

## Making links to the UCSC Genome Browser. Part 3: Composite tracks, custom tracks and spreadsheets

Welcome to the third and final part of the video series showing how to make links to the UCSC Genome Browser.

In the first installment of the series, we saw how to pick apart the UCSC Browser URL, how to turn on specific tracks of interest and control their visibility level, how to hide undesired tracks and how to navigate to locations in the genome using gene names or SNP identifiers.

In the second video of the series we looked at how to navigate to specific locations within genes without knowing their genomic coordinates, how to set highlights, how to find obscure configuration parameters, and how to use other useful features for making customized links to the Browser.

In Part Three, we will discuss how to load composite tracks, how to access your remotely hosted custom tracks and hubs and how to make useful links in your spreadsheets.

[ [0:51](#) ] Why make links?

It is instructive to review the reasons you may want to make links to the Browser by direct URL. Many of these have already been detailed in the previous two videos of this series.

If you simply want to share a link with a colleague that shows some location with certain tracks turned on, you can quickly construct it directly in the URL. It's fairly simple, does not make a permanent record in your Saved Sessions and it need not hide tracks that your colleague was already viewing.

You may wish to have some links in your bookmarks to key places in the genome. Because URLs tend to be additive, if you have data tracks turned on and click a link that simply points to some gene, URL links can be designed go to the other location, leaving all other datasets in your display in the same configuration.

You may be developing a web site that generates information about locations, genes, or other annotations. Your software can generate a link to the Browser on the fly that shows exactly what you wish to show. You control the configuration. You can embed the Browser graphic image alone.

You may have custom tracks or a track hub on a server somewhere and you wish to easily load the data.

Or you may have a spreadsheet with many rows of coordinates, gene names or reference SNP identifiers. You can easily make a link that you copy down an entire column, creating custom links for each spreadsheet row.

[ [2:18](#) ] Links to composite tracks

Let's set our location where we left off in the second video of this series about links using the following URL:

```
https://genome.ucsc.edu/cgi-bin/hgTracks?hideTracks=1&db=hg38
&knownGene=pack&singleSearch=knownCanonical&position=DDX21
&encodeCcreCombined=pack
```

Some of the tracks on the Browser are actually several tracks under one pulldown menu. A good example of that is the new format of the SNP tracks, snp153, as is the popular set of gnomAD tracks. For each, it is best to think of the track as a container with several tracks inside. We call this setup a composite track. Although the interface for composite tracks is under active development, in the current video we will describe the feature as it exists now, which will familiarize you with the concept, even if some details of the implementation change.

The track control for this type of track will often offer "show" as a visibility option. You can see this in the gnomAD track. That is a signal that there are tracks can be configured separately. The ellipses (...) after the Short Label on the track control, also seen on the gnomAD track, indicate that there are likely *three* levels of containers (though there are a few exceptions). We call a three-level track a supertrack. Although we will confine our conversation about making links to the case of dbSnp153, which has just two levels, the ideas are generalizable to the third level.

For the SNP tracks, beginning in release 132 on hg19, UCSC has pulled out several sub-tracks from the dbSNP release, mostly to highlight that some variants were common in the population (typically defined as 1% minor allele frequency in the population), but not very common in the database (only a very small and falling fraction of the database now). We had learned that some researchers believed that dbSNP contained only actual polymorphisms – that is, they are segregating in the population at 1% MAF or more. We wanted to help them understand that the dbSNP database, despite the "P" in the name, contained not only formal polymorphisms, but also many rare variants. The Common SNPs track pulls out the 15 million Common SNPs into a separate dataset.

Starting with release 153, dbSNP radically changed their way of distributing data, and we took the re-engineering task as an opportunity to change the way we present the data. snp153 is now a composite track, having several data tracks beneath a single track control below the Browser graphic.

As we showed in Part One of this series, you can learn the name of the track using the mouseover and looking for the "g" parameter in the URL.

dbSnp153Composite

In this case, because it is a composite track, we are seeing the name of the outer wrapper track, not the name of a database table. This was true of the RefSeq track in an earlier video as well, but we chose not to make an issue out of that at the time.

Let's click into the link above the track controls and see the composition of the subtracks. You can see that there are five tracks in total, and only the Common SNPs track comes on by default. There are also ClinVar, Multiply mapping SNPs and All SNPs. Let's just go back to the Browser and not mess with any of these right now.

When making a URL for a composite track, you can turn on the track with an ampersand and equal sign, just as for other tracks, as we showed in the previous videos. Let's just add the

```
&dbSnp153Composite=pack
```

to the URL. You see that a track comes on in dense mode. But we put "pack." What gives? The short answer is, the visibility comes on at the default setting. We see the track, "Short Genetic Variants from dbSNP release 153." At the left, you see the Short Label, "Common dbSNP (153)." As we saw, the Common SNPs track is one of several tracks in the composite. If you want to see the other subtracks, click into the configuration bar at the left. This is the same as clicking the link above the track pulldown menu that we visited a moment ago.

Here on the configuration page, we see that the visibility of "Common SNPs" is set to "dense." So even though we set the outer wrapper track "Composite" to "pack" on the main page, the visibility is limited by this default setting. We could set it to "pack" here, but let's learn the name of the table and set it to "pack" via the URL we are building.

Note that below the configuration box for the Variants track, we see four other tracks not turned on, including ClinVar Variants and All SNPs. We can learn the names for each of these tracks using the "Schema" link. Clicking into the schema for Common SNPs, we see:

```
Primary Table: dbSnp153Common
```

That means we can set the track to pack using

```
&dbSnp153Common=pack
```

and adding it to the URL. In general, if you wish to turn on a track in a composite using a URL, it is better to turn on the individual tracks instead, unless you are content with the default visibilities. Let's discard the Schema page and return to the configuration page for the track.

Let's get the names for ClinVar SNPs and All SNPs as well. Let's just do that by mouseover here, though you cannot copy/paste as easily as if you click into the Schema link.

```
dbSNP153ClinVar  
dbSnp153
```

The All SNPs track in previous releases of our display of dbSNP has always been a stand-alone track named `snp<version number>`. That convention is broken here, now that it is part of a composite, with the addition of the characters "db" in front of the characters "Snp".

It will get very busy on the screen if we turn all of these on to "pack," so let's set two of them to "dense" and add to our URL.

Let's go back the Browser using the Genome Browser link in the top bluebar and hack the URL some more.

```
&dbSnp153Common=pack  
&dbSnp153ClinVar=dense  
&dbSnp153=dense
```

You see that the three tracks have come on as designated: Common SNPs on “pack,” ClinVar and All SNPs on “dense.”

[ 9:05] Loading custom tracks in links.

Many people have data that they wish to display alongside resident data in the Browser. These Custom Tracks, if small enough, can be loaded manually directly from your local machine. If you wish to share your data via a direct link to the Browser, however, it gets more complicated.

The Genome Browser does not have access to your computer, so to load Custom Track data into the Browser via a link, the data must be stored somewhere that is web-accessible for the Browser to see it. Your data can be anywhere on the web addressable by a URL. If you have difficulty identifying suitable storage, we suggest you look into the CyVerse project. Read more about hosting the data in our documentation on the following page:

<https://genome.ucsc.edu/goldenpath/help/hgTrackHubHelp.html#Hosting>

This page can be found from the top bluebar by selecting Help > Browser Documentation and typing “hosting” or “CyVerse” into the search box.

To demonstrate remote hosting and calling by URL, we have sample data in barChart format among our track data examples.

<https://genome.ucsc.edu/goldenPath/help/examples/barChart/barChartExampleOne.txt>

We can have a look at the file before we load it.

Note that the parameter db=hg38 is encoded in the custom track header line to ensure that we not load it to the wrong assembly by mistake. An error message will inform us if we try to load it onto the wrong genome assembly.

We can add the track to a URL using the parameter:

`&hgt.customText=<url>`

OR

`&hgct_customText=<url>`

`&hgct_customText=  
https://genome.ucsc.edu/goldenPath/help/examples/barChart/barChartExampleOne.txt`

The Browser opens to chr14 because of a line in the custom track:

```
browser position chr14:<chromStart>-<chromEnd>
```

directed the Browser to the location of the three items in the track.

Note that the position parameter in this full URL specifies chrX, while the “browser” line in the file specified chr14:

```
https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&position=chrX:15560138-15602945&hgt.customText=https://genome.ucsc.edu/goldenPath/help/examples/barChart/barChartExampleOne.txt
```

If you have conflicting position information in the custom track and explicitly listed in the URL, then the position specified in the custom track will be honored.

### [11:31] Links to track hubs

You can also host a track hub on your own server. Track hubs are binary indexed files designed for fast access to very large datasets and engineered to load only the data needed for display in the current Browser window instead of uploading the entire file. We will not get into any details here about how to address a track hub in a URL, other than to point you to the Browser documentation:

```
https://genome.ucsc.edu/goldenPath/help/hgTrackHubHelp.html
```

and to point out the syntax for loading:

```
&hubUrl=<url>
```

There are many other parameters useful for loading and configuring your own custom tracks from remote servers and for loading track hubs at the following location:

```
https://genome.ucsc.edu/goldenPath/help/customTrack.html#SHARE
```

### [12:22] Spreadsheets

Many biologists and clinicians have a significant amount of data in spreadsheets. They may have many rows of data with coordinates or gene names and would benefit from being able to quickly access a Browser session specific for the data in a certain row. We have prepared a template spreadsheet that can be adapted for use in this context. Let’s start by finding the sheet on our genomewiki site.

```
https://genomewiki.ucsc.edu
```

Type “excel” in the wiki search box on the left, then click into the first item in the list, “Spreadsheet links to Genome Browser views.” This page has some abbreviated instructions about making links not unlike some of the information in these videos about links.

Let’s download the template Excel file and have a look. Click “ucscLinks.xlsx” and open the file. This sheet has template links for a number of different data formats, such as sequence coordinates, gene names, cytobands and SNP IDs.

The links for hg38 can be seen in the right-most column, and if you mouseover the cell, you see the formula that generates the link. You can see from the mouseover that two tracks are designated in the URL: knownGene and omimGene2. Let’s just click into the first link and see

it open a Browser to three million bases of chr3 on hg38, as designated by the columns of the spreadsheet. As we saw in the first video, tracks that were already turned on stay on unless the "hideTracks=1" parameter is set, which it is not in these links.

The formula in the spreadsheet is keyed to refer to data in a fixed relationship to the formula; specifically in this case, the chrom column is five to the left and chromStart and chromEnd are four and three to the left, respectively. You could copy this formula into your own spreadsheet as long as the relationship to the coordinates is the same.

I have a small spreadsheet here to use as an example. We can copy the formula in the right relationship to the coordinates and the link will go to the right place. We must be sure to use the formula that matches the format of the coordinates in the chromosome column. Depending on whether your data have "chr" in the chromosome name or a bare number, you'll need to use the proper formula in the excel link.

But what if your spreadsheet does not have room in the proper column, so that your link cannot be put in the proper relationship to the coordinates? In a case like this, you can do that by copying the formula into the right spot relative to the data columns, and use the "move" command to put it in the right relationship to the coordinate columns. Then we use that formula to copy down the page. You can execute this move command on the original spreadsheet before copying, if you prefer.

This formula should be in the proper relationship to the coordinate data.

Because the Browser had the SNP 153 track turned on when we clicked the link earlier, the many annotations at the full-gene level really crowded the screen. We can turn it off by changing the URL in the link. Let's turn them off now using

```
&dbSnp153Composite=hide
```

We'll copy it down the page now, and you can see by clicking up here in the formula bar, that our link is referring to the proper columns in the spreadsheet. Let's test one them. We'll click into the POU5F1 gene's link right here and it should go to chr6, and open a Browser with the SNP 153 track turned off.

So what if we want to make a column of links using the gene names instead of coordinates from the template spreadsheet? We can copy this formula, and using the same paradigm as before, you can see that the formula making the link in this case should be three columns to the right of the gene name. We copy the formula into the proper relationship to the columns in our target spreadsheet, and once again we use the move command, dragging it to the column we wish to populate.

So this new column will refer to the values in the gene name column, and the one to the left refers to the coordinates. If we copy the gene name formula down the page, each formula will each refer to the gene in its row. You can see by clicking up here in the formula bar, that the reference is the gene name, not the coordinates as it was before. Now we should jump directly to the right gene. Let's go to NF1.

By inspecting the formula in the spreadsheet and reading the wiki page, combined with what you have learned in the previous two videos, you can see how you might configure the links to your liking, showing the tracks that you wish to see.

This concludes our three-part conversation about making links to the Genome Browser. We hope you have seen that the process allows much flexibility and a wide range of capabilities.

Don't forget to subscribe to the Browser video channel and watch for further installments in our tutorial series.

Visit:

<https://bit.ly/ucscVideos>

This video was released during the global coronavirus pandemic. Please note that when it is safe to travel again, we will restart our on-location training program. You can contact us to host a one-day or two-day in-person training at

<https://bit.ly/ucscTraining>

We will also entertain remote-learning scenarios.

Thanks for listening and thanks for your interest in the Genome Browser.

Stay safe.