots (e.g. heat maps and bar graphs) aligned and interconnected via a common X or Y axis. This form of grid-like graphical summary is particularly suitable for presenting data with an intricate and associative nature. iCoMut aims to further transform the static image into an interactive data exploring experience that is flexible and versatile, and through a web-enabled user interface, provides greater public access.

Step 1: Access firebrowse.org

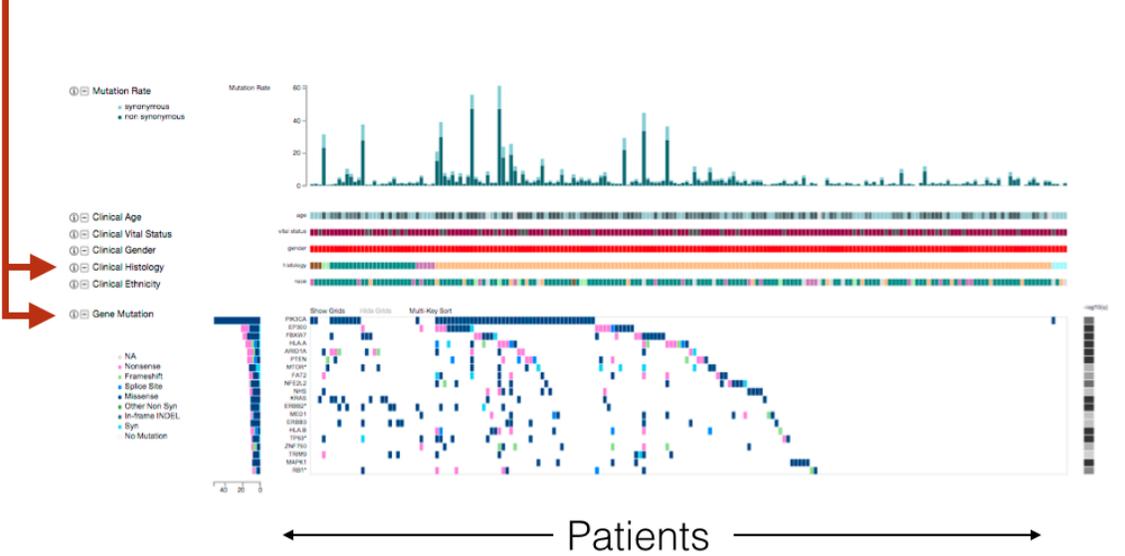
<http://firebrowse.org>

Step 2: Type 'CESC' in the box next to 'View Analysis Profile' and click on 'View Analysis Profile' box. You will now visualize the iCoMut plot for all CESC samples in TCGA, with various tracks for clinical and genomic data associated with these samples

Step 3: Explore different ways of sorting the tracks

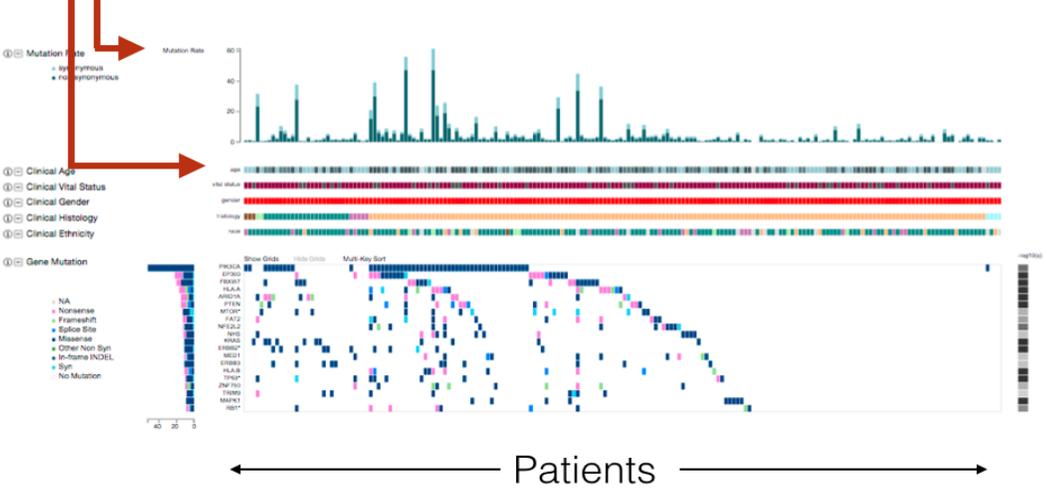
By default, patients in a cancer cohort are sorted first by the histological subtypes (if available), then by the gene mutations.

By default, patients are sorted by histology and gene mutation

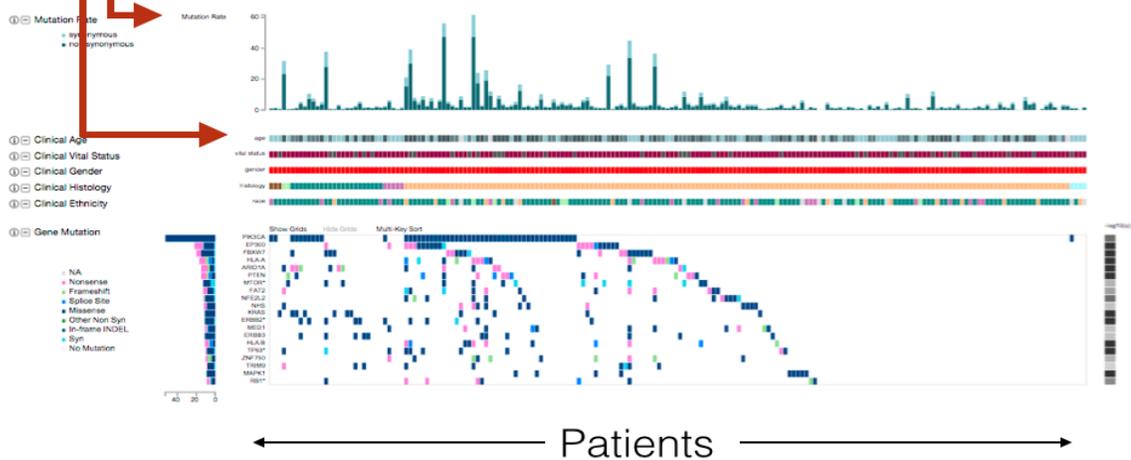


Data can be sorted by clicking on the text labels of the axes and the heat maps. To perform a multi-key sorting, **alt-click** on a text label to add a key.

Alt-Click on text labels to add key to sorting

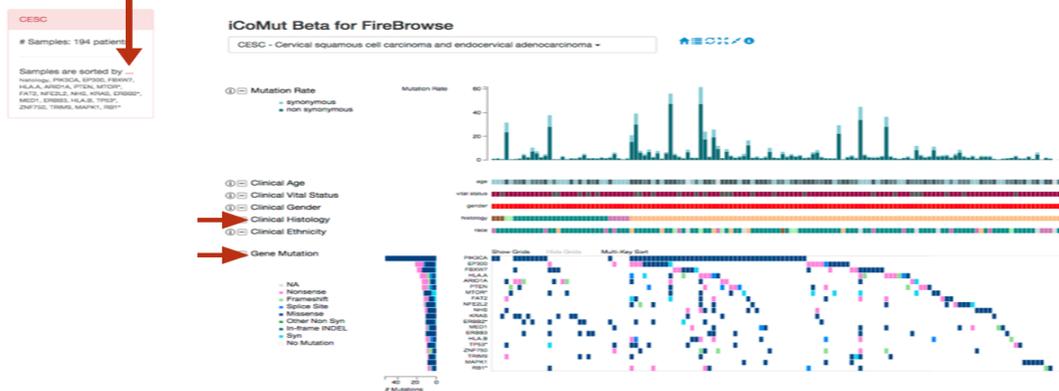


Click on the text labels to change sorting.



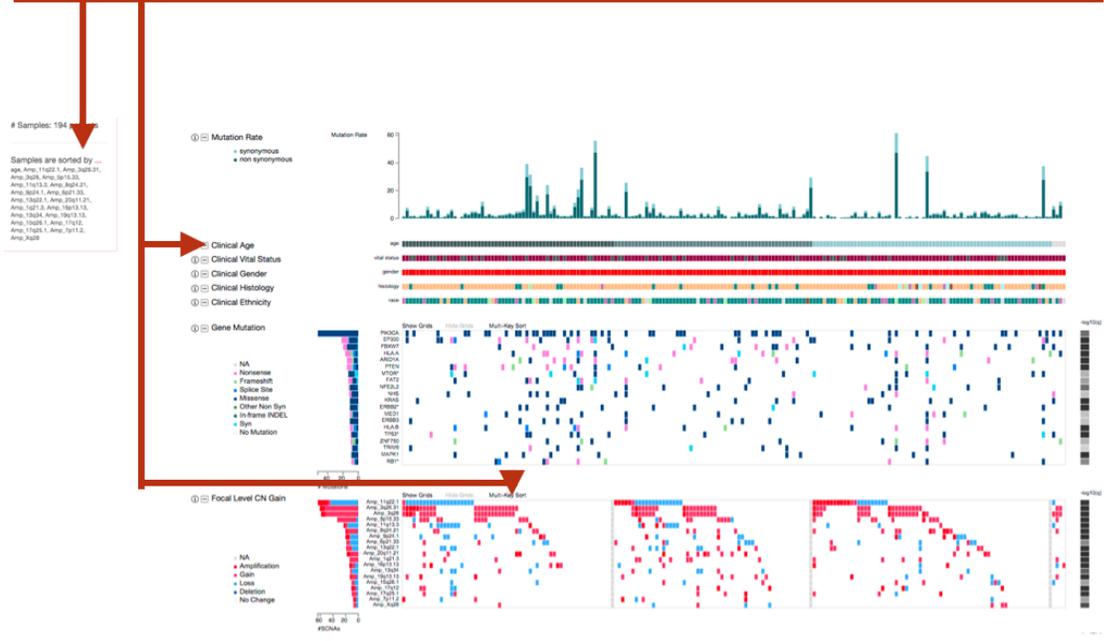
The info box located on the left reports the sorting status of the patients.

The sorting status of samples is reported in the info box on the left



Many options to sort patients can be done by mouse clicking. Here, for example, patients are sorting by age (click) then by focal-level copy number gain (alt-click on "Multi-Key Sort" located above the copy number heat map).

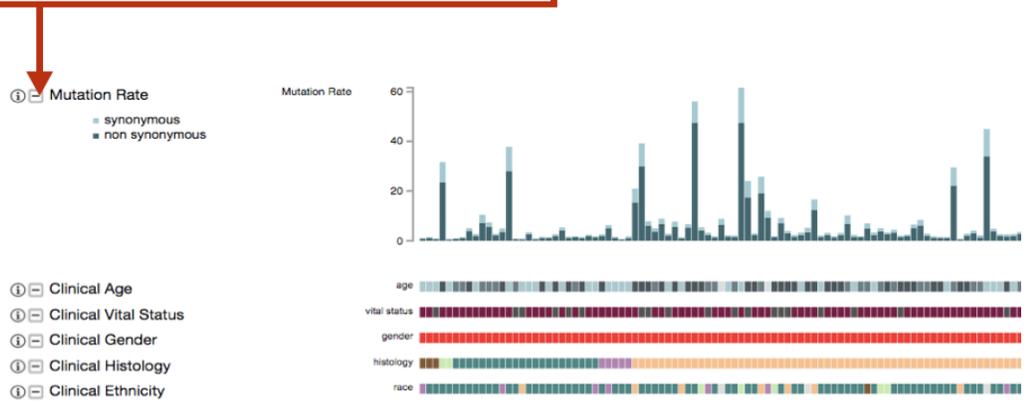
Patients sorted by age and focal-level copy number gain panel



Step 4: Explore rearranging the panel layout

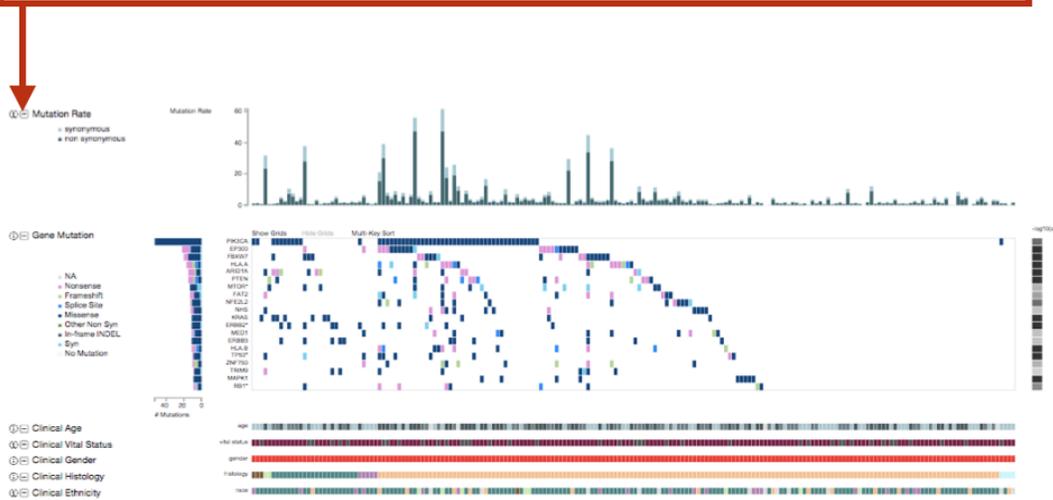
A panel can be collapsed by clicking on the minus icon.

Click on  to collapse a panel



Horizontal panels can be rearranged by drag and drop using the minus/plus icons located on the left of the panel titles.

Drag and drop the  or  icon to rearrange the panels



Rearranged panels

Collapse and Expand All

Control buttons, for collapsing and expanding all panels, are conveniently provided at the top of the comut plot. Additionally, by first collapsing all panels, the rearrangement of the layout order of panels can be done more easily by drag and dropping the plus sign next the panel title.

Click  to collapse all panels

iCoMut Beta for FireBrowse

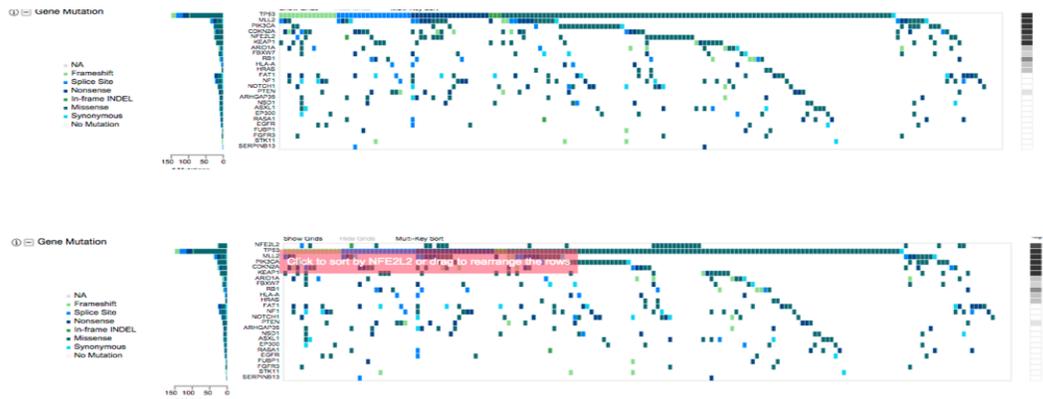
CEC - Cervical squamous cell carcinoma and endocervical adenocarcinoma

- Mutation Rate
- Clinical Age
- Clinical Vital Status
- Clinical Gender
- Clinical Histology
- Clinical Ethnicity
- Gene Mutation
- Focal Level CN Gain
- Focal Level CN Loss
- CN cNMF
- Methylation cNMF
- RPPA cNMF Clusters
- RPPA cHierarchical
- mRNAseq cNMF
- mRNAseq cHierarchical
- miRseq cNMF
- miRseq cHierarchical
- miRseq Mature cNMF
- miRseq Mature cHierarchical

Step 5: Explore sorting rows within an interactive Heat Map

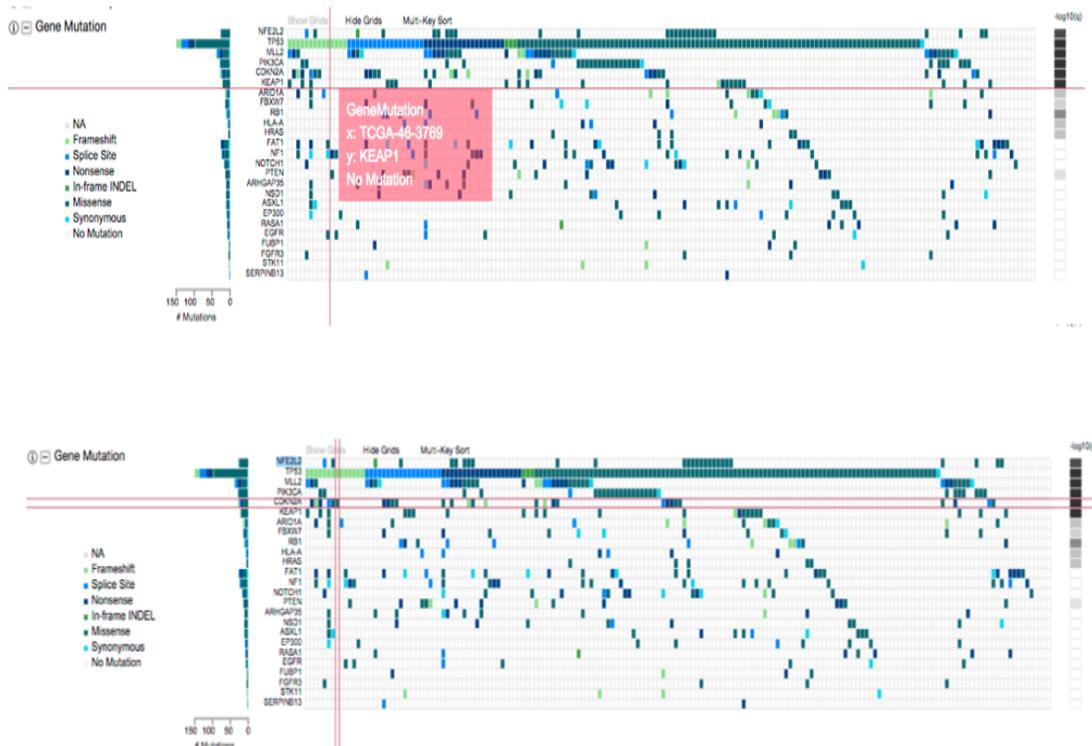
Clicking (or alt-clicking) a row name will sort the columns (i.e. patients) by the row. The sorting order is toggled by clicking the same row name again. Additionally, users can easily rearrange the order of rows by dragging and dropping the row names.

Drag and drop the row names to rearrange the row order



Clicking 'Show Grids' will show all the rectangles in a heat map, including the ones with no data. When in this mode, the cross-hair cursor is also shown. The cursor can be locked in at a rectangle by double-clicking. The locked cross-hair cursor will appear double-lined. To unlock the cursor, click again. Click 'Hide Grids' to hide the grids and cross-hair cursor. By not showing all the graphical elements in a heat map, the visual rendering is faster.

Cross hair mode



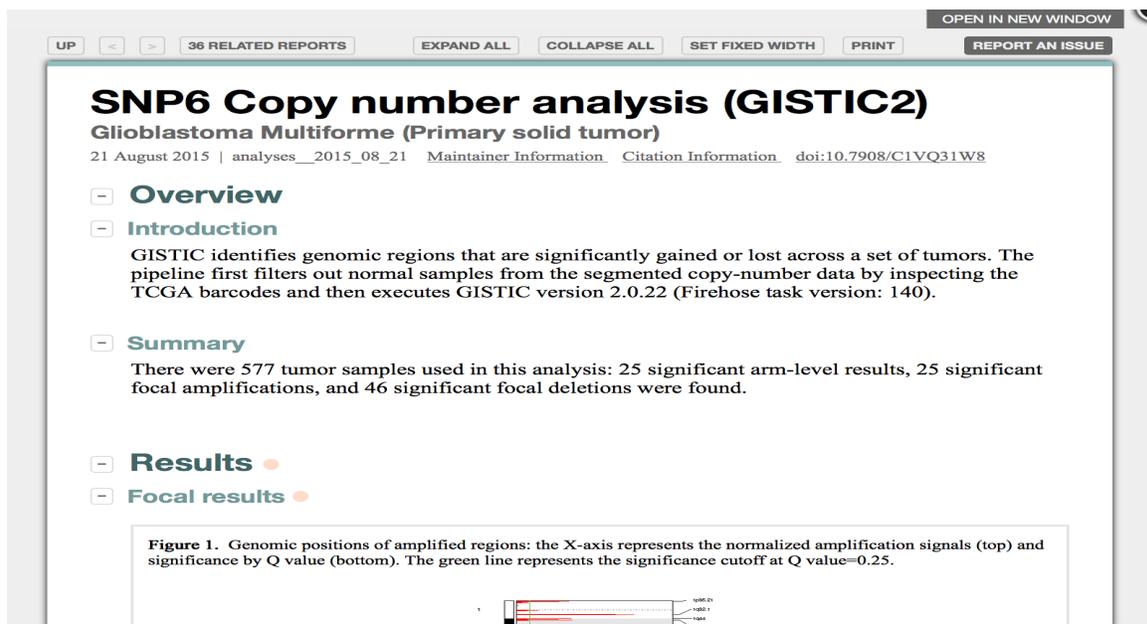
Exercise 2: For the Glioblastoma multiforme (GBM) cohort, find the genes under the top 2 most significantly amplified and deleted regions based on Copy number analysis using Gistic2

Step 1: Select ‘Glioblastoma multiforme (GBM)’ in the SELECT COHORT drop-down list on the front page

Step2: Click on CopyNumber Analyses, and it will show you all the analyses for that data-type

Step 3: Click on ‘CopyNumber Gistic2’ to select the analysis in which GISTIC algorithm was used to identify genomic regions that are significantly gained or lost across the set of all tumors.

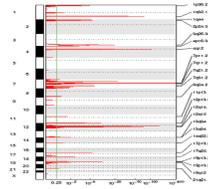
Step 4: You should see a window open with details of the SNP6 Copy number analysis (GISTIC2). Click on ‘Expand All’ to expand all the report sections



Step 5: Under Results -> Focal results, you should see Figure 1 and Table 1. Figure 1 is the graphical view of the genomic positions of the amplified regions; Table 1 lists the 25 most significant amplifications found.

- Results ●
- Focal results ●

Figure 1. Genomic positions of amplified regions: the X-axis represents the normalized amplification signals (top) and significance by Q value (bottom). The green line represents the significance cutoff at Q value=0.25.



[GET FULL TABLE](#)

Table 1. Amplifications Table - 25 significant amplifications found. Click the link in the last column to view a comprehensive list of candidate genes. If no genes were identified within the peak, the nearest gene appears in brackets.

Cytoband	Q value	Residual Q value	Wide Peak Boundaries	# Genes in Wide Peak
7p11.2	0	0	chr7:54964317-54973195	1
12q14.1	1.9251e-150	8.9579e-124	chr12:58135797-58152580	4
4q12	4.6936e-106	4.6936e-106	chr4:55139083-55159166	2
1q32.1	5.0601e-76	7.513e-76	chr1:204443069-204529807	2
12q15	8.296e-75	1.3168e-31	chr12:69178021-69260755	3

Step 6: For the first region (7p11.2), click on the number under the column ‘# Genes in Wide Peak’. You should now see another window open up showing SNORA73 gene as amplified under the 7p11.2 peak.

Table 1. Amplifications Table - 25 significant amplifications found. Click the link in the last column to view a comprehensive list of candidate genes. If no genes were identified within the peak, the nearest gene appears in brackets.

Cytoband	Q value	Residual Q value	Wide Peak Boundaries	# Genes in Wide Peak
7p11.2	0	0	chr7:54964317-54973195	1
12q14.1	1.9251e-150	8.9579e-124	chr12:58135797-58152580	4
4q12	4.6936e-106	4.6936e-106	chr4:55139083-55159166	2
1q32.1	5.0601e-76	7.513e-76	chr1:204443069-204529807	2
12q15	8.296e-75	1.3168e-31	chr12:69178021-69260755	3
3q26.33	1.5076e-20	1.5076e-20	chr3:181256574-181508514	1
12p13.32	5.0468e-11	5.0468e-11	chr12:4325317-4436676	2
19p13.2	1.0528e-11	3.8819e-10	chr19:7300056-7398264	0 [INSR]
7q31.2	3.2182e-08	5.8541e-08	chr7:116318088-116493406	2

Genes in Wide Peak

This is the comprehensive list of amplified genes in the wide peak for 7p11.2.

Table S1. Genes in bold are cancer genes as defined by [The Sanger Institute: Cancer Gene Census](#)[7].

Genes
[SNORA73|ENSG00000252054.1]

Step 7: Similarly by clicking on the second region with 4 genes (12q14.1), you should see these 4 genes as amplified.

Table 1. Amplifications Table - 25 significant amplifications found. Click the link in the last column to view a comprehensive list of candidate genes. If no genes were identified within the peak, the nearest gene appears in brackets.

Cytoband	Q value	Residual Q value	Wide Peak Boundaries	# Genes in Wide Peak
7p11.2	0	0	chr7:54964317-54973195	1
12q14.1	1.9251e-150	8.9579e-124	chr12:58135797-58152580	4
4q12	4.6936e-106	4.6936e-106	chr4:55139083-55159166	2
1q32.1	5.0601e-76	7.513e-76	chr1:204443069-204529807	2
12q15	8.296e-75	1.3168e-31	chr12:69178021-69260755	3
3q26.33	1.5076e-20	1.5076e-20	chr3:181256574-181508514	1
12p13.32	5.0468e-11	5.0468e-11	chr12:4325317-4436676	2
19p13.2	1.0528e-11	3.8819e-10	chr19:7300056-7398264	0 [INSR]
7q31.2	3.2182e-08	5.8541e-08	chr7:116318088-116493406	2

Genes in Wide Peak

This is the comprehensive list of **amplified** genes in the wide peak for 12q14.1.

Table S2. Genes in bold are cancer genes as defined by [The Sanger Institute: Cancer Gene Census](#)[7].

Genes
CDK4
TSPAN31
MARCH9
AGAP2

Step 8: For finding the genes under the top 2 significantly deleted regions, follow the same steps using Table 2. You will see important genes like CDKN2A, CDKN2B and PTEN amongst the top 2 deleted regions.

Table 2. Deletions Table - 46 significant deletions found. Click the link in the last column to view a comprehensive list of candidate genes. If no

[GET FULL TABLE](#)

Cytoband	Q value	Residual Q value	Wide Peak Boundaries	# Genes in Wide Peak
9p21.3	0	0	chr9:21227468-23690570	18
10q23.31	6.5075e-64	1.1173e-60	chr10:89617158-89755074	3
1p36.23	2.112e-58	2.9977e-43	chr1:7908902-8336254	4
6q26	1.1419e-34	3.0061e-30	chr6:163736503-164264583	3
13q14.2	1.5865e-26	7.5581e-20	chr13:48875329-49064807	2
10q26.3	1.3777e-26	2.0282e-18	chr10:135190263-135214343	1
19q13.41	2.3326e-15	2.4075e-15	chr19:51485195-51915080	21
1p32.3	2.2509e-15	2.0343e-12	chr1:50879767-51529064	4
15q14	4.1062e-12	4.1062e-12	chr15:33485762-45387898	149
22q13.32	7.681e-12	7.708e-12	chr22:48668761-51304566	42
14q13.1	4.5717e-13	9.0946e-12	chr14:33297556-34394000	1
4p34.2	1.4106e-09	1.4106e-09	chr4:170280063-182752110	4

Table S25. Genes in bold are cancer genes as defined by [The Sanger Institute: Cancer Gene Census](#)[7].

Genes
DMRTA1
RN7SL151P
SNORD39I|ENSG00000264379.1
MIR31HG
IFNA6
KLHL9
CDKN2A
CDKN2B
IFNA1
IFNA2
IFNA5
IFNA8

Exercise 3: Viewing expression of two genes across cancer types

A 2014 report (see abstract below) states that AML cells are exceptionally reliant on the cell-cycle regulator CDK6, but not its functional homolog CDK4, and that the preferential growth inhibition induced by CDK6 depletion is mediated through enhanced myeloid differentiation. Check if it is true that the expression of CDK6 is higher in LAML (Acute Myeloid Leukemia) and check whether CDK4 shows the same expression pattern as CDK6 in LAML

Blood. 2014 Jul 3;124(1):13-23. doi: 10.1182/blood-2014-02-558114. Epub 2014 Apr 24.

Requirement for CDK6 in MLL-rearranged acute myeloid leukemia.

Placke T¹, Faber K², Nonami A³, Putwain SL⁴, Salih HR⁵, Heidel FH⁶, Krämer A⁷, Root DE⁸, Barbie DA⁹, Krivtsov AV¹⁰, Armstrong SA¹⁰, Hahn WC⁹, Huntly BJ⁴, Sykes SM¹¹, Milsom MD¹², Scholl C², Fröhling S¹.

Author information

Abstract

Chromosomal rearrangements involving the H3K4 methyltransferase mixed-lineage leukemia (MLL) trigger aberrant gene expression in hematopoietic progenitors and give rise to an aggressive subtype of acute myeloid leukemia (AML). Insights into MLL fusion-mediated leukemogenesis have not yet translated into better therapies because MLL is difficult to target directly, and the identity of the genes downstream of MLL whose altered transcription mediates leukemic transformation are poorly annotated. ~~We used a functional genetic approach to uncover that AML cells driven by MLL-AF9 are exceptionally reliant on the cell-cycle regulator CDK6, but not its functional homolog CDK4, and that the preferential growth inhibition induced by CDK6 depletion is mediated through enhanced myeloid differentiation.~~ CDK6 essentiality is also evident in AML cells harboring alternate MLL fusions and a mouse model of MLL-AF9-driven leukemia and can be ascribed to transcriptional activation of CDK6 by mutant MLL. Importantly, the context-dependent effects of lowering CDK6 expression are closely phenocopied by a small-molecule CDK6 inhibitor currently in clinical development. These data identify CDK6 as critical effector of MLL fusions in leukemogenesis that might be targeted to overcome the differentiation block associated with MLL-rearranged AML, and underscore that cell-cycle regulators may have distinct, noncanonical, and nonredundant functions in different contexts.

© 2014 by The American Society of Hematology.

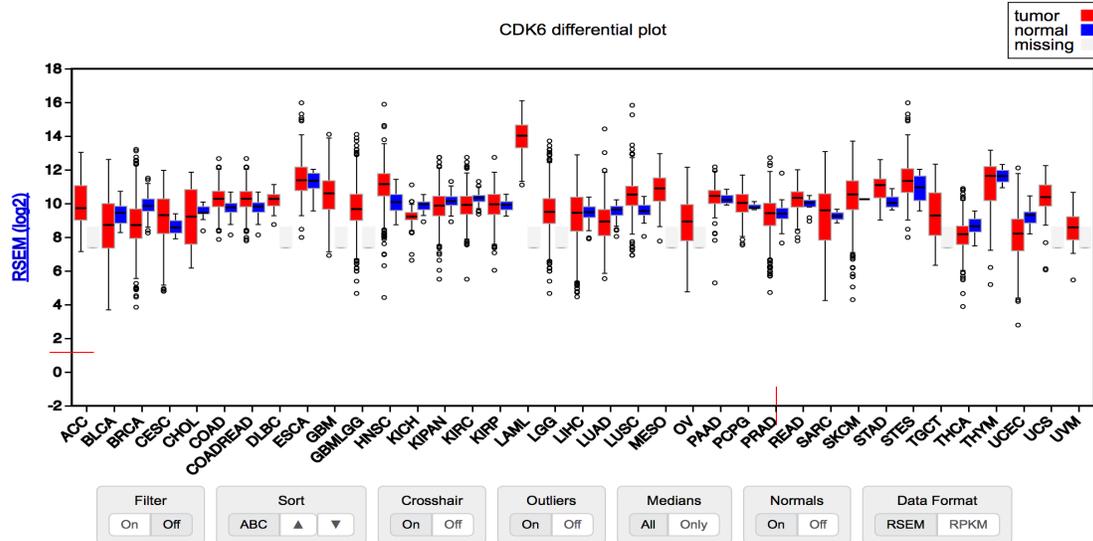
Step 1: On the homepage (firebrowse.org), type 'CDK6' in the search box and click on 'View Expression Profile'.

Search analysis

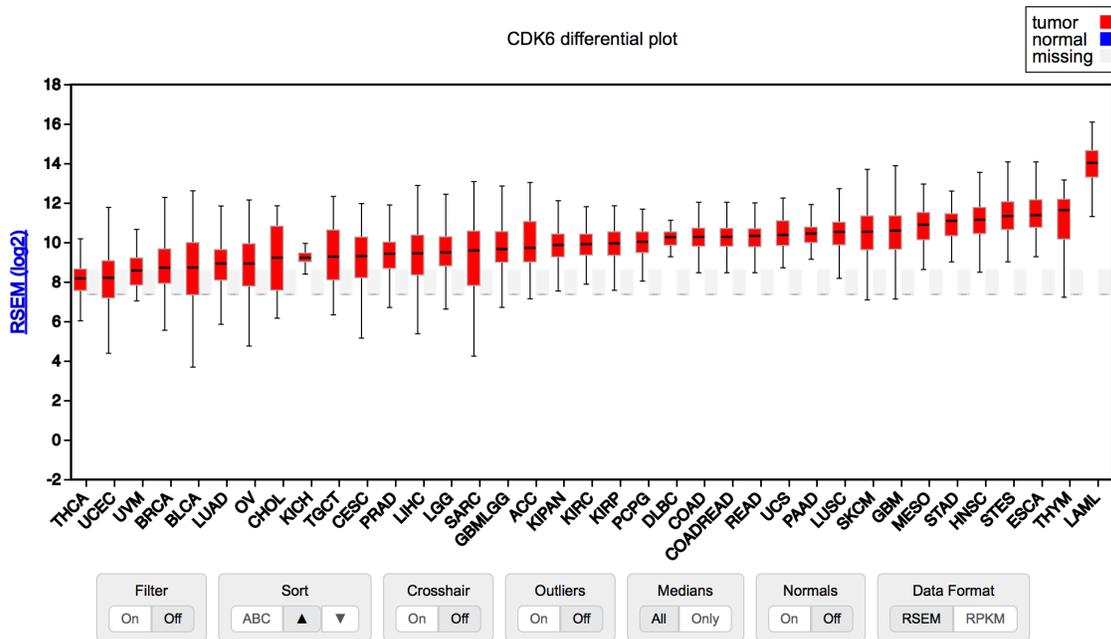
HOME BROAD GDAC WEB API TUTORIAL RELEASE NOTES ANALYSES GRAPH FAQ CONTACT

View Expression Profile CDK6 Enter cohort abbrev View Analysis Profile

Step 2: This will open another window with the following boxplots showing the expression of CDK6 across various cancer types. Red bars are for tumor samples and blue bars are for normal samples.



Step 3: As we would like to view only Tumor samples, turn OFF display of normal samples on the control panel below the boxplot. Also turn OFF Outliers display and Sort by Low to High expression. This will change the boxplot and you should see this figure:



Step 4: We can now clearly see that the expression of CDK6 is the highest in LAML (Acute Myeloid Leukemia) and it verifies the statement made in the abstract above. Further, type CDK4 in the search box on top of the page and visualize the expression boxplot for CDK4 across cancer types. Again, sort the view by Low to High expression. You should see the figure shown below. We can now confirm that in contrast to CDK6 expression, the expression of CDK4 is not high in LAML, in line with the statement made in the abstract above.

