**Author: Jacob R. Porter**

**Date: 11/10/2015**

**This guide to the 3DChrom Program will show you what you need to know to start building**

**File Structure**

* **The 3Dchrom folder holds our 3Dchrom project.**
* **css** - Inside, our css folder holds our css. For css, we are using custom themes and Bootstrap. Bootstrap documentation is available at <http://getbootstrap.com/>
* **files** - Our files folder has all of our data files for the project
* **HGDF** - Human Gene Definition Files. Inside of this folder there are a few files.
* **Human\_x\_gene\_file.txt** - A backup file with just some ensemble fields
* **hgdf\_x\_flat\_file** - All of the current fields. Split into HGDF\_SEGMENTS.
* **HGDF\_SEGMENTS** – This folder holds all of our HGDF segments.
* **PDB** - A folder of Program Database Files.
* **POS** - A folder of Gene positions.
* **fonts** - A folder of fonts. (You will never need to touch this.)
* **j2s** - A jmol folder. (You shouldn’t have to worry about this folder.)
* **java** - A jmol Java folder. (You shouldn’t have to worry about this folder.)
* **js** - All of our JavaScript. Some Bootstrap, some jmol.
* **php** - A jmol PHP folder with one file. (You shouldn’t have to worry about this folder.)
* **plugins** - Contains our DataTables and Jmol plugins.
* **homoSapienChrom.php** - This file contains the view and code of the page that allows you to search for a domain.
* **homoSapienDomain.php** - Follows the homoSapienChrom.php file. Once you select a domain, it displays it with this code.
* **Index.php** - Contains the initial page code.
* **hgdf\_x\_files - A comprehensive list of the files that made up the hgdf\_flat\_file.**
* **PERL Scripts - Contain a number of PERL scripts that manipulate files.**

**Installation and starting**

To install and start working with the above code, you will need a XAMPP installation. Once XAMPP is installed, navigate to the htdocs folder. In my Windows system, it is available at C:\xampp\htdocs\

Although a lot of the above files mention Jmol, we are actually using JSmol. Documentation is available at <http://wiki.jmol.org/index.php/Jmol_JavaScript_Object> and many great guides can be found. Most Jmol is compatible with JSmol, but JSmol is newer, so I would Google for Jmol if you need help with styling or documentation. Here is a quick guide <http://wiki.jmol.org/index.php/Jmol_Tutorials>

**Deployment**

Dr. Zheng Wang has likely already provided you an ID and access, but just in case you haven’t here is what you need. You will need to log on with SSH or whatever else you prefer to log on to the USM system. Here is the connection info.  
  
--**REDACTED--**

You will need to deploy your code into **/www/html/3Dchrom**

Once deployed, it will be available at <http://dna.cs.usm.edu/3Dchrom/>

**Pages So Far**

So far, there are three webpages that can be accessed: <http://dna.cs.usm.edu/3Dchrom/>, <http://dna.cs.usm.edu/3Dchrom/homoSapienChrom.php>, and <http://dna.cs.usm.edu/3Dchrom/homoSapienDomain.php>.

**Index**

The 3Dchrom/**index** page is a simple stationary page with an introduction. Neither the About or Contact buttons do anything as of yet, but in the future clicking those would provide a contact page and an about page. On the page the gene set we are currently working with is also present, and some other information for display purpose only. **The entirety of this page can be changed if needed.**

Click **View Here** to continue to the **homoSapienChrom.php** page.

**homoSapienChrom**

The 3Dchrom/**homoSapienChrom** page uses DataTables to display all of our available Program Database files. The file names contain a domain. For instance, in human\_X\_100520000\_102320000.pdb the domain start would be 100520000 and the domain end would be 102320000. Inside of each of these files is a number of individual locations called ATOM that help programs know how to draw the domain. You will not have to do much with the .pdb file except for load it whenever you want JSmol to draw the domain.

By default, ten entries are showing, but this amount can easily be changed with the DataTables configuration. <https://www.datatables.net/>

There is also a Search to narrow down the results. Pagination is occurring, so by clicking 2, 3, 4, etc., you can go to another set of 10 entries.

Home goes to the index.php page and clicking Search will refresh this page, as this is the Search page. **Clicking a file name** will lead to a page displaying information and graphics about that particular **domain**.

**HomoSapienDomain**

The 3Dchrom/**homoSapienDomain** page uses DataTables to display all of the geneID associated with the particular domain that you clicked on. When you click the check box associated with the geneID, the information about that gene will show up under Selected Gene Information. It will also highlight the gene on the graphic that is being displayed by JSmol.

The Selected Gene Information contains information about the gene and a link to more information about it.

The Home and Search buttons work as they did on the previous page. There is also a spin display button, but that’s just for fun. **Also, feel free to take off the Jacob Porter 2015 on all of the pages.**

**The Files**

I’ve already mentioned the **homoSapienDomain, homoSapienSearch, and index** files to some degree. There are also the **PDB** files, the **POS** files, and the **HGDF** files, and finally some **PERL scripting files** for combining and separating data from the files that we get from the Ensembl Biomart store.

**The PDB files are Program Database files.** These are used by the JSmol plugin to draw the domain graphic. Other than that, we never have to do much with these files.

**The POS files break down the domain into pieces.** These pieces correlate with the pieces that make up the genes on our **homoSapienDomain** pages. You will see why these files are handy in the code.

**The HGDF file/files is/are downloaded from Ensembl Biomart.** <http://www.ensembl.org/biomart/martview/d4f7f54b278945d5037553e78cd35de5>

To get a new copy of the HGDF file with more fields, navigate to the above URL and click – Choose Database – Select Ensembl Genes 82, and click – CHOOSE DATASET – Select Homo Sapiens Genes (GRCh38.p3)

Now, you can click Attributes to add more fields or remove them. You can also use Filters to make sure that we only are working with the human X chromosome. At this time, we are only working with the human X chromosome. Once you have selected the attributes and filters you like, click Results. You can then save the file in different formats. **My code uses TSV.**

**The Perl scripts are custom, but probably not efficient.** I built them on the fly and they were used to combine multiple HGDF files. They are not using the best logic because I was short on time, but they work. I figured I would include them so that you could at least use them as reference. I have named them by their function.

Also are the hgdf\_x\_files. These are base files and flat files that I used in conjunction with the PERL scripts.

**Dive into the code**

Now, you have a decent understanding of what is going on. It’s time to dive into the code! The code is commented, so it should be pretty easy to understand. However, if you need any help feel free to reach out.