

Tutorial Content

- Summary of Viewers and data sets on CGWB page 2-3
- Getting started page 4-5
- Navigation page 6-10
- View somatic mutation page 11-32
- View integrated genomic data page 33-42
- View correlation plot page 43-48

Cancer Genome Workbench (CGWB)

A summary of visualization tools

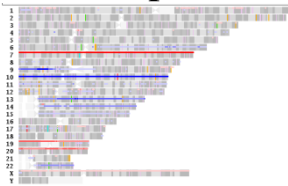
1. Landscape View: Identify target regions such as mutation hotspots, recurrent and/or focal copy number changes
2. Genome View: See how the mutation/copy number changes affect the reference human genome. Evaluate the relationship between somatic mutation and copy number change using the integrated copy-number view. Are these two types of somatic changes exclusive or inclusive? You can also compare the mutations discovered by TCGA project with the published mutations and find out how many new mutations are discovered in this project
3. Heatmap View: View genomic data along chromosome, gene, gene list or pathway
4. Mutation View: Report of mutation frequency, functional changes, assay info and history (known or novel)
5. Protein View: view mutations on a protein sequence
6. 3D Structure View: View the impact of mutation on protein 3D structure
7. Trace View: Evaluate the quality of a putative mutation
8. Correlation plot: see how methylation data correlates with gene expression, how microRNA regulate gene expression

Data and Viewers @ CGBW

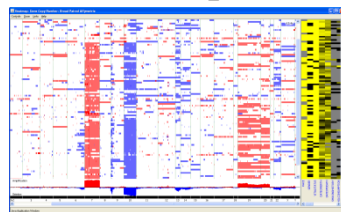
Project	Disease	Genomic Data				
		Copy Number	Gene Expression	Methylation	Clinical	Mutation
TCGA	GBM, Ovarian	+	+	+	+	+
TSP	lung	+			+	+
TARGET	ALL, NBL	+	+		+	+
JHU	GBM, Pan					+
Rembrandt	GBM	+				
GSK cell line	>30 tissues	+	+			
COSMIC	>30 tissues					+

Viewers for individual & integrated genomic Data

Landscape



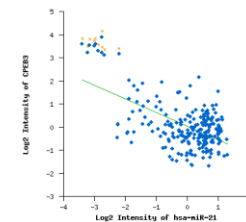
Heatmap



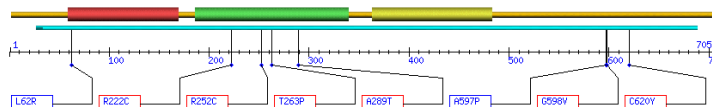
Genome



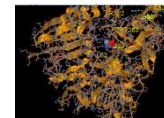
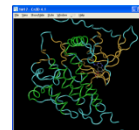
Correlation Plot



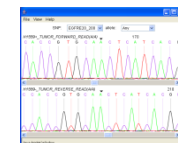
Protein



3D Structure Viewer



Trace



Getting Started in CGWB

The screenshot shows the CGWB homepage with a navigation bar containing links: [CGWB Home](#), [Tutorial](#), [Datadump](#), [Help](#), [Log In](#), and [Heatmap](#). Below the navigation bar is a search input field labeled "Enter Genomic Coords or Gene Name:". Four callout boxes with arrows point to the navigation links: "View Tutorial" points to the Tutorial link, "Get Mutation Summary" points to the Datadump link, "Login to access protected data" points to the Log In link, and "Click to view heatmaps" points to the Heatmap link.

The screenshot shows a Mozilla Firefox browser window displaying a mutation summary table. The browser's address bar shows the URL <https://cgwb-test.nci.nih.gov>. The page content includes a note: "CGWB Home - Note: this output is project focused, for all genes: [Click Here.](#)". Below the note is a table with the following data:

Gene	Project	Tissue	#Tumors	#Mutations	%MutatedTumor	Mutation/Mb	
TP53	TCGAGSC	GBM_Valid	Brain	181	57	25	266
PTEN	TCGAGSC	GBM_Valid	Brain	181	40	20	182
EGFR	TCGAGSC	GBM_Valid	Brain	181	24	12	36
NF1	TCGAGSC	GBM_Valid	Brain	230	23	8	11
RB1	TCGAGSC	GBM_Valid	Brain	181	12	6	23
PIK3R1	TCGAGSC	GBM_Valid	Brain	230	13	5	25
FKBP9	TCGAGSC	GBM_Valid	Brain	135	6	4	25
SYNE1	TCGAGSC	GBM_Valid	Brain	170	10	4	2
PIK3CA	TCGAGSC	GBM_Valid	Brain	147	7	4	14
DST	TCGAGSC	GBM_Valid	Brain	180	7	3	2
BCL11A	TCGAGSC	GBM_Valid	Brain	159	5	3	12
CHEK2	TCGAGSC	GBM_Valid	Brain	152	5	3	18
ERBB2	TCGAGSC	GBM_Valid	Brain	230	11	3	12
ROR2	TCGAGSC	GBM_Valid	Brain	147	3	2	7
TFV	TCGAGSC	GBM_Valid	Brain	170	4	2	6

The browser's status bar at the bottom shows "Done" and the URL cgwb-test.nci.nih.gov.

Click Heatmap and Go To Gene-based or Genome-based (Scroll down) Heatmap Data for TCGA

TCGA GBM project:

Extent	Data type	Submitter	Platform	Title	View
Gene-based view	copy number	combined	combined	Gene Copy Number: Combined	Launch viewer
				Gene Copy Number: Combined Paired	Launch viewer
				Gene Copy Number: Combined Unpaired	Launch viewer
	copy number	Broad	Affymetrix	Gene Copy Number: Broad Paired Affymetrix	Launch viewer
				Gene Copy Number: Broad Unpaired Affymetrix	Launch viewer
				Gene Copy Number: Broad Affymetrix	Launch viewer
		HMS	Agilent	Gene Copy Number: HMS Agilent	Launch viewer
				Gene Copy Number: HMS Paired Agilent	Launch viewer
				Gene Copy Number: HMS Unpaired Agilent	Launch viewer
		MSKCC	Agilent	Gene Copy Number: MSKCC Agilent	Launch viewer
				Gene Copy Number: MSKCC Paired Agilent	Launch viewer
				Gene Copy Number: MSKCC Unpaired Agilent	Launch viewer
				Gene Copy Number: SUSM Illumina	Launch viewer

Done cgwb.nci.nih.gov

Navigate Cancer Genome Workbench

- Gene Name or Genomic Region
- Tissue
- Project
- Exploration of the Landscape view of a Project

Navigation: 1) Navigate a Specific Gene or Region

[CGWB Home](#) | [Tutorial](#) | [Datadump](#) | [Help](#) | [Log In](#) |

The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz. Software Copyright (c) The Regents of the University of California. Additional Copyright (c) by NIH. All rights reserved.

Enter Genomic Coords or Gene Name:

EGFR

Go

Tissue Browse

[Click here to reset](#) the browser us

[click here to add](#) a custom track

Type a gene name or a region
Then hit GO

TCGA

TARGET

TSP

COSMIC

Rembrandt

BCAC

OCAC

JenLab

Somatic Alteration Hotspots

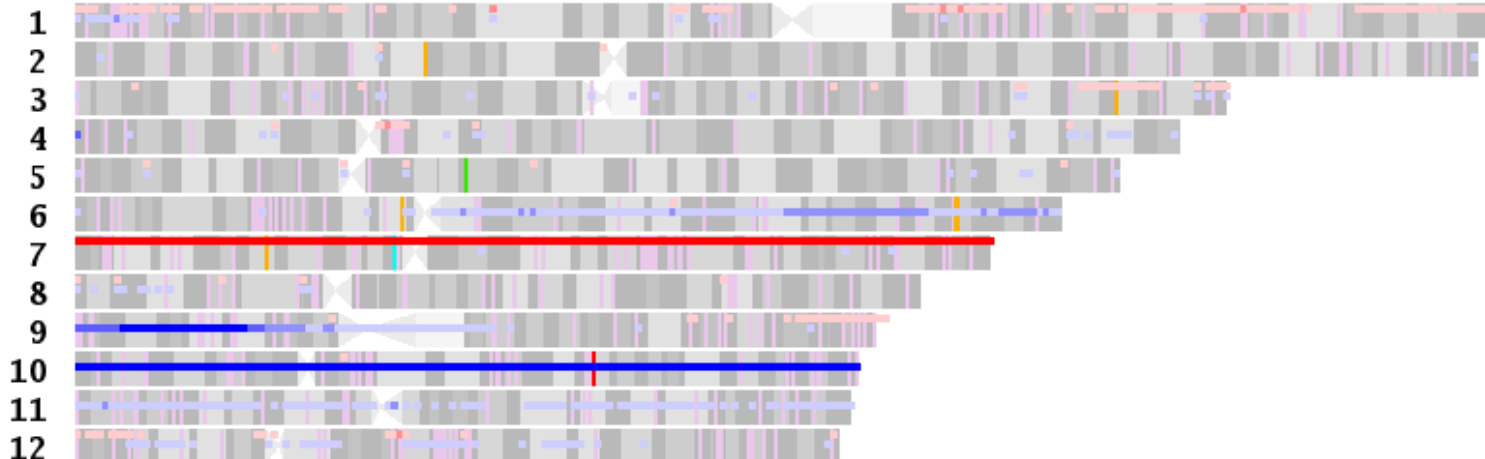
(you can mouse over and click)

Percentage of samples that have somatic mutations (horizontal bars)

0-3 4-5 6-10 10-19 >=20

Percentage of samples that have copy number change (vertical bars)

Amplification >= 40% >= 30% >= 20% >= 10%
Deletion >= 40% >= 30% >= 20% >= 10%



Navigation: 2) Navigate a Gene or Region from a Specific Tissue

The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz. Software Copyright (c) The Regents of the University of California. Additional Copyright (c) by NIH. All rights reserved.

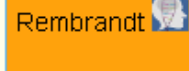
Enter Gene Name and Pick Tissue:

EGFR

Blood
Bone
Brain
Breast

Go

[Click here to reset](#) the browser user interface settings to their defaults. [Click here to add](#) a custom track



Select Brain from the Tissue Browser to access all somatic mutation data from brain tissue. Hit Go

Somatic Alteration Hotspots

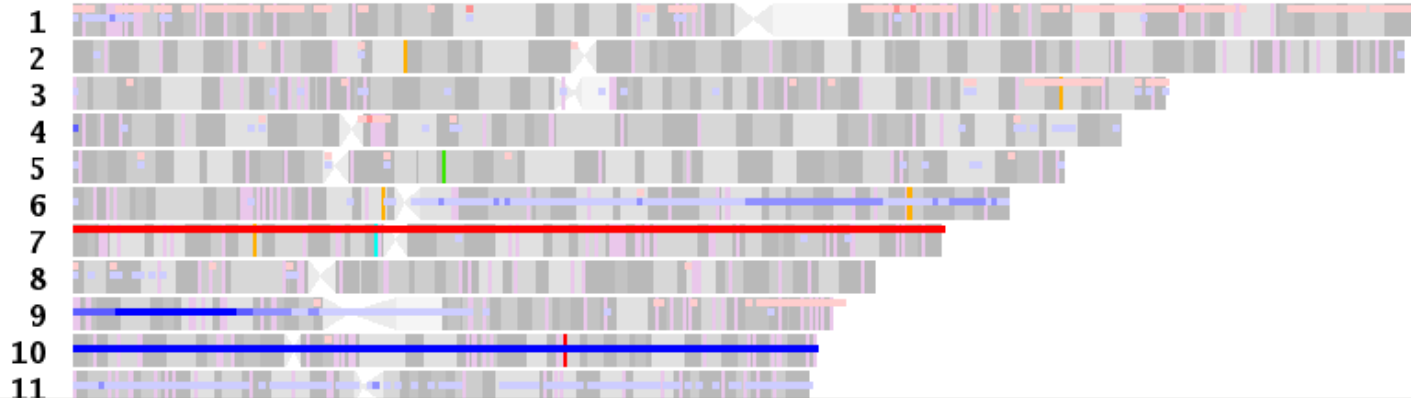
(you can mouse over and click)

Percentage of samples that have somatic mutations (horizontal bars)

0-3 4-5 6-10 10-19 ≥ 20

Percentage of samples that have copy number change (vertical bars)

Amplification $\geq 40\%$ $\geq 30\%$ $\geq 20\%$ $\geq 10\%$
Deletion $\geq 40\%$ $\geq 30\%$ $\geq 20\%$ $\geq 10\%$



Navigation: 3) Navigate a Gene in a Specific Project

1. Click the “Super Project” to open the sub-project. A default sub project will be selected automatically

2. Click to select a sub-project to get the somatic mutation summary, full project title and the landscape view.

3. Click on a Gene

“Super” Project

Project

Project Title

Mutation Summary

LandScape View

TCGA TARG JenLab

TCGAWIBR_GBM TCGAWUGSC_GBM TCGAGSC_GBM_Valid TCGABCM_Ovarian

TCGAWIBR_Ovarian TCGAWUGSC_Ovarian

Current Project: **Validated Somatic Mutations of The Cancer Genome Atlas Project Submitted by Genome Sequencing Center**

List of Genes ordered by Frequency of Samples with Somatic Mutations (with percentage of mutated samples in parentheses):
[TP53 \(25%\)](#) [PTEN \(20%\)](#) [EGFR \(12%\)](#) [NF1 \(8%\)](#) [RB1 \(6%\)](#) [PIK3R1 \(6%\)](#) [FKBP9 \(4%\)](#) [PIK3CA \(4%\)](#) [SYNE1 \(4%\)](#) [BCL11A \(3%\)](#) [CHEK2 \(3%\)](#) [DST \(3%\)](#) [ERBB2 \(3%\)](#) [ADAMTSL3 \(2%\)](#) [ANK2 \(2%\)](#) [COL1A1 \(2%\)](#) [COL6A2 \(2%\)](#) [DDR2 \(2%\)](#) [EPHA7 \(2%\)](#) [GSTM5 \(2%\)](#) [ITGB3 \(2%\)](#) [ITPR3 \(2%\)](#) [MSH6 \(2%\)](#) [PSMD13 \(2%\)](#) [PIK3C2G \(2%\)](#) [PIK3C1 \(2%\)](#)

Click **to GET 467 MORE GENES**

Somatic Alteration Hotspots (you can mouse over and click)

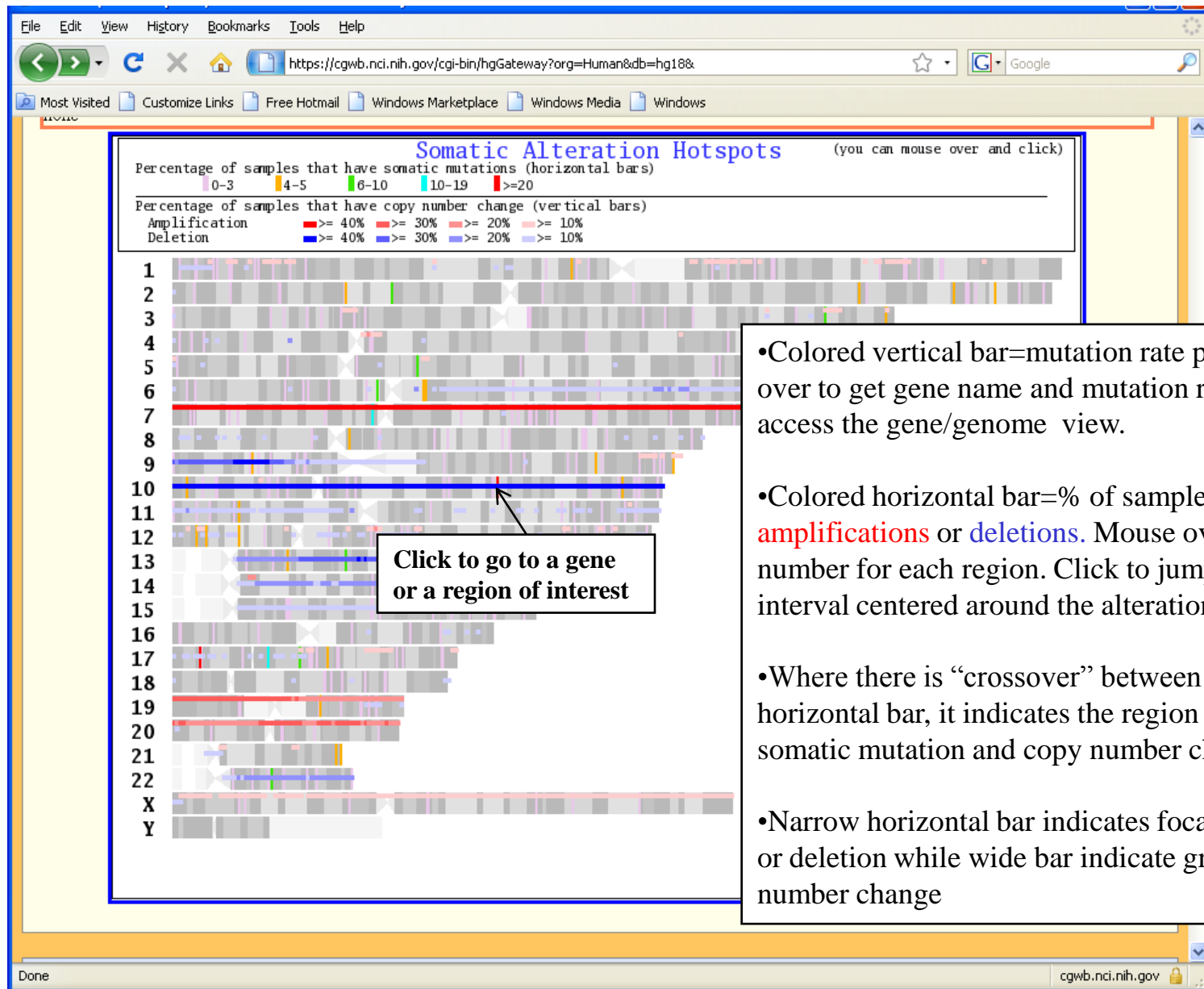
Percentage of samples that have somatic mutations (horizontal bars)
0-3 4-5 6-10 10-19 >=20

Percentage of samples that have copy number change (vertical bars)
Amplification >= 40% >= 30% >= 20% >= 10%
Deletion >= 40% >= 30% >= 20% >= 10%

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

<https://cgwb-test.nci.nih.gov/cgi-bin/hgTracks?db=hg18&position=chr7:55054218-55242524&fpj=41> cgwb-test.nci.nih.gov

Navigation: 4) Explore the LandScape View



- Colored vertical bar = mutation rate per gene. Mouse over to get gene name and mutation rate. Click to access the gene/genome view.

- Colored horizontal bar = % of samples that have **amplifications** or **deletions**. Mouse over to see the number for each region. Click to jump to a 1-Mb interval centered around the alteration

- Where there is “crossover” between vertical bar and horizontal bar, it indicates the region has both somatic mutation and copy number change.

- Narrow horizontal bar indicates focal amplification or deletion while wide bar indicate gross copy number change

View Somatic Mutation

Genome/Gene View Has Three Components

Cancer Genome Workbench - Produced by Laboratory of Population Genetics / National Cancer Institute / Bethesda Maryland - based on UCSC Genome Browser

CGWB Home CGWB Help Variation Table Genotype Table CGWB Log In UCSC Link Tables DNA PDF/PS Help

UCSC Genome Browser on Human Mar 2006 Assembly

move <<< << < > >> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x half left right

position/search chr7:55,054,218-55,242,524 jump clear size 188,307 bp. configure

chr7 (p11.2) 21.8 21.11 22.1 31.1 33.3 34 35

chr7: 55100000 55150000 55200000

RefSeq Genes Human mRNAs from GenBank Vertebrate Multiz Alignment & PhastCons Conservation (26 Species) Mammal Cons RefSeq Motifs Validated Somatic Mutations of The Cancer Genome Atlas Project Submitted by Genome Sequencing Center

motifs 50% 40% 30% 20% 10% 0%

HEAT MAP

Help Button

SNP INFO (mouseover and click)

a) Reference Genome View

b) Tumor Genome View

Cancer Genome Workbench - TCGA refresh

TCGABCM GBMPrimer *hide	TCGABCM GBM *hide	TCGAWIBR GBMPrimer *hide	TCGAWIBR GBM *hide	TCGAWUGSC GBMPrimer *hide
TCGAWUGSC GBM *hide	TCGAGSC GBM Valid *dense	TCGA GBM Significant Region hide	TCGABCM OvarianPrimer *hide	TCGABCM Ovarian *hide
TCGA GBM Integrated hide	TCGAWIBR OvarianPrimer *hide	TCGAWIBR Ovarian *hide	TCGAWUGSC OvarianPrimer *hide	TCGAWUGSC Ovarian *hide
TCGA Ovarian Integrated *hide				

+ Cancer Genome Workbench - TARGET	refresh
+ Cancer Genome Workbench - TSP	refresh
+ Cancer Genome Workbench - COSMIC	refresh
+ Cancer Genome Workbench - JenLab	refresh
+ Cancer Genome Workbench - Johns_Hopkins	refresh
+ Cancer Genome Workbench - GSK_Cell_Lines	refresh
other	refresh

*=CGWB Track has items in current range

Combine Tracks refresh

c) Control Panel at the Bottom for Selecting Style, Marker (silent, missense, etc) and Data sets.

Genome View: a “Dense” View Summarizes Validated Mutations in EGFR

Cancer Genome Workbench - Produced by Laboratory of Population Genetics / National Cancer Institute / Bethesda Maryland - based on UCSC Genome Browser

CGWB Home CGWB Help Variation Table Genotype Table CGWB Log In UCSC Link Tables DNA PDF/PS Help

Export Variation Info Export Genotype

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x half left right

position/search chr7:55,054,218-55,242,524 jump clear size 188,307 bp. configure

chr7 (p11.2) 21.3 21.11 22.1 31.1

chr7: 55100000 55150000 RefSeq Genes

Human MENC from GenBank

Verteb

Conservation (28)

RefSeq Motifs

Validated Somatic Mutations of The Cancer Genome Atlas Project Submitted by Genome Sequencing Center

Protein View

Heatmap View

HEAT MAP

Frequency 50% 40% 30% 20% 10% 0%

Help Button

SNP INFO (mouseover and click)

Toggle to Protein motif “pack view” by clicking

Red dot indicates there are multiple mutations at this pixel and the frequency is summary of all mutations

Gray means novel variation
Magenta means “missense”
Click “Help button” for all the color legend

A289D_21685 missense substitution Sc 0.0389 click for more

Mouse-over to get a brief description. Click to get the variation report view

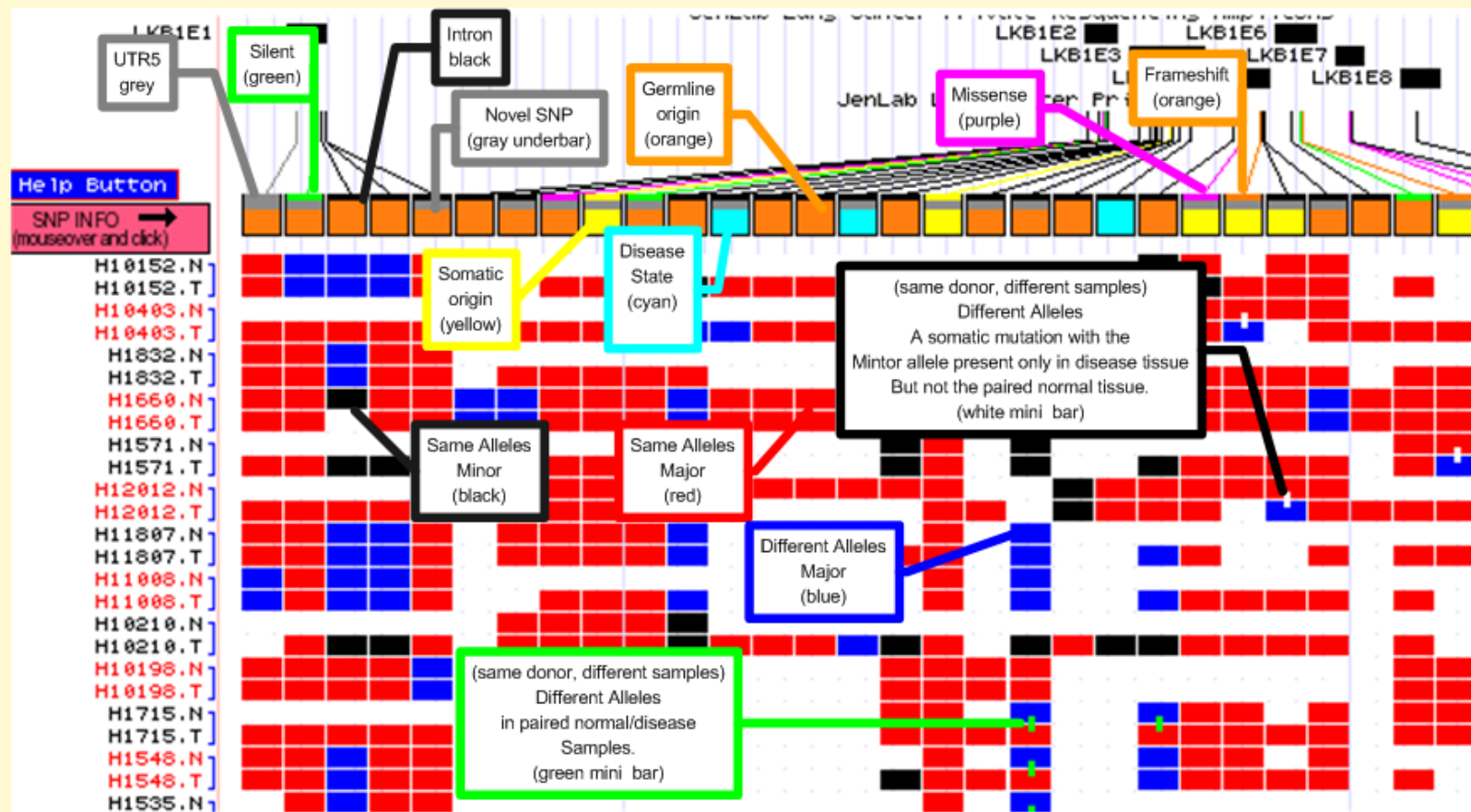
Variation Color Legend After Clicking "Help Button"

Cancer Genome Workbench - Produced by Laboratory of Population Genetics / National Cancer Institute / Bethesda Maryland - based on UCSC Genome Browser

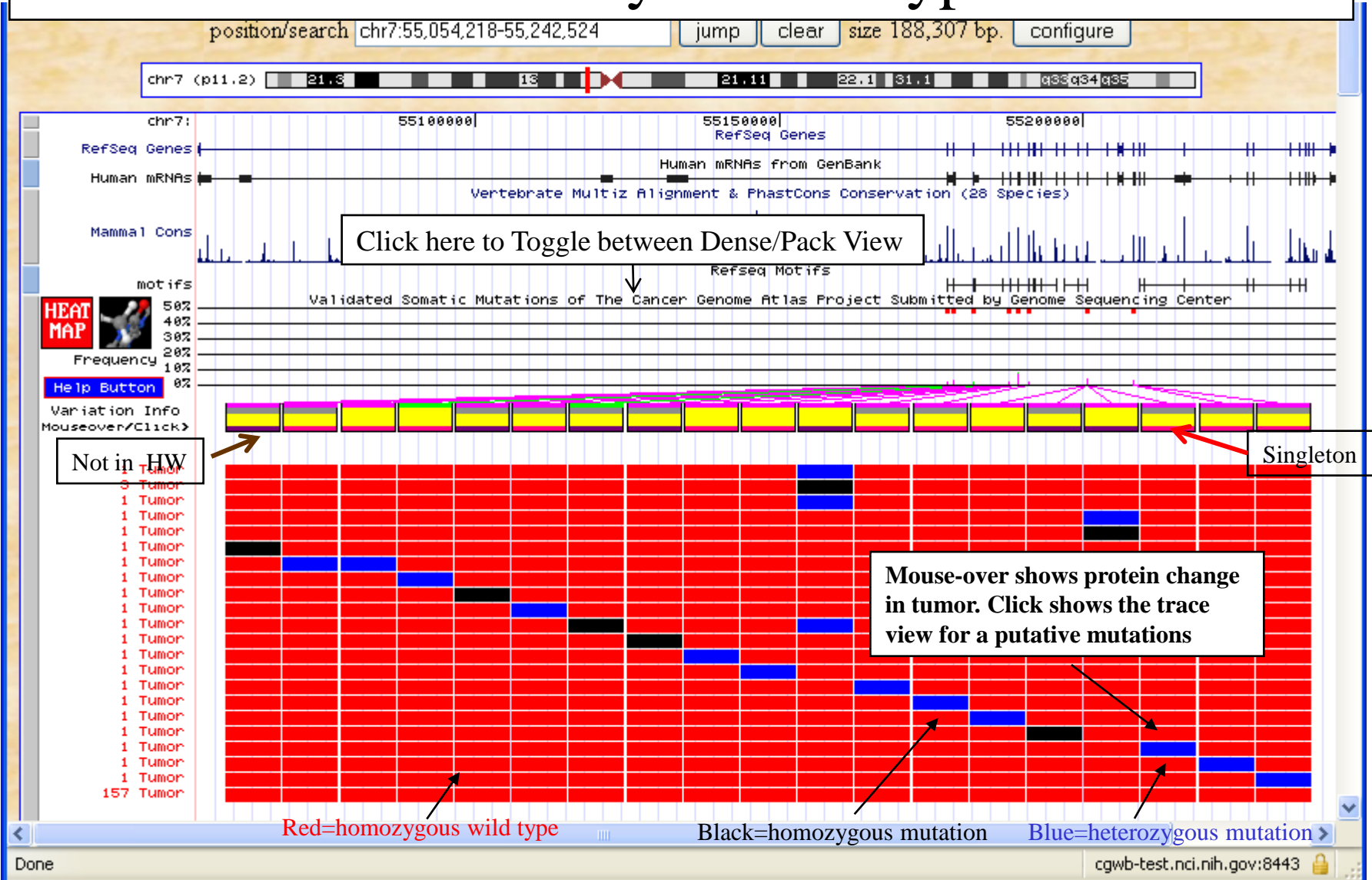


[CGWB Home](#) [Genomes](#) [Tables](#) [Gene Sorter](#) [PCR](#) [FAQ](#) [Help](#)

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#). Software Copyright (c) The Regents of the University of California. All rights reserved.



Mutations can be Displayed in “Pack” View that Shows a Summary of Genotype Profile



Configure Genome Display by Scrolling to the Control Panel at the Bottom

- 1) Select a Subset of Variations: missense, silent, somatic, novel?
- 2) Select a Display Style: Dense (default), Pack or Full

The Cancer Genome Atlas Project by Washington University

Display mode: **dense**

LPG C **"SNPMapping" Type**

Any type of data can be excluded from view by deselecting the checkbox below.

SNP Type:

intergenic UTR_5 missense silent intron UTR_3 frameshift splice exon nonsense

proteinDel proteinIns known Germline Somatic DiseaseState

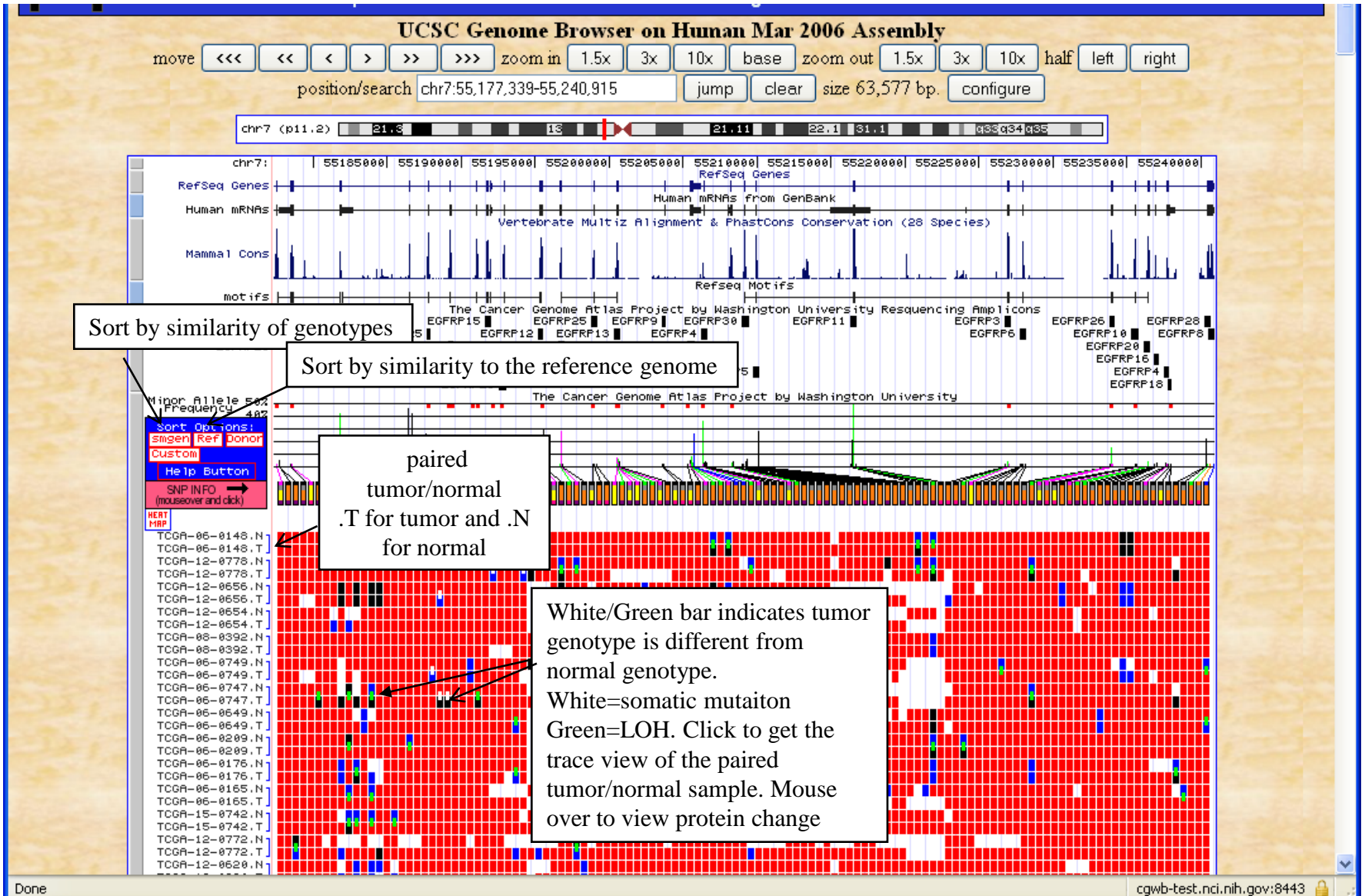
Found in tumor samples that do not have matching normal

[View table schema](#)

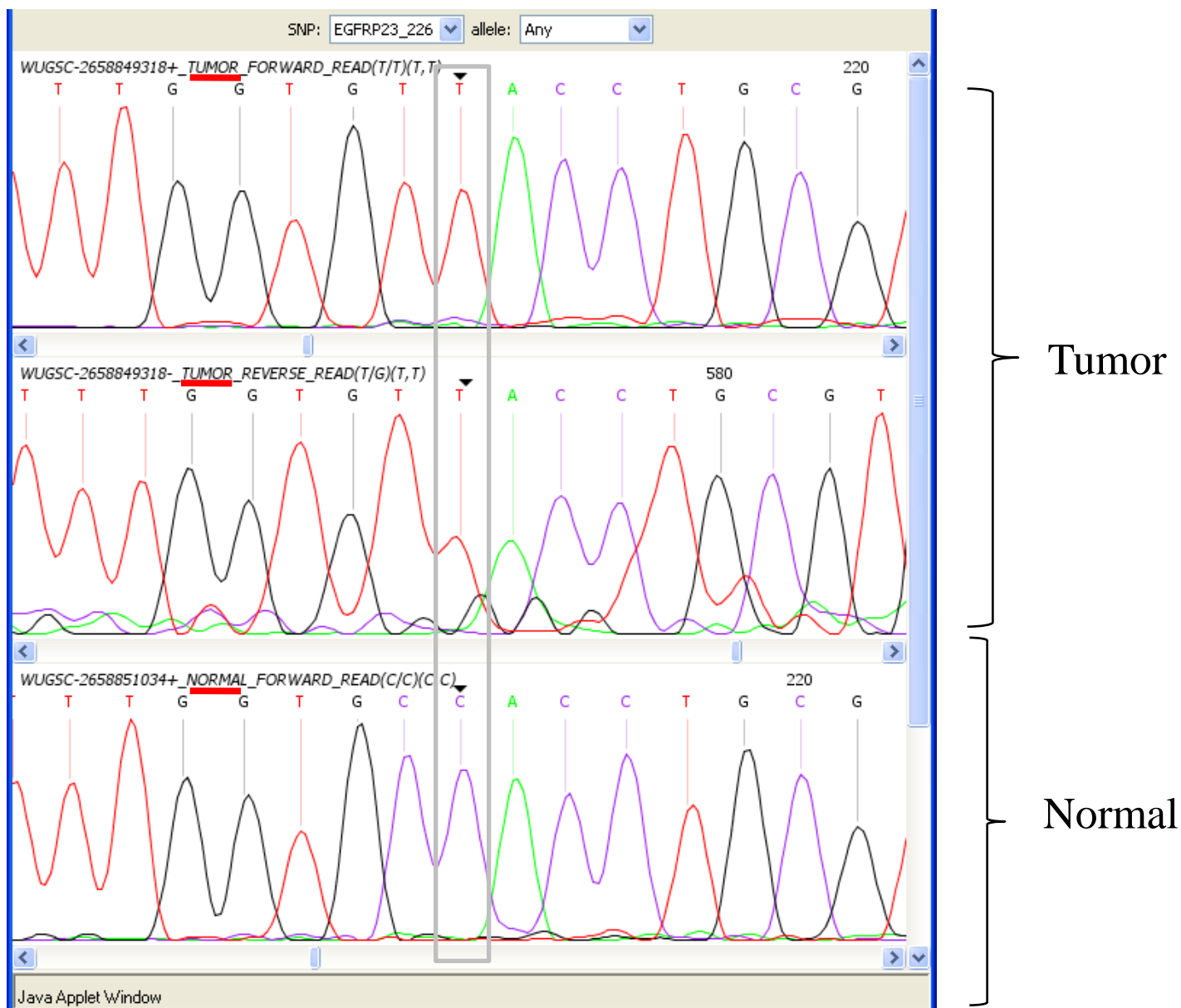
Done cgwb-test.nci.nih.gov:8443

“Full” View Shows Paired Tumor/Normal Genotype for Each Sample

This view takes a while to generate for unvalidated sequence variations and it is meaningless unless you have applied for access control for the protected data to the real sample name.



Trace View centered around a putative mutation with paired tumor/normal sample



Protein Structure Viewer

The screenshot displays the Protein Structure 3D viewer interface for the EGFR gene. The browser window title is "Protein Structure 3D viewer - Mozilla Firefox" and the URL is "https://cgwb.nci.nih.gov/cgi-bin/3dViewer/Gene.cgi?proj=valid&sym=EGFR". The page header includes the National Cancer Institute logo and the text "Protein Structure Viewer". Below the header, the gene information is displayed: "Gene : EGFR GI : 29725609 1210 aa Also see GI : 41327736 41327732 41327734". A callout box points to the "Also see GI" numbers with the text "Access proteins from alternative splicing". The main content area shows a protein domain diagram with a cyan bar representing the 3D structure viewing region (residues 25 to 638). Below the diagram, a list of alternative splicing variants is shown, each with a unique amino acid sequence. A callout box points to the first variant with the text "Click to get 3D viewer". The status bar at the bottom shows "Done" and the URL "cgwb.nci.nih.gov".

Protein Structure 3D viewer - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://cgwb.nci.nih.gov/cgi-bin/3dViewer/Gene.cgi?proj=valid&sym=EGFR

Human chr7:55,053,218-55,243,524 - ... Protein Structure 3D viewer

National Cancer Institute
Protein Structure Viewer
National Institutes of Health | www.cancer.gov

Gene : EGFR GI : 29725609 1210 aa Also see GI : 41327736 41327732 41327734

Access proteins from alternative splicing

Protein Motifs : 4 Solid : Pfam domain Pdb hits : 1 cyan 3D Structure Viewing region(s) : 25 to 638

SNP LogE & SIFT: Red : Predicted Deleterious Blue : Predicted Tolerant Black : Undecided Gray : Not Analyzed

Phosphorylate : 29 Pubmed Tabular format

Click to get 3D viewer

Done cgwb.nci.nih.gov

3D Structure Viewer

Protein Structure 3D Viewer - Mozilla Firefox

https://cgwb.nci.nih.gov/cgi-bin/3dviewer/ViewAA.cgi?proj=valid&gi=29725609&id=722&pdb=1yy9:A&sim=0.995&gstart=25&glen=613&phos=

National Cancer Institute U.S. National Institutes of Health | www.cancer.gov

GI: 29725609 Viewing region 25 - 638 99.5 % similar to Pdb: 1yy9 chain A 2 - 614

1yy9

More help

Help

Big, highlighted atoms refer to the mutated amino acids (shown in red in the bottom panel). You can also click on the mutated amino acid (shown in red) to turn on or off a specific mutation

Jmol

Sequence : Click on a letter once will turn on the spacefill. Double-click on the same letter will turn the spacefill off.

	end
EEKVKVCGGTSNKLTLQGLTFEDHFLSLQRMFNNEVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLIALNT	71
72 VERIPLLENLQIIRGNMYEENSALAVLSNYDANKTGLKELPMRNLQEIHLGAVRFSNNPALCNVESIQWRD	142
143 IVSSDFLSNMSMDFTQNHLSGSCQKCDPSPNGSCWGWAGEENCQKLTKIICAQQCSGRCRGKSPSDCCHNQCA	213
214 AGCTGPRESDCLVCRKFRDEATCKDTCPPLMLNYPTTYQMDVNPPEGKYSFGATCVKCKPRNYVVTDHGSCV	284
285 RACGADSYEMEEDGVRKCKKCEGFCRKCNGIGIGEFKDSLINATNIKHFKNCTSISSGDLHILPVAFRGD	355
356 SFHTPPPLDPPQELDILKTVKEITGFLLIQAWPENRTDLHAFENLEIIRGRTKQHGGQFSLAVVSLNITSLGL	426
427 RSLKEISDGDVVISGNKNLCYANTINWKKLFGTSGQKTKIISNRGENKCKATGQVCHALCSPGECWGPPEPR	497
498 DCVSCRNVSRGRECVDKCKLLEGEPREFVENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCV	568
569 KTCPAGVMGENNTLVWQKYADAGHVCHLCHPNCTYGCTGPGRLRGCT	614

Jmol script completed

cgwb.nci.nih.gov

View Somatic Mutations from Multiple Projects with the “Combine Tracks” Feature

The screenshot displays the Cancer Genome Workbench interface with several tracks grouped by project:

- Cancer Genome Workbench - JenLab:** Includes tracks for COSMIC (stomach, testis, thyroid, urinary tract, endometrium) and JenLab (Primer, JenLab).
- Cancer Genome Workbench - Johns_Hopkins:** Includes tracks for JHU (GBM, pancreas).
- Cancer Genome Workbench - GSK_Cell_Lines:** Includes tracks for GSK (Bladder, Blood, Bone, Brain, Breast, CNS, Cervix, Colon, Connective Tissue, Esophagus, Eye, Kidney, Liver, Lung, Muscle, Ovary, Pancreas, Pharynx, Placenta, Prostate, Rectum, Sarcoma, Skin, Stomach, Synovium, Thyroid, Uterus, Vulva).
- other:** Includes tracks for BCAC (Primer, BCAC), OCAC (Primer, OCAC), and Rembrandt.

Each track has a "hide" button. A red text annotation reads: ***=CGWB Track has items in current range**. A callout box points to the "Combine Tracks" button with the text: **Click to set up option for combining data from different projects. Then make “Combined” Track visible**. A "refresh" button is present for each track group.

Scroll to bottom

Step 1) Select Data from Different Projects/Groups

Combine Tracks - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://cgwb-test.nci.nih.gov:8443/cgi-bin/cmbntrks?org=Human&db=hg18&position=chr7:550542

<input type="checkbox"/> COSMIC_testis Catalogue Of Somatic Mutations In Cancer COSMIC_testis	<input type="checkbox"/> Catalogue Of Somatic Mutations In Cancer COSMIC_thyroid	<input type="checkbox"/> COSMIC_urinary_tract Catalogue Of Somatic Mutations In Cancer COSMIC_urinary_tract	* <input type="checkbox"/> JenLab JenLab Lung Cancer Genes
* <input type="checkbox"/> JHU_GBM Validated Somatic Mutations of Glioblastoma Multiforme by Johns Hopkins University	<input type="checkbox"/> JHU_pancreas Validated Somatic Mutations of Pancreatic Cancer by Johns Hopkins University	<input type="checkbox"/> Ovarian OCAC OCAC Familial Ovarian Cancer Project	<input type="checkbox"/> Target_ALL Validated Mutations of NCI TARGET project of Childhood Acute Lymphoblastic Leukaemia
* <input checked="" type="checkbox"/> TCGABCM_GBM The Cancer Genome Atlas Project by Baylor College of Medicine	* <input type="checkbox"/> TCGABCM_Ovarian The Cancer Genome Atlas Project (Ovarian Cancer) by Baylor College of Medicine	* <input checked="" type="checkbox"/> TCGAGSC_GBM_Valid Validated Somatic Mutations of The Cancer Genome Atlas Project Submitted by Genome Sequencing Center	* <input checked="" type="checkbox"/> TCGAWIBR_GBM The Cancer Genome Atlas Project by Broad Institute
* <input type="checkbox"/> TCGAWIBR_Ovarian The Cancer Genome Atlas Project by Broad Institute	* <input checked="" type="checkbox"/> TCGAWUGSC_GBM The Cancer Genome Atlas Project by Washington University	* <input type="checkbox"/> TCGAWUGSC_Ovarian The Cancer (Ovarian) Genome Atlas Project by Washington University	* <input type="checkbox"/> TSPBCM_Lung The Tumor Sequencing Project by Baylor College of Medicine
* <input type="checkbox"/> TSPGSC_Lung_Valid Validated Somatic Mutations of The Tumor Sequencing Project Submitted by Genome Sequencing Centers			

Submit Set All Clear All

* = Project has SNPs in the current range : [chr7:55054217-55242524](#)

Done cgwb-test.nci.nih.gov:8443

Selected the 3-TCGA projects that have putative mutations and 1 validated TCGA project Hit Submit to set up the projects and you will be directed back to the control page.

Step 2: Set up the display option of Combined

You need to go the middle of control panel to find the Combined display option. This is ugly

The screenshot shows the Human Genome Workbench interface with several track configuration panels. The 'Integrated Tracks' panel is highlighted, showing a list of tracks with their respective display options. A callout box points to the 'Combined' link, and another callout box points to the 'dense' option in the dropdown menu. A third callout box points to the 'refresh' button.

Comparative Genomics

- Conservation: squish
- 17-Way Cons: hide
- 17-Way Most Cons: hide
- Cons Indels MmCf: hide

Variation and Repeats

- SNPs (129): hide
- SNP Arrays: hide
- HapMap SNPs: hide
- HapMap LD Phased: hide
- DGV Struct Var: hide
- Segmental Dups: hide
- RepeatMasker: hide
- Interrupted Rpts: hide
- Simple Repeats: hide
- Self Chain: hide

Cancer Genome Workbench - Integrated Tracks

- motifs: dense
- Combined: hide (dropdown menu open showing: hide, dense, squish, pack, full)
- TCGABCM GBMPrimer: *hide
- TCGAWUGSC GBM: *hide
- TCGAGSC GBM Valid: *dense
- TCGA GBM Significant Region: hide
- TCGABCM OvarianPrimer: *hide
- TCGABCM Ovarian: *hide
- TCGA GBM Integrated: full
- TCGAWIBR OvarianPrimer: *hide
- TCGAWIBR Ovarian: *hide
- TCGAWUGSC OvarianPrimer: *hide
- TCGAWUGSC Ovarian: *hide
- TCGA Ovarian Integrated: *hide

Callouts:

- Hit the top link to select variation of interest (points to 'Combined')
- Select the "Dense" mode and then click refresh button. (points to 'dense' in dropdown and 'refresh' button)

Use the “Dense” Mode for Combined Track and Select Only the non-Silent Somatic Mutations

Cancer Genome Workbench (CGWB) - View Cancer Genomic Data



[CGWB Home](#) [Genome Browser](#) [Tutorial](#) [Help](#) [Log In](#)

Combined Track Settings

Combined CGWB Track

Display mode:

LPG CGWB "SNPMapping" Type (for Combined Tracks)

Any type of data can be excluded from view by deselecting the checkbox below.

SNP Type:

intergenic UTR_5 missense silent intron UTR_3 frameshift splice exon nonsense proteinDel proteinIns known Germline Somatic DiseaseState

[View table schema](#)

Combined view for all four TCGA GBM projects

Cancer Genome Workbench - Produced by Laboratory of Population Genetics / National Cancer Institute / Bethesda Maryland - based on UCSC Genome Browser

CGWB Home CGWB Help Variation Table Genotype Table CGWB Log In UCSC Link Tables Gene Sorter DNA PDF/IPS Help

UCSC Genome Browser on Human Mar 2006 Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search chr7:55,053,217-55,243,524 jump clear size 190,308 bp. configure

chr7 (p11.2) 21.8 18 21.11 22.1 31.1 q35q34q35

chr7: 55100000 55150000 55200000 RefSeq Gene

RefSeq Genes Human mRNAs Mammal Cons motifs HEAT MAP COMBO 50% 40% 30% 20% 10% 0% Frequency 10% 0% Help Button SNP INFO (mouseover and click)

Number indicates how many groups found this variation

Click to view this novel mutation only found by WUGSC

found in TCGAGSC_GBM_Valid click for more info

Each colored bar represents one center Mouse over to get the center name

Click gray/blue bars on

move start < 2.0 > move end < 2.0 >


default tracks hide all add custom tracks configure refresh

Use drop-down controls below and press refresh to alter tracks displayed.

Done cgwb-test.nci.nih.gov:8443

Detailed Mutation Report Generated by Clicking on Variations at the Genome/Gene View

National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov

Laboratory of Population Genetics 

Project: TCGABCM_GBM/TCGAWUGSC_GBM/TCGAGSC_GBM_Valid/TCGAWIBR_GBM

EGFRP23_233: A novel Somatic substitution variation that is missense in [EGFR](#)

Variation Summary

HELP

Alleles

Genotype Summary









G T GG 310 GT 1

Fasta Sequence

```

GTGCTGGCT TCTCCGTGT GTGGCGCTGA GTGACTTAC CTCACCTGCC
CAGCGTGTCC TCTCTCCTCC ATAGGCTCTGC CGCAAATTCC GAGACGAAGC
CACGTGCAAG GACACCTGCC CCCCACTCAT GCTCTACAAC CCCACCACGT
ACCAATGGA TGTGAACCCG GAGGGCAAAT ACAGCTTTGG TGCCACCTGC
G/T
TGAGAAGTG TCCCGGTGAG TCTCTCTCTG TGGGCGCTCT AACTGGTCAG
GCATCCTTG CCCGCTCTGT CTCCTGCTGA GCGCTGGAGT ATCCCATCTT
GGAGAGTCTT TGGGTGGATG TGTTCCTT GCTTGGAGGA GCGCACCTTG
TGCCCGTCCA GGCACACAGG CGAGGGGAGG GCGTGGCTTG CTACCGAGGA
                    
```

Position and Function

Genomic		Transcript			Protein			Structure			Functional Change		
Chr	Pos	RefSeq	Pos	Strand	GI	AA Change	Flanking	MMDB	Pos	Cn3D	Jmol	LogE ?	Sift ?
chr7	55189324	NM_005228	1120	+	29725609	V(GTG)292L(TTG)	FGATC[V/L]KKCPR	20809	268			1.6706	↓ 0.0166 ↓
chr7	55189324	NM_201284	1120	+	41327736	V(GTG)292L(TTG)	FGATC[V/L]KKCPR	20809	268			1.6706	↓ 0.0024 ↓
chr7	55189324	NM_201282	1120	+	41327732	V(GTG)292L(TTG)	FGATC[V/L]KKCPR	20809	268			1.6706	↓ 0.0106 ↓
chr7	55189324	NM_201283	1120	+	41327734	V(GTG)292L(TTG)	FGATC[V/L]KKCPR	20809	268			1.6706	↓ 0.0023 ↓

Motifs/Domains/Sites

MOTIF: [CDD:85007](#) Furin-like Furin-like cysteine rich region; pfam00757

deleterious

You need to install NCBI's Cn3D to make this work.

Jmol structure view requires Java.

Scroll down to view individual genotype



Done

cgwb-test.nci.nih.gov:8443 

Jmol View of the Selected Mutation: no plug-in, but does require Java.

Protein Structure 3D

File Edit View History

Human chr7:55,053,215-55,243,524 - ... Protein Structure 3D Viewer

National Cancer Institute U.S. National Institutes of Health | www.cancer.gov

GI: 29725609 Viewing region 25 - 638 9900 % similar to Pdb

1yy9

Jmol View: Selected Mutation shown in enhanced atoms

Jmol

start	Sequence : Click on a letter once will turn on the spacefill. Double-click on the same letter will turn the spacefill off.	end
2	EEKKVCGGTSNKLTLGLTFEDHFLSLQRMFNNEVVLGNLEITYVORNYDLSFLKTIQEVAGYVLIALNT	71
72	VERIPLLENLQIRGNMYEENSALAVLSNYDANKTGLKELPMRNLQEIILHGAVRFSNNPALCNVESIQWRD	142
143	IVSSDFLSNMSMDFQNHLSGCOCKDPSCPNGSCWGAEEENCKLTKIICAQQCSGRCRGKSPSDCCCHNQCA	213
214	AGCTGPRESDCLVCRKFRDEATCKDTCPLMLLYNPPTYQMDVNP EGKYSFGATCVKKCPRNYVVTDHGSCV	284
285	RACGADSYEMEEDGVRCKKCEGPKRKCVCNGIGIGEFKDSLSINATNIKHFKNCTSSISGDLHILPVAFRGD	355
356	SFTHTPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAFENLEIIRGRTKQHGQFSLAVVSLNITSLGL	426
427	RSLKEISDGDVIISGNKNLGYANTINWKKLFGTSGQKTKIISNRGENCKKATGQVCHALCSP EGCWGP EPR	497
498	DCVSCRNVSRGRECVDKCKLLEGEPREFVENSECICQHP ECLPQAMNITCTGRGPDNCIQCAHYIDGPHCV	568
569	KTCPAGVMGENNTLVWKYADAGHVHCLCHPNCTYGCCTGPGLRGCP T	614

National Cancer Institute FIRSTGOV

Jmol script completed cgwb.nci.nih.gov

How to Highlight Mutation on Cn3D View?

1. Copy the flanking

Protein	Structure	Functional Change	
Change	Flanking	MMDB Pos Cn3D LogE Sift	
Y(TAC)	HVCHL[C/Y]HPNCT	20809 596	unknown 0
Y(TAC)	HVCHL[C/Y]HPNCT	20809 596	unknown 0
Y(TAC)	HVCHL[C/Y]HPNCT	20809 596	unknown 0

2. Hit Cn3D

3. Go to View and Select "Find Pattern"

4. Paste in your flanking. Add the 6th mutation aa. Then look for highlight

Input

Enter a pattern using ProSite syntax:

HVCHL

OK Cancel

Slide 20 of 22

Default Design

start Re... 6 R... 2 M... Co... tcg... 2 W... Mic... FO... unt... 2 F... 1V... 2:50 PM

Primers and Genotypes: Select Samples for Trace View

Primers

Name	Forward	Reverse
EGFRP27	TAACTTGGGCTTTCTGACGGG	CACATCCACCCAAAGACTCTCC
EGFRP7	AGCTGGGTTTTCCACACTA	GACAACCTGAATACCCACGGC
EGFRP5	ACTTGGGCTTTCTGACGGGA	GCAAGGCAAACACATCCACC
EGFRP25	TAATTTCAATCCATAGTCACCGC	CTCGGTAGCAAGCCAGCC
EGFRP23	GTAACCTGGGCTTTCTGACGGG	CACATCCACCCAAAGACTCTCC

Genotypes

(Select samples and then click on the "Submit" button)

Hit for trace view
You can select as many samples you like

[Traceviewer Help File.](#)

Sample Name	Disease	Normal
WUGSC-2377692378	<input checked="" type="checkbox"/> WUGSC-2377692378 (T/G) [quality:medium]	
WUGSC-2051260781		<input checked="" type="checkbox"/> WUGSC-2051260781 (G/G) [quality:medium]
WUGSC-2051260789		<input type="checkbox"/> WUGSC-2051260789 (G/G) [quality:high]
WUGSC-2051260911	<input type="checkbox"/> WUGSC-2051260911 (G/G) [quality:medium]	
WUGSC-2051260942		<input type="checkbox"/> WUGSC-2051260942 (G/G) [quality:medium]
WUGSC-2051261056		<input type="checkbox"/> WUGSC-2051261056 (G/G) [quality:medium]
WUGSC-2051261068	<input type="checkbox"/> WUGSC-2051261068 (G/G) [quality:medium]	
WUGSC-2051261072	<input type="checkbox"/> WUGSC-2051261072 (G/G) [quality:medium]	
WUGSC-2051261081		<input type="checkbox"/> WUGSC-2051261081 (G/G) [quality:medium]
WUGSC-2051261086	<input type="checkbox"/> WUGSC-2051261086 (G/G) [quality:medium]	
WUGSC-2051261351		<input type="checkbox"/> WUGSC-2051261351 (G/G) [quality:medium]

Mutations in Published Literature are Linked

Getting Sta...

Human chr7:55,110,706-55,299,014 - ...

SNP Information Page - A289D_2...



Project: TCGABCM_GBM/TCGAWIBR_GBM/TCGAWUGSC_GBM/TCGAGSC_GBM_Valid

A289D_21685: A Somatic substitution variation that is missense in [EGFR](#)

This variation also matches [COSMIC:A289D](#) [COSMIC:A289V](#)

HELP

Variation Summary

Alleles	Genotype Summary		
C T	CC 598	CT 8	TT 5

This is a known mutation deposited in Sanger Center's COSMIC database. Link to COSMIC

Fasta Sequence

```
TCAACACCGT GCTGCGCTTC CTCCTGTGT GCGCGTGAAT GTACTTACCT
CACTTGCCCA GCGTGTCTC TCTCCTCCAT AGGCTGCGCG CAAATTCCGA
GACGAAGCCA CGTGCAAGGA CACCTGCCCC CCACTCATGC TCTACAACCC
CACCCAGTAC CAGATGGATG TGAACCCCGA GGGCAAAATAC AGCTTTGGTG
C/T
CACCTGCGTG AAGAAGTGTG CCCGTGAGTC CTCCTGTGTG GGCCTCTAA
CTGGTCAGGC ATCCTTGTCC CGCTCTGTCT CCGCTGAGC CCTGGAGTAT
CCCATCTTGG AGAGTCTTTG GGTGGATGTG TTTGCCTTGC TTGGAGGAGG
CGACCCCTGT CCCGTCCAGG CACACAGGGC AGGGGAGGGG CTGGCTTGTCT
```

Position and Function

Genomic		Transcript			Protein			Structure			Functional Change	
Chr	Pos	RefSeq	Pos	Strand	GI	AA Change	Flanking	MMDB	Pos	Cn3D	LogE	Sift
chr7	55189316	NM_201283	1112	+	41327734	A(GCC)289V(GTC)	KYSFG[A/V]TCVKK	20809	265		1.3328	0.0011

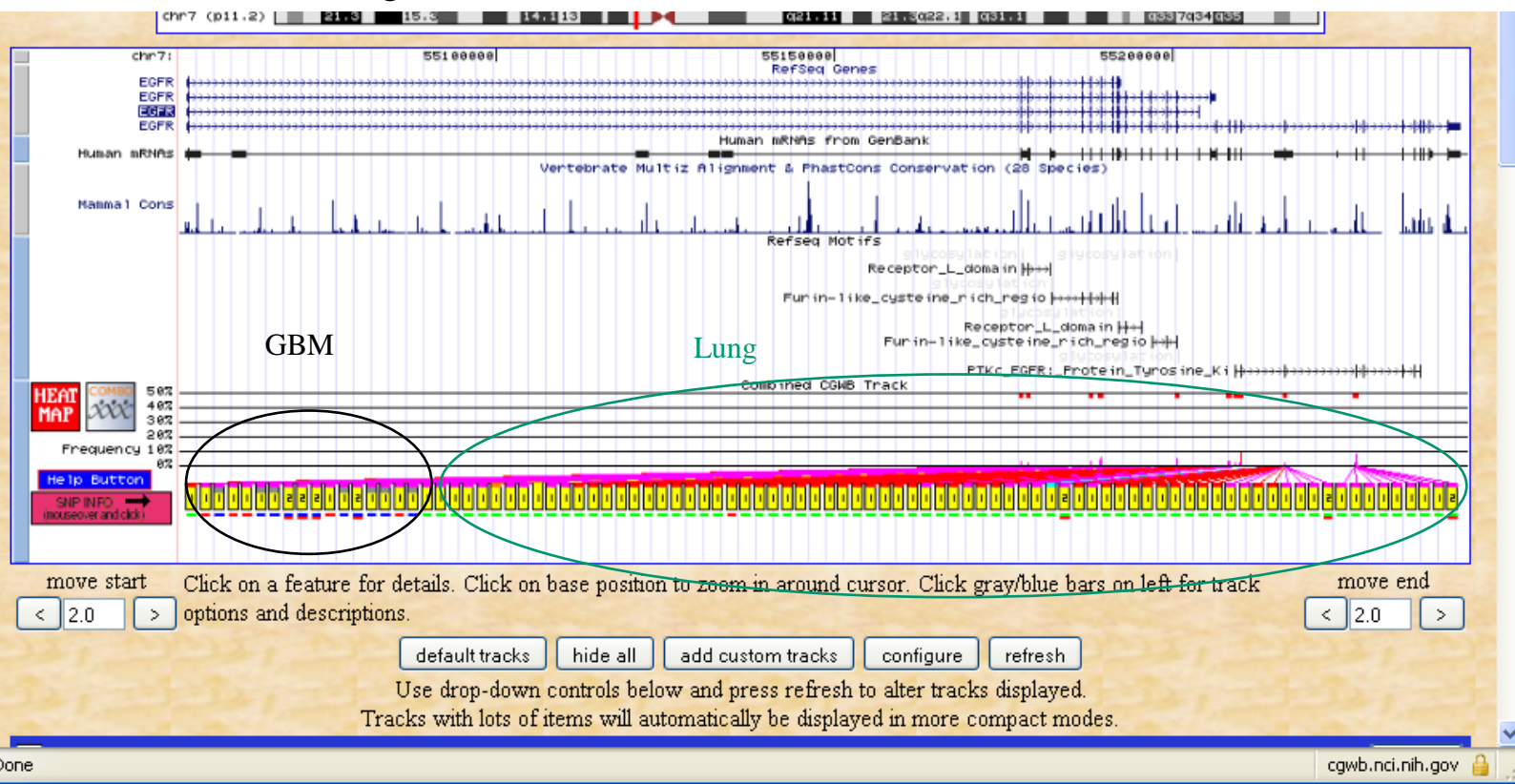
Combined View of EGFR mutation in Lung versus in GBM

COSMIC_Lung in green underline

TCGA_GBM in blue underline

COSMIC_GBM in red underline

This combined view shows that the EGFR mutations in GBM rarely map to the kinase domain while those in lung always go to the kinase domain (the first green bar is the EGFRvIII mutations found in lung)



View Integrated Data of Copy Number, Exon
Expression, Somatic Mutation, Methylation
and MicroRNA

Integrated genomic data should be on by default.
You can always scroll down to the control panel to turn on or off this option

Cancer Genome Workbench - Integrated Tracks refresh

motifs Combined
dense hide

Cancer Genome Workbench - TCGA refresh

TCGABCM GBMPrimer *hide	TCGABCM GBM *hide	TCGAWIBR GBMPrimer *hide	TCGAWIBR GBM *hide	TCGAWUGSC GBMPrimer *hide
TCGAWUGSC GBM *hide	TCGAGSC GBM Valid *dense	TCGA GBM Significant Region hide	TCGABCM OvarianPrimer *hide	TCGABCM Ovarian *hide
TCGA GBM Integrated full	TCGAWIBR OvarianPrimer *hide	TCGAWIBR Ovarian *hide	TCGAWUGSC OvarianPrimer *hide	TCGAWUGSC Ovarian *hide

Cancer Genome Workbench - TARGET refresh

Cancer Genome Workbench - TSP refresh

Cancer Genome Workbench - COSMIC refresh

Cancer Genome Workbench - JenLab refresh

Cancer Genome Workbench - Johns Hopkins refresh

Cancer Genome Workbench - GSK_Cell_Lines refresh

other refresh

*=CGWB Track has items in current range

Combine Tracks
refresh

Transferring data from cgwb-test.nci.nih.gov... cgwb-test.nci.nih.gov:8443

TCGA_GBM_Integrated Track Settings - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://cgwb-test.nci.nih.gov:8443/cgi-bin/hgTrackUI?hgsid=55931&c=chr7&g=

Cancer Genome Workbench (CGWB) - View Cancer Genomic Data

CGWB Home Genome Browser Tutorial Help Log In

TCGA_GBM_Integrated Track Settings

TCGA_GBM_Integrated

Display mode:

LPG CGWB "Genomic Data" Type

Any type of data can be excluded from view by deselecting the checkbox below.

Genomic Data:

Deletion Amplification CopyNumber Neutral ExonArray Methylation SomaticMutation Putative Mutation ExpAffyHT_HG-U133 MicroRNA_Expression ExpAgilentG4502A_07

Gene Expression measured by Affymetrix U133 platform

Gene Expression measured by Agilent Platform

Exon array and putative mutation are available via login

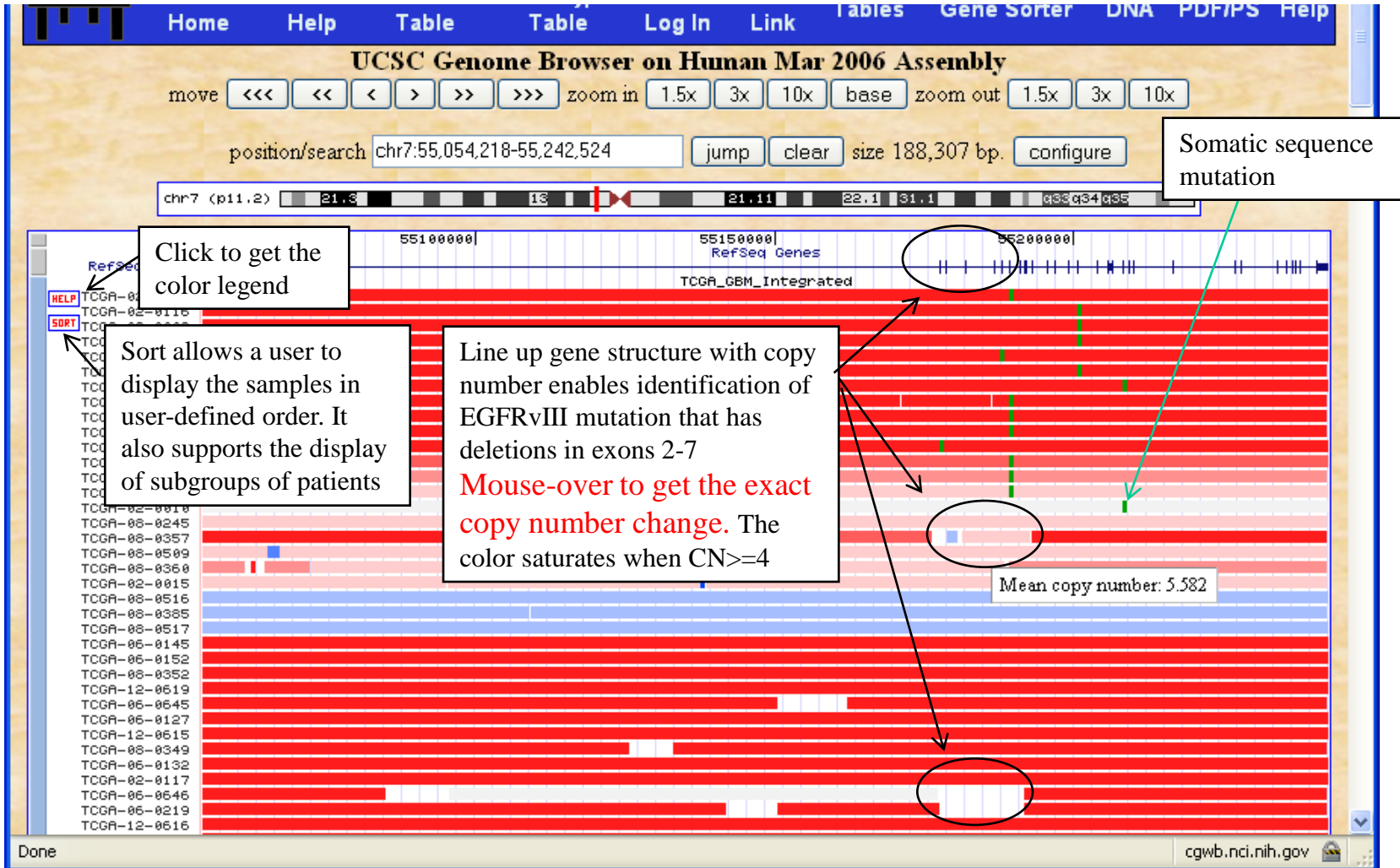
[View table schema](#)

Data last updated: 2009-04-02

Done cgwb-test.nci.nih.gov:8443

Integrated view of somatic sequence mutation and copy number change

Each line represents copy-number alteration in tumor superimposed with somatic mutations (green bars)
Mouse-over to get the exact copy number value or the amino-acid changes in each mutation



Create Customized Sample Display Order and Sample Subgroups

Custom Sample Ordering (sorting)

Paste any sample or subset of samples.
Use '>' for Custom User Labels which will appear on the track output.
Enter only one sample or label per line

```
>EGFRvIII wt-predominant
TCGA-06-0646
TCGA-08-0357
TCGA-08-0244
TCGA-08-0518
TCGA-08-0354
TCGA-08-0358
TCGA-06-0137
TCGA-02-0064
>EGFRvIII vIII-predominant|
TCGA-08-0360
```

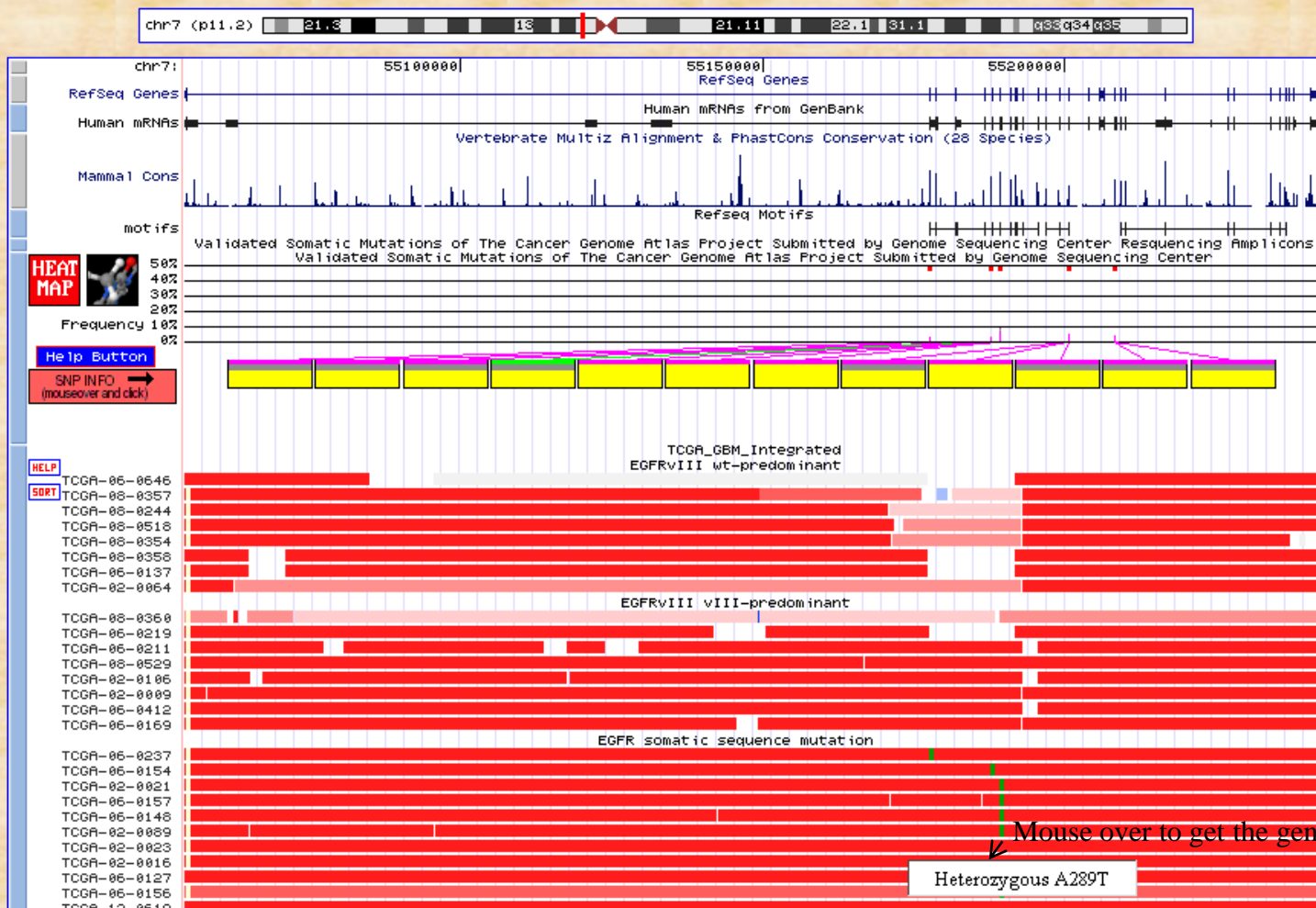
Label for subgroups

The sub-groups and the display order will be applied to all subsequent displays. Click "Clear All Custom Sorting" to revert to the default display which shows the samples that have somatic mutations at the top followed by focal deletions/focal amplifications

Submit Clear All Custom Sorting

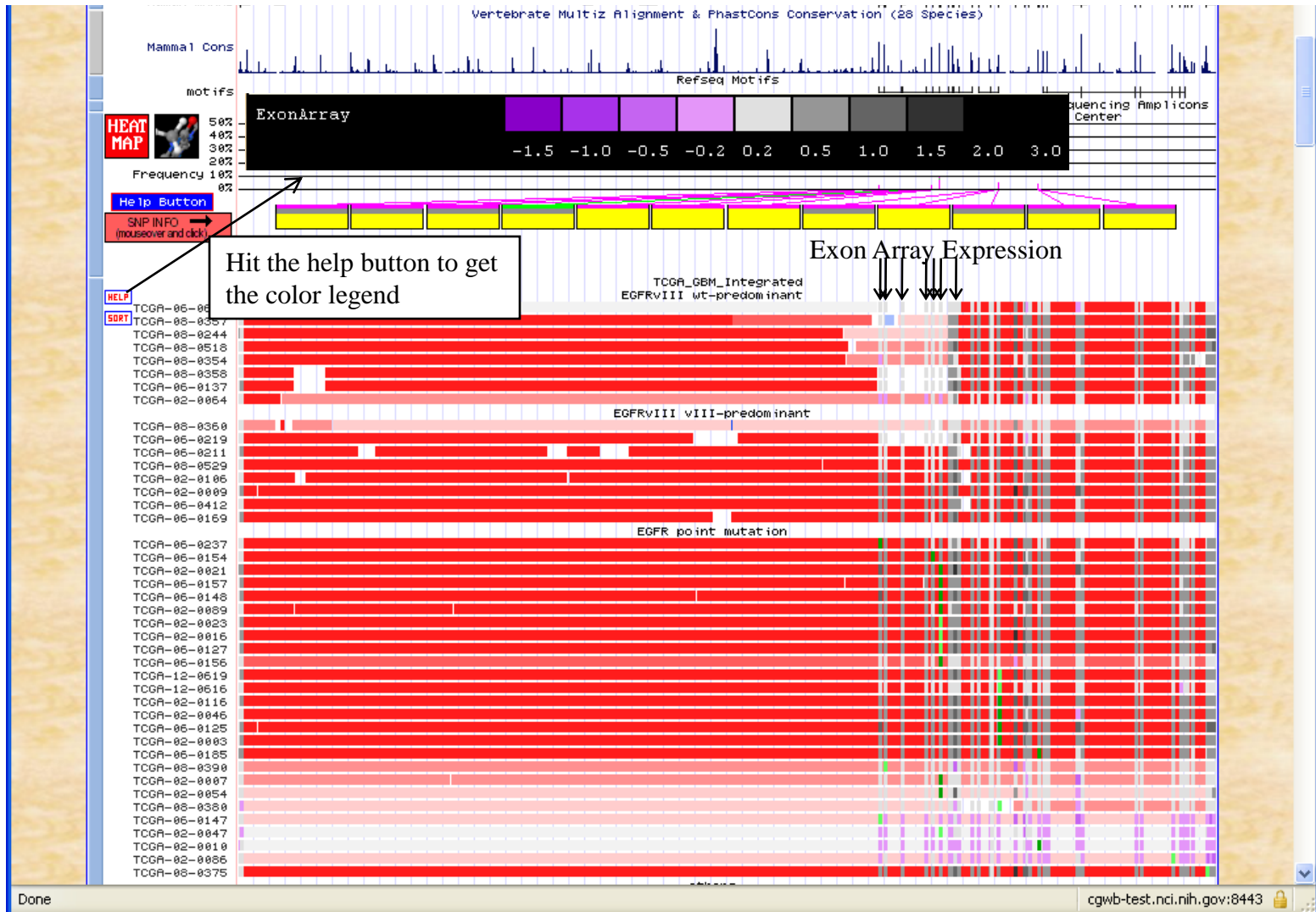
```
TCGA-02-0001
TCGA-02-0002
TCGA-02-0003
TCGA-02-0004
TCGA-02-0006
TCGA-02-0007
TCGA-02-0009
TCGA-02-0010
TCGA-02-0011
TCGA-02-0014
TCGA-02-0015
TCGA-02-0016
TCGA-02-0021
```

Integrated view of Copy Number and Somatic Sequence Mutation Displayed using the Custom-defined Order and Sample Subgroups



Mouse over to get the genotype
Heterozygous A289T

Integrated view with Exon Array and Putative Mutations Superimposed to the Copy Number/Validated Mutations (only users with login can access this view)



Integrated Data in "Pack" View (Scroll to bottom to make the change)

Top row

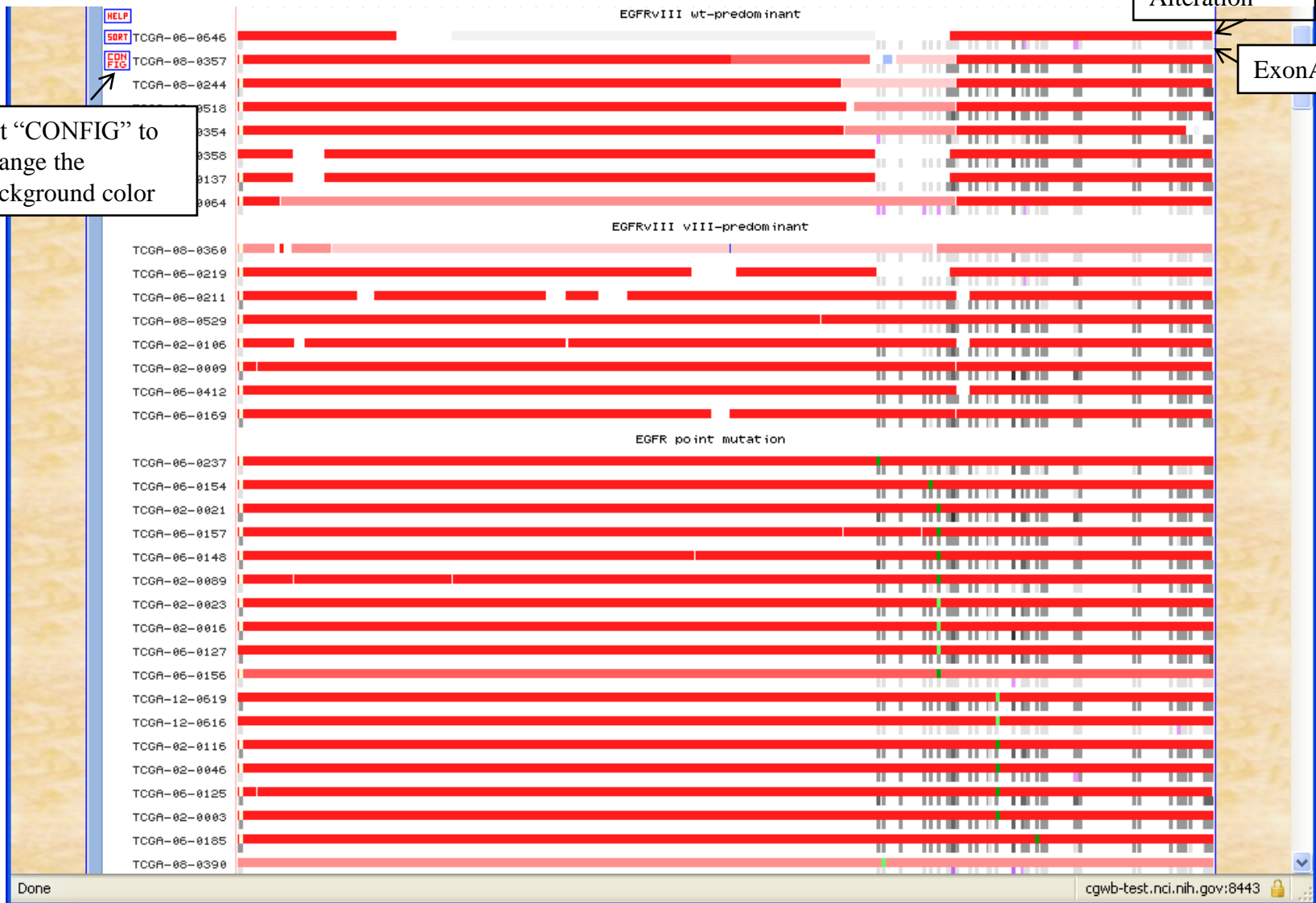
Copy Number +Somatic mutation+Methylation

Bottom row

Exon Array

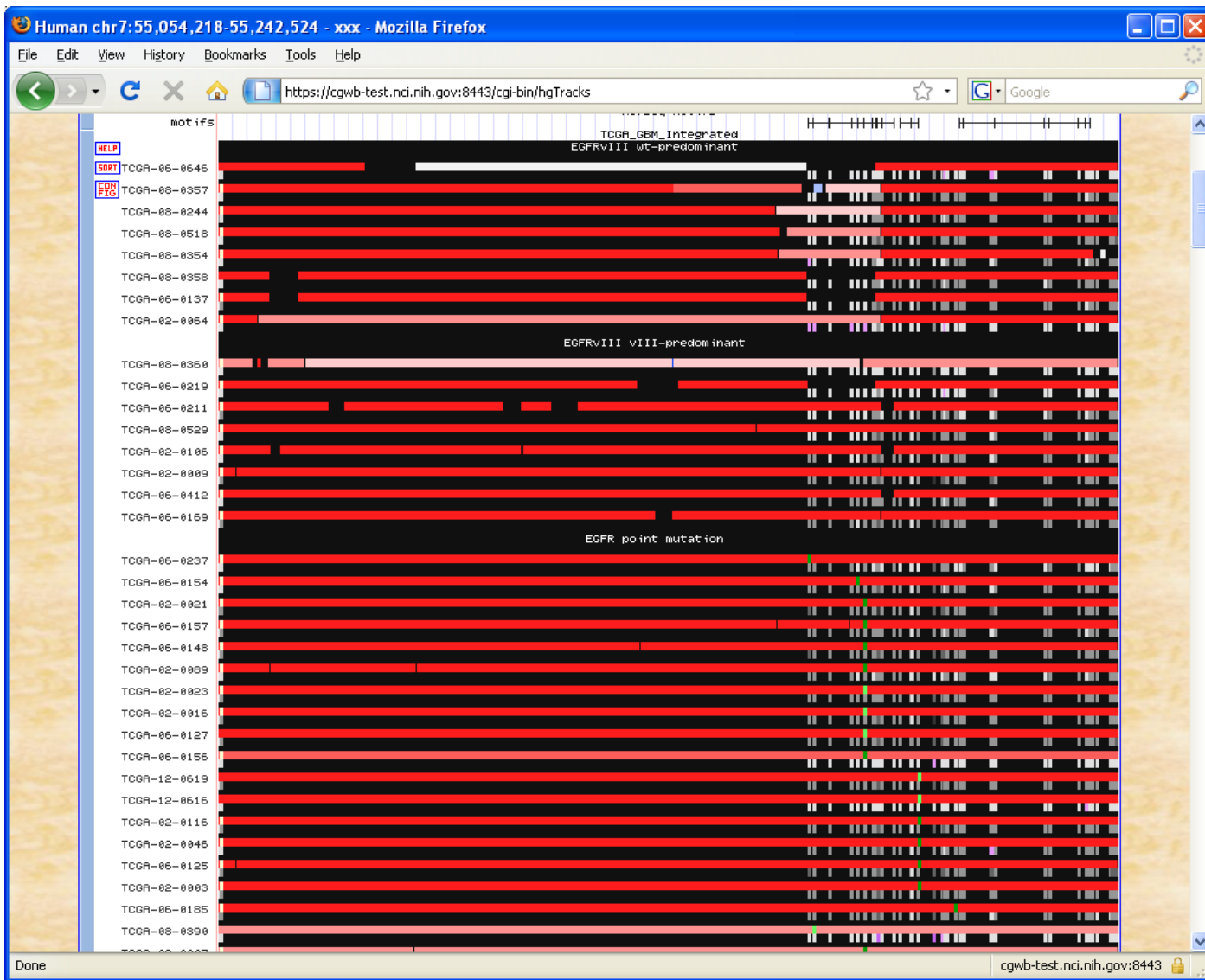
Hit "CONFIG" to change the background color

Copy Number Alteration
ExonArray

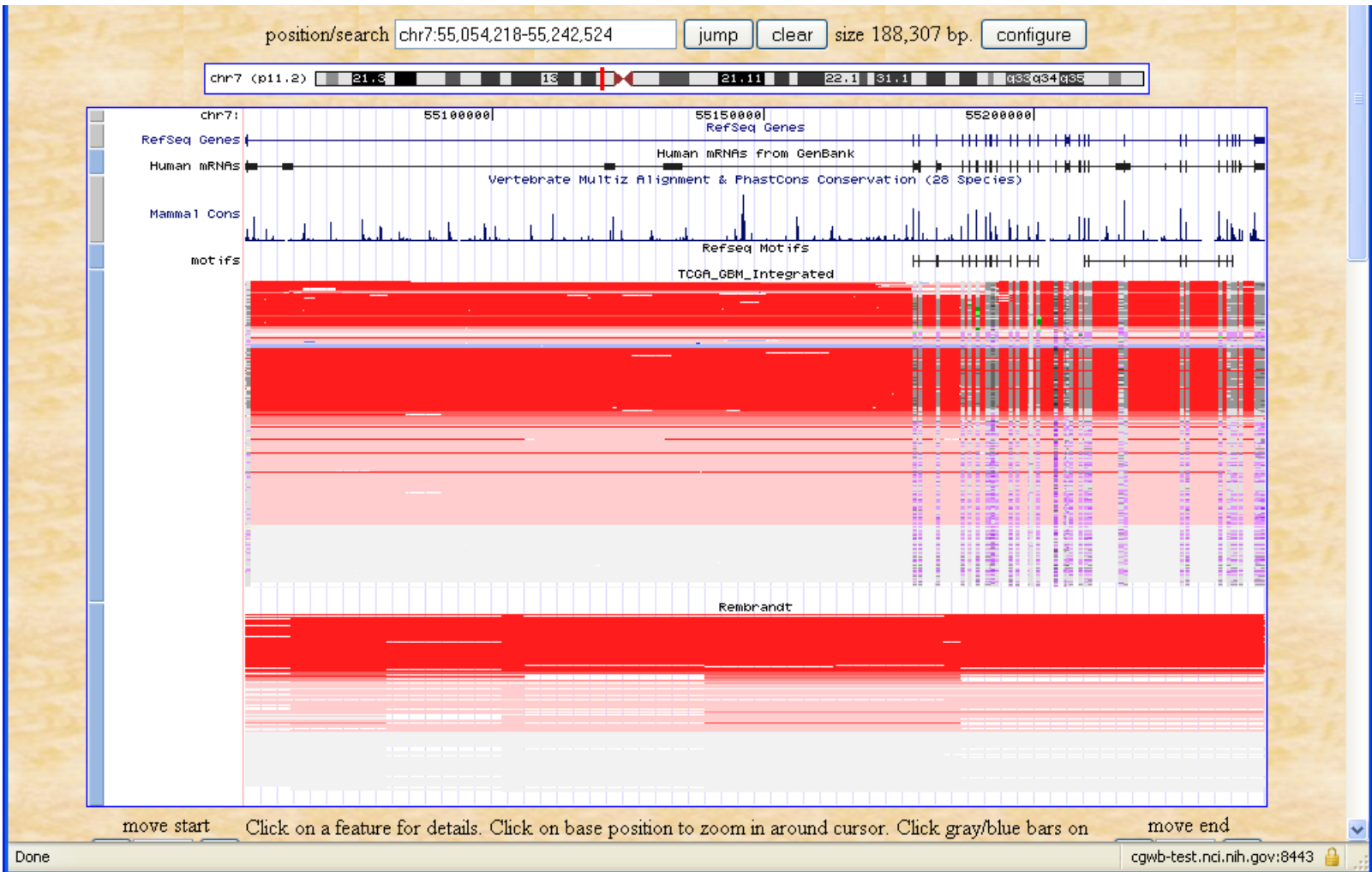


Same Region/Same Data Viewed with Black Background

Black background looks good for regions with a lot of amplification but not those with a lot of deletion



“Dense” View that Shows Both TCGA data (Top) and the Rembrandt Data (Bottom, no exon array) at EGFR locus
“Dense” View is Good For Cross-Project Comparison



View Correlation Plot of Methylation vs
Gene Expression; MicroRNA vs Gene
Expression

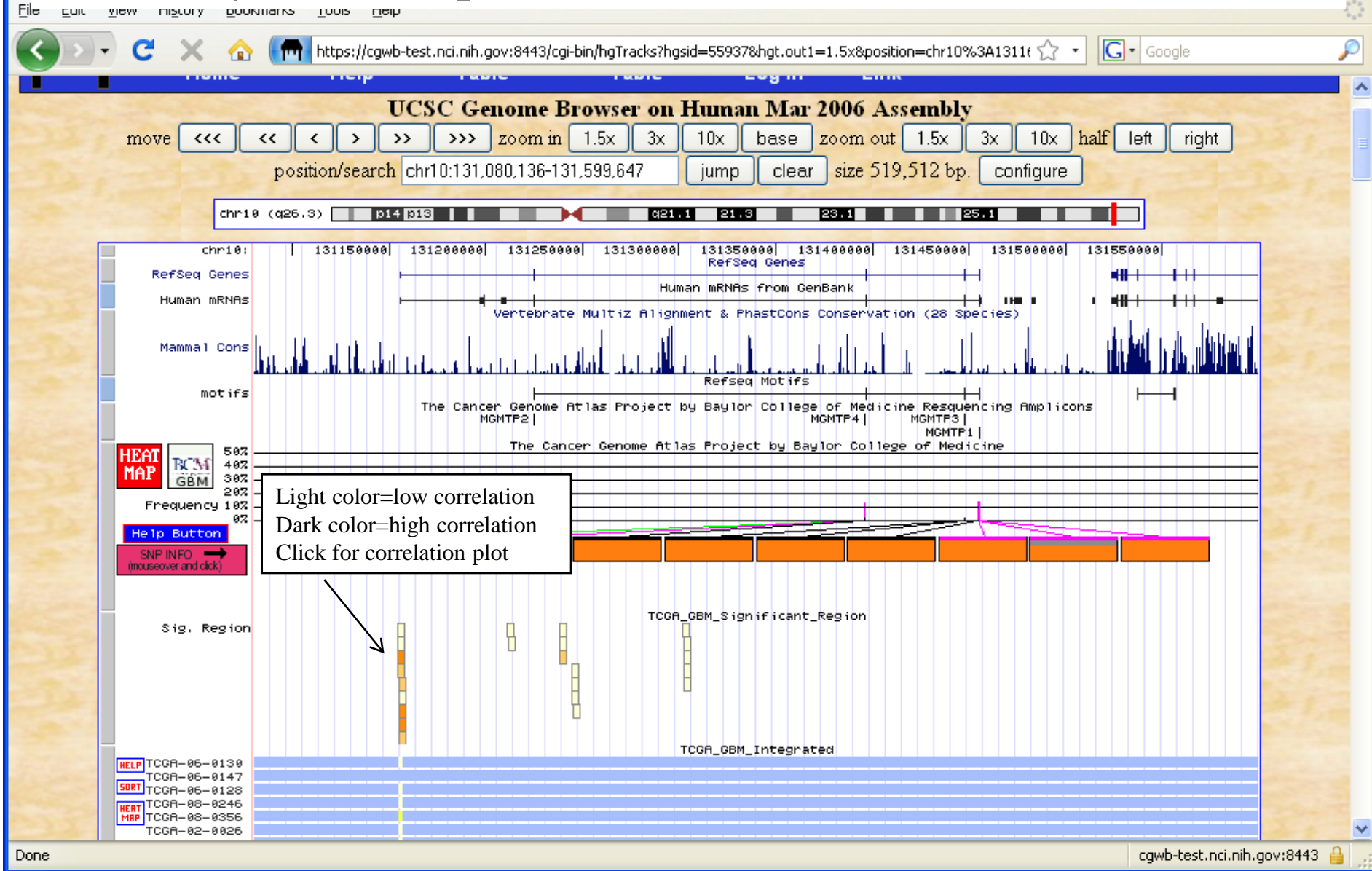
Turn on display of correlation plot at the control panel

The screenshot shows the Human Genome Workbench interface for Human chromosome 10, region 131,224,445-131,455,338. The interface is organized into several sections:

- Comparative Genomics:** Includes tracks for Conservation (squish), 17-Way Cons (hide), 17-Way Most Cons (hide), and Cons Indels MmCf (hide).
- Variation and Repeats:** Includes tracks for SNPs (129) (hide), SNP Arrays (hide), HapMap SNPs (hide), HapMap LD Phased (hide), DGV Struct Var (hide), Segmental Dups (hide), RepeatMasker (hide), Interrupted Rpts (hide), Simple Repeats (hide), and Self Chain (hide).
- Cancer Genome Workbench - Integrated Tracks:** Includes motifs (dense) and Combined (hide).
- Cancer Genome Workbench - TCGA:** This section contains multiple tracks with dropdown menus for display options:
 - TCGABCM GBMPrimer: * pack
 - TCGABCM GBM: * dense
 - TCGAWIBR GBMPrimer: hide
 - TCGAWIBR GBM: hide
 - TCGAWUGSC GBMPrimer: hide
 - TCGAWUGSC GBM: hide
 - TCGAGSC GBM Valid: hide
 - TCGA GBM Significant Region: hide (with a dropdown menu open showing options: hide, dense, squish, pack, full. An arrow points to 'pack' with a text box: "Select 'Significant' Region")
 - TCGABCM OvarianPrimer: * hide
 - TCGABCM Ovarian: * hide
 - TCGA GBM Integrated: full
 - TCGAWIBR OvarianPrimer: hide
 - TCGAWUGSC OvarianPrimer: hide
 - TCGAWUGSC Ovarian: hide
 - TCGA Ovarian Integrated: * hide
- Cancer Genome Workbench - TARGET:** Includes a refresh button.

The browser address bar shows the URL: <https://cgwb-test.nci.nih.gov:8443/cgi-bin/hgTracks?hgsid=55937&hgt.out1=1.5x&position=chr10>

Methylation/Expression Correlation at MGMT locus



Light color=low correlation
Dark color=high correlation
Click for correlation plot



HEAT MAP **TCM GBM**

50%
40%
30%
20%
10%
0%

Frequency

Help Button

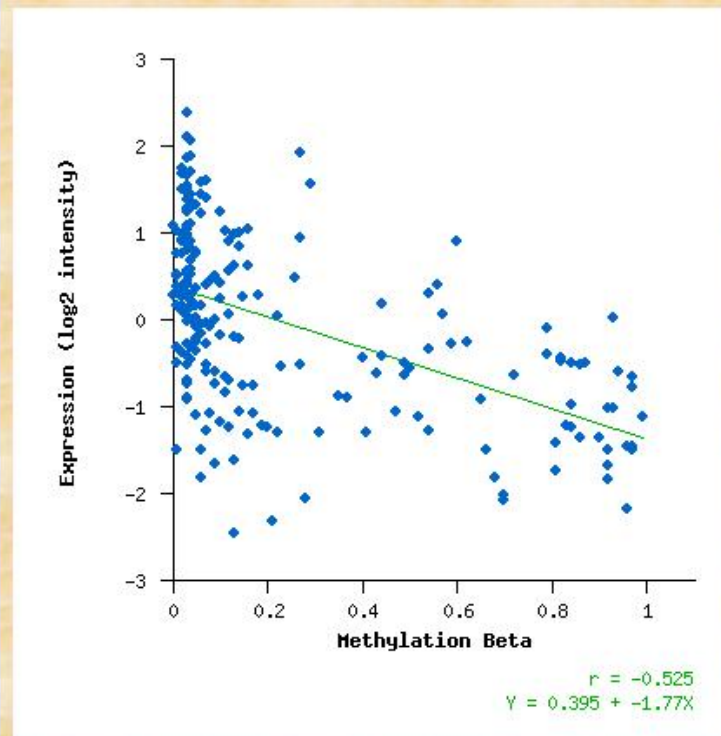
SNP INFO (mouseover and click)

HELP TCGA-06-0130
SDRT TCGA-06-0147
TCGA-06-0128
TCGA-08-0246
HEAT MAP TCGA-08-0356
TCGA-02-0026



[CGWB](#)
[Home](#)

Pearson correlation between MGMT expression levels and methylation beta value



Number of Samples is 215

Sample (X)Methylation Beta (Y)Expression (log2 intensity)

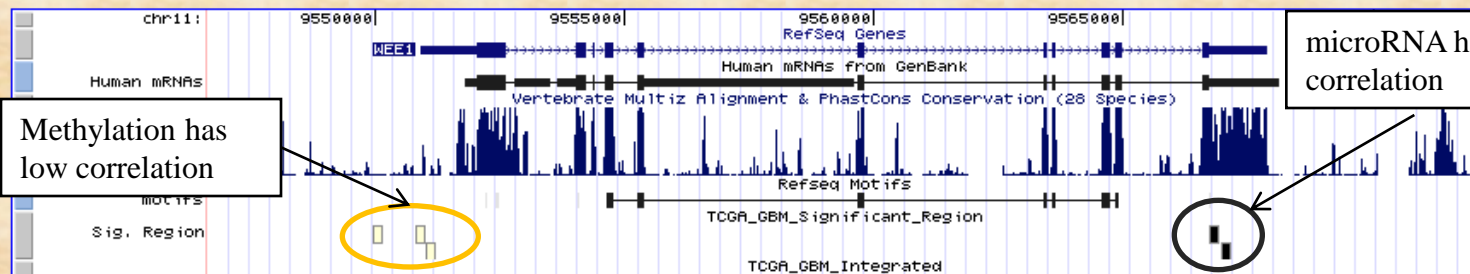
MicroRNA to Gene Expression at WEE1 locus

Cancer Genome Workbench - Produced by Laboratory of Population Genetics / National Cancer Institute / Bethesda Maryland - based on UCSC Genome Browser

CGWB Home CGWB Help Variation Table Genotype Table CGWB Log In UCSC Link Tables DNA PDF/PS Help

UCSC Genome Browser on Human Mar 2006 Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x half left right
position/search chr11:9,546,676-9,572,114 jump clear size 25,439 bp. configure



Methylation has low correlation

microRNA has high correlation

- TCGA-12-0616
- TCGA-02-0059
- TCGA-02-0048
- TCGA-06-0409
- TCGA-16-0848
- TCGA-08-0351
- TCGA-02-0086
- TCGA-06-0146
- TCGA-12-0619
- TCGA-06-0410
- TCGA-16-0846
- TCGA-02-0069
- TCGA-02-0058
- TCGA-02-0084
- TCGA-02-0047
- TCGA-06-0414
- TCGA-02-0099
- TCGA-12-0691
- TCGA-06-0128
- TCGA-02-0010
- TCGA-08-0517
- TCGA-06-0221
- TCGA-06-0184
- TCGA-14-0817
- TCGA-06-0129
- TCGA-02-0114
- TCGA-08-0245
- TCGA-06-0213
- TCGA-14-0867
- TCGA-02-0087
- TCGA-14-0871
- TCGA-16-0850
- TCGA-06-0185
- TCGA-02-0054
- TCGA-06-0354

Reduced expression of MIR-128b in tumor may have caused increased expression of WEE1

