Cancer Genome Workbench (CGWB) 4 TCGA

Views

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

Cancer Genome Workbench (CGWB) 4 TCGA

Data for the TCGA Project

1. Validated mutations from all three genome centers collected in July, 2008
2. Putative mutations computed from traces submitted by the three genome centers
3. Somatic mutations of GBM from published Literature (collected from Sanger Center)
4. Copy number data summarized from 4 institutions: Broad, MSKCC, HMS and Stanford

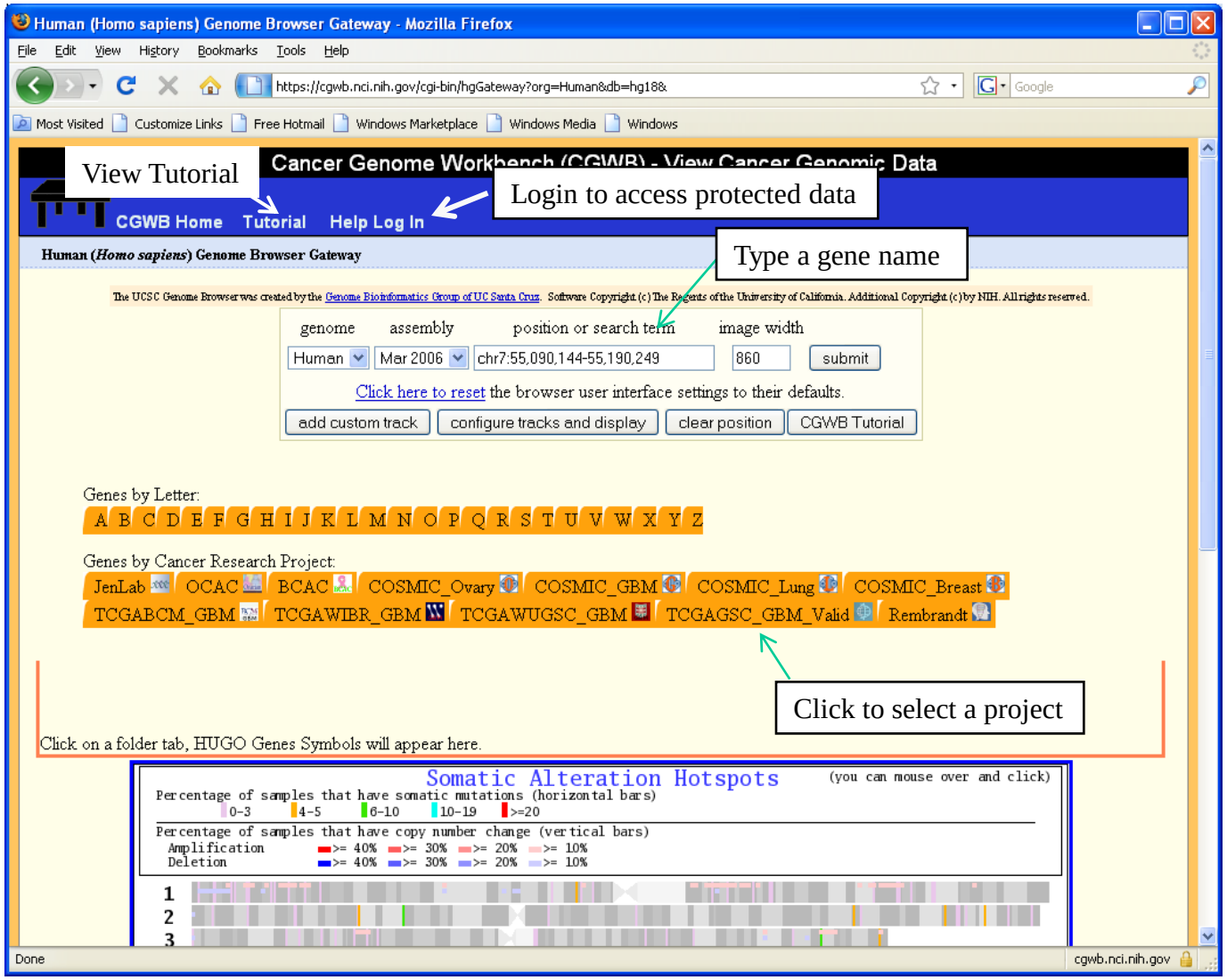
	Paired	Unpaired
Broad	120	61
MSKCC	37	0
Harvard	1	0
Stanford	1	1
Total	159	62

5. Methylation data collected in July, 2008
6. Exon array data collected in July, 2008 (visible only with a login)

Other related projects

1. Somatic mutations of GBM from published Literature (COSMIC from Sanger Center)
2. GBM Rembrandt data from Howard Fine (51 paired and 79 unpaired)
3. Tumor Sequencing Project in Lung cancer (will be available once the paper is published)

Navigate Cancer Genome Workbench



Select a Project on Cancer Genome Workbench

Human (Homo sapiens) Genome Browser Gateway - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://cgwb.nci.nih.gov/cgi-bin/hgGateway?org=Human&db=hg18&

Most Visited | Customize Links | Free Hotmail | Windows Marketplace | Windows Media | Windows

Cancer Genome Workbench (CGWB) - View Cancer Genomic Data

CGWB Home Tutorial Help Log In Heatmap

Heatmap button visible for projects that have Heatmap data

Human (*Homo sapiens*) Genome Browser Gateway

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.

genome: Human | assembly: Mar 2006 | position or search term: chr7:55,090,144-55,190,249 | image width: 860 | submit

Click here to reset the browser user interface settings to their defaults.

add custom track | configure tracks and display | clear position | CGWB Tutorial

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

Project title

Genes by Letter:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Genes by Cancer Research Project:

JenLab | OCAC | BCAC | COSMIC_Ovary | COSMIC_GBM | COSMIC_Lung | COSMIC_Breast | TCGABCM_GBM | TCGAWIBR_GBM | TCGAWUGSC_GBM | TCGAGSC_GBM_Valid | Rembrandt

A2M ABCA3 ABCC4 ADAM12 ADAM15 ADAMTSL3 ADM AIFM1 ALK ANK2 ANXA1 ANXA7 AOC3 APBB1IP ARNT ASPM ASXL1 ATM
ATP6V1E1 ATR AURKB AVIL AXL BAI1 BAI3 BAMBI BAX BCAR1 BCL11A BCL2 BCL2L13 BCR BMPR1A BRCA1 BRCA2 BTK C18orf25
C20orf160 C22orf24 CASP9 CAST CBL CD46 CDC123 CDKL5 CDKN2A CENPF CENTG1 CES3 CFTR CHAT CHD5 CHEK1 CHEK2 CHI3L2
CHIC2 CHL1 CNTFR COL11A1 COL1A1 COL1A2 COL3A1 COL6A2 CRLF1 CSF3R CSMD3 CSNK1E CTNNB1 CTSH CYLD CYP27B1
CYP3A4 DCX DDR1 DDR2 DES DGKD DHTKD1 DMBT1 DOCK1 DOCK8 DPYSL4 DST DTX3 EGFR ELAVL2 EP300 EP400 EPHA2 EPHA4
EPHA6 EPHA7 EPHA8 EPHB1 ERBB2 ERBB3 EYA1 F13A1 FBXW7 FGFR1 FGFR2 FGFR3 FLI1 FLJ10324 FLT1 FLT4 FN1 FOXO3 FOXO4
FRAP1 FURIN GARNL3 GATA3 GCLC GDF10 GLI1 GLI3 GNAI1 GNAS GPR78 GRIA2 GRLF1 GRN GSTM5 GYPC HABP2 HSP90AA1 ID3
IFITM3 IFNA21 IGF1R IL1RL1 IL31 ILK ING4 INHBE IQGAP1 IRAK3 IRS1 ISL1 ITGB2 ITGB3 JAG1 KIAA1632 KIF3B KIT KLF4 KLF6 KLK8
KPNA2 KRAS KTN1 LAMP1 LATS2 LAX1 LCK LDHA LGALS3BP LRRN2 LTF LYN MAG MAP3K6 MAPK13 MAPK7 MAPK8IP2
MAPK8IP3 MAPK9 MDM2 MDM4 MEOX2 MET MLH1 MLL4 MN1 MSH2 MSH6 MSI1 MTAP MYCN MYLK2 MYO1E MYO3A MYST4 NBN

Done | cgwb.nci.nih.gov

Click a gene here (sorted alphabetical order) or Select a mutation hot spot by scrolling down to see the "landscape view" of somatic mutations and copy number

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

Most Visited | Customize Links | Free Hotmail | Windows Marketplace | Windows Media | Windows

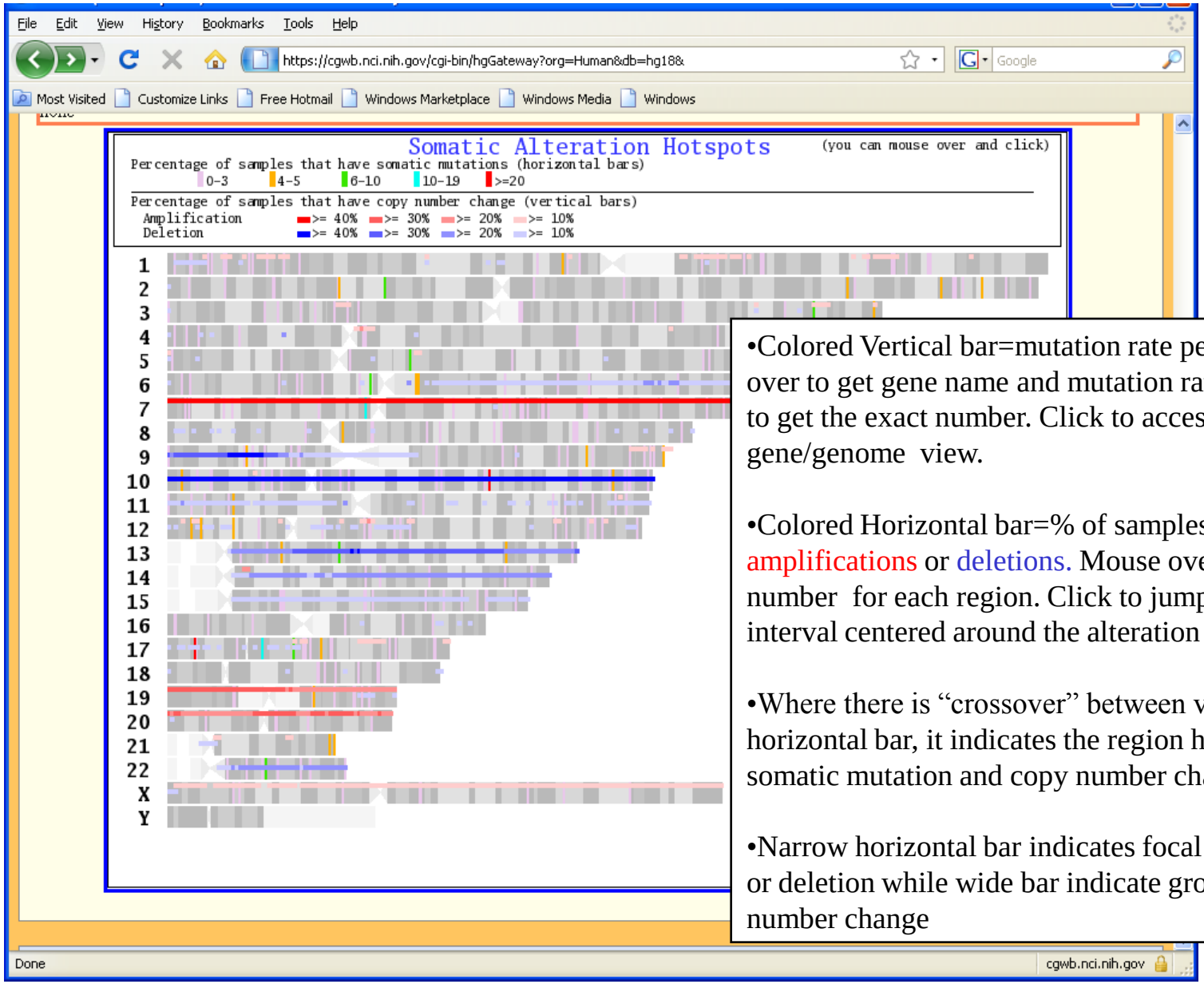
Human (Homo sapiens) Genome Brows... | CGWB heatmap viewer

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		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

Copy Number & Mutation in Landscape View



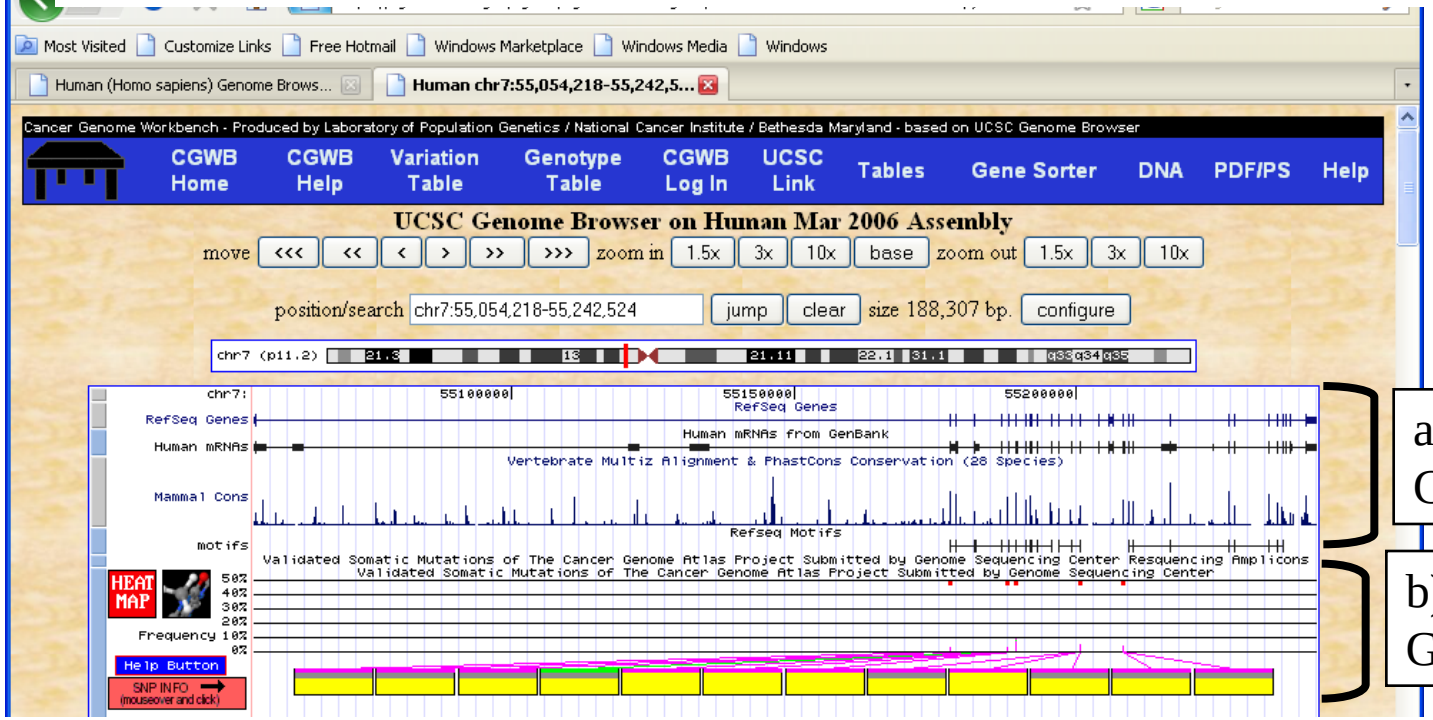
- Colored Vertical bar=mutation rate per gene. Mouse over to get gene name and mutation rate. Mouse over to get the exact number. Click to access the gene/genome view.

- 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

Genome/Gene View Has Three Components

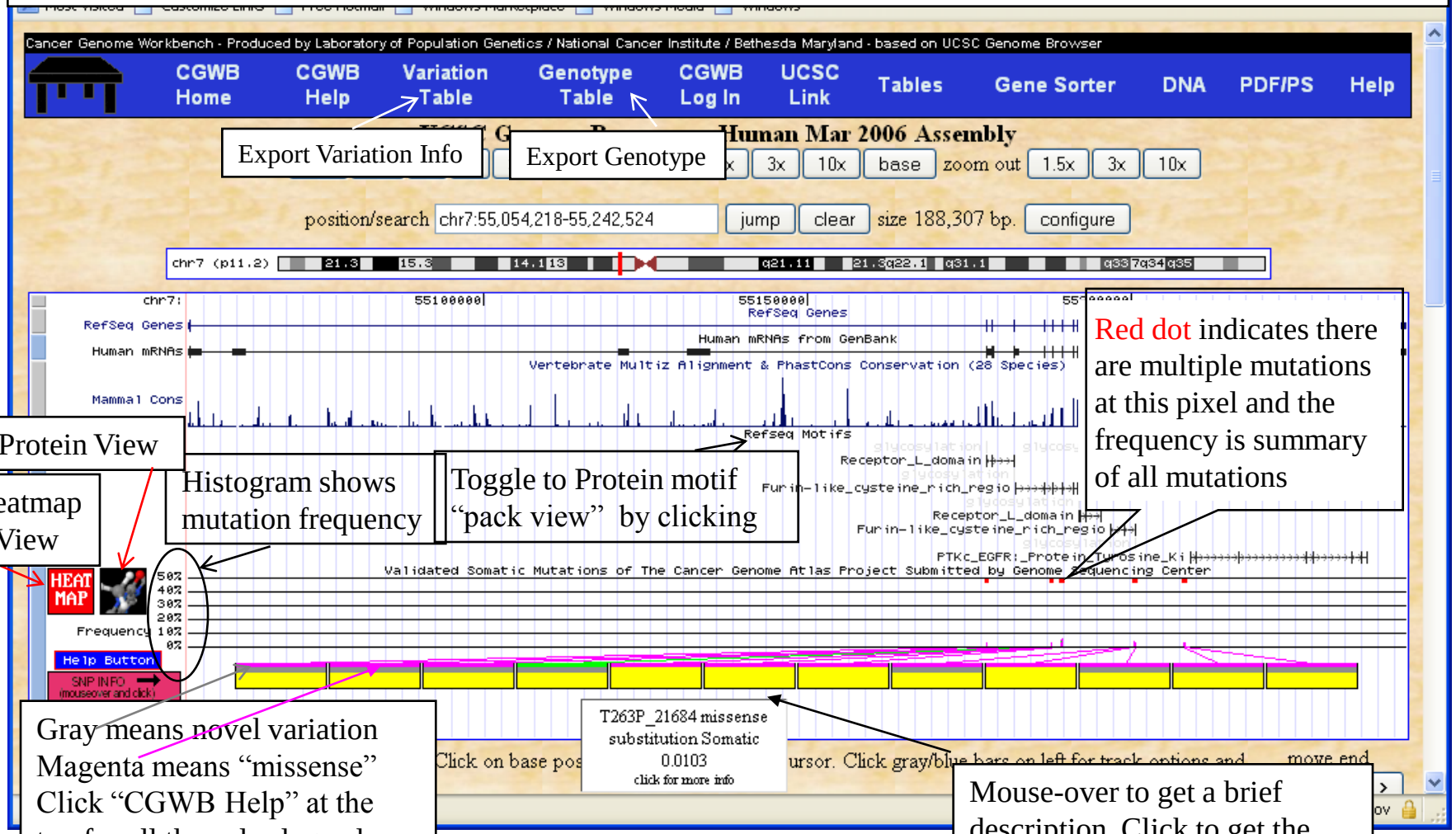


a) Reference Genome View

b) Tumor Genome View

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

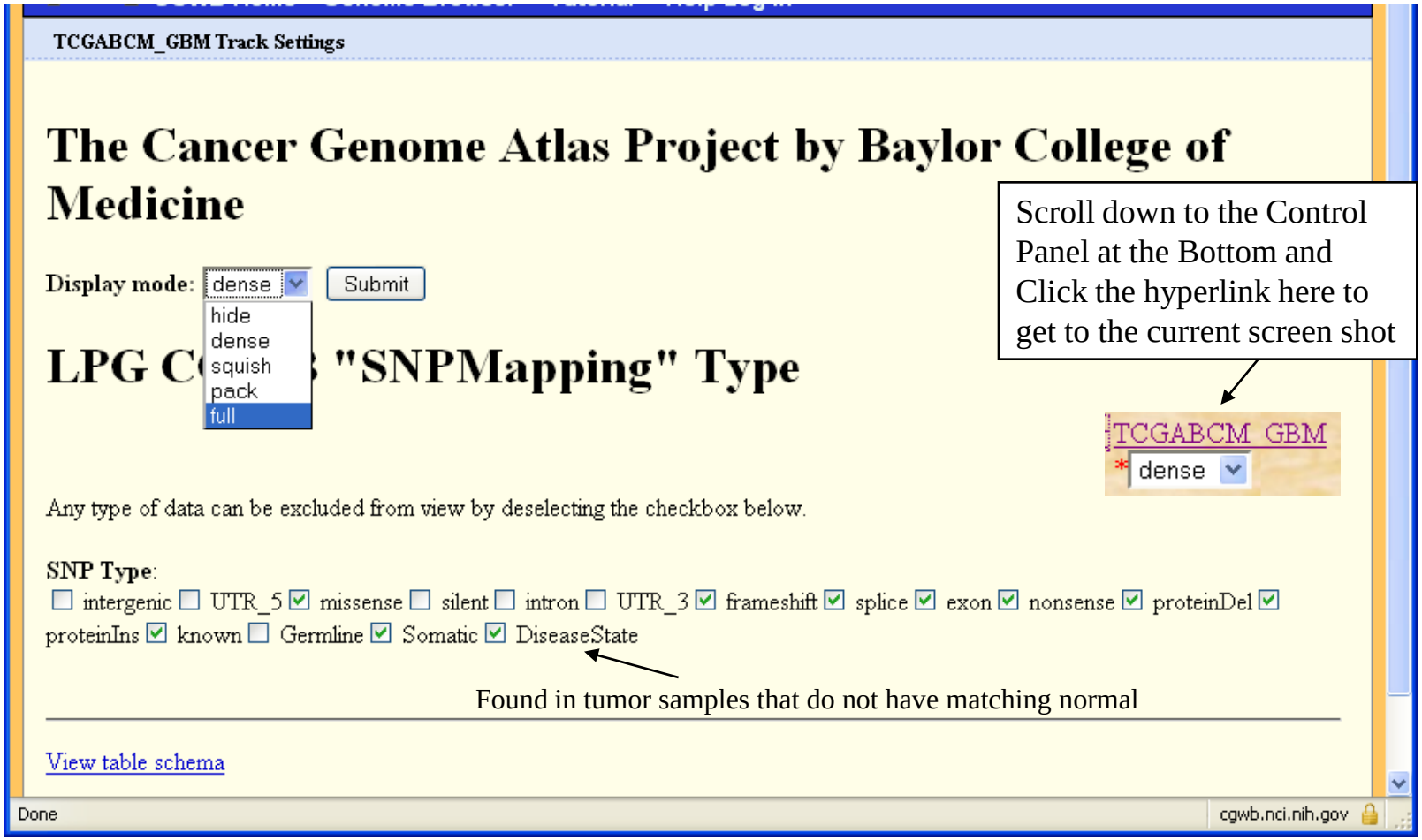
Genome View: a "Dense" View Summarizes Validated Mutations in EGFR



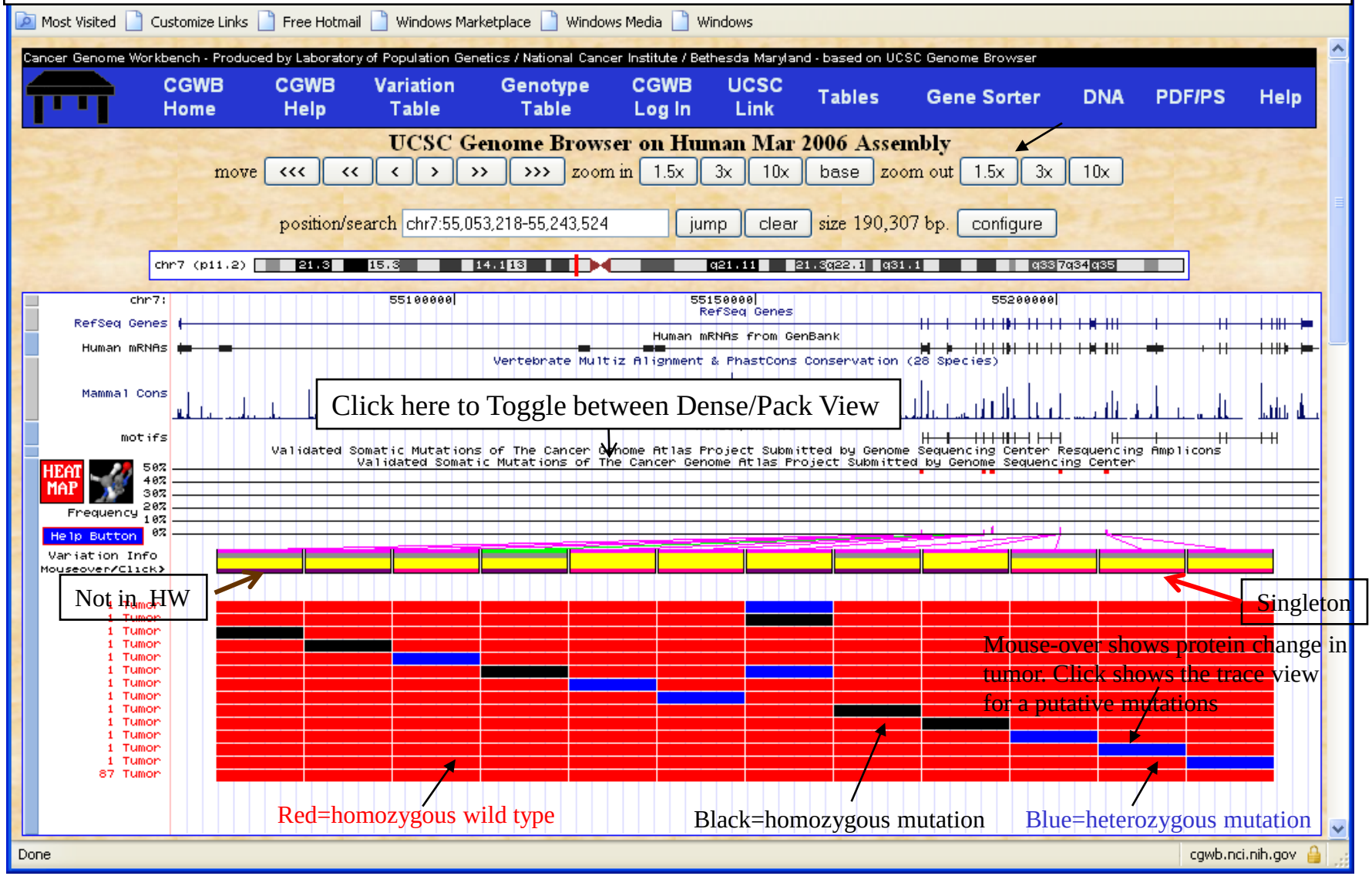
Configure Genome Display by Scrolling to the Control Panel at the Bottom

1)Select a Subset of Variations: missense, silent, somatic, novel?

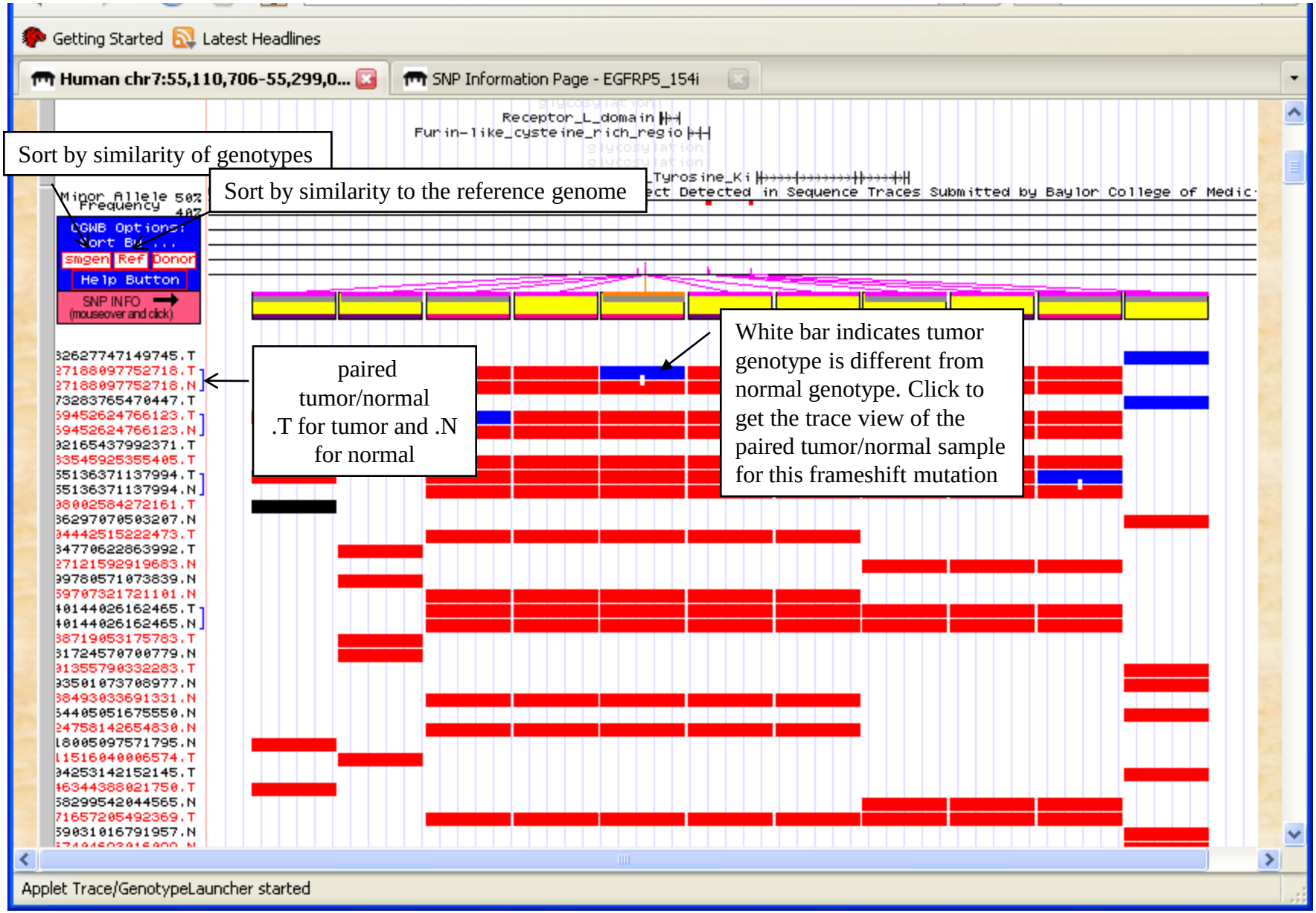
2)Select a Display Style: Dense (default), Pack or Full



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



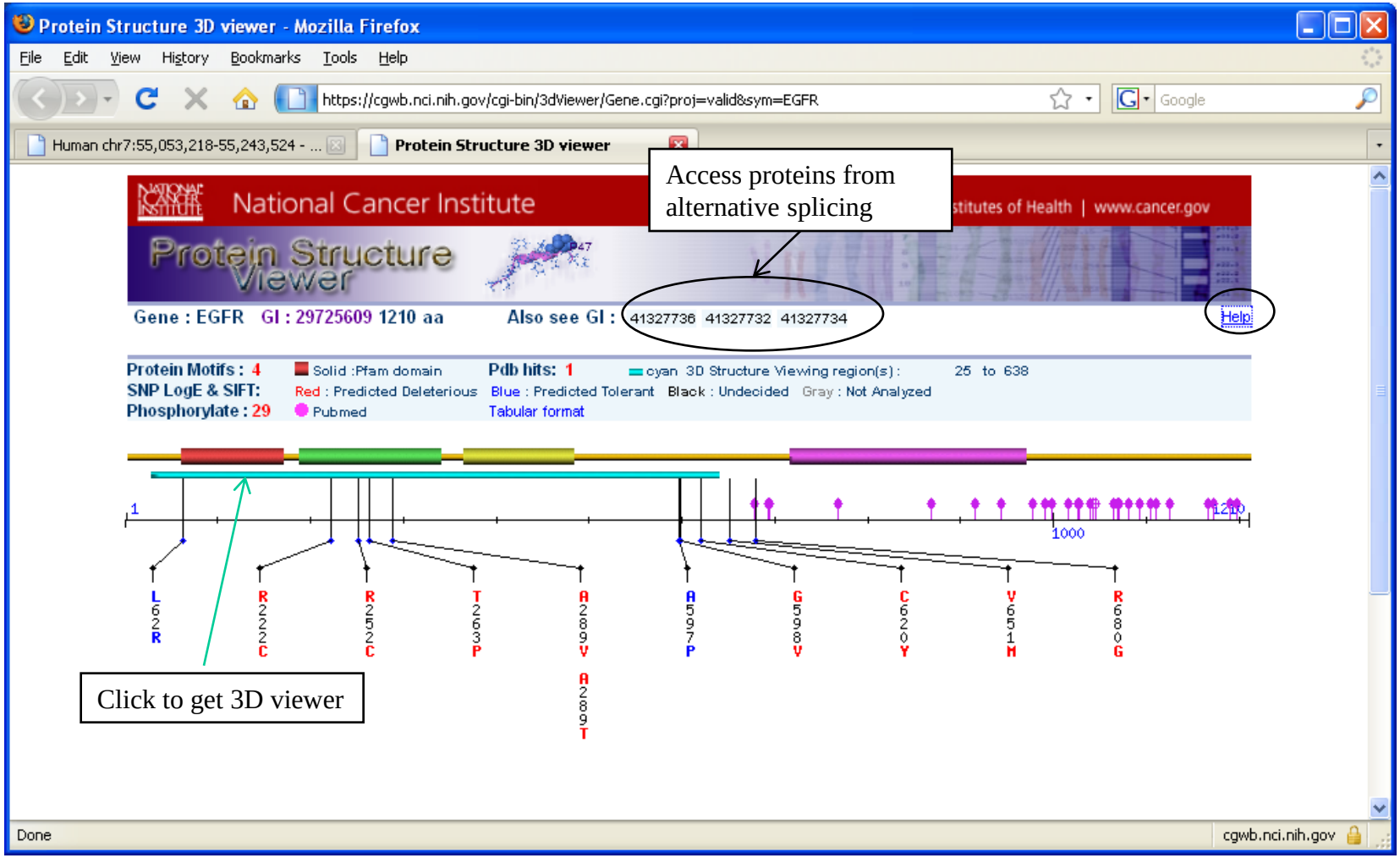
“Full” View Shows Paired Tumor/Normal Genotype for Each Sample



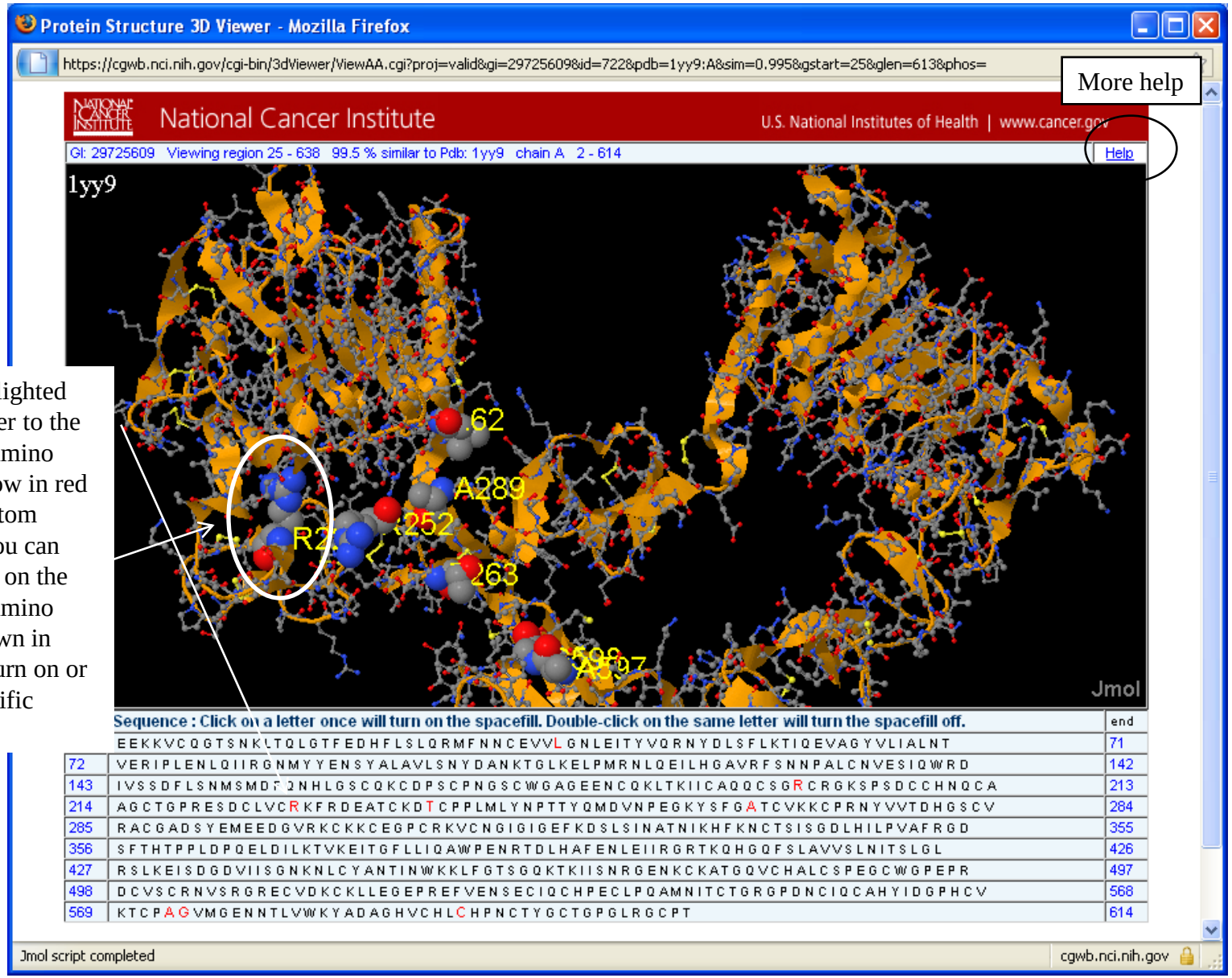
Trace View of a frameshift mutation in paired tumor/normal sample. The tumor has an insertion or a deletion compared to the normal

(This is the putative mutation presented on Figure 1D of the paper)

Protein Structure Viewer

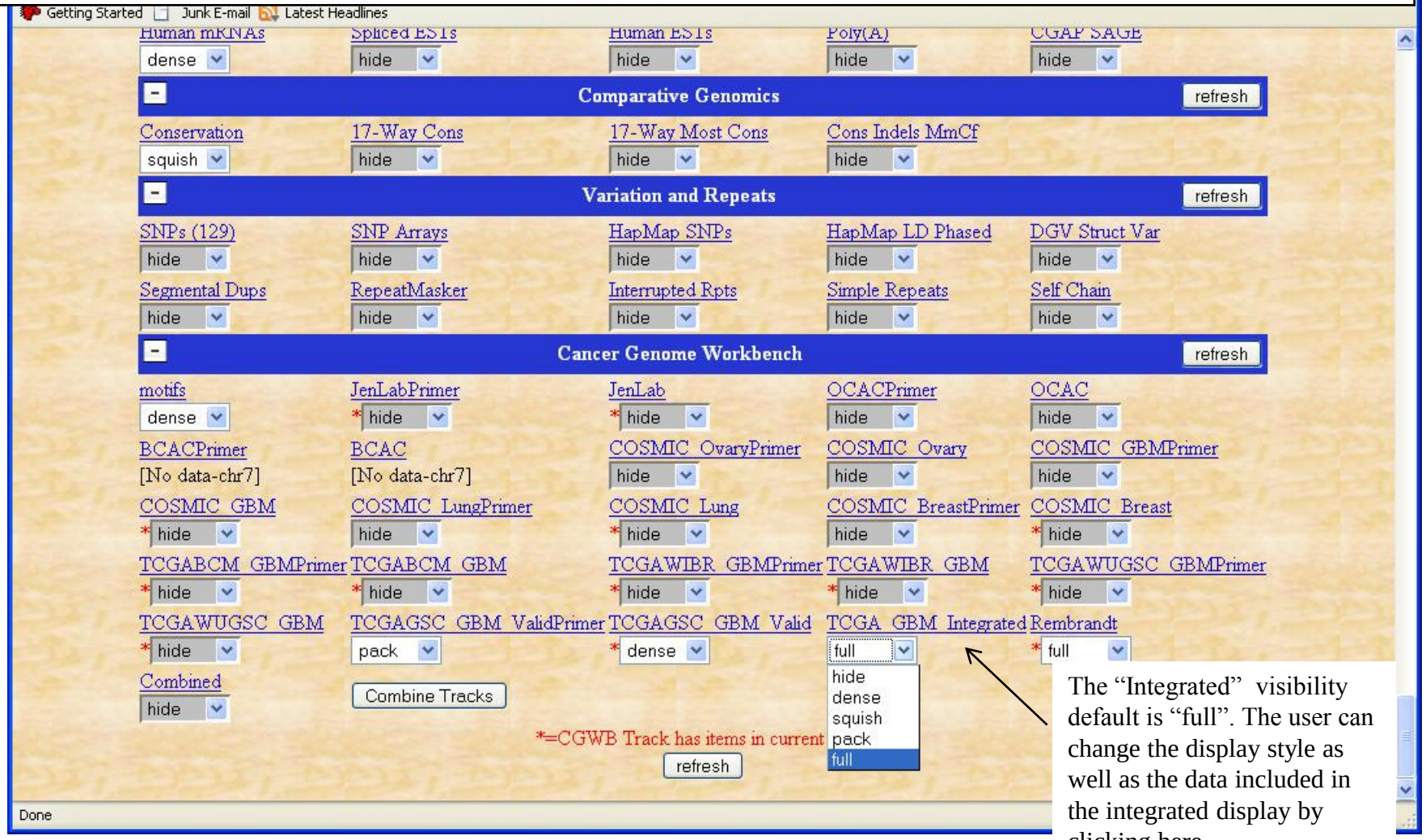


3D Structure Viewer



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

Scroll Down to the Control Panel and Select the View of the Copy Number Data



The "Integrated" visibility default is "full". The user can change the display style as well as the data included in the integrated display by clicking here.

TCGA_GBM_Integrated Track Settings - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://cgwb.nci.nih.gov/cgi-bin/hgTrackUi?hgsid=80144&c=chr7&g=TCGA_GBM_Integrated

Getting Started Junk E-mail Latest Headlines

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: full Submit

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

---

View table schema

Data last updated: 2008-05-31

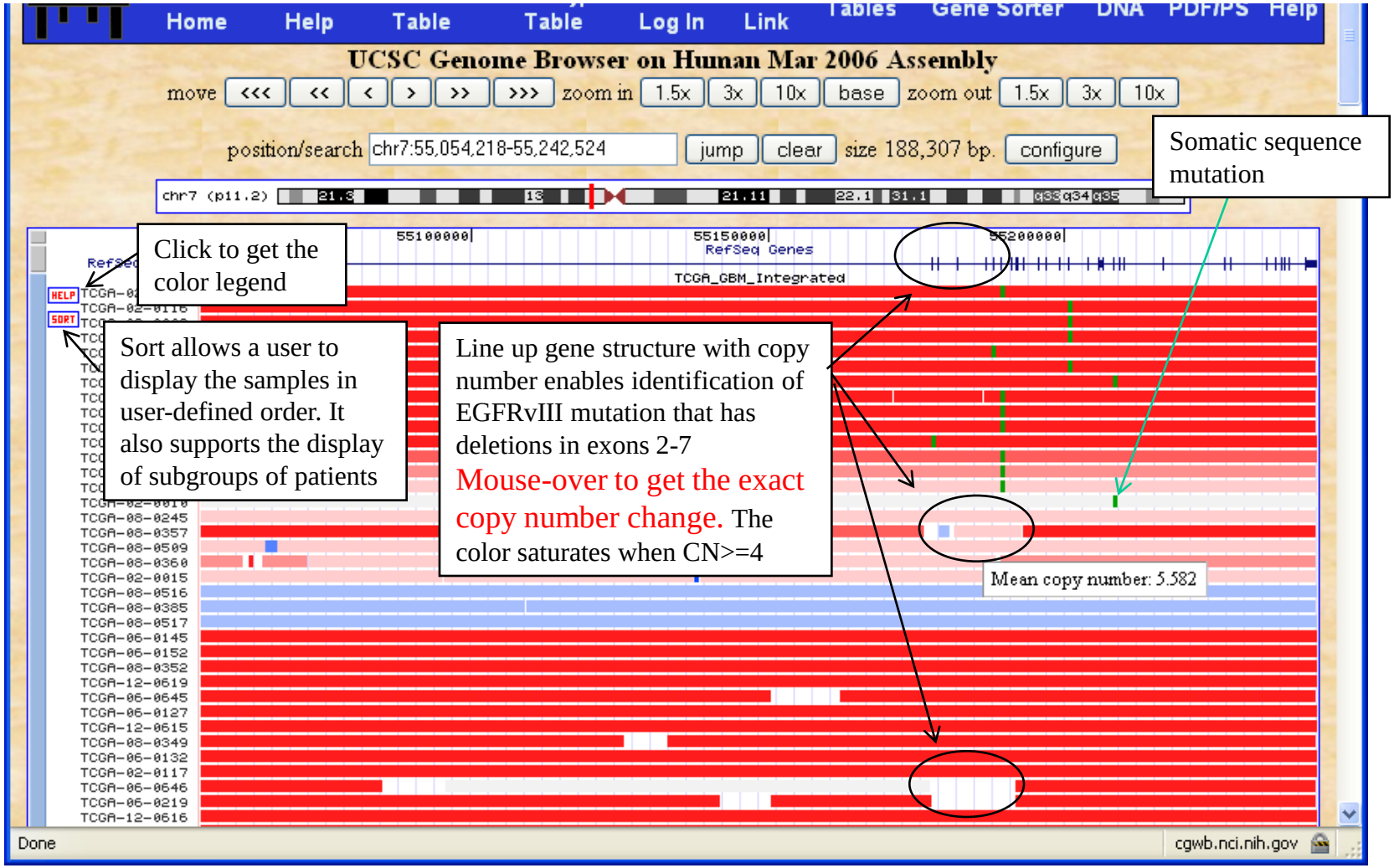
Exon array and putative mutation are available via login

Done cgwb.nci.nih.gov

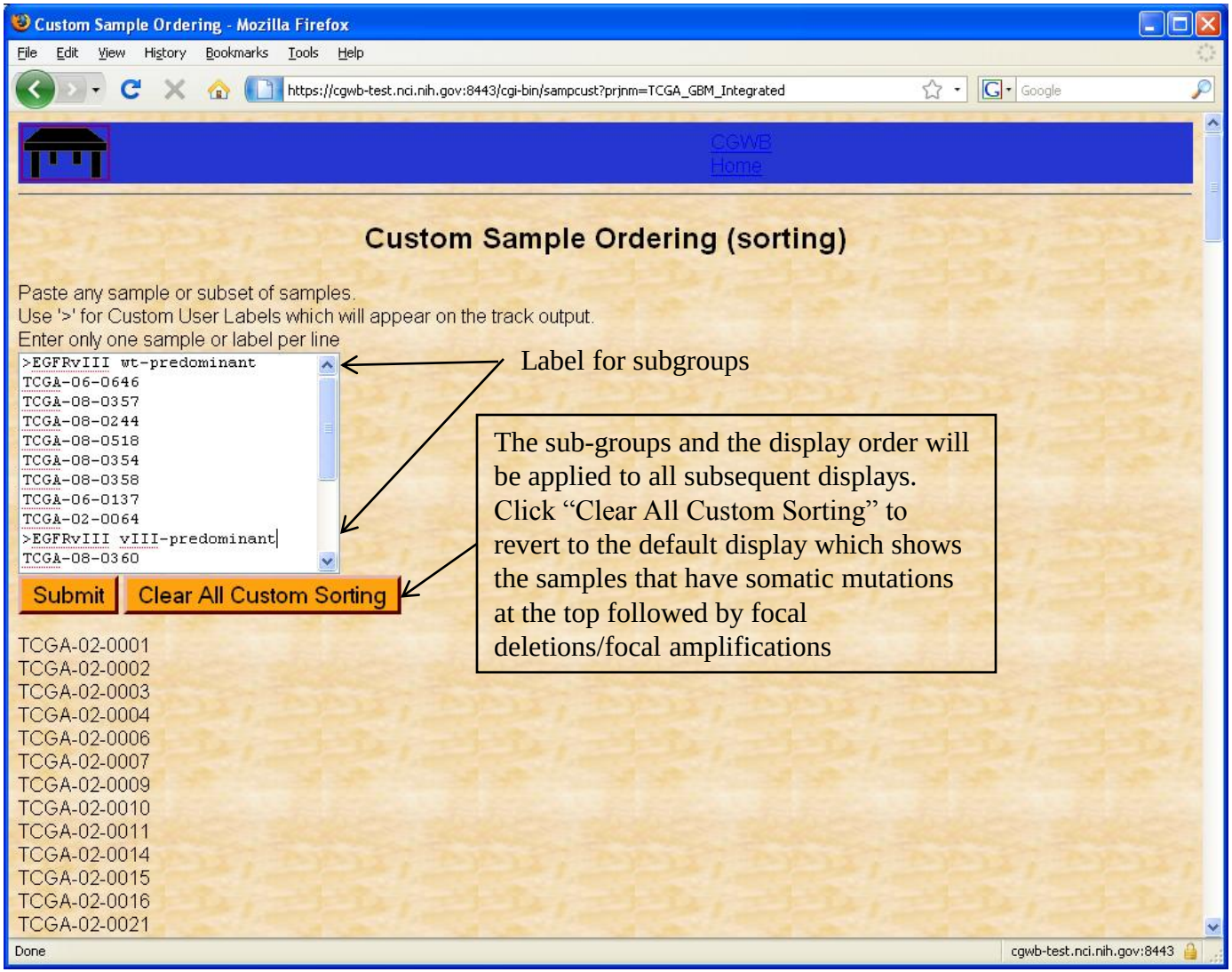
Integrated view of somatic sequence mutation and copy number change

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



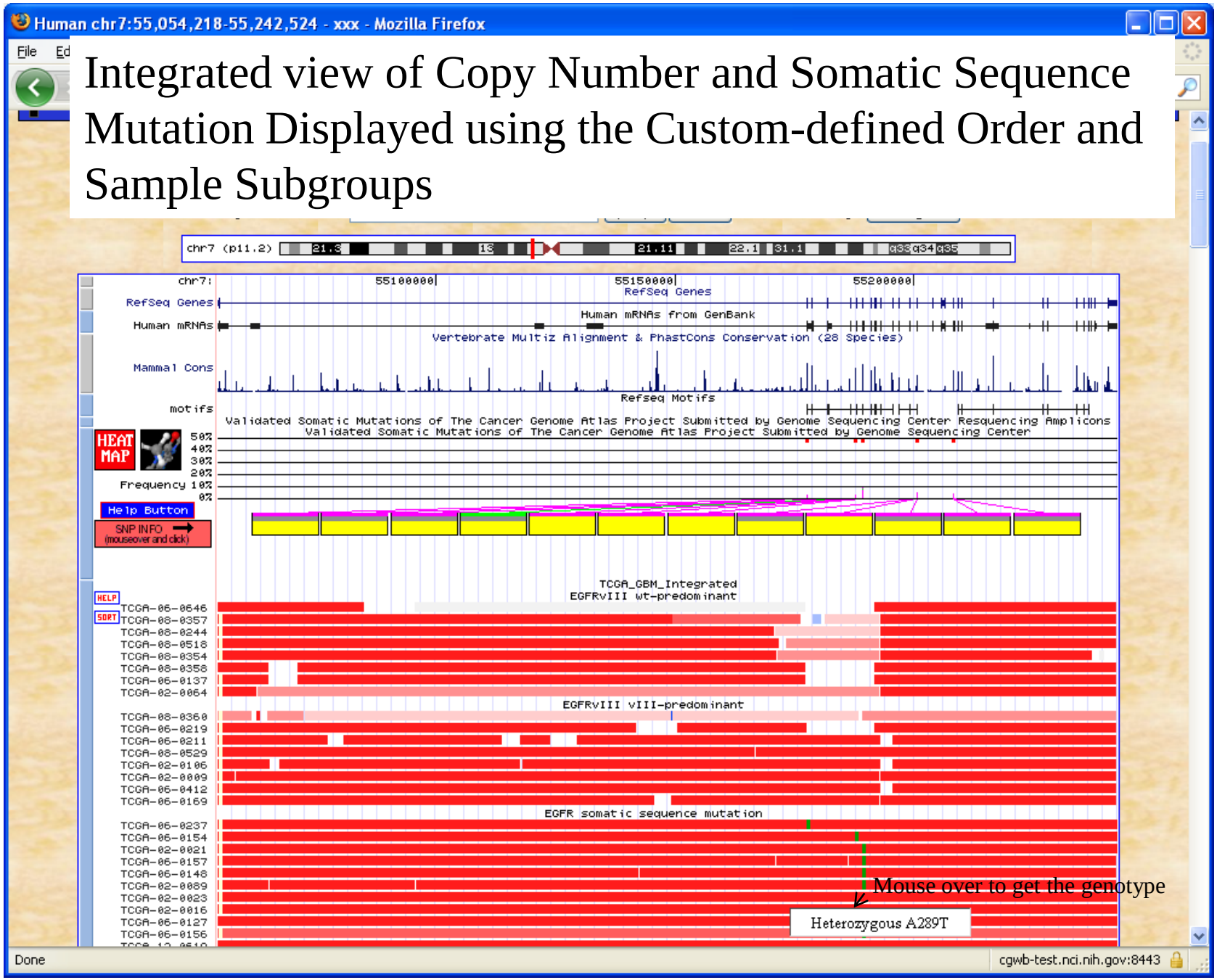
Create Customized Sample Display Order and Sample Subgroups



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

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

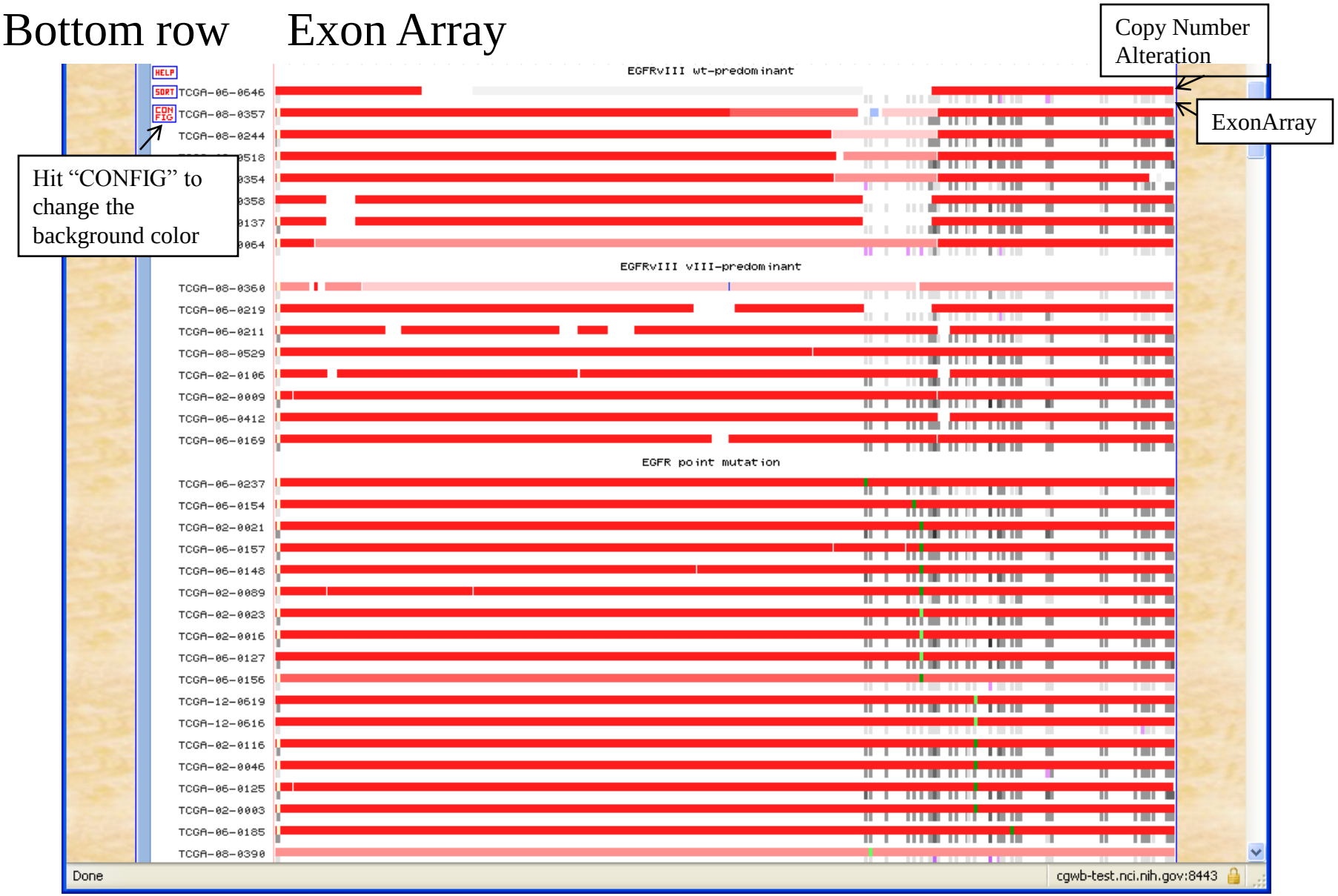


Integrated view with Exon Array and Putative Mutations Superimposed to the Copy Number/Validated Mutations (only users with login can access this view)

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

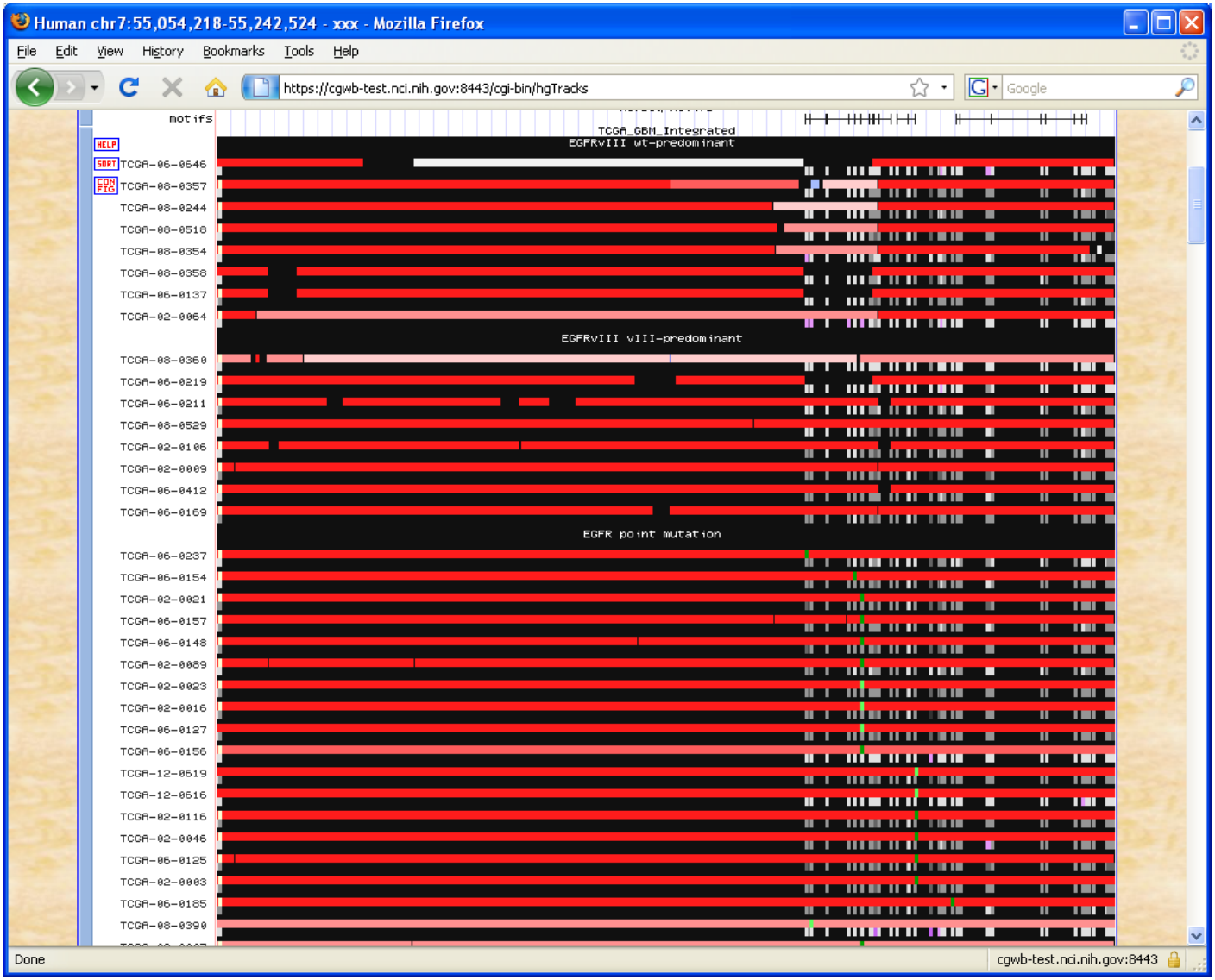
Top row Copy Number +Somatic mutation+Methylation

Bottom row Exon Array



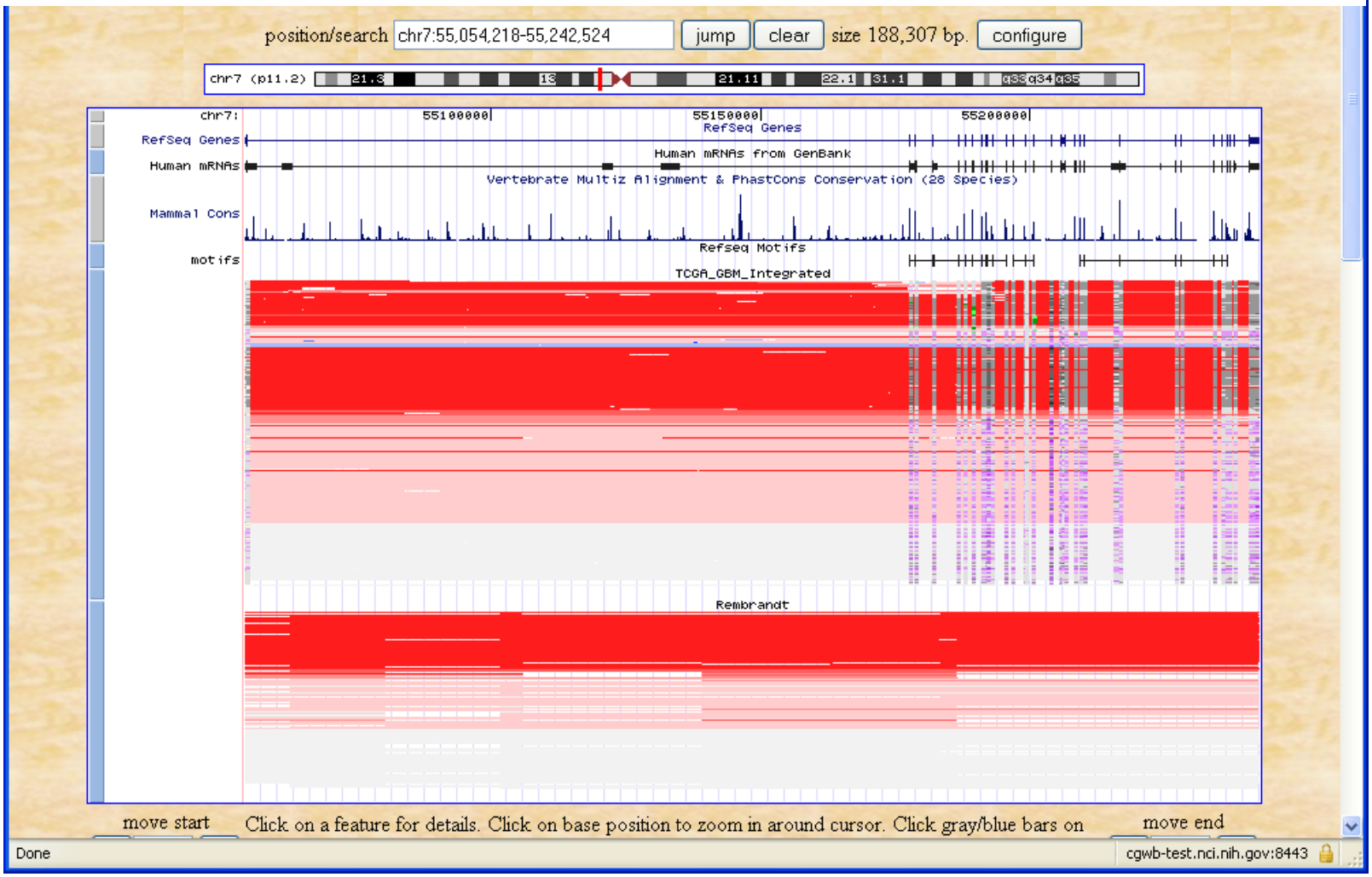
Same Region/Same Data Viewed with Black Background

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 Somatic Mutations from Multiple Projects with the "Combine Tracks" Feature

Menu for "Combo"
Select Data from Different Projects/Groups

CGWB
Home

Combine SNP Tracks

Select Any Combination of Projects to Create a Combined Track

☐ : BCAC BCAC Breast Cancer Association Consortium	* ☐ : COSMIC_Breast COSMIC_Breast Catalogue Of Sanger COSMIC Variations In Breast Cancer	* ☐ : COSMIC_GBM COSMIC_GBM Catalogue Of Sanger COSMIC Variations In Glioblastoma Multiforme	* ☐ : COSMIC_Lung COSMIC_Lung Catalogue Of Sanger COSMIC Variations In Lung Cancer
☐ : COSMIC_Ovary COSMIC_Ovary Catalogue Of Sanger COSMIC Variations In Ovarian Cancer	* ☐ : JenLab JenLab JenLab Lung Cancer Genes	☐ : OCAC OCAC OCAC Familial Ovarian Cancer Project	* ☑ : TCGABCM_GBM TCGABCM_GBM The Cancer Genome Atlas Project by Baylor College of Medicine
* ☑ : TCGAGSC_GBM_Valid TCGAGSC_GBM_Valid Validated Somatic Mutations of The Cancer Genome Atlas Project Submitted by Genome Sequencing Center	* ☑ : TCGAWIBR_GBM TCGAWIBR_GBM The Cancer Genome Atlas Project by Broad Institute	* ☑ : TCGAWUGSC_GBM TCGAWUGSC_GBM The Cancer Genome Atlas Project by Washington University	

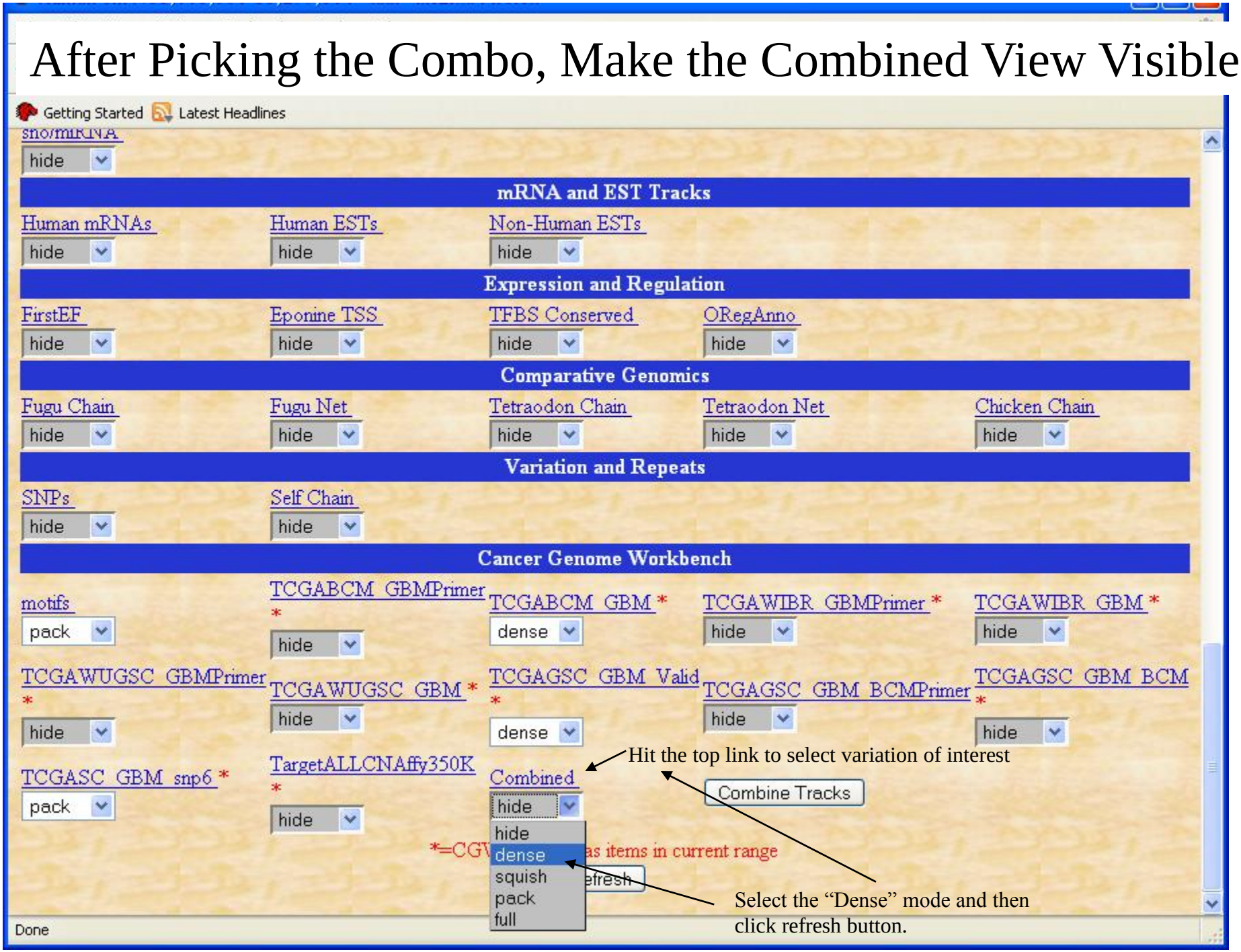
Submit | Set All | Clear All

* = Project has SNPs in the current range : chr7:55053217-55243524

Selected the 3-TCGA projects that have putative mutations and 1 validated TCGA project

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

After Picking the Combo, Make the Combined View Visible



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

Combined Track Settings - Mozilla Firefox

CGWB Home Genome Browser Tutorial Help Log In

Combined Track Settings

Combined CGWB Track

Display mode: dense Submit

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

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

Human chr7:55,053,217-55,243,524 - xxx - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://cgwb-test.nci.nih.gov:8443/cgi-bin/hgTracks?org=Human&db=hg18&position=chr7:55053217-55243524&c

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/PS | 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) 21.3 13 21.11 22.1 31.1 q33 q34 q35

chr7: 55100000 55150000 55200000

RefSeq Genes

Human mRNAs

Vertebrate Multiz Alignment & PhastCon

Mammal Cons

Refseq Motifs

Combined CGWB Track

motifs

HEAT MAP COMBO

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

Help Button

SNP INFO (mouseover and click)

Click to view this novel mutation only found by WUGSC

Number indicates how many groups found this variation

found in TCGAGSC_GBM_Valid click for more info

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

move start < 2.0 >

Click gray/blue bars on

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

HELP

Variation Summary

Alleles	Genotype Summary	
G T	GG 310	GT 1

Fasta Sequence

```
GTGCTGCGCT TCCTCCGTGT GTGGCGCTGA GTGTACTTAC CTCACTTGCC
CAGCGTGTCC TCTCTCCTCC ATAGGTCTGC CGCAAATTCC GAGACGAAGC
CACGTGCAAG GACACCTGCC CCCCACTCAT GCTCTACAAC CCCACCACGT
ACCAGATGGA TGTGAACCCC GAGGGCAAAT ACAGCTTTGG TGCCACCTGC
G/T
TGAAGAAGTG TCCCCGTGAG TCCTCCTCTG TGGGCCCTCT AACTGGTCAG
GCATCCTTGT CCCGCTCTGT CTCCTGCTGA GCCCTGGAGT ATCCCATCTT
GGAGAGTCTT TGGGTGGATG TGTTTGCCTT GCTTGGAGGA GGCGACCCTG
TGCCCGTCCA GGCACACAGG CGAGGGGAGG GGCTGGCTTG CTACCGAGGA
```

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

Jmol structure view requires Java.

Scroll down to view individual genotype

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 ↓

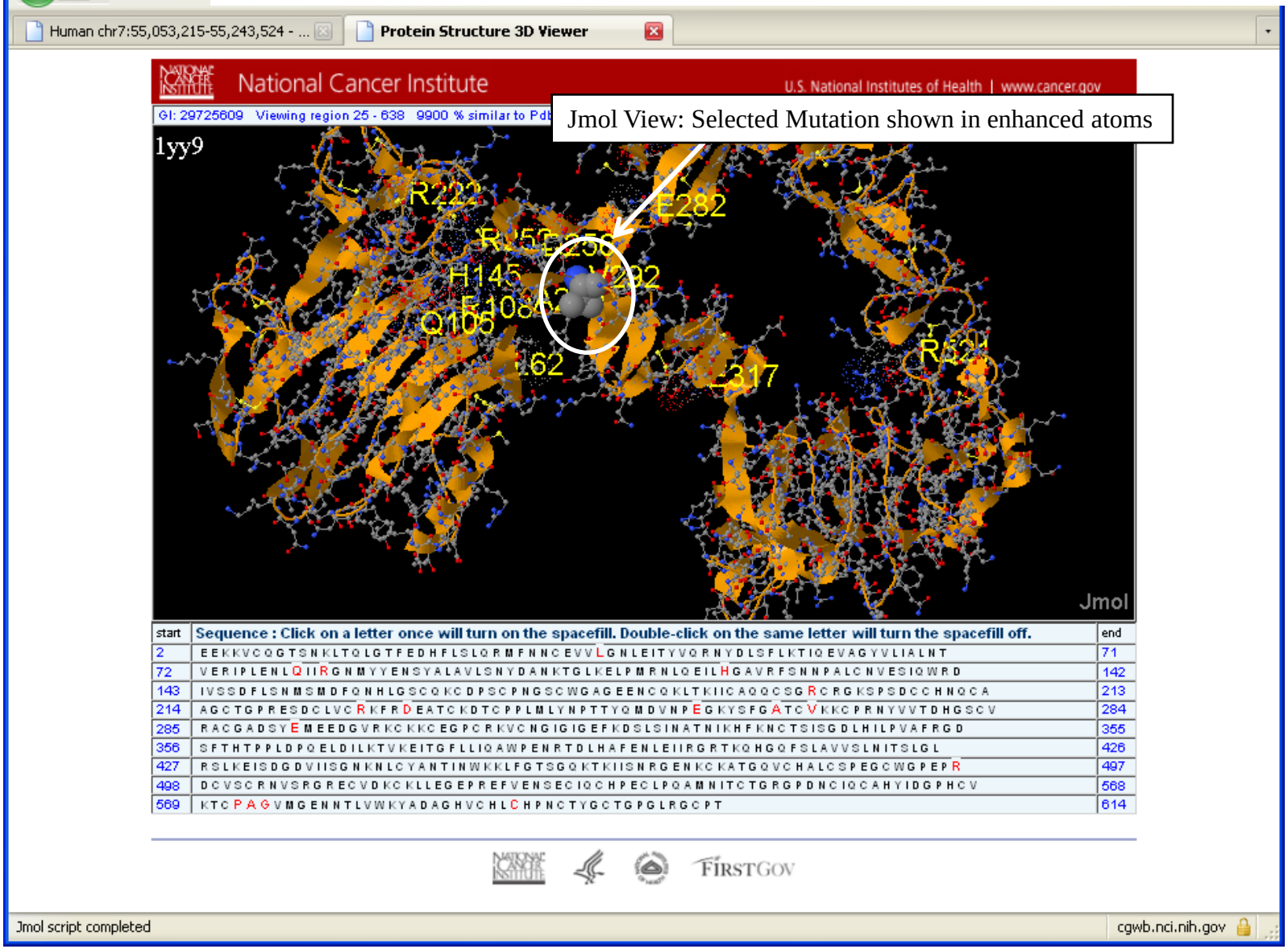
deleterious

Motifs/Domains/Sites

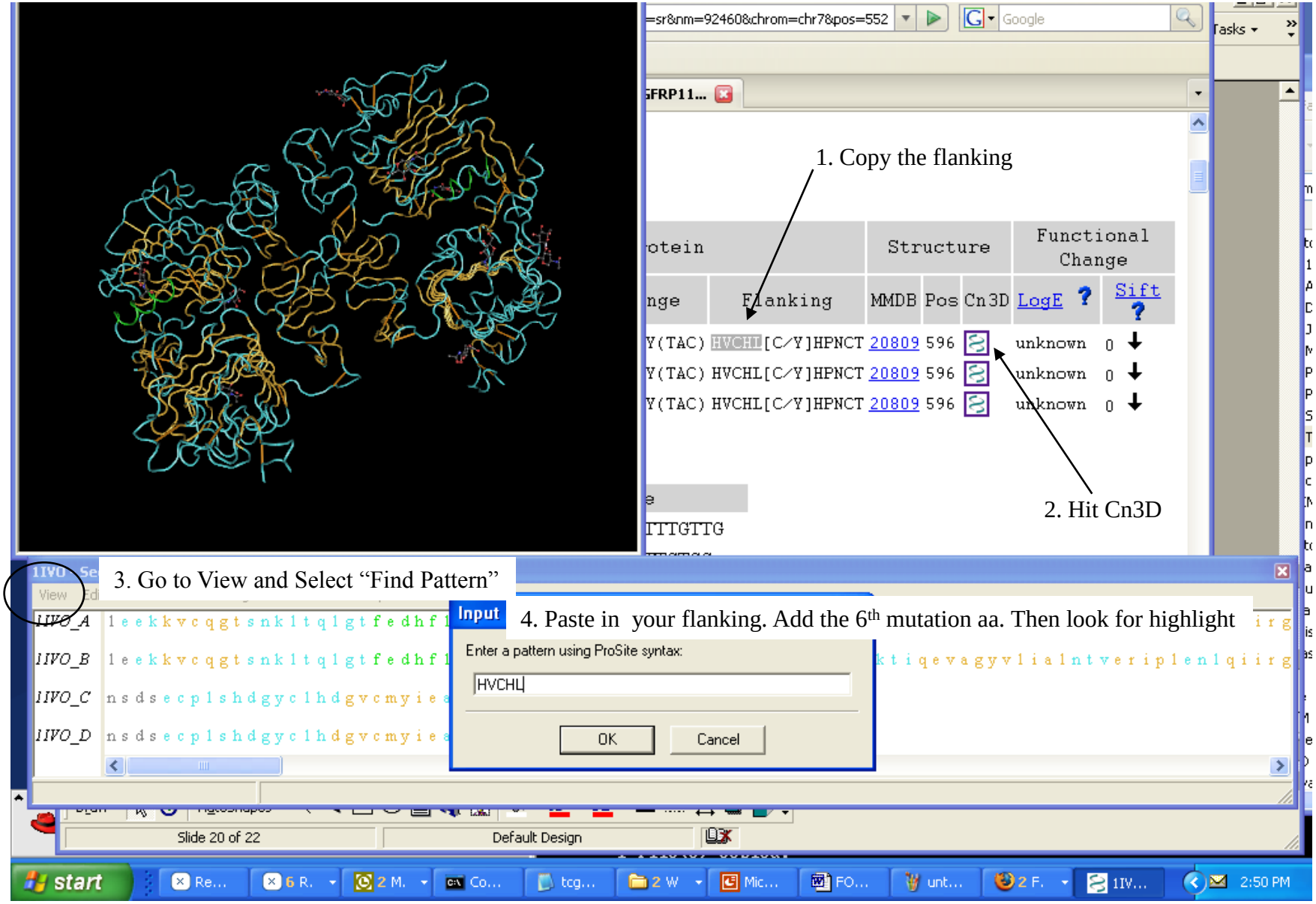
MOTIF: CDD:85007 Furin-like Furin-like cysteine rich region; pfam00757

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

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



How to Highlight Mutation on Cn3D View?



Primers and Genotypes: Select Samples for Trace View

SNP Information Page - EGFRP23_233 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Human chr7:55,053,217-55,243,524 - ... | SNP Information Page - EGFRP23...

Primers

Name	Forward	Reverse
EGFRP27	TAACTTGGGCTTTCTGACGGG	CACATCCACCCAAAGACTCTCC
EGFRP7	AGCTGGGTTTTCCCACACTA	GACAACTGAATACCCACGGC
EGFRP5	ACTTGGGCTTTCTGACGGGA	GCAAGGCAAACACATCCACC
EGFRP25	TAATTTCATTCCATAGTCACCGC	CTCGGTAGCAAGCCAGCC
EGFRP23	GTAACTTGGGCTTTCTGACGGG	CACATCCACCCAAAGACTCTCC

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

Genotypes

(Select samples and then click Submit to view the traces for each genotype.)

Submit | Select All | Clear All | Traceviewer Help File.

Sample Name	Disease	Normal
WUGSC-2377692378	☑ WUGSC-2377692378 (T/G) [quality:medium]	
WUGSC-2051260781		☑ WUGSC-2051260781 (G/G) [quality:medium]
WUGSC-2051260789		☐ WUGSC-2051260789 (G/G) [quality:high]
WUGSC-2051260911	☐ WUGSC-2051260911 (G/G) [quality:medium]	
WUGSC-2051260942		☐ WUGSC-2051260942 (G/G) [quality:medium]
WUGSC-2051261056		☐ WUGSC-2051261056 (G/G) [quality:medium]
WUGSC-2051261068	☐ WUGSC-2051261068 (G/G) [quality:medium]	
WUGSC-2051261072	☐ WUGSC-2051261072 (G/G) [quality:medium]	
WUGSC-2051261081		☐ WUGSC-2051261081 (G/G) [quality:medium]
WUGSC-2051261086	☐ WUGSC-2051261086 (G/G) [quality:medium]	
WUGSC-2051261351		☐ WUGSC-2051261351 (G/G) [quality:medium]

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

Mutations in Published Literature are Linked

File Edit View History Bookmarks Tools Help

Human chr7:55,110,706-55,299,014 -... | SNP Information Page - A289D_2...

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

Laboratory of Population Genetics

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

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

Variation Summary

Alleles	Genotype Summary		
C T	CC 598	CT 8	TT 5

Fasta Sequence

```
TCAACACCGT GCTGCGCTTC CTCCGTGTGT GGCGCTGAGT GTACTTACCT
CACTTGCCCA GCGTGTCCTC TCTCCTCCAT AGGTCTGCCG CAAATTCCGA
GACGAAGCCA CGTGCAAGGA CACCTGCCCC CCACTCATGC TCTACAACCC
CACCACGTAC CAGATGGATG TGAACCCCGA GGGCAAATAC AGCTTTGGTG
C/T
CACCTGCGTG AAGAAGTGTC CCCGTGAGTC CTCCTCTGTG GGCCCTCTAA
CTGGTCAGGC ATCCTTGTCC CGCTCTGTCT CCTGCTGAGC CCTGGAGTAT
CCCATCTTGG AGAGTCTTTG GGTGGATGTG TTTGCCTTGC TTGGAGGAGG
CGACCCTGTG CCCGTCCAGG CACACAGGCG AGGGGAGGGG CTGGCTTGCT
```

Position and Function

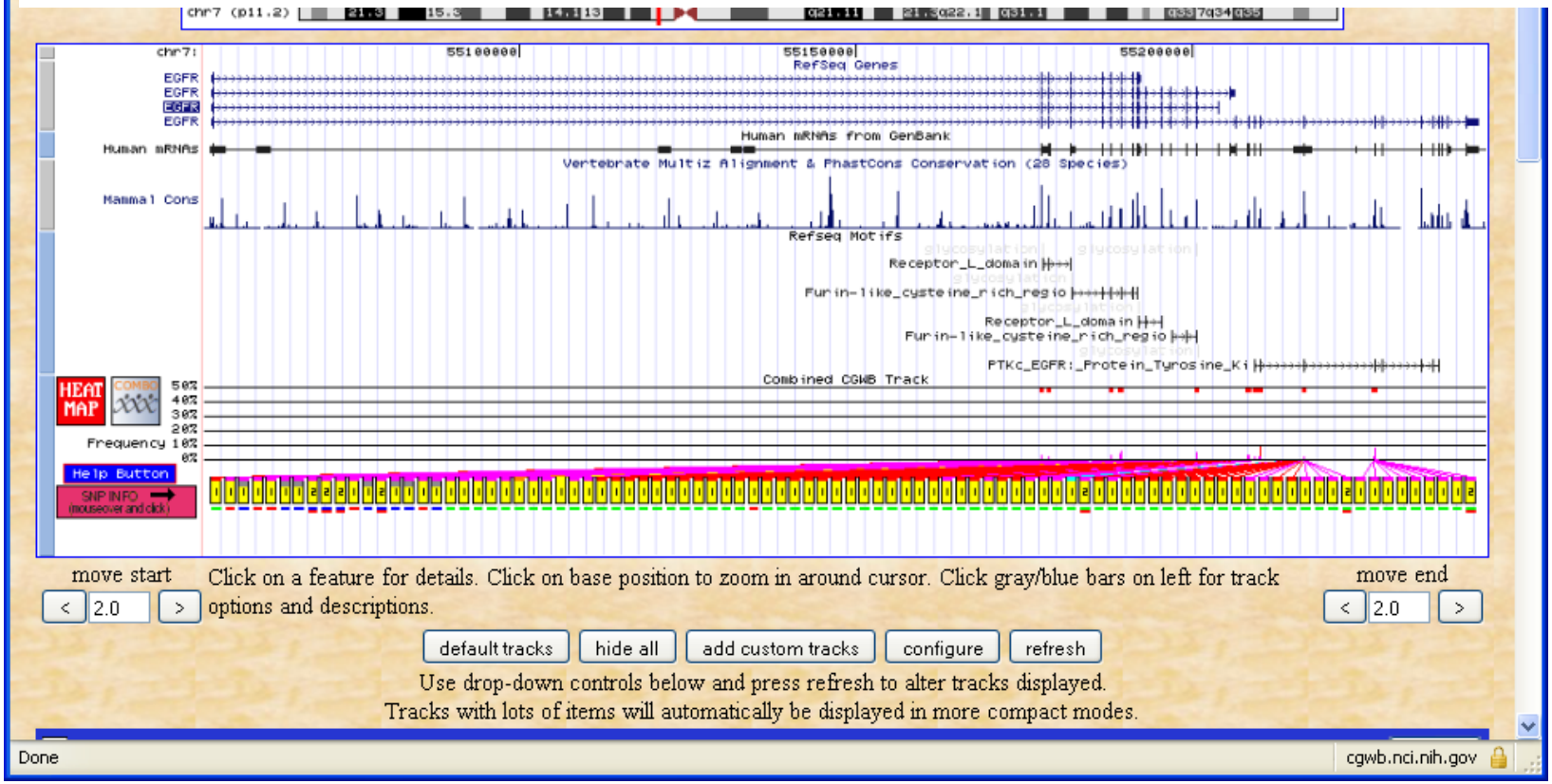
Genomic		Transcript			Protein			Structure			Functional Change	
Chr	Pos	RefSeq	Pos	Strand	GI	AA Change	Flanking	MMDB	Pos	Cn3D	LogE ?	Sift ?
chr7	55189316	NM_201283	1112	+	41327734	A(GCC)289V(GTC)	KYSFG[A/V]TCVKK	20809	265		1.3328 ↓	0.0011 ↓

Done

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)