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

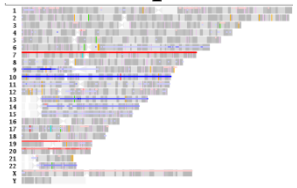
1. Landscape View: Identify target regions such as mutation hotspots, recurrent and/or focal copy number changes
2. Genome View: See how the mutation/copy number changes affect the reference human genome. Evaluate the relationship between somatic mutation and copy number change using the integrated copy-number 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

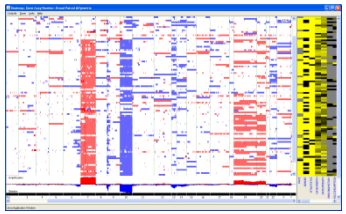
Project	Disease	Genomic Data				
		Copy Number	Gene Expression	Methylation	Clinical	Mutation
TCGA	GBM, Ovarian	+	+	+	+	+
TSP	lung	+			+	+
TARGET	ALL, NBL	+	+		+	+
JHU	GBM, Pan					+
Rembrandt	GBM	+				
GSK cell line	>30 tissues	+	+			
COSMIC	>30 tissues					+

Viewers for individual & integrated genomic Data

Landscape

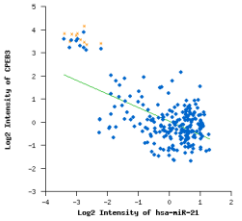


Heatmap

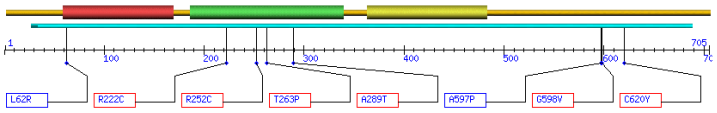


Genome

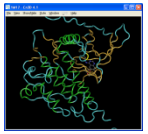
Correlation Plot

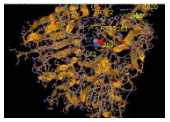


Protein

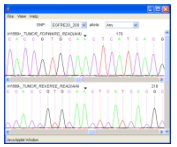


3D Structure Viewer

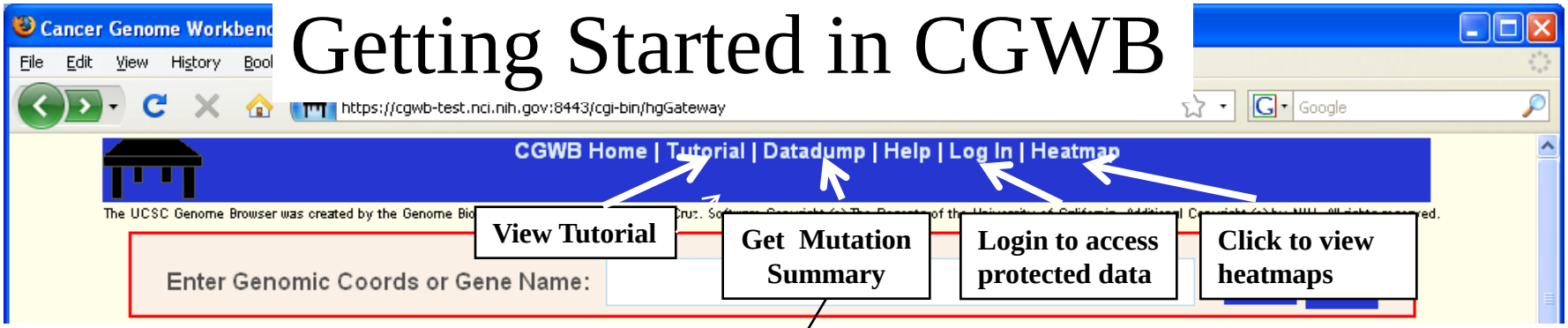


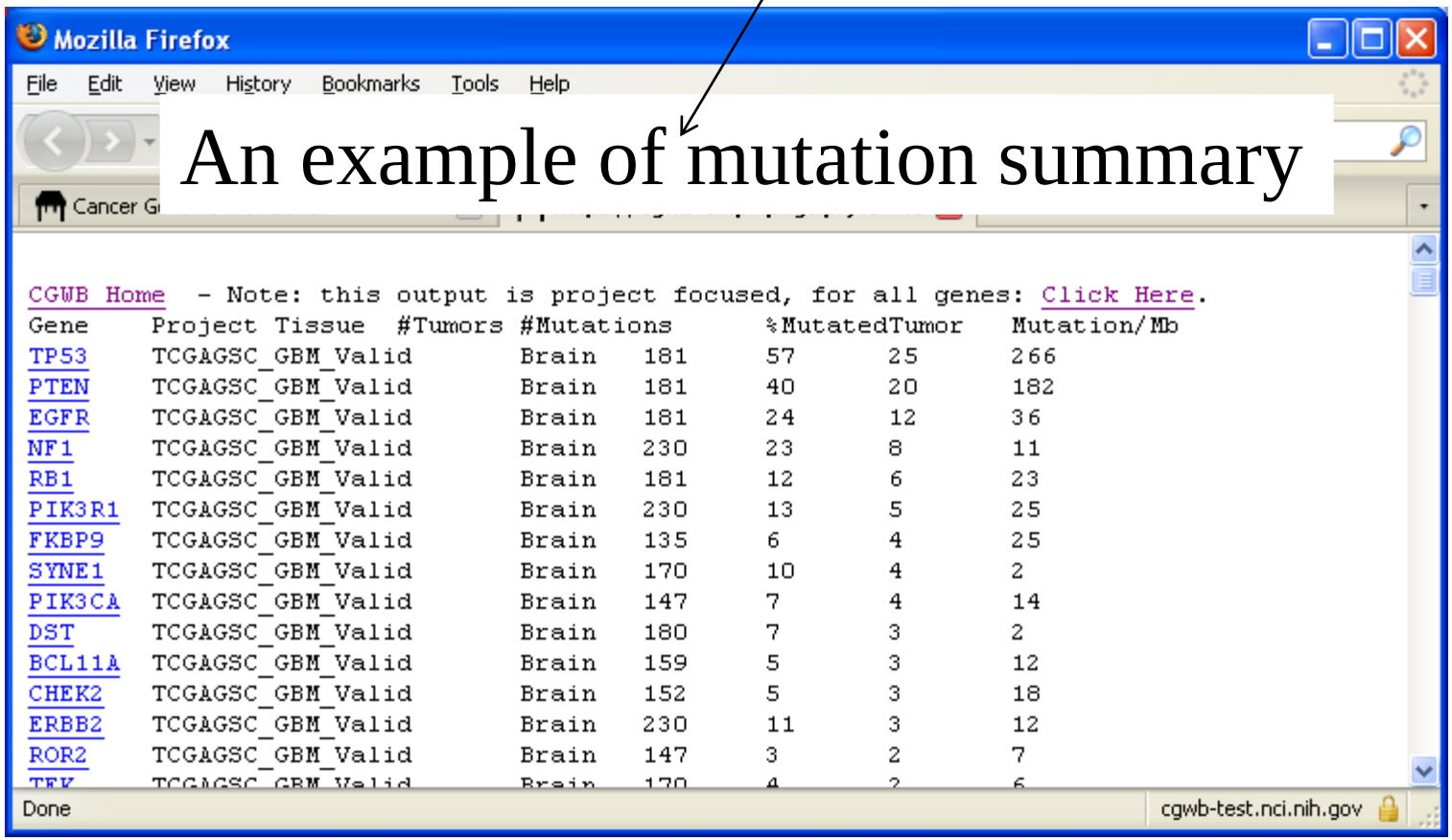


Trace

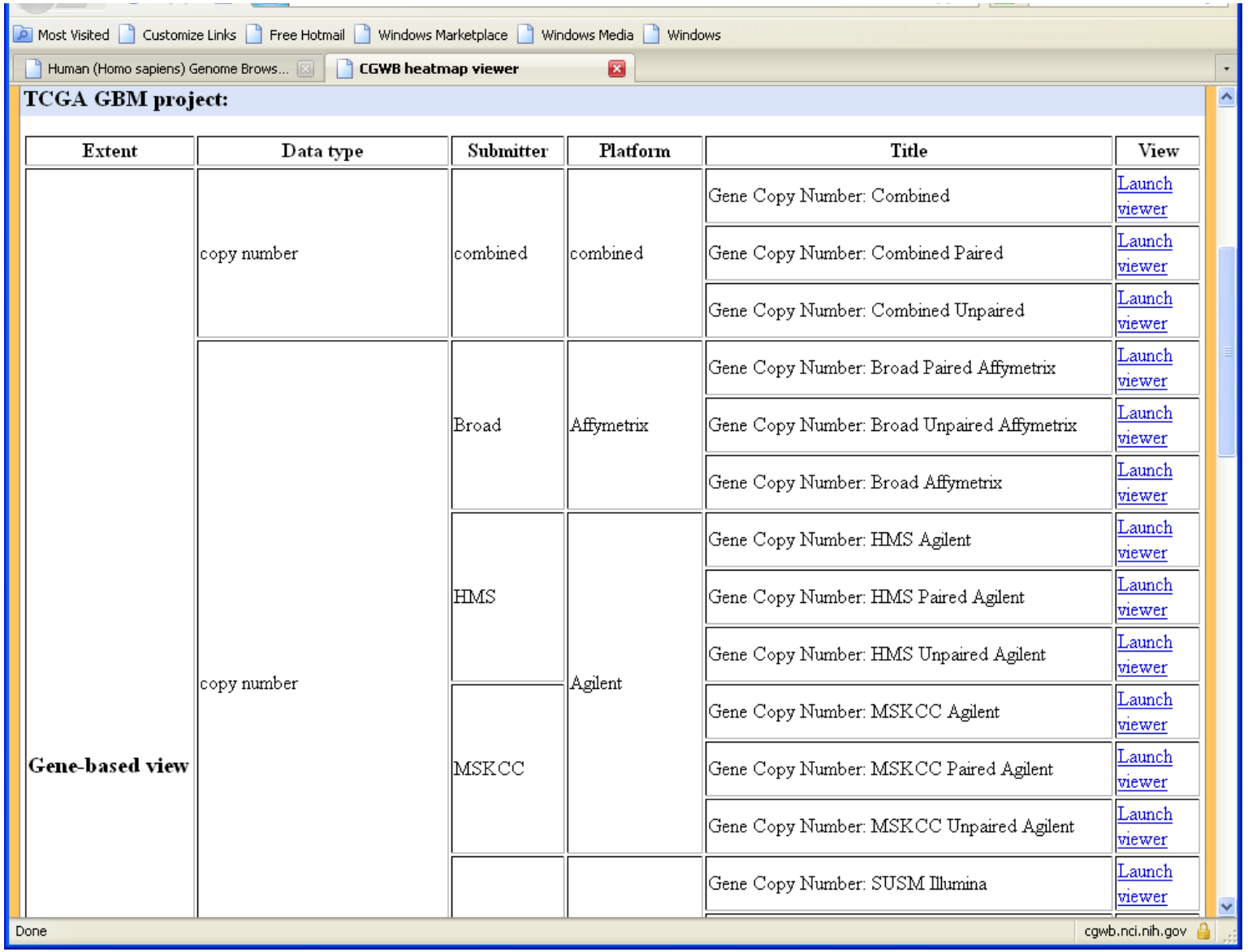


Getting Started in CGWB





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

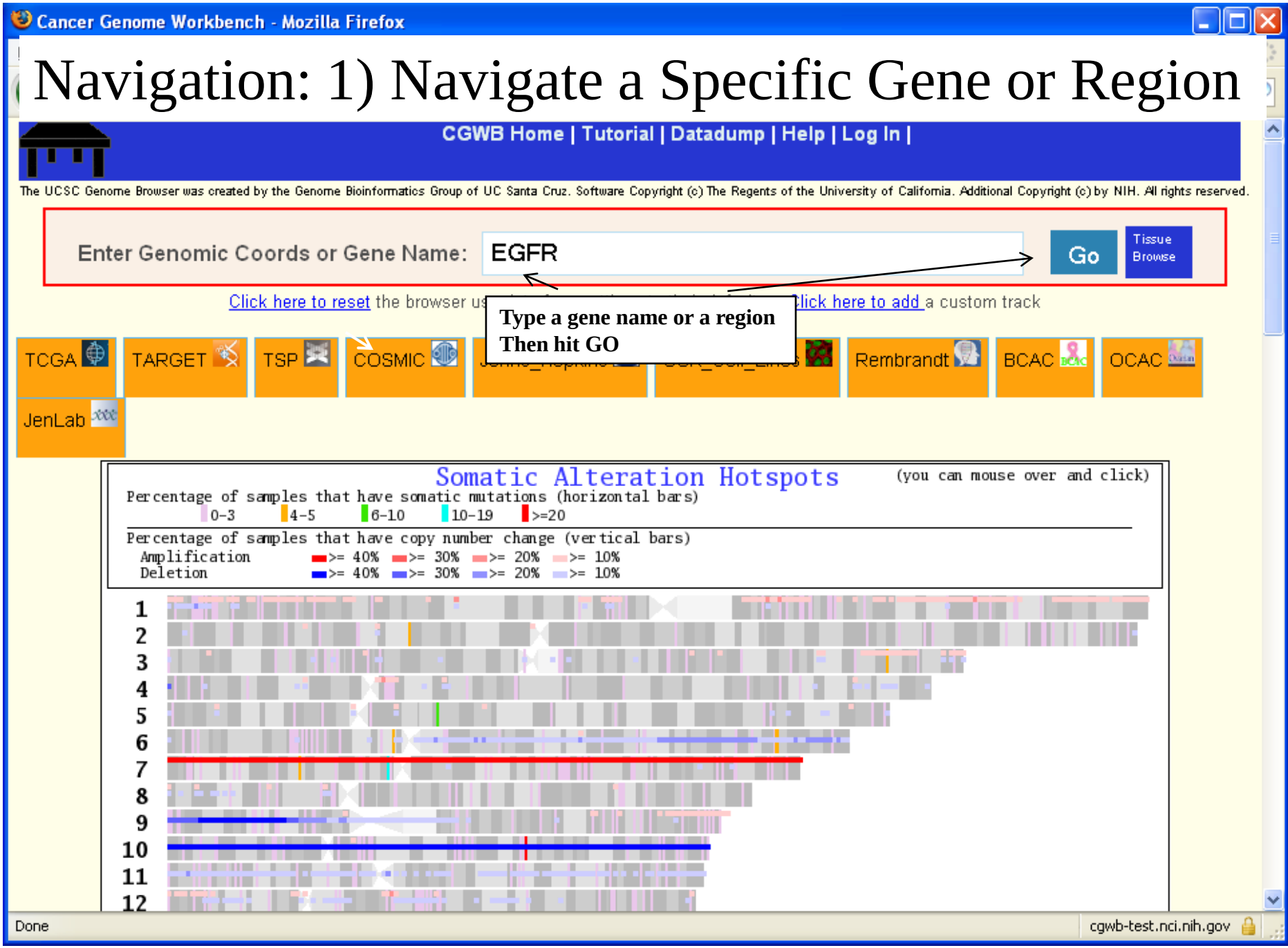


Extent	Data type	Submitter	Platform	Title	View
Gene-based view	copy number	combined	combined	Gene Copy Number: Combined	Launch viewer
				Gene Copy Number: Combined Paired	Launch viewer
				Gene Copy Number: Combined Unpaired	Launch viewer
		Broad	Affymetrix	Gene Copy Number: Broad Paired Affymetrix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

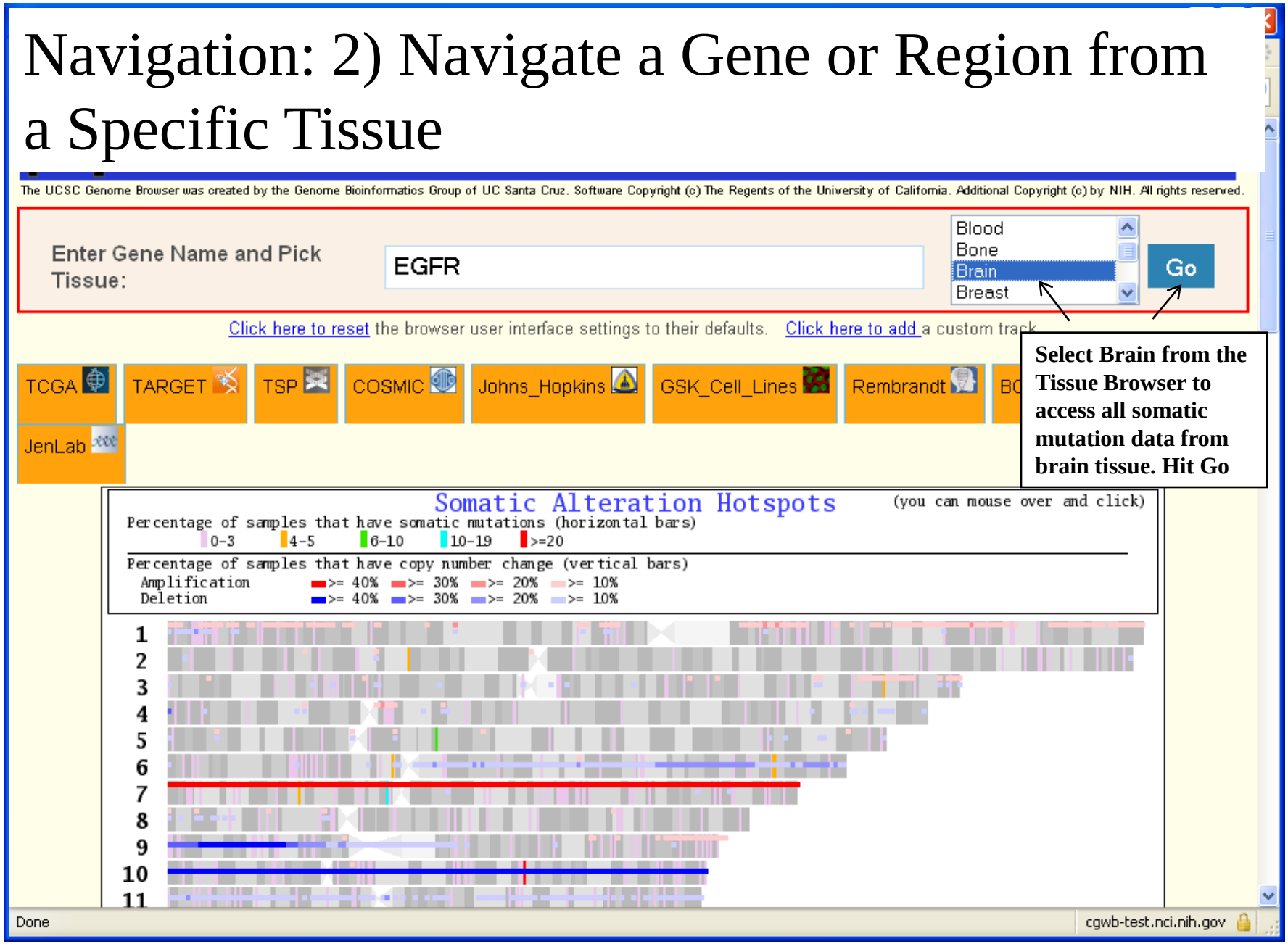
Navigate Cancer Genome Workbench

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

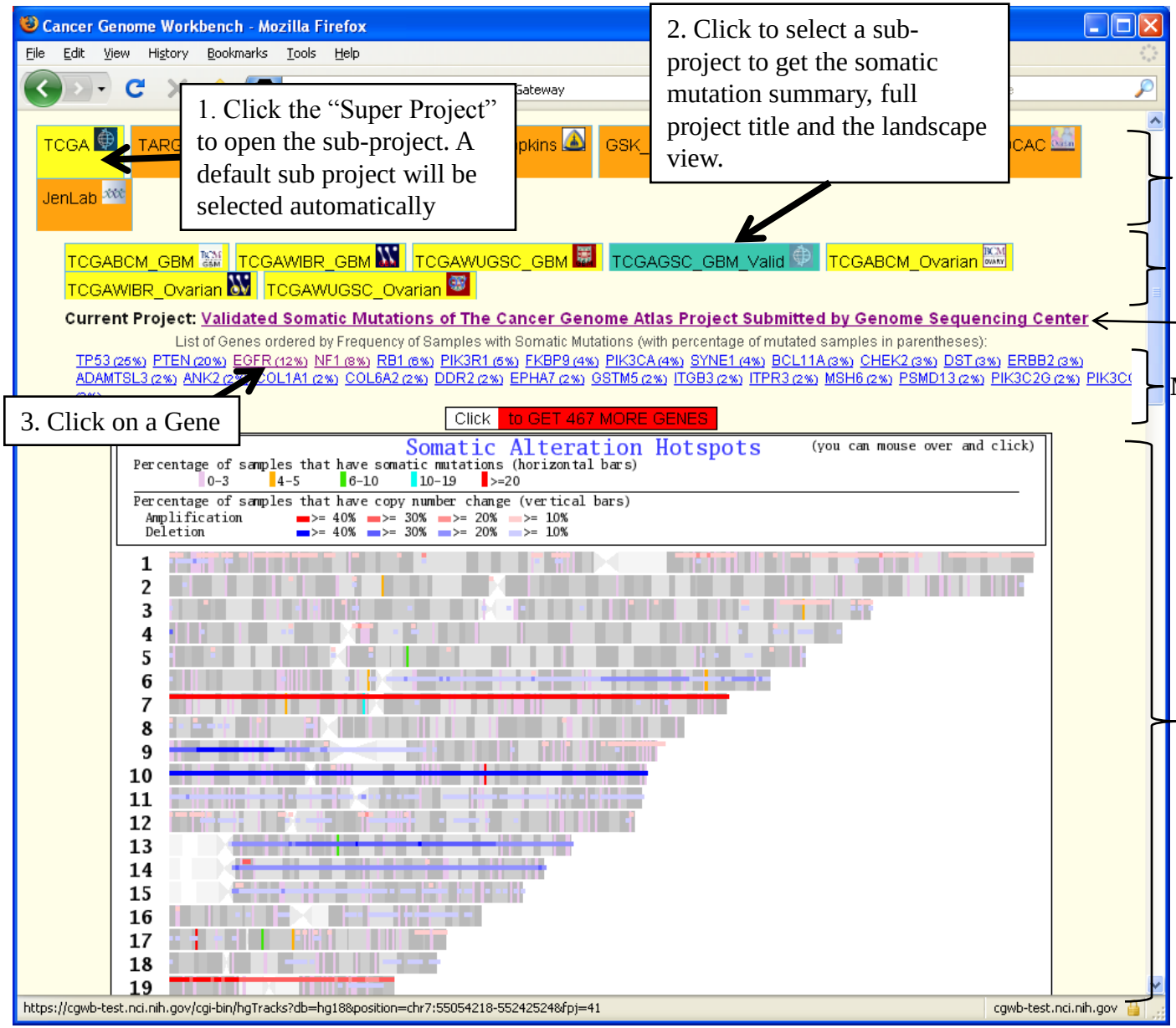
Navigation: 1) Navigate a Specific Gene or Region



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



Navigation: 3) Navigate a Gene in a Specific Project



"Super" Project

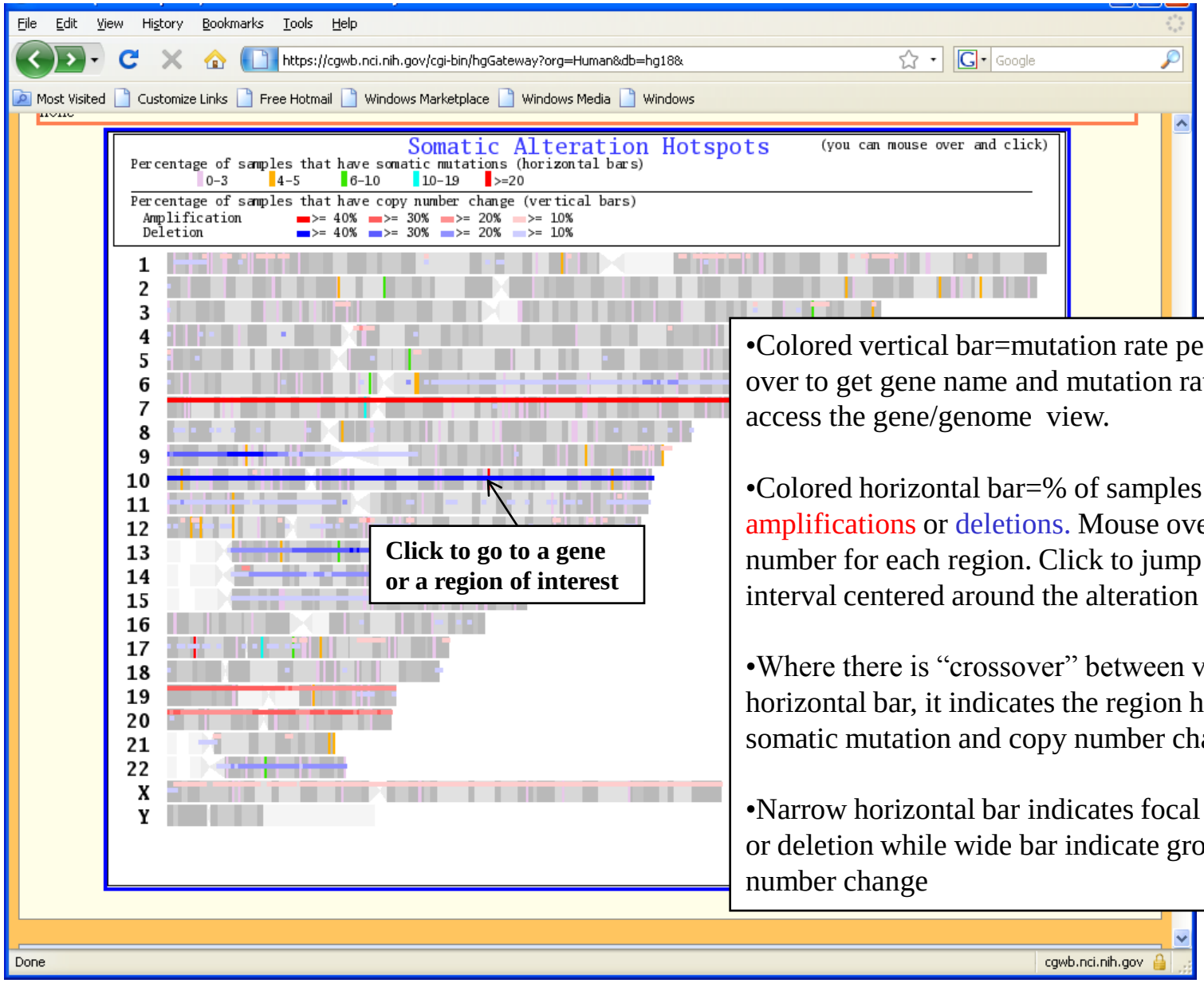
Project

Project Title

Mutation Summary

LandScape View

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.

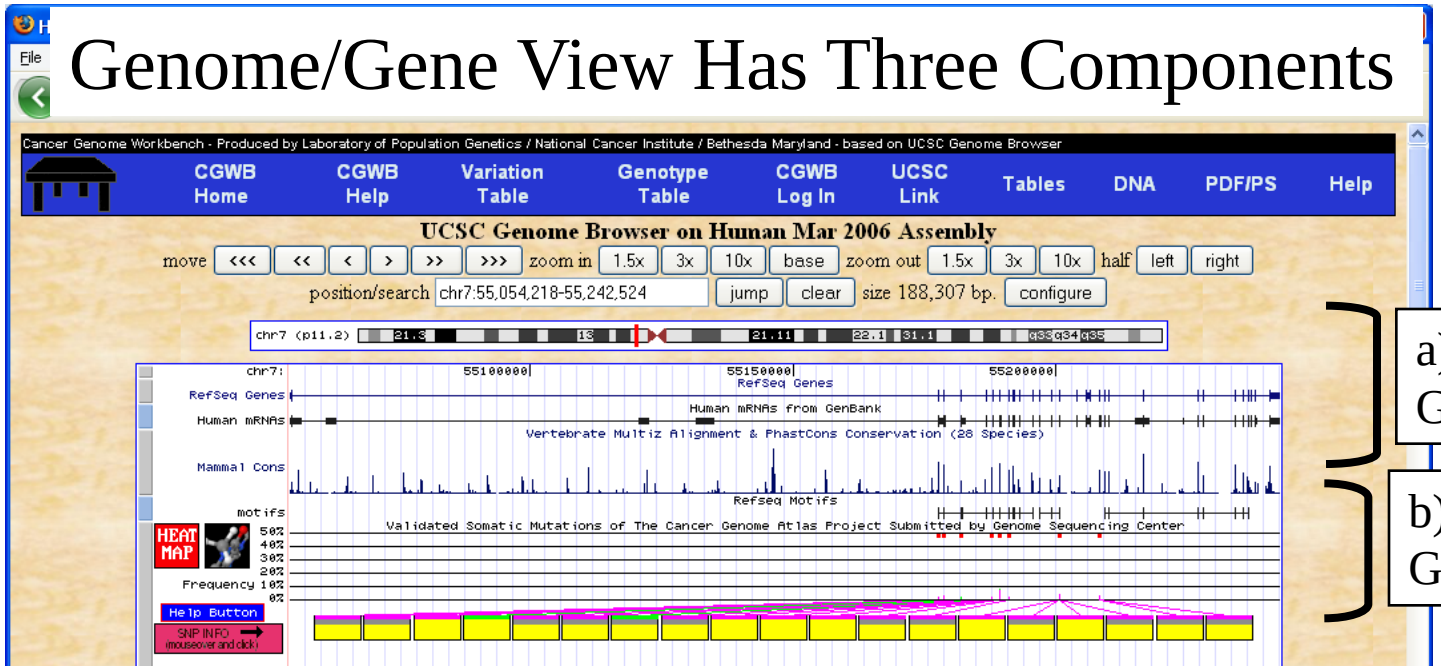
- Colored horizontal bar=% of samples that have amplifications or deletions. Mouse over to see the number for each region. Click to jump to a 1-Mb interval centered around the alteration

- 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

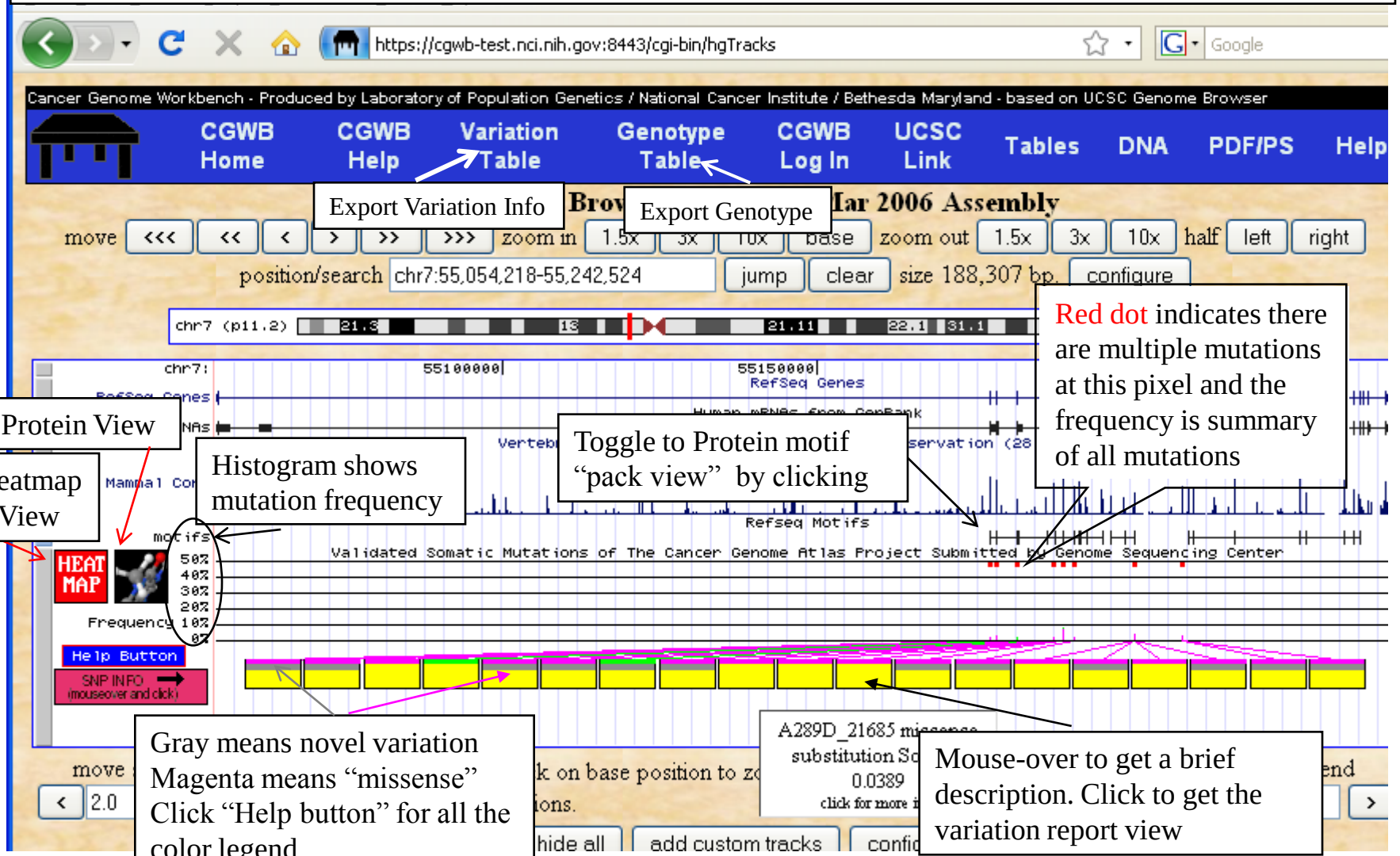


a) Reference Genome View

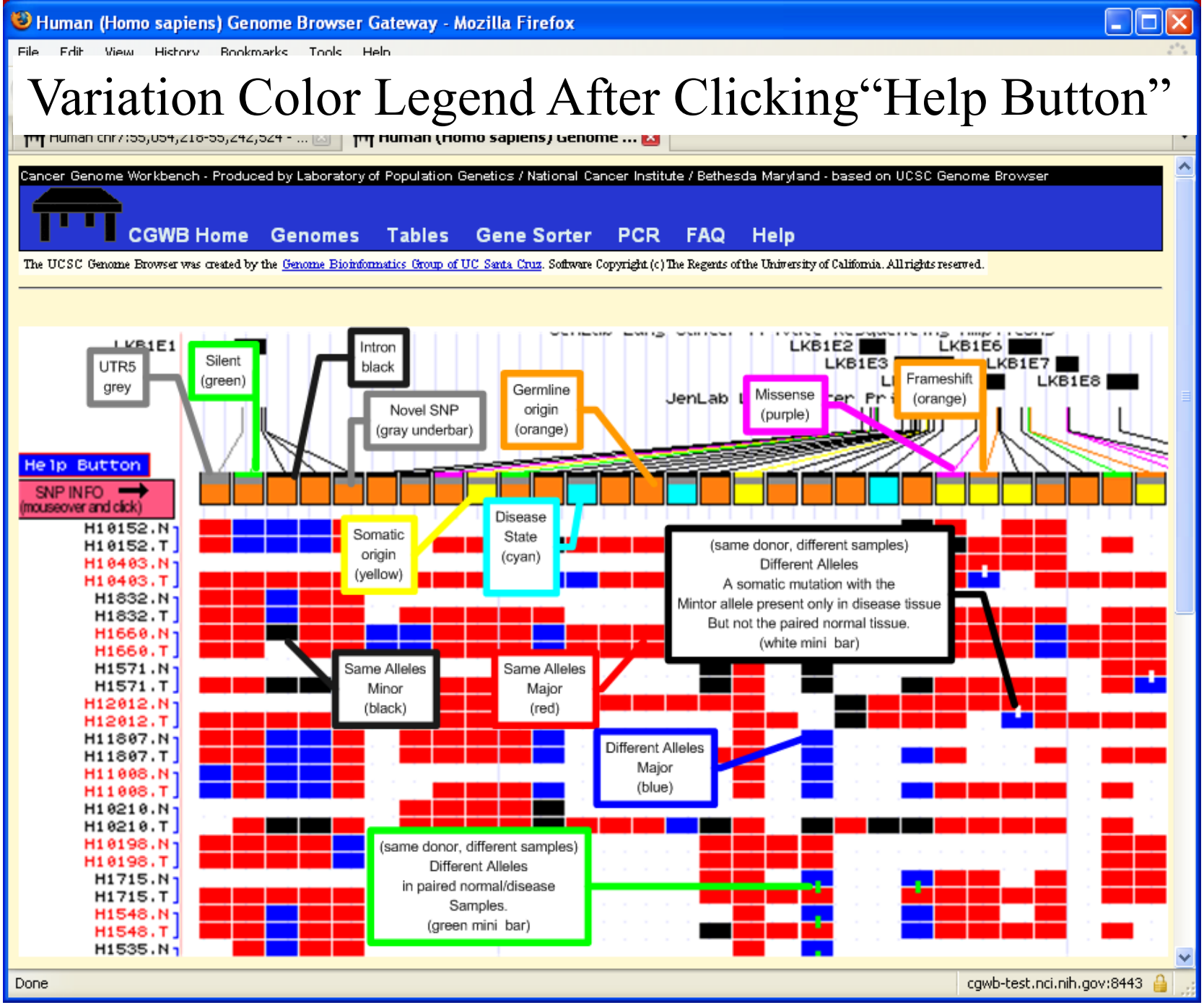
b) Tumor Genome View

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

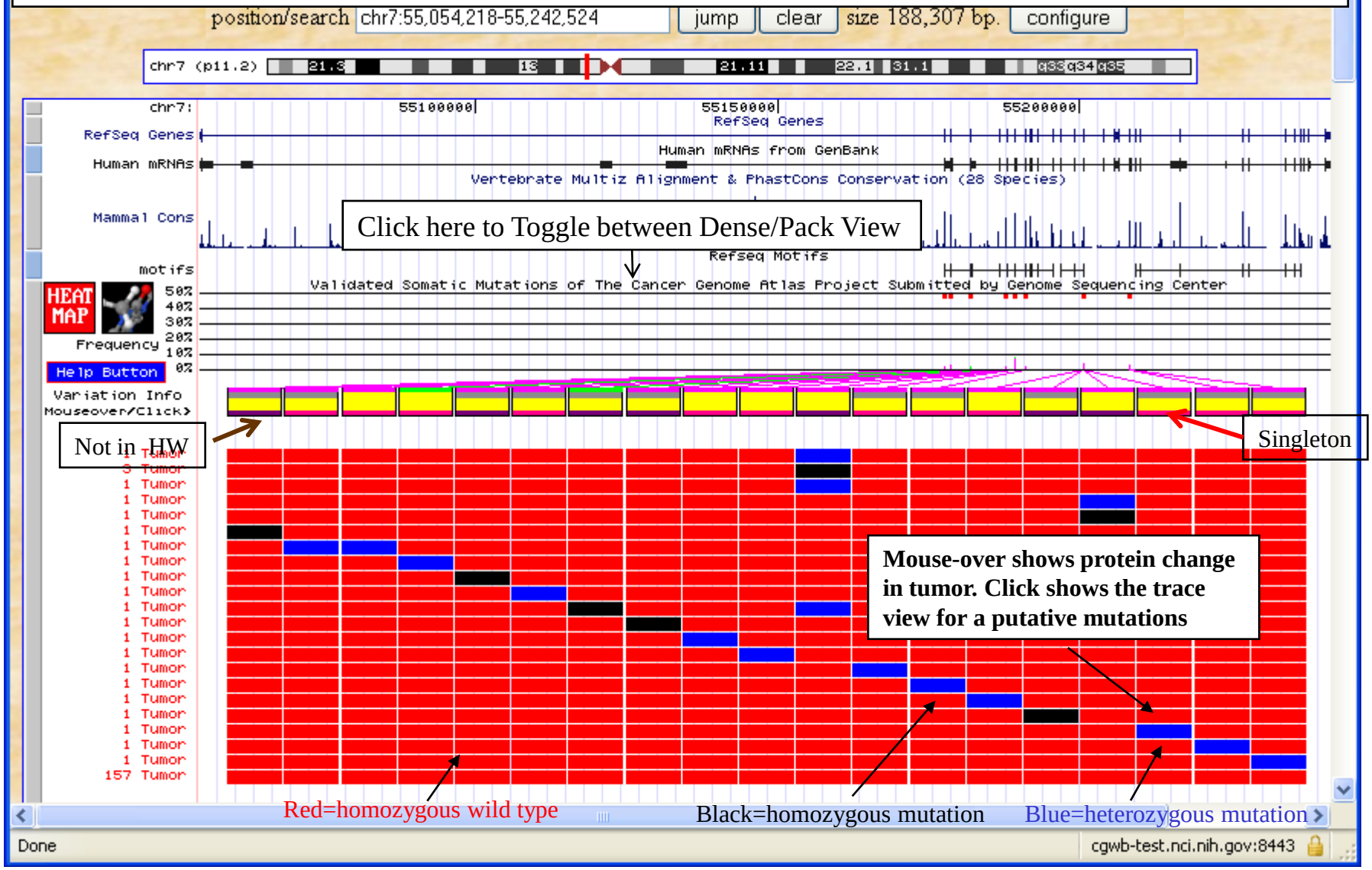
Genome View: a "Dense" View Summarizes Validated Mutations in EGFR



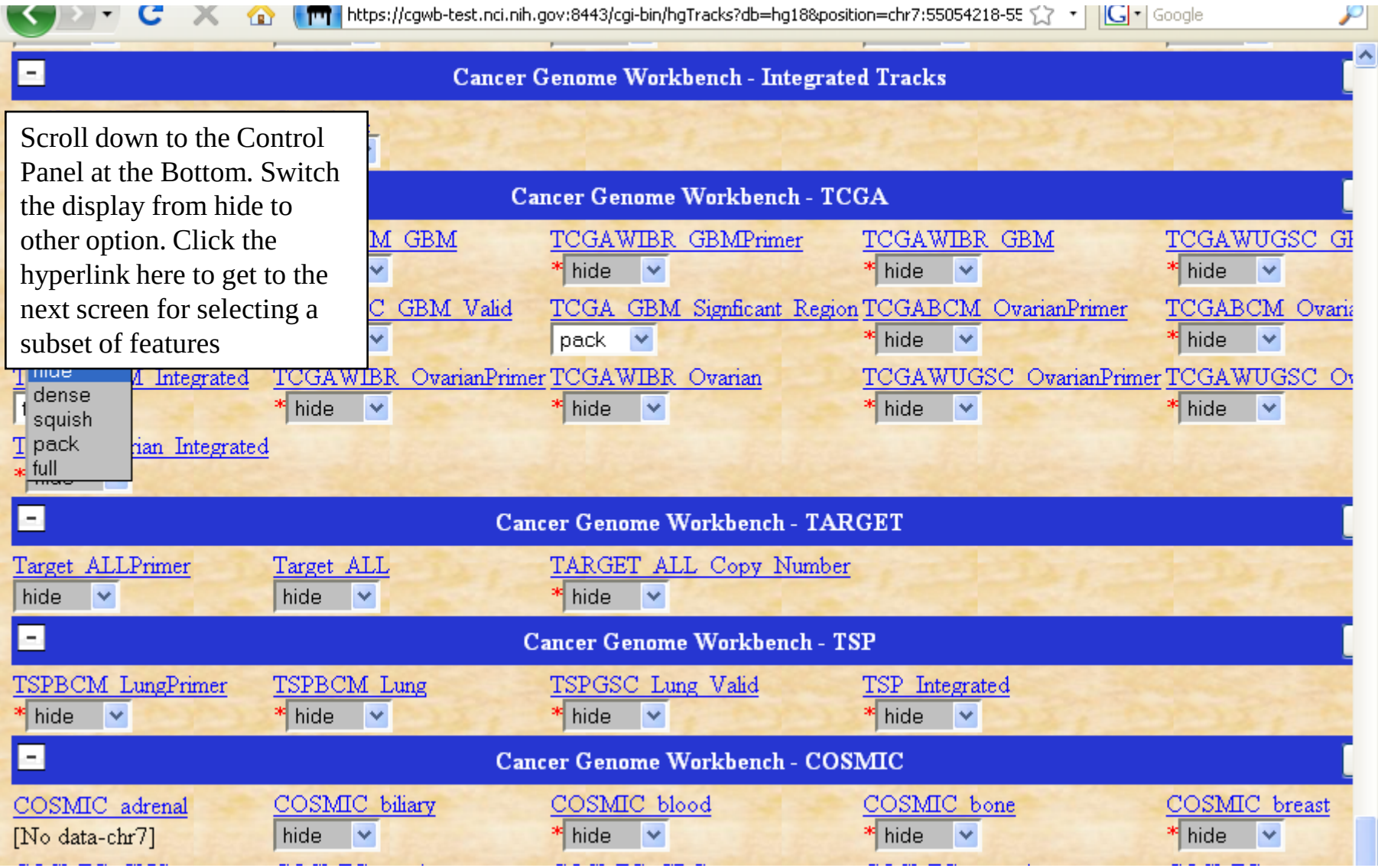
Variation Color Legend After Clicking"Help Button"



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



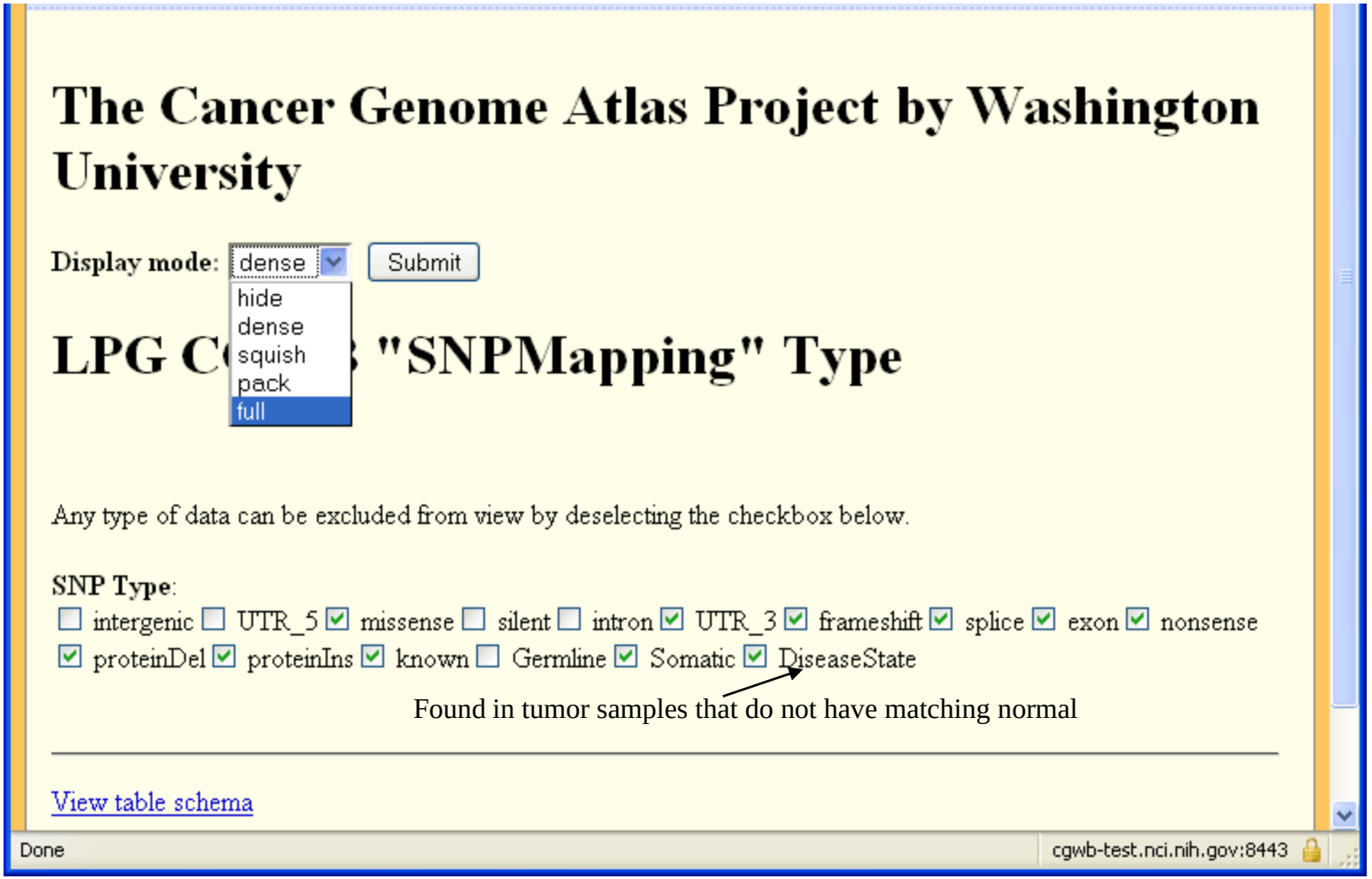
Select a Project or a Subset Data in the Browser View



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

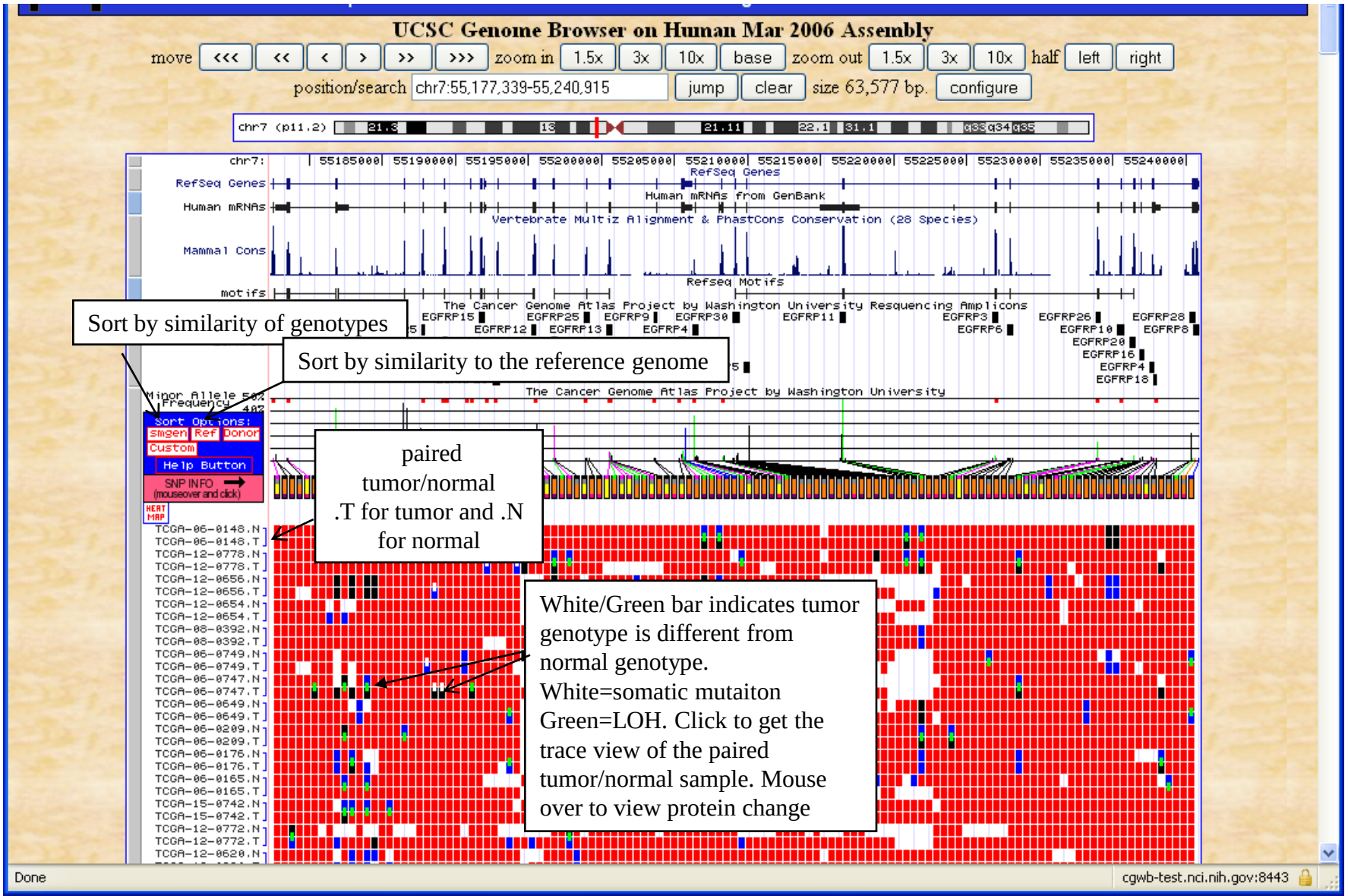
1)Select a Subset of Variations: missense, silent, somatic, novel?

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

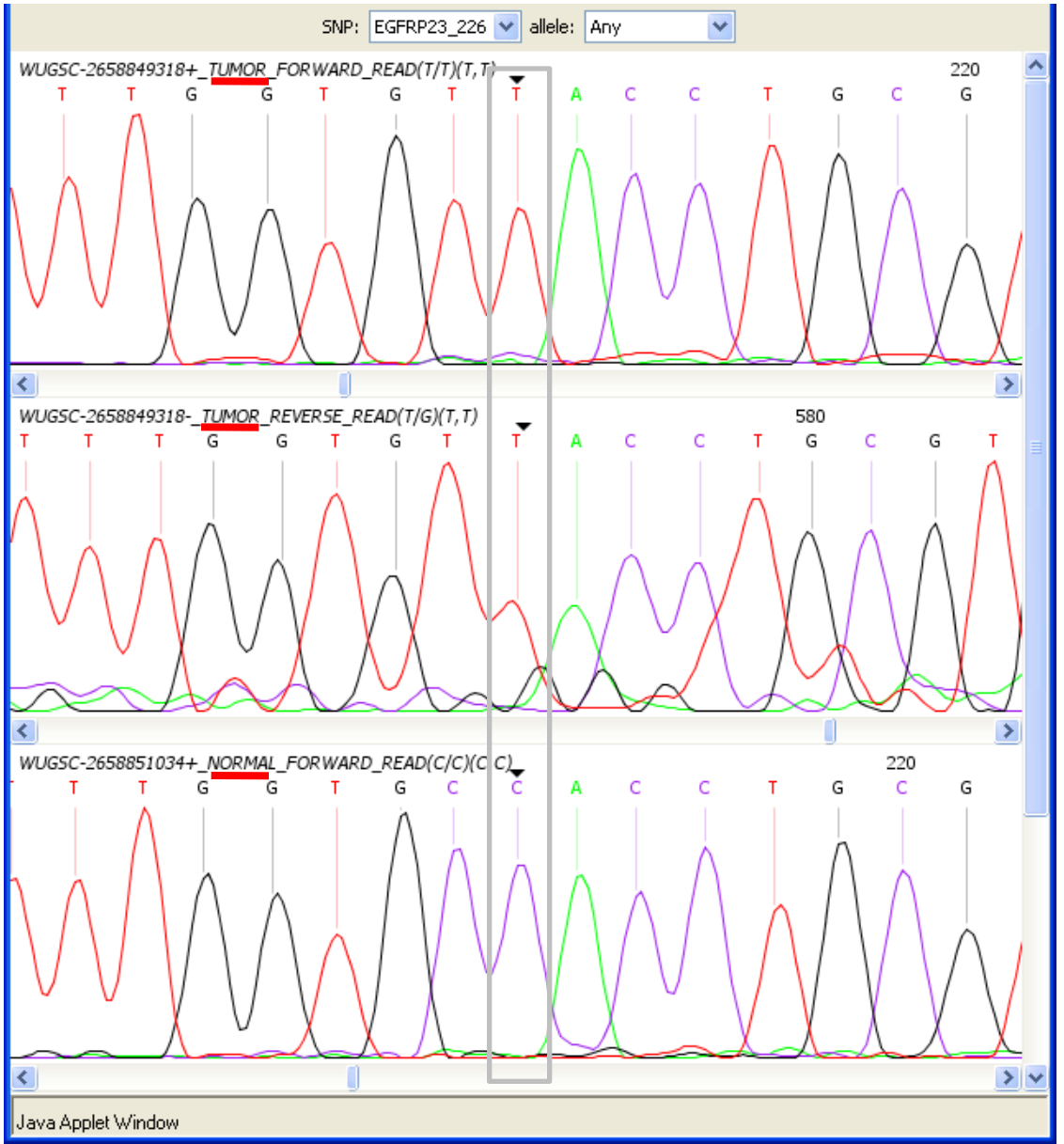


“Full” View Shows Paired Tumor/Normal Genotype for Each Sample

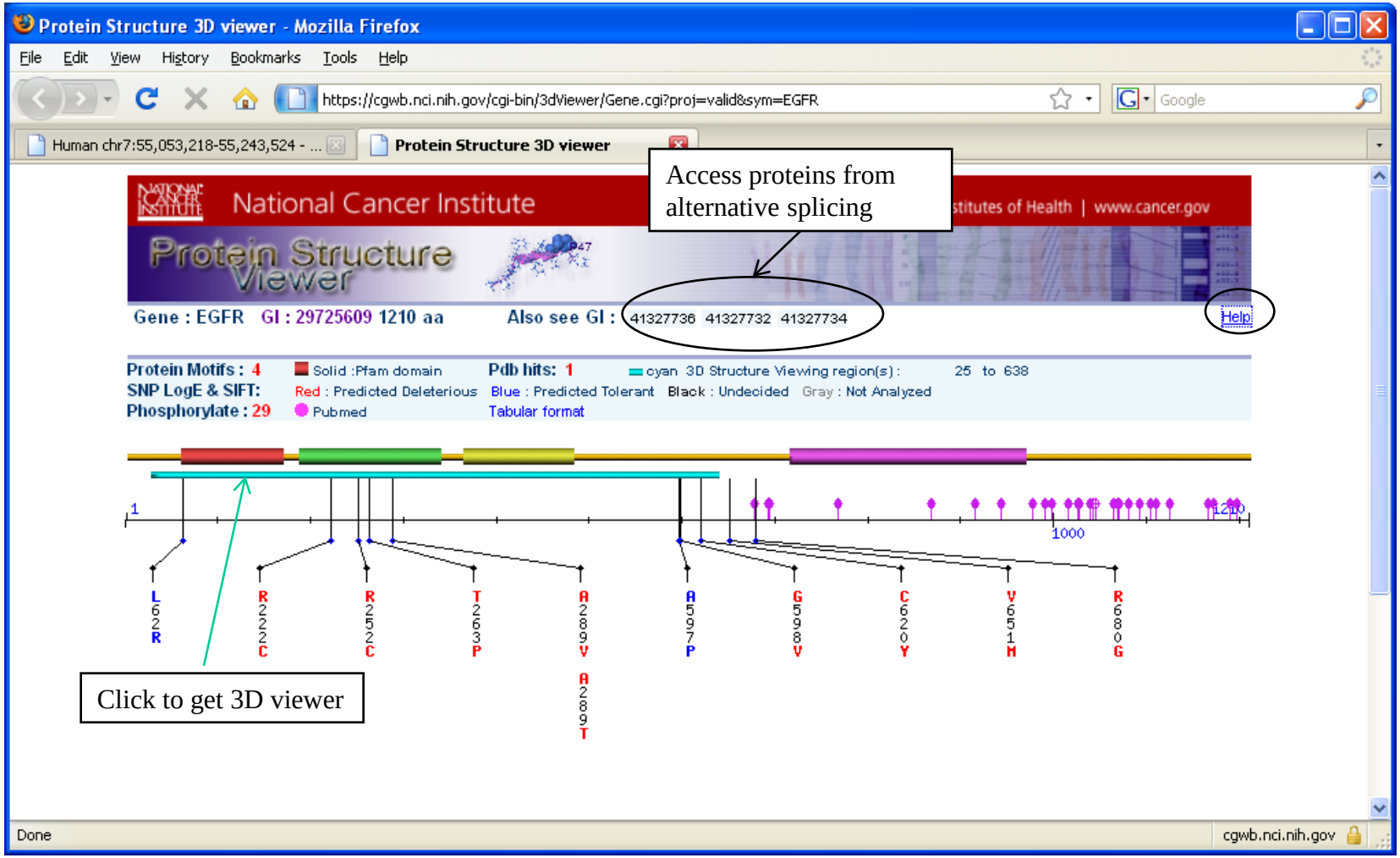
This view takes a while to generate for unvalidated sequence variations and it is meaningless unless you have applied for access control for the protected data to the real sample name.



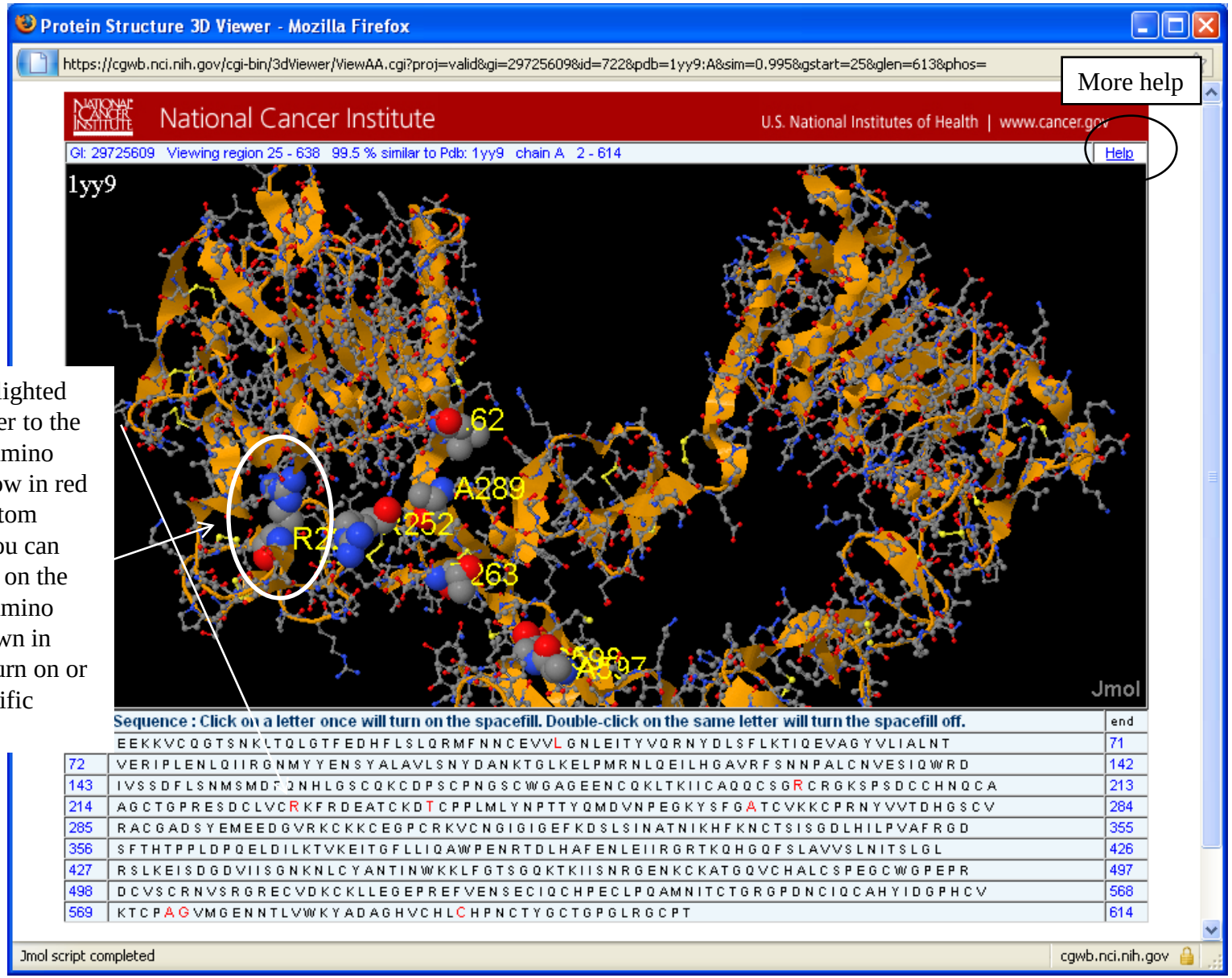
Trace View centered around a putative mutation with paired tumor/normal sample



Protein Structure Viewer

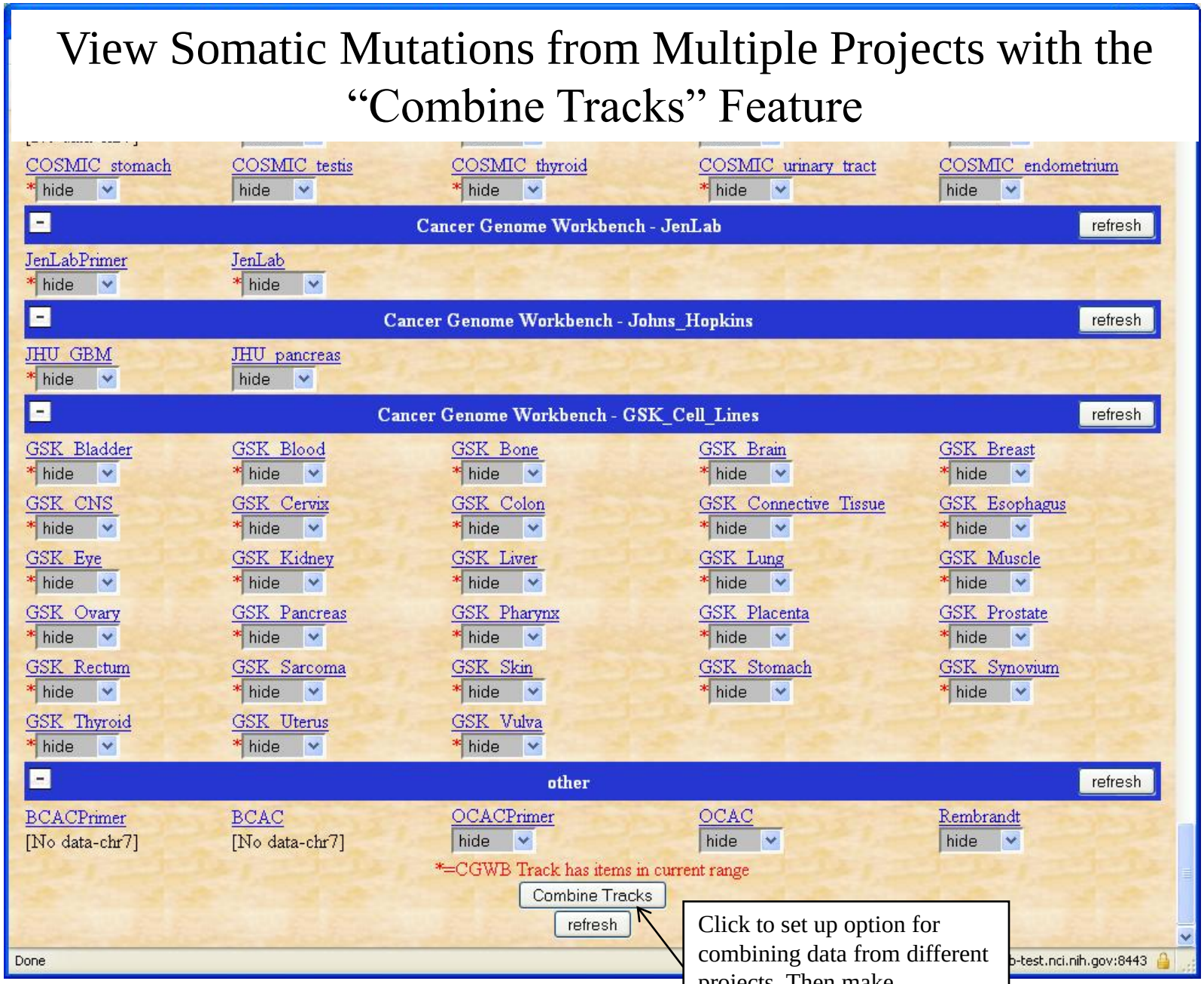


3D Structure Viewer



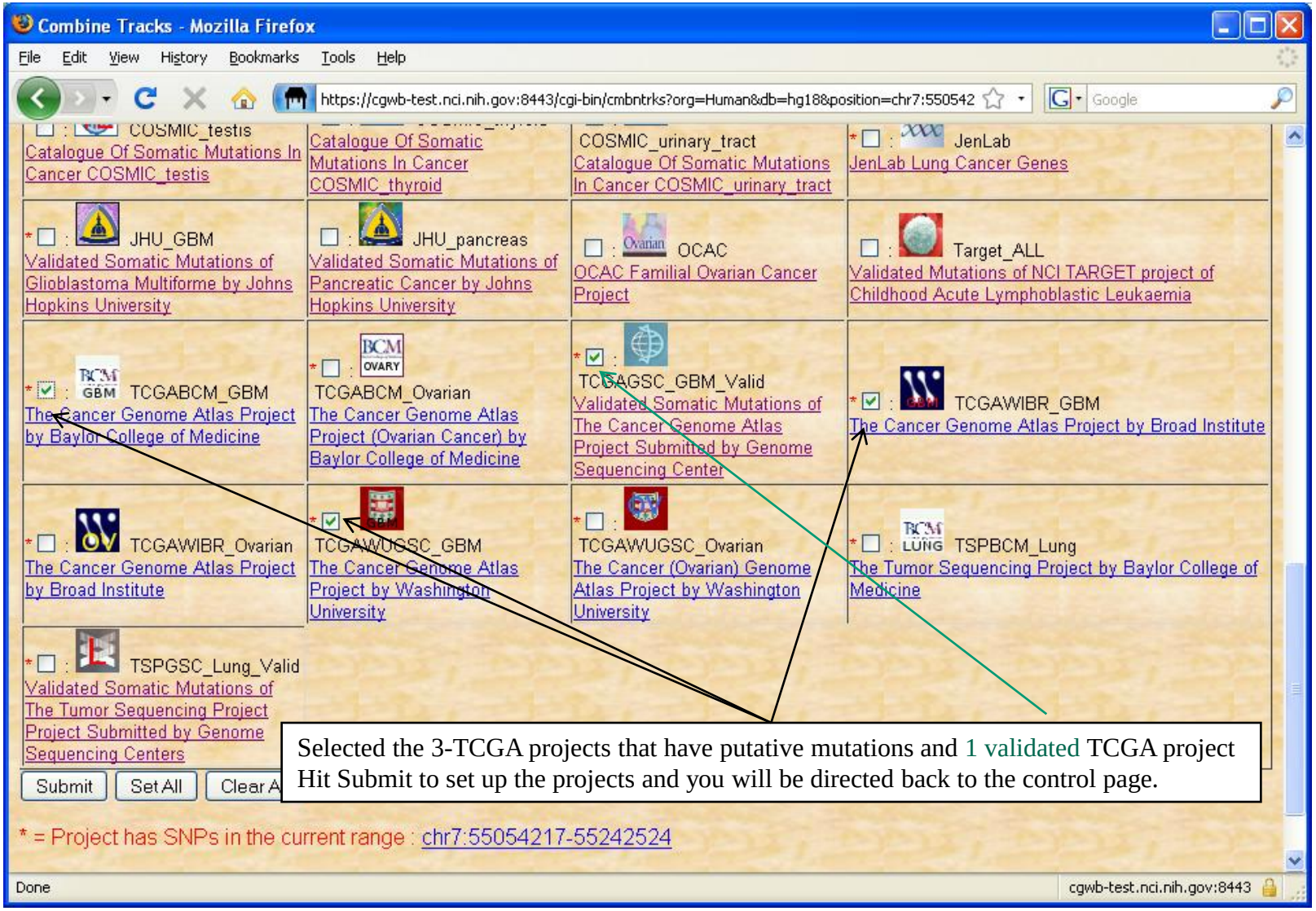
Big, highlighted atoms refer to the mutated amino acids (show in red in the bottom panel). You can also click on the mutated amino acid (shown in read) to turn on or off a specific mutation

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



Scroll to bottom

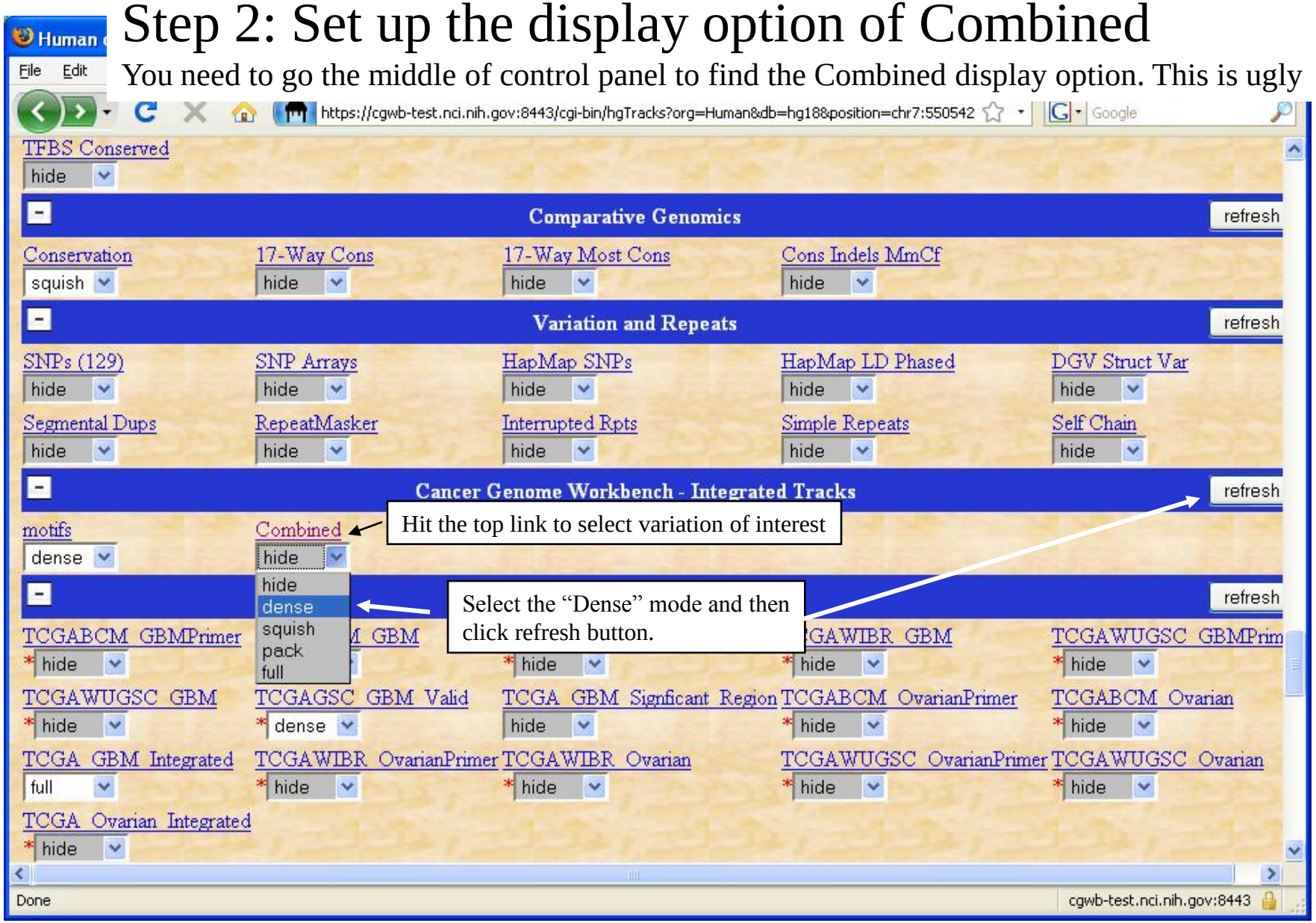
Step 1) Select Data from Different Projects/Groups



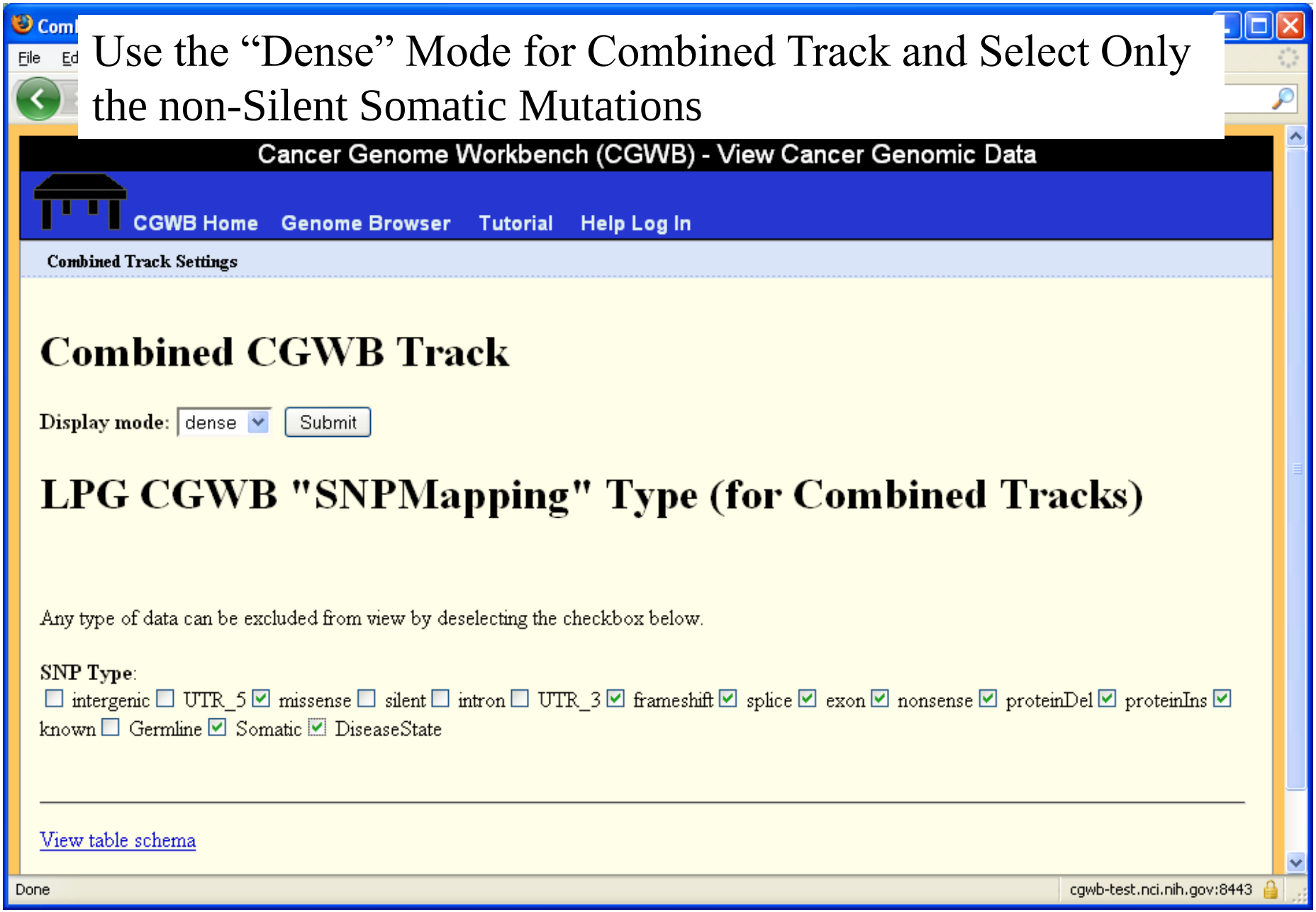
Selected the 3-TCGA projects that have putative mutations and 1 validated TCGA project
Hit Submit to set up the projects and you will be directed back to the control page.

Step 2: Set up the display option of Combined

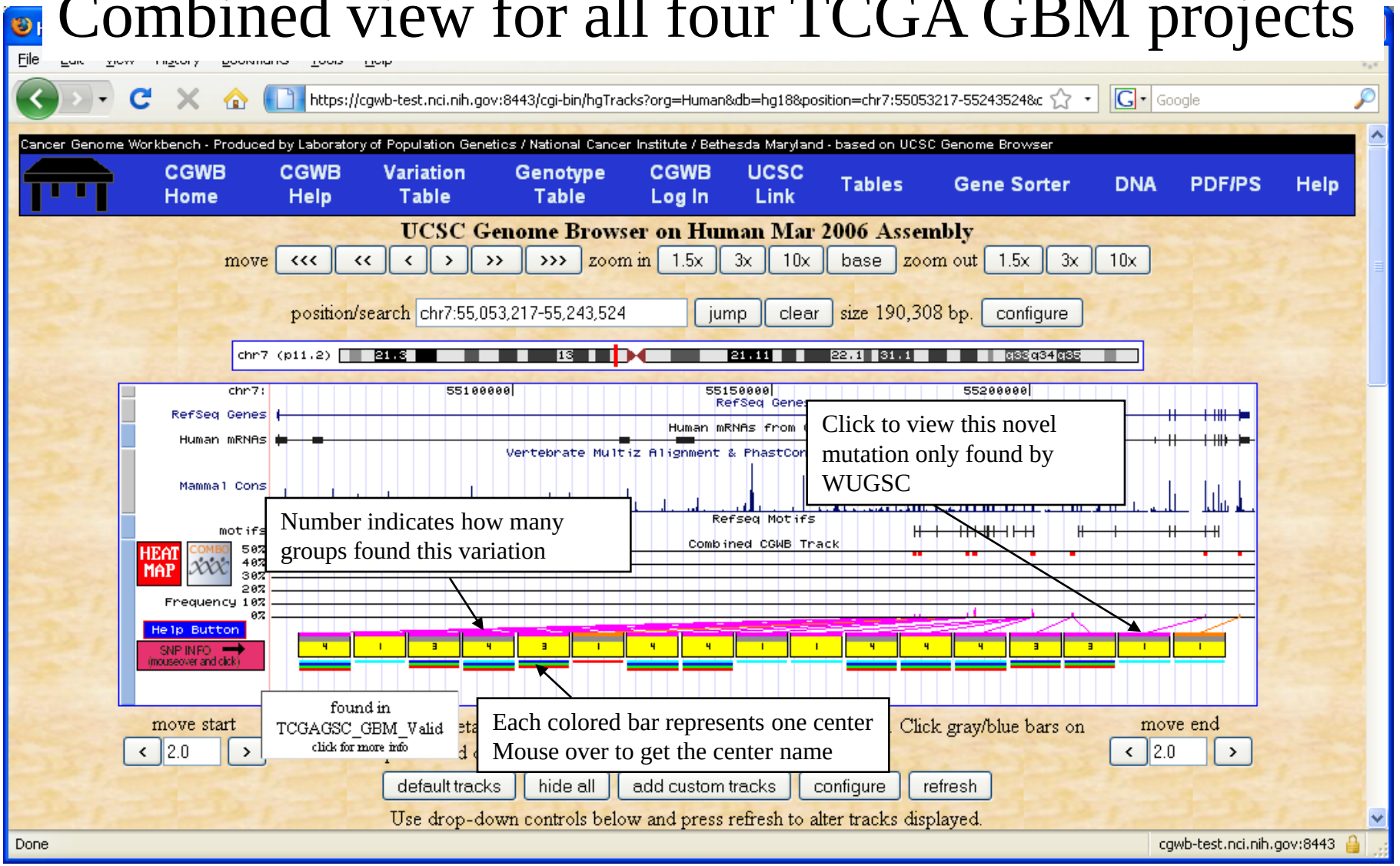
You need to go the middle of control panel to find the Combined display option. This is ugly



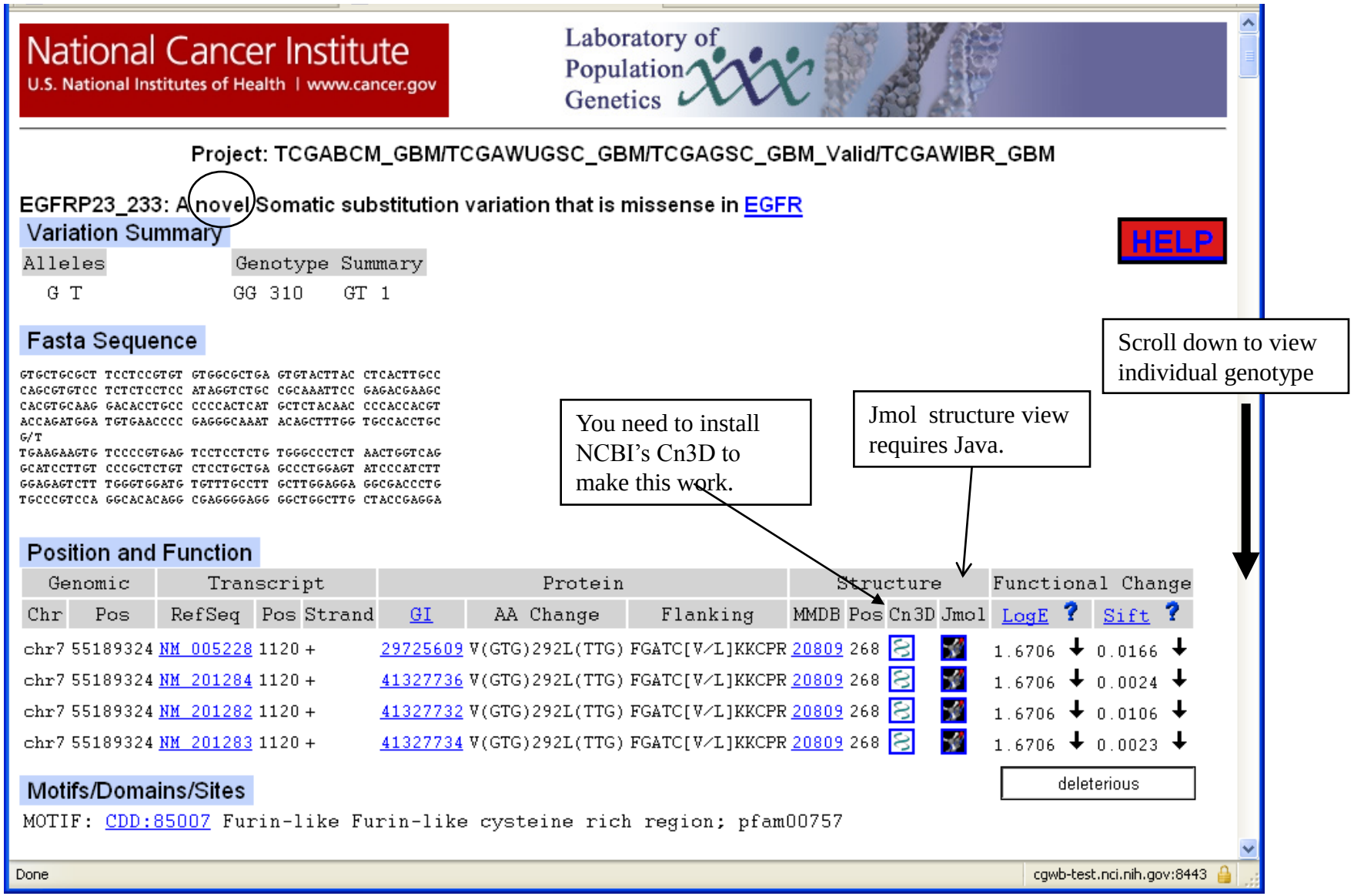
Use the “Dense” Mode for Combined Track and Select Only the non-Silent Somatic Mutations



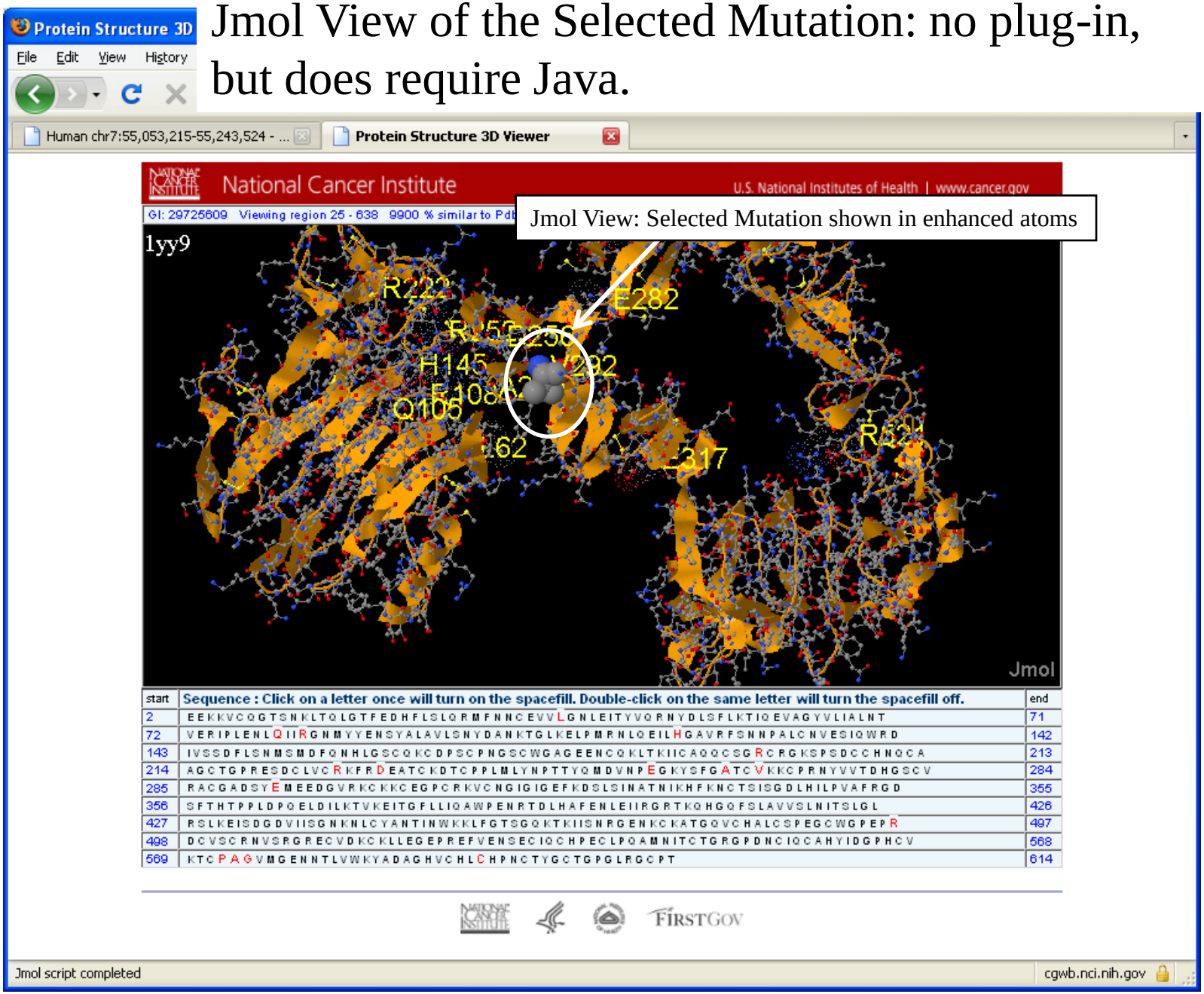
Combined view for all four TCGA GBM projects



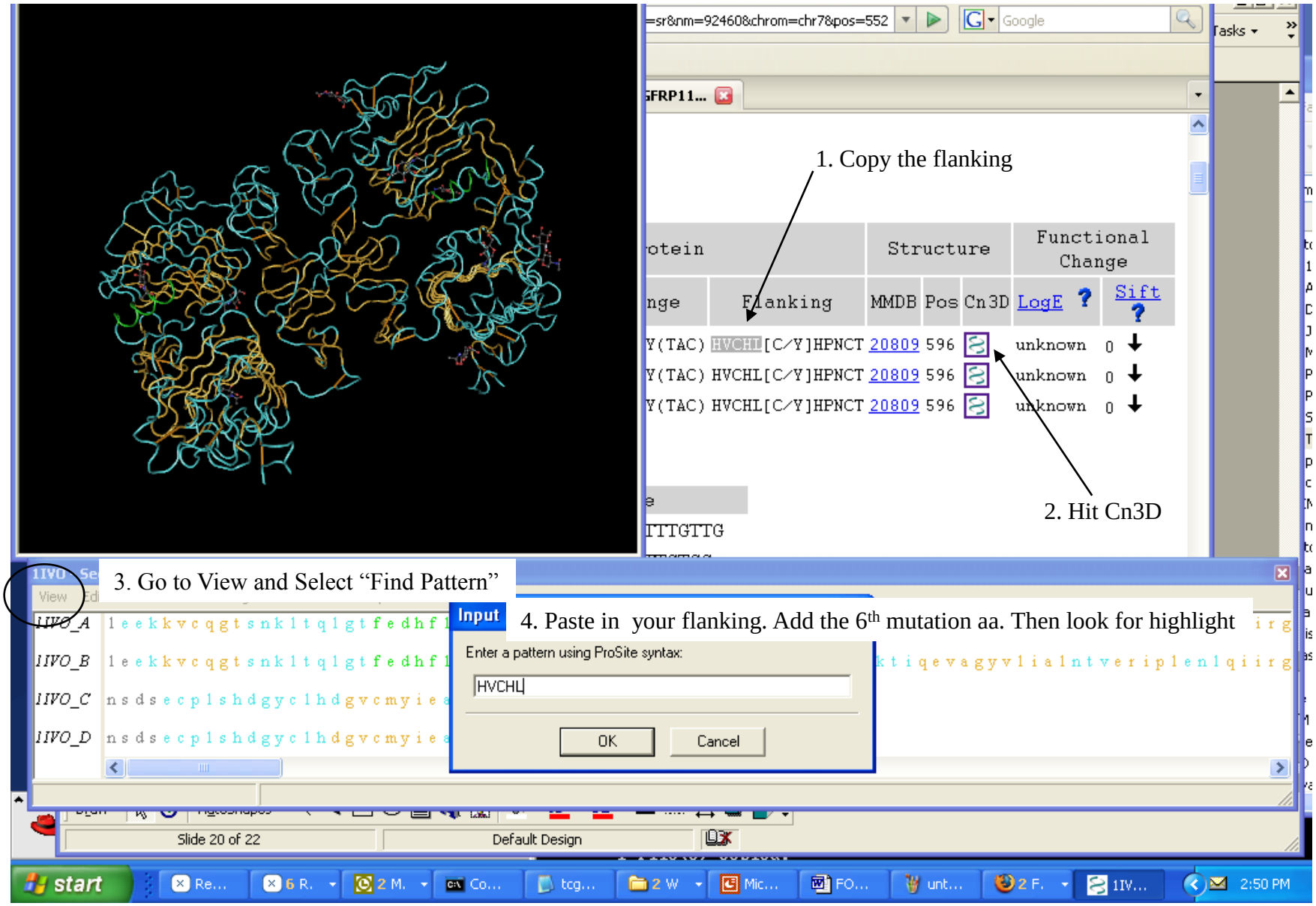
Detailed Mutation Report Generated by Clicking on Variations at the Genome/Gene View

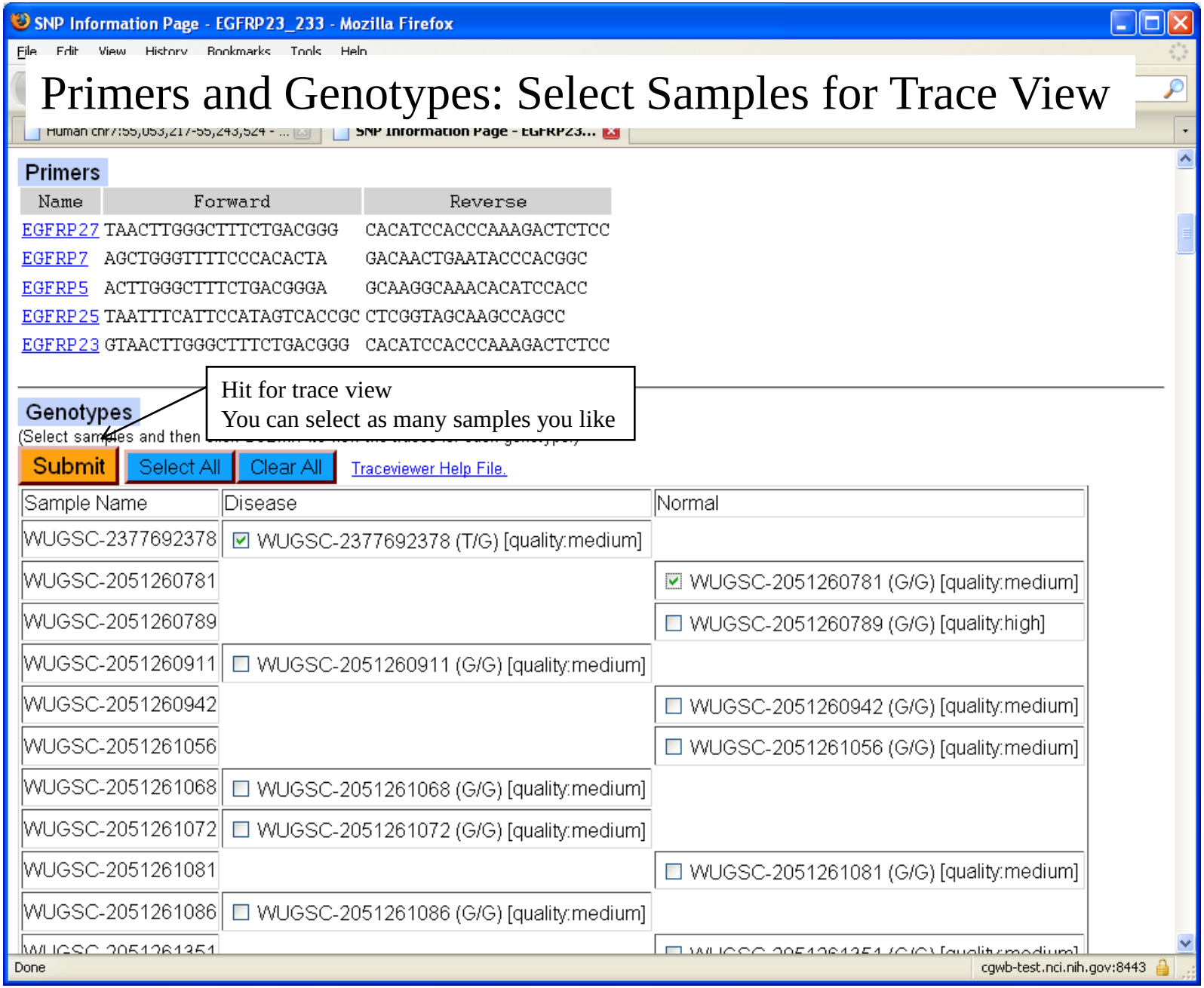


Jmol View of the Selected Mutation: no plug-in, but does require Java.



How to Highlight Mutation on Cn3D View?




SNP Information Page - EGFRP23_233 - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Primers and Genotypes: Select Samples for Trace View
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
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]
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

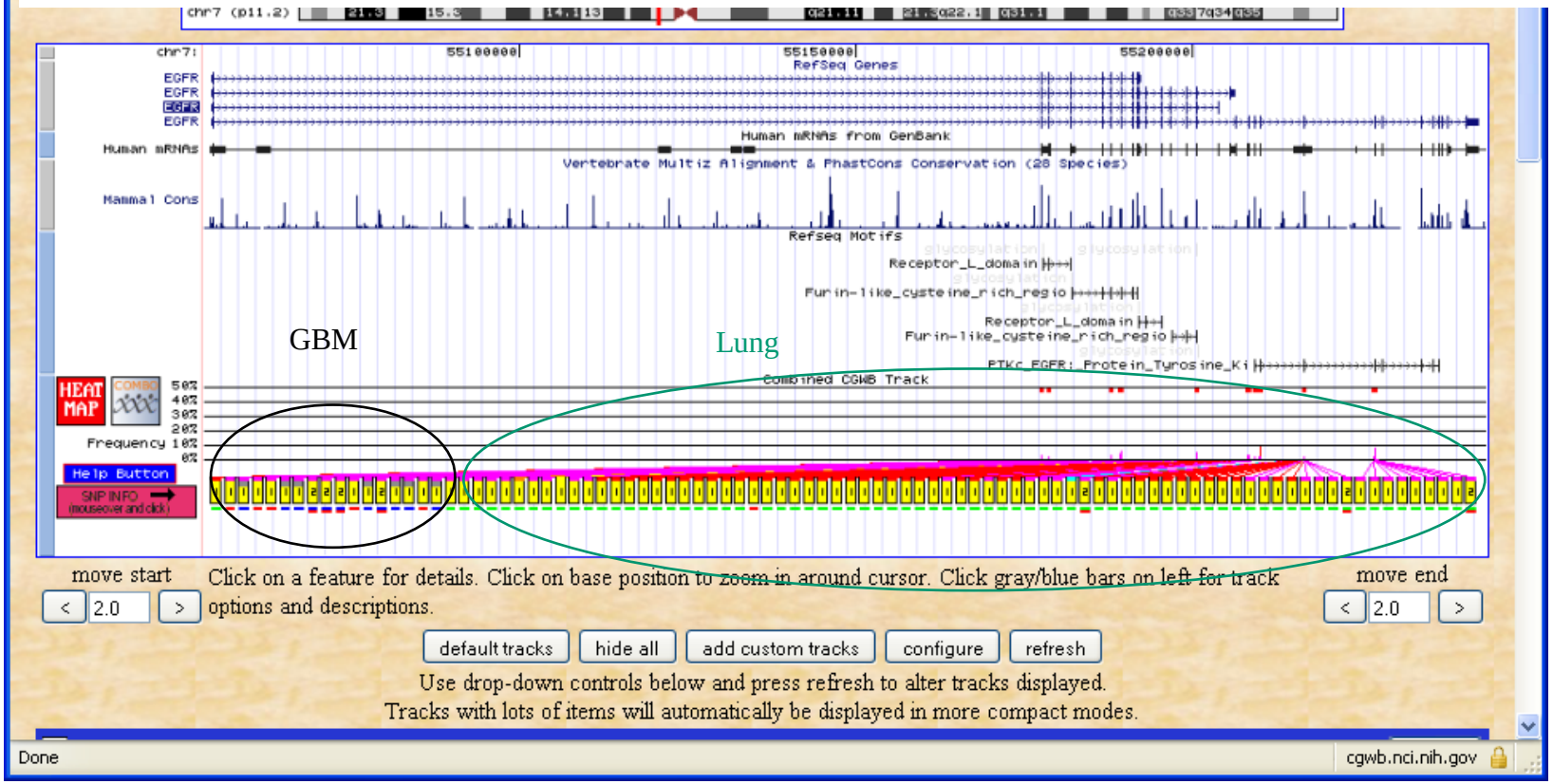
Combined View of EGFR mutation in Lung versus in GBM

COSMIC_Lung in green underline

TCGA_GBM in blue underline

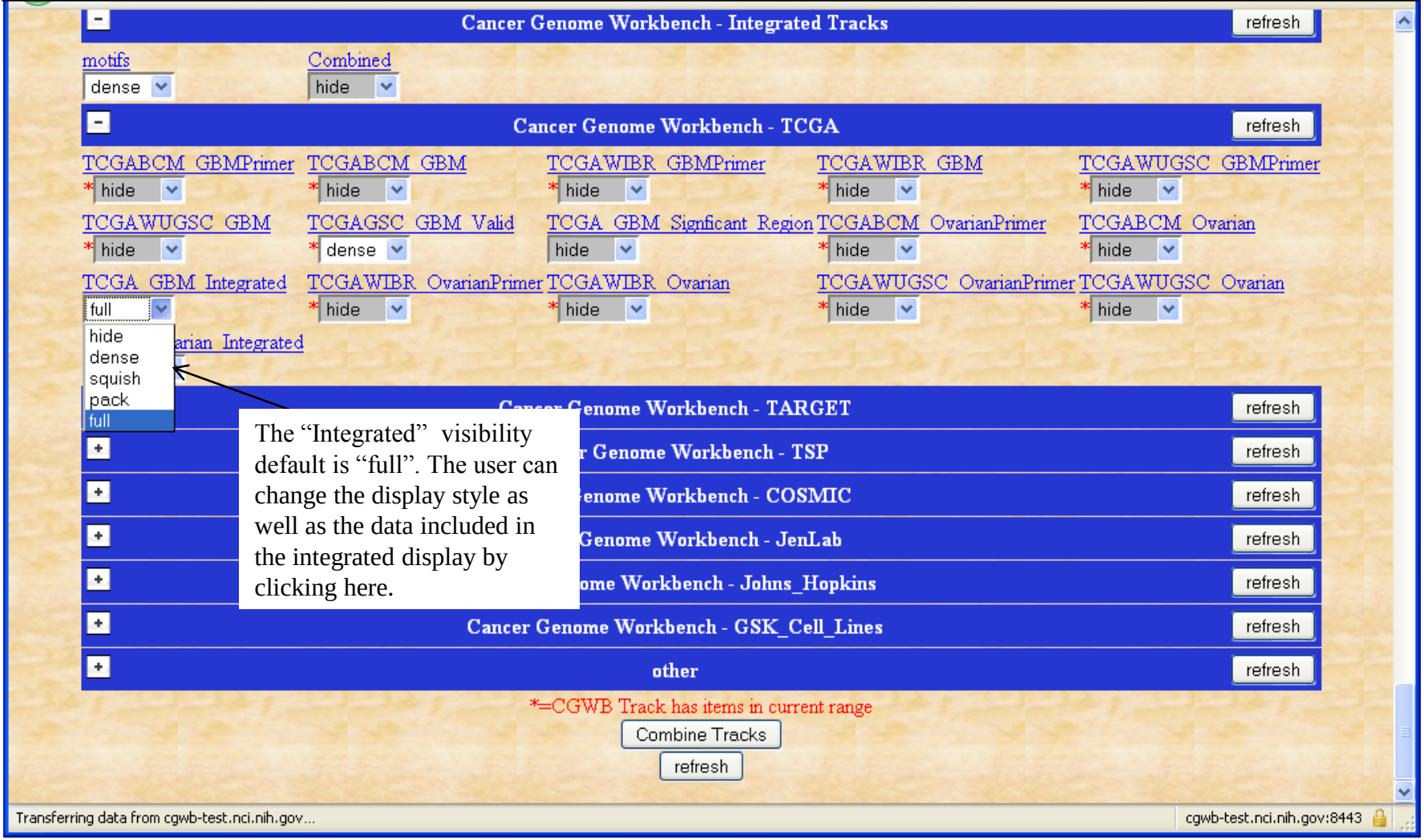
COSMIC_GBM in red underline

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)

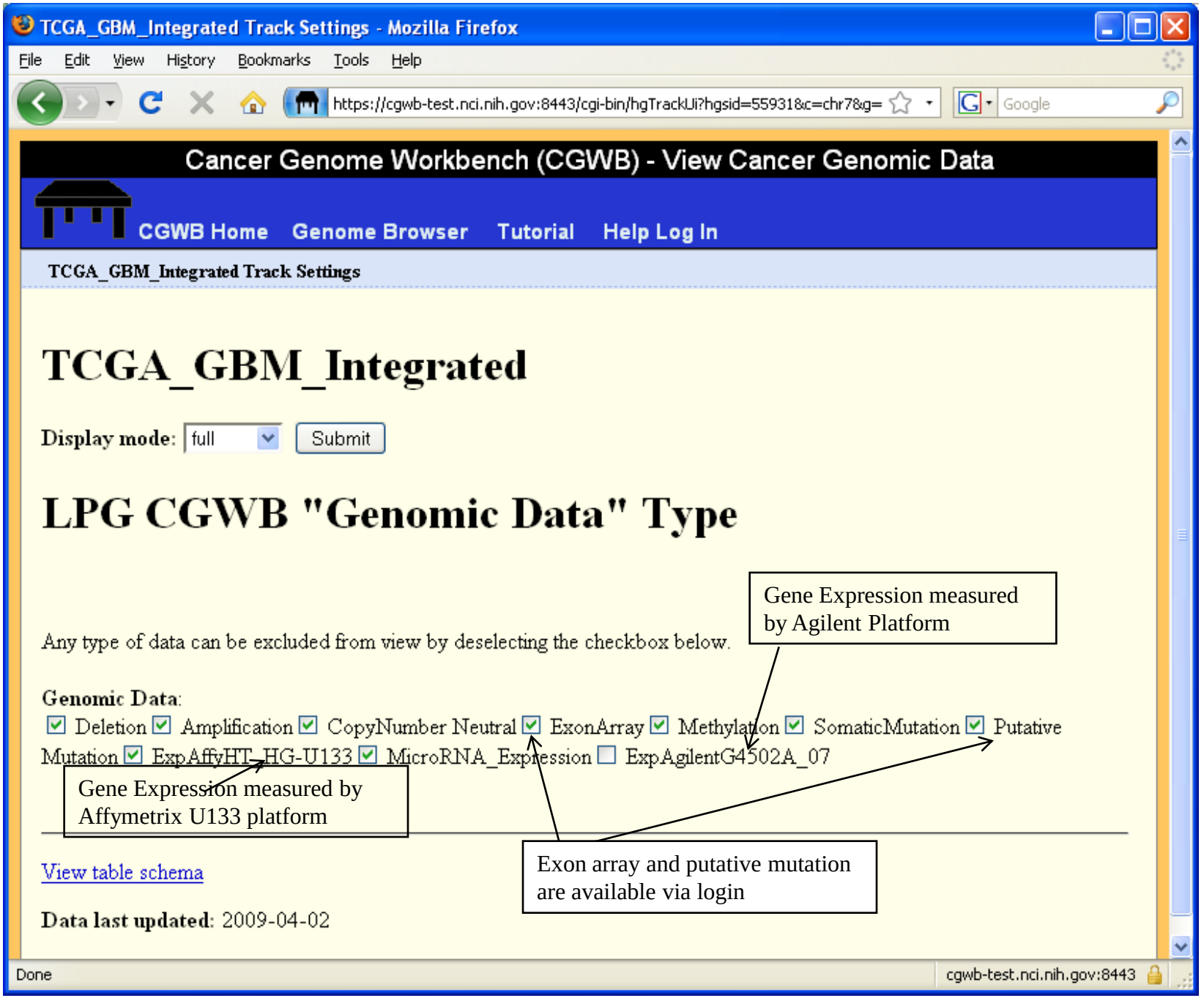


View Integrated Data of Copy Number, Exon Expression, Somatic Mutation, Methylation and MicroRNA

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



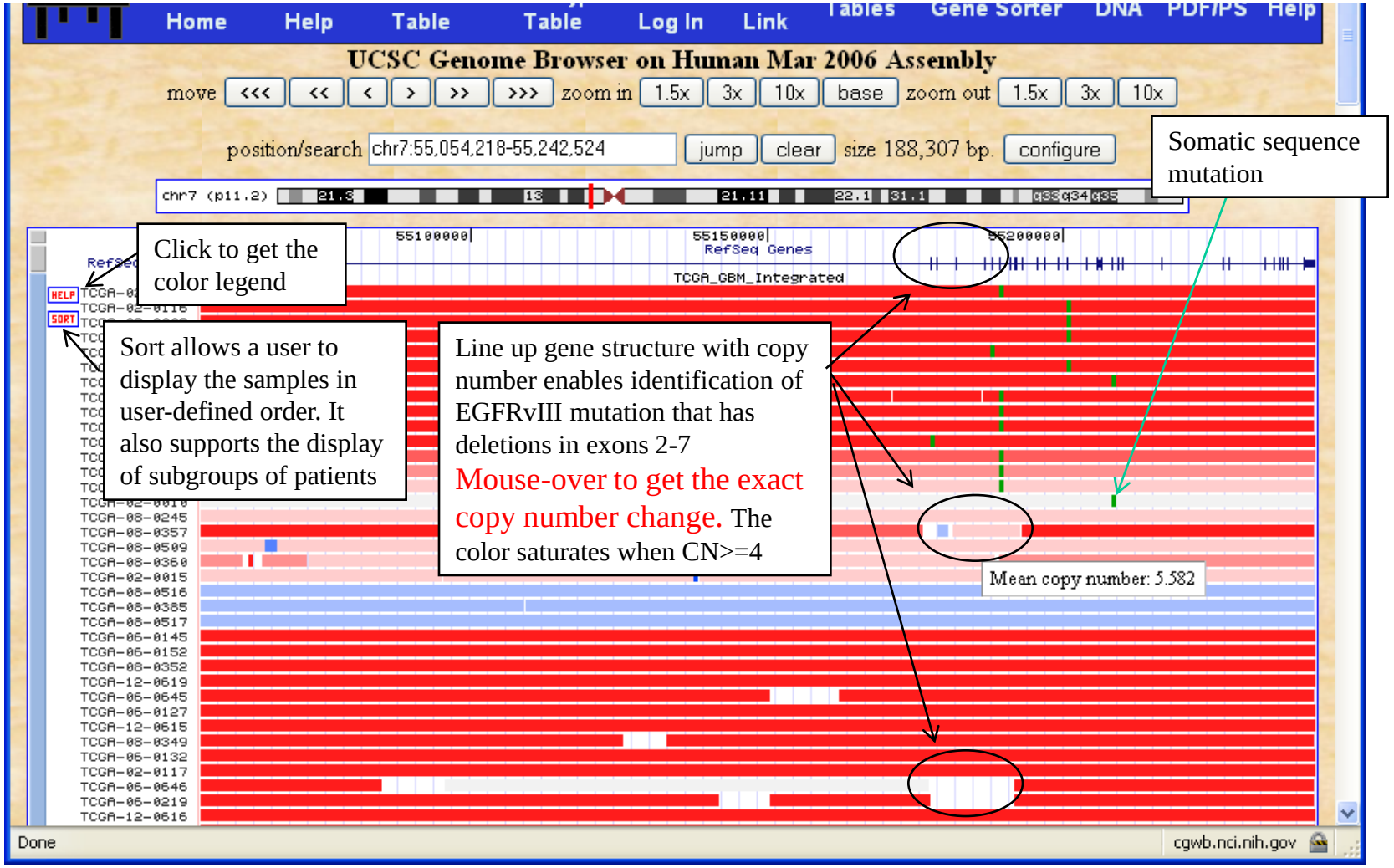
The "Integrated" visibility default is "full". The user can change the display style as well as the data included in the integrated display by clicking here.


TCGA_GBM_Integrated Track Settings - Mozilla Firefox
File Edit View History Bookmarks Tools Help
https://cgwb-test.nci.nih.gov:8443/cgi-bin/hgTrackUi?hgsid=55931&c=chr7&g=
Google
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 ExpAffyHT_HG-U133 MicroRNA_Expression ExpAgilentG4502A_07
View table schema
Data last updated: 2009-04-02
Gene Expression measured by Agilent Platform
Gene Expression measured by Affymetrix U133 platform
Exon array and putative mutation are available via login
Done
cgwb-test.nci.nih.gov:8443

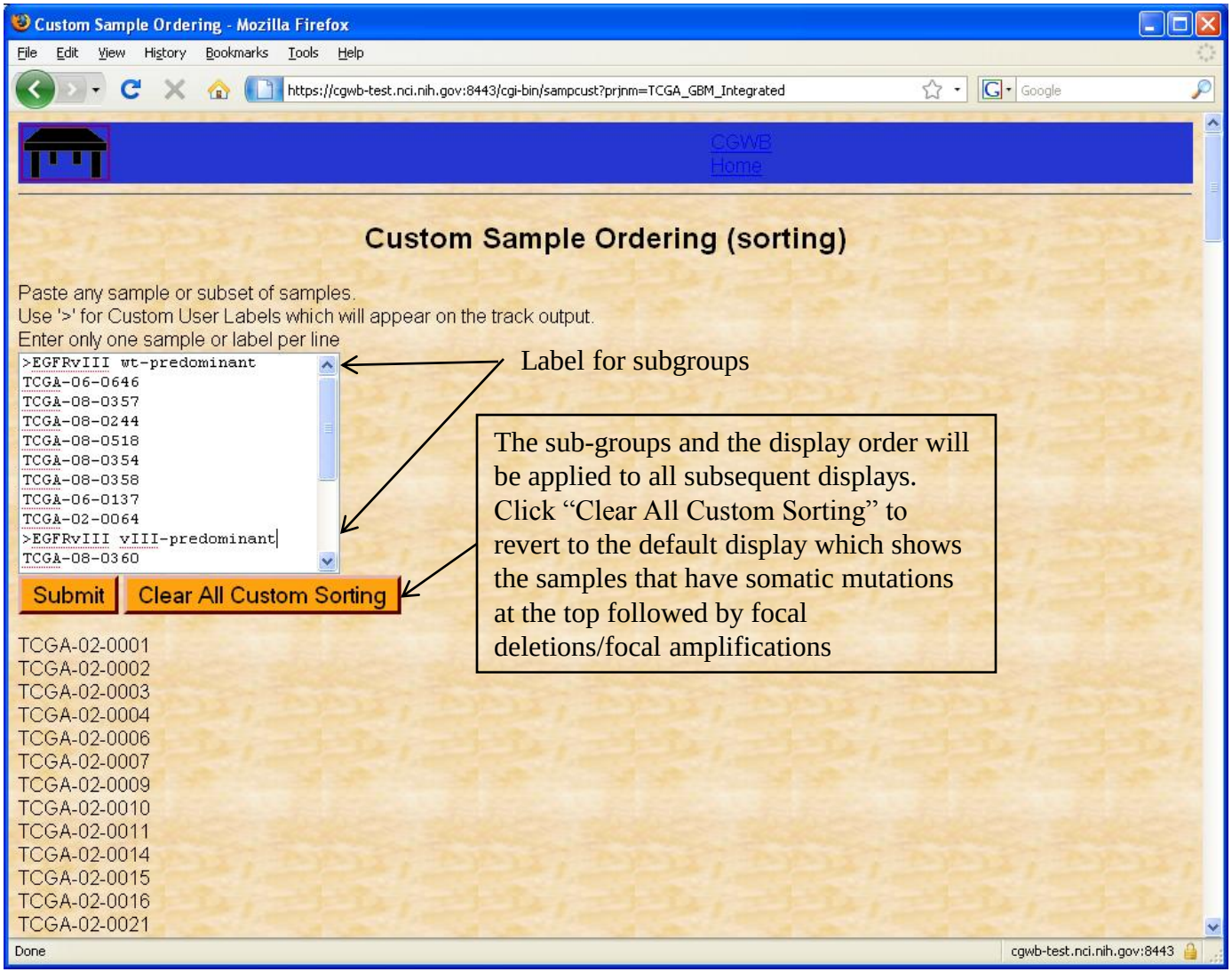
Integrated view of somatic sequence mutation and copy number change

Each line represents copy-number alteration in tumor superimposed with somatic mutations (green bars)

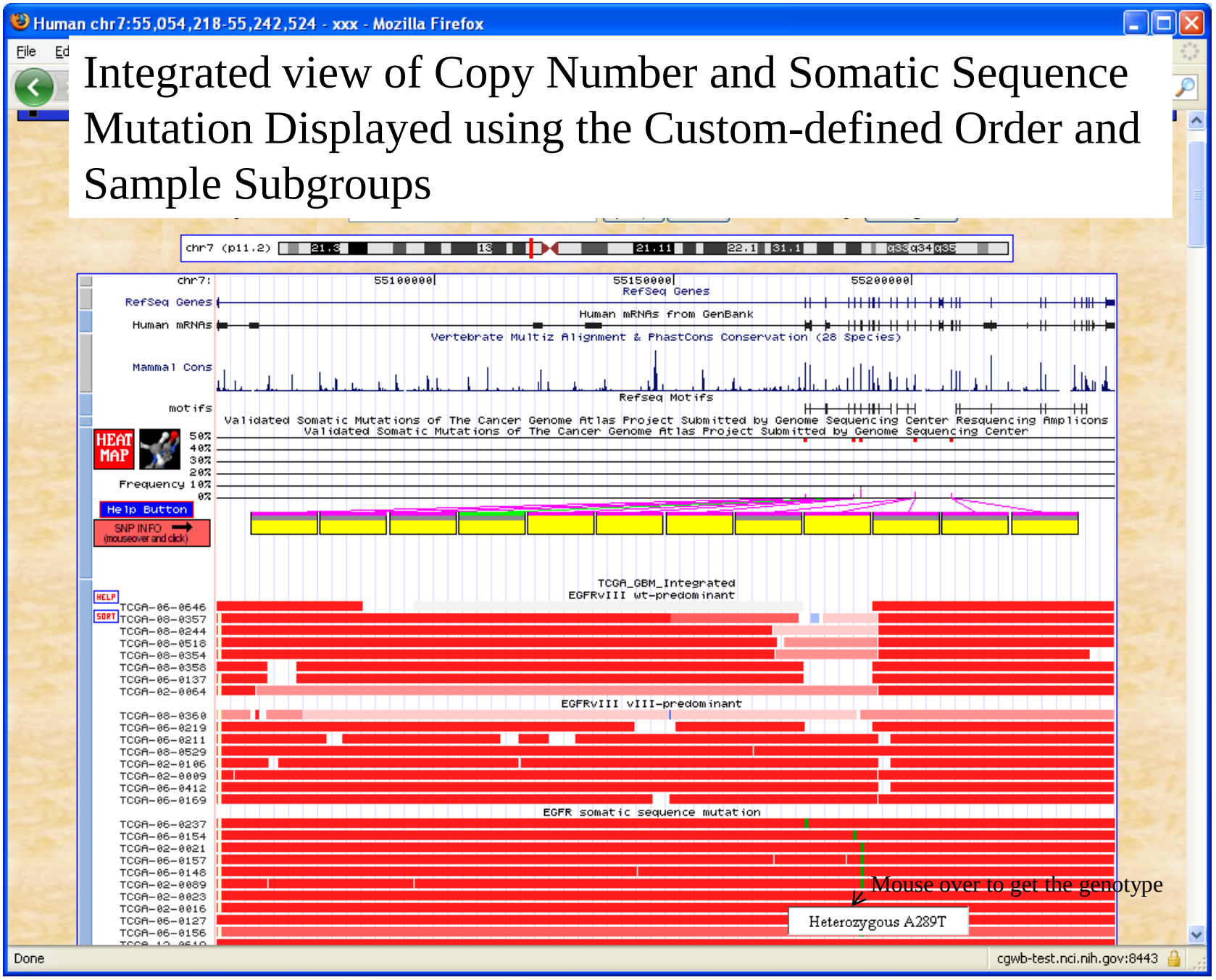
Mouse-over to get the exact copy number value or the amino-acid changes in each mutation



Create Customized Sample Display Order and Sample Subgroups



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

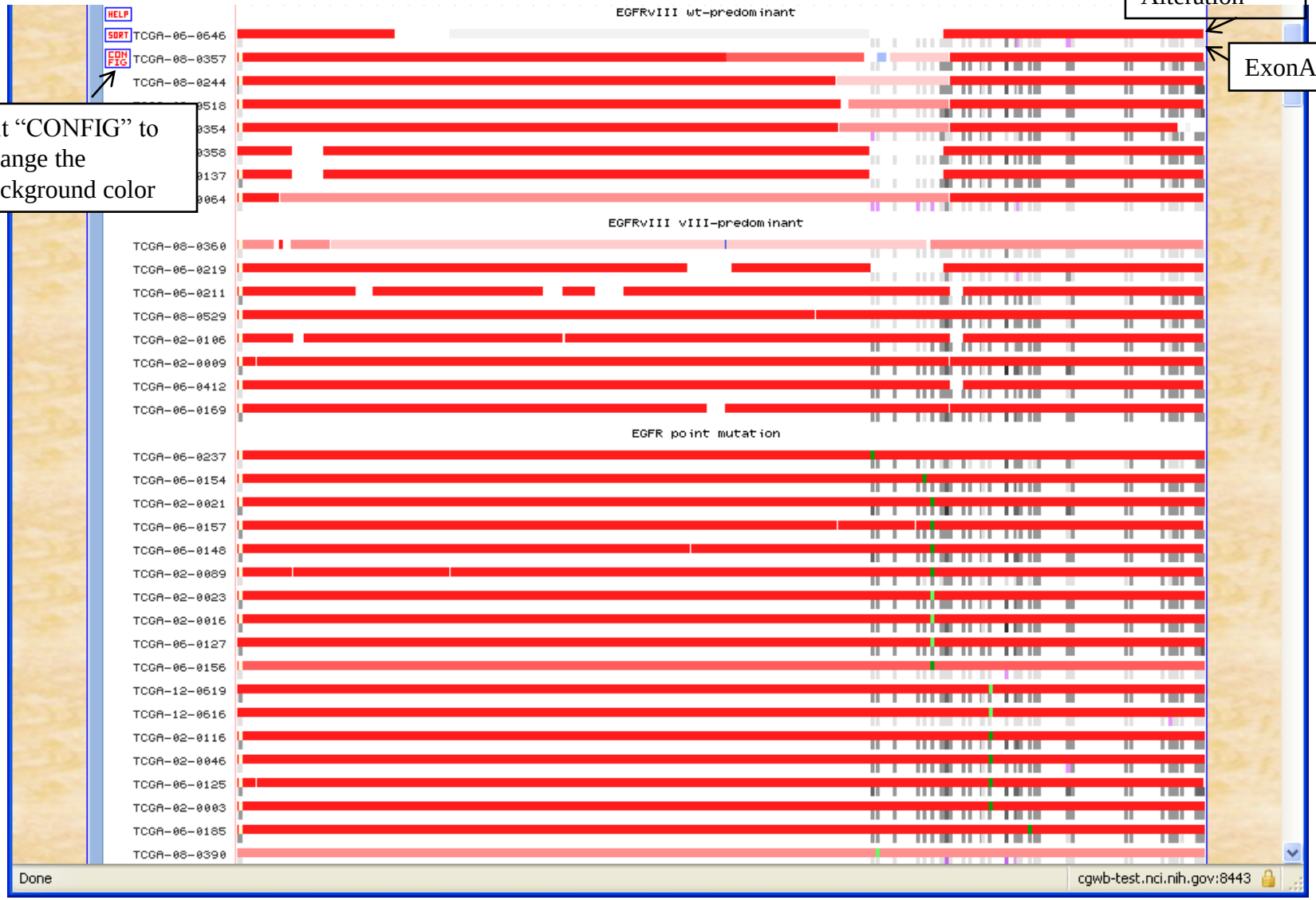


Integrated view with Exon Array and Putative Mutations Superimposed to the Copy Number/Validated Mutations (only users with login can access this view)

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

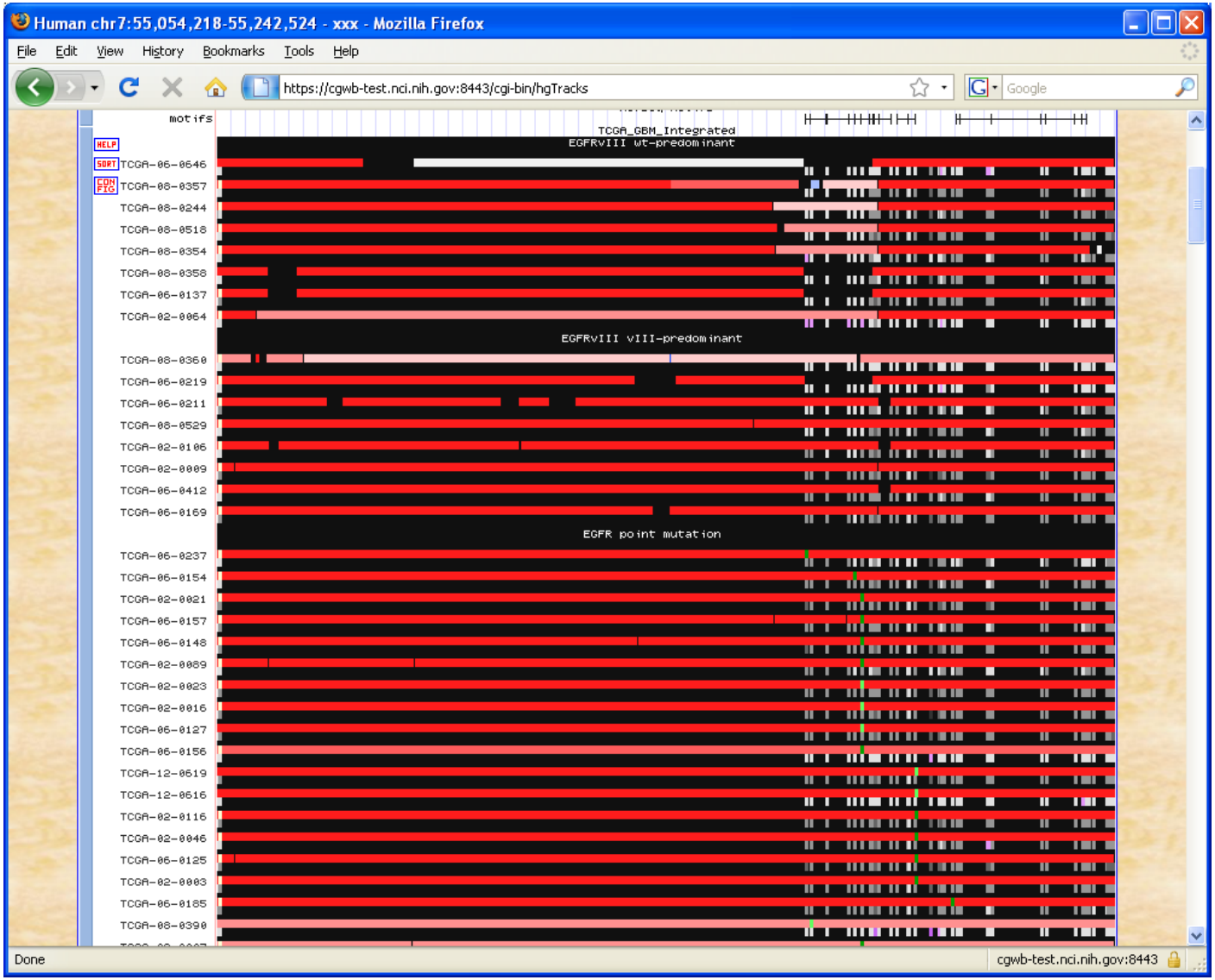
Top row Copy Number +Somatic mutation+Methylation

Bottom row Exon Array



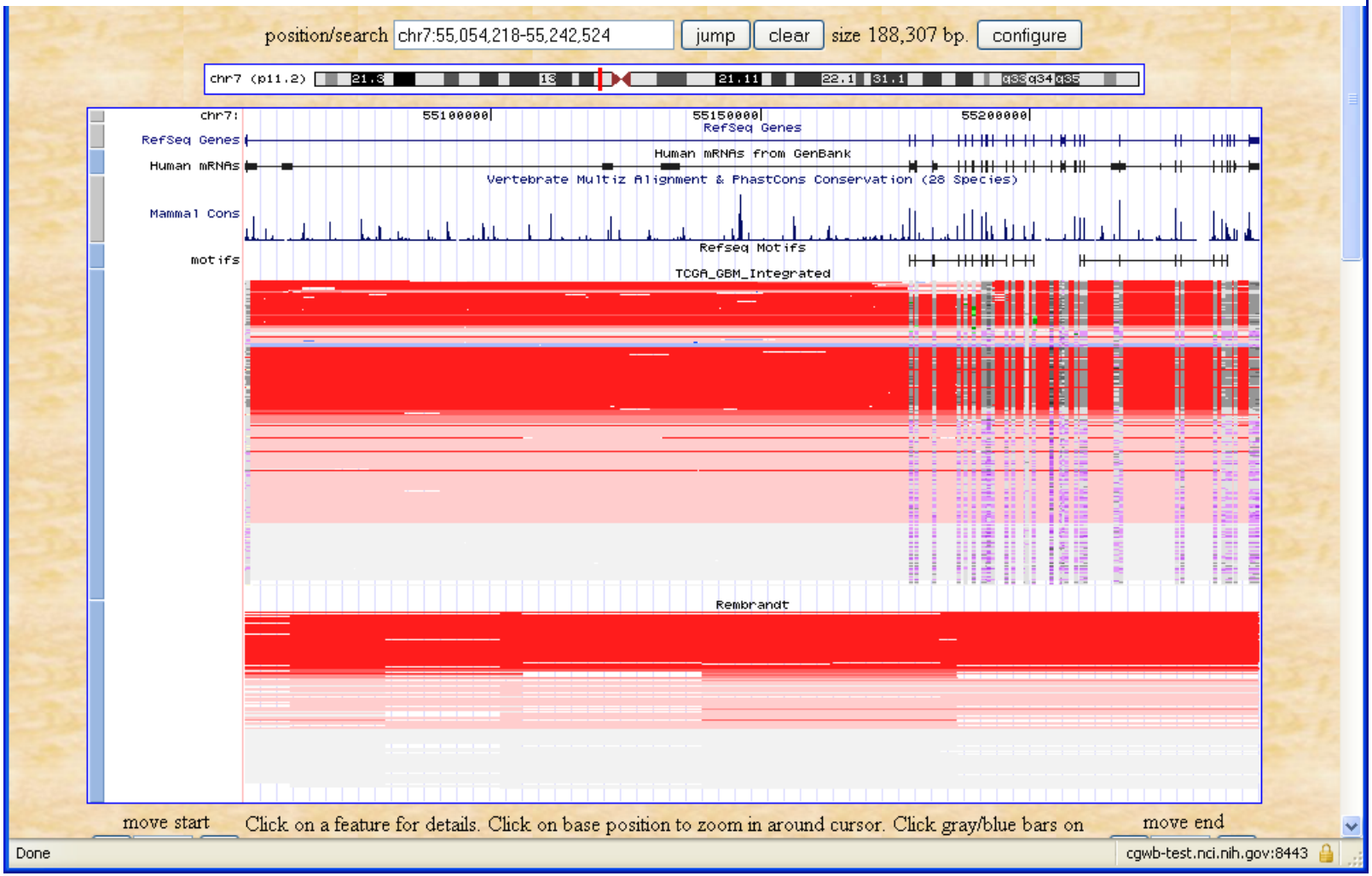
Same Region/Same Data Viewed with Black Background

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



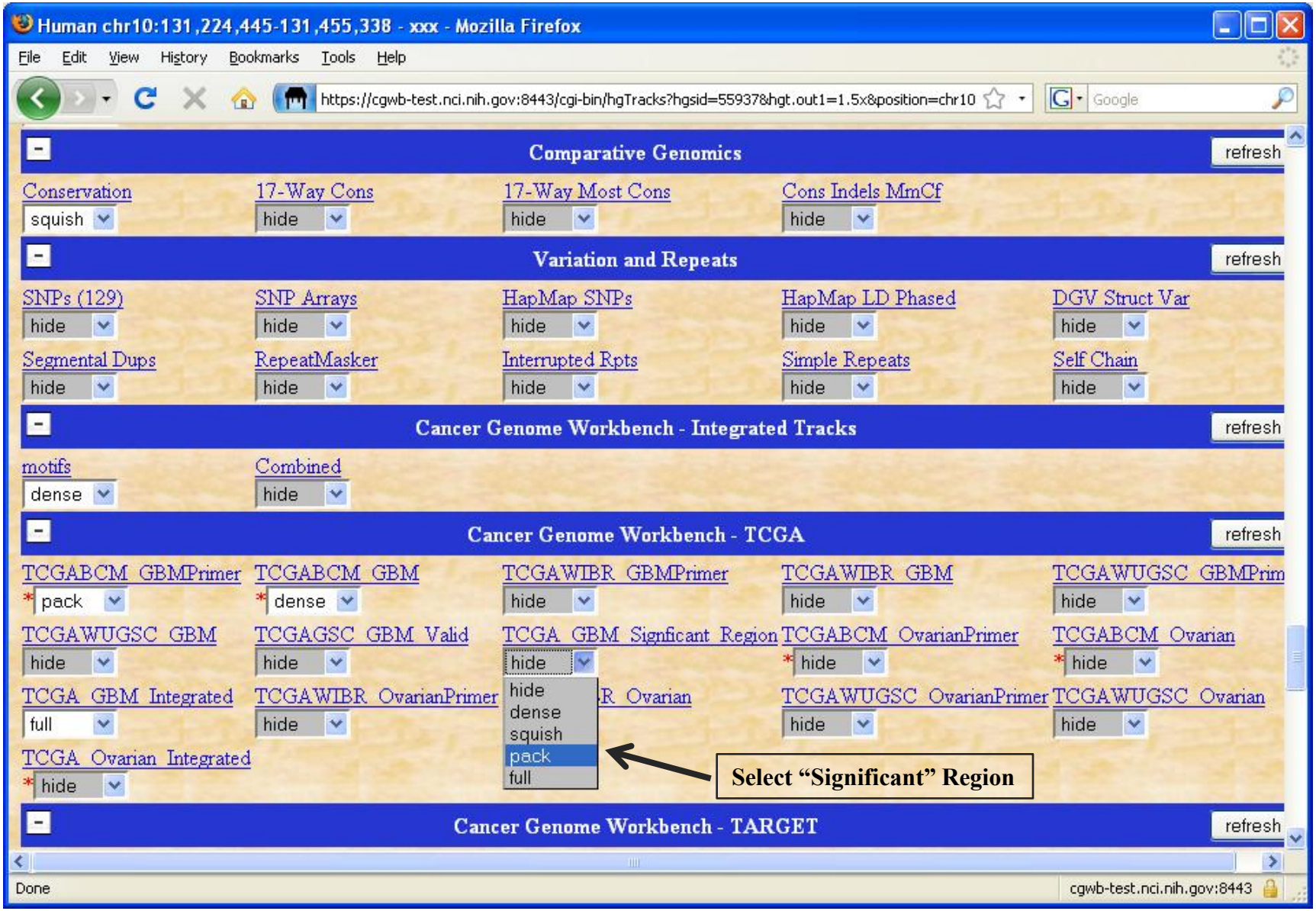
“Dense” View that Shows Both TCGA data (Top) and the Rembrandt Data (Bottom, no exon array) at EGFR locus

“Dense” View is Good For Cross-Project Comparison

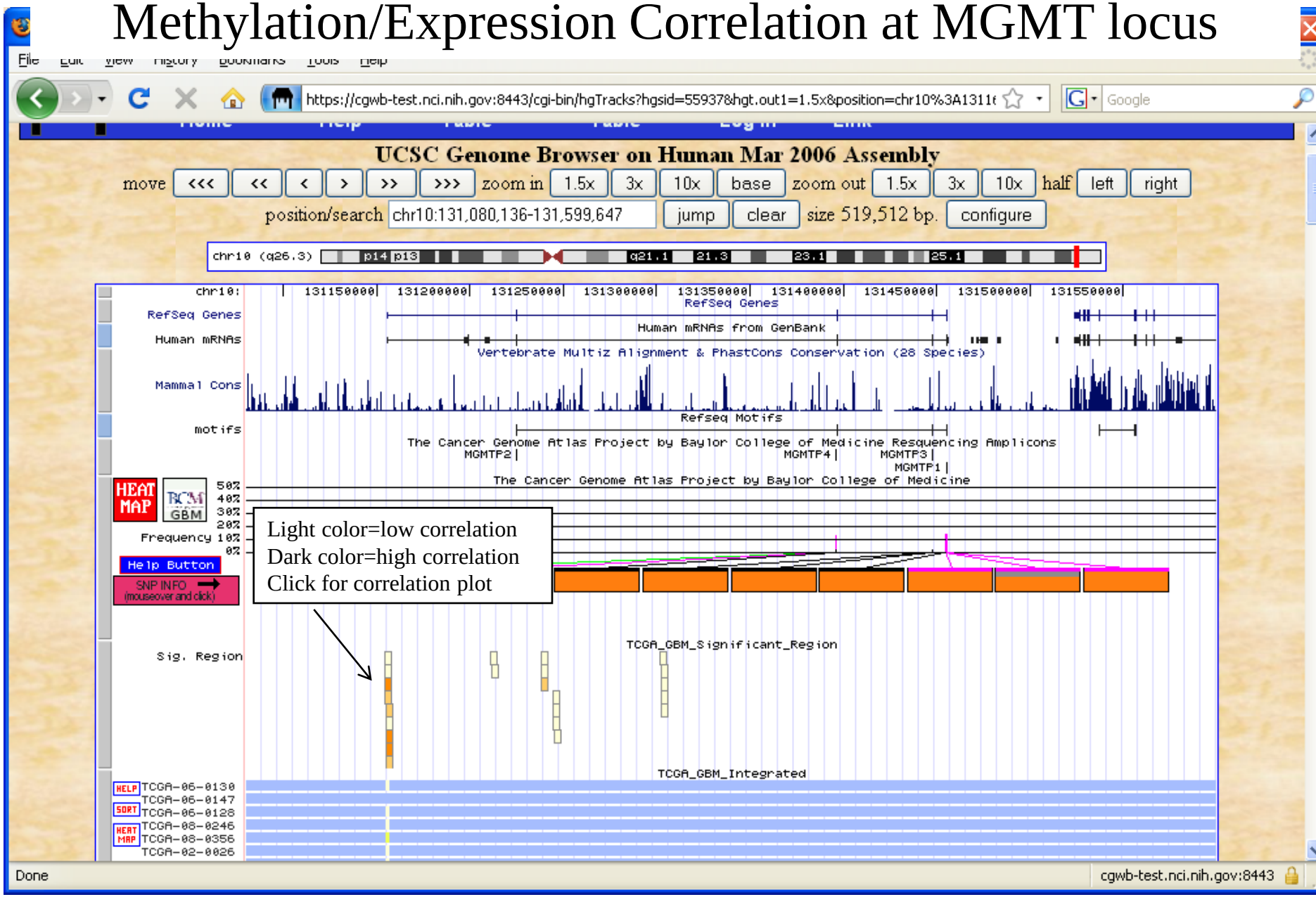


View Correlation Plot of Methylation vs Gene Expression; MicroRNA vs Gene Expression

Turn on display of correlation plot at the control panel



Methylation/Expression Correlation at MGMT locus



Correlation Plot - Mozilla Firefox

File Edit View History Bookmarks Tools Help

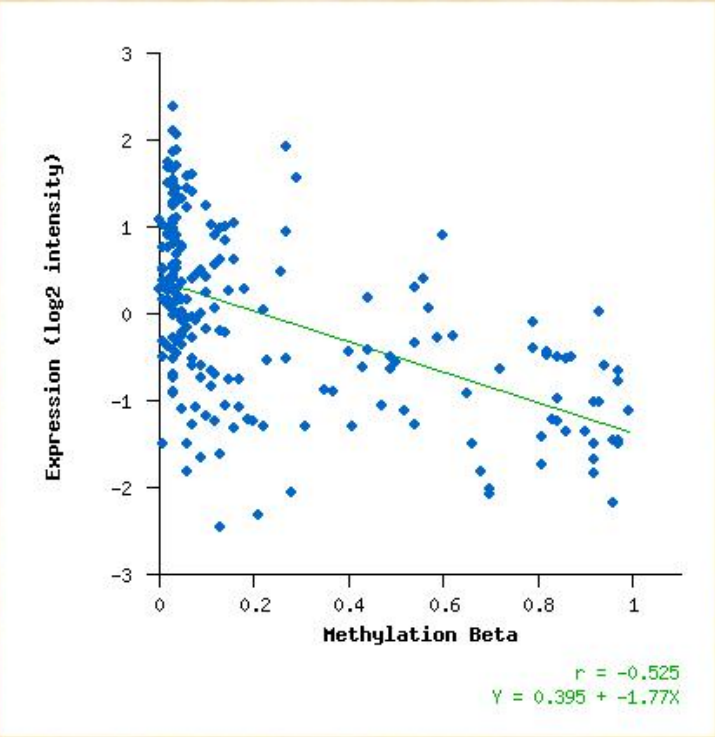
https://cgwb-test.nci.nih.gov:8443/cgi-bin/plottwist?srid=3261

Google

Human chr10:131,166,722-131,513,06... | Correlation Plot

CGWB Home

Pearson correlation between MGMT expression levels and methylation beta value



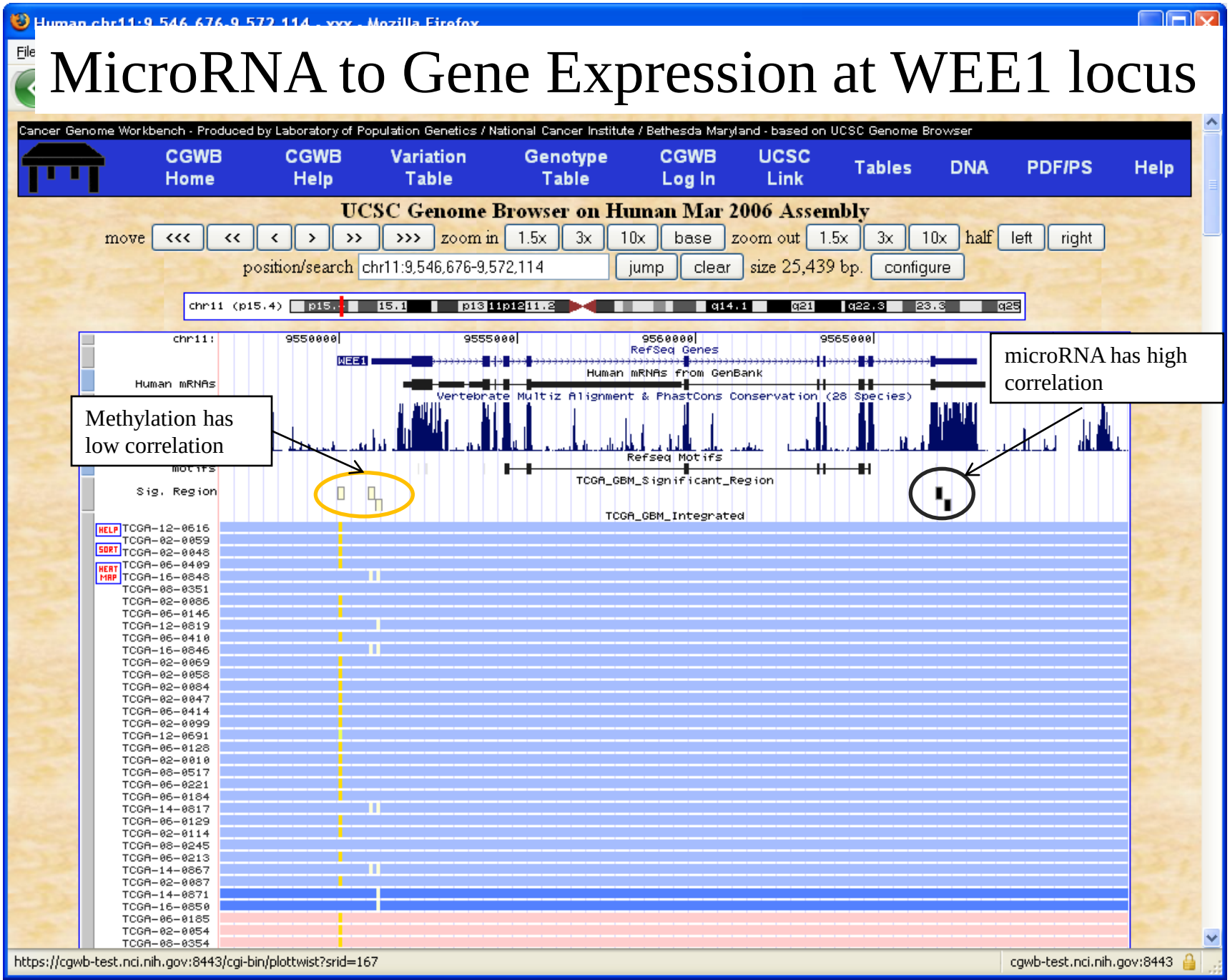
Number of Samples is 215

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

Done

cgwb-test.nci.nih.gov:8443

MicroRNA to Gene Expression at WEE1 locus



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