Chromosome Visualization Software+ User Manual

Ariadne Dulchinos

All screenshots are directly taken from the Mac version, v3.0.

Table of Contents

How to Install & Run 3

The Home Screen 4

The Top Bar 5

The Idiogram 8

The Search Bar 9

Troubleshooting 11

The Buttons 12

The Quick Find Bar 13

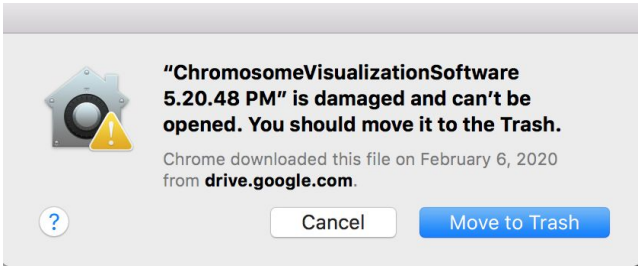
The Loci 14

How to Install & Run

Mac

1. Download ChromosomeVisualizationSoftware.zip
2. Extract ChromosomeVisualizationSoftware.app from the .zip file
3. Run the .app file.

Note, sometimes, this popup is displayed when trying to open the app.



If this occurs, please follow these instructions.

1. Take the damaged app "ChromosomeVisualizationSoftware.app" and move it to your applications folder.

2. Open the Terminal (you can search for it in Launchpad) and copy/paste this command:

    xattr -cr /Applications/ChromosomeVisualizationSoftware.app

3. Open "ChromosomeVisualizationSoftware" it should work

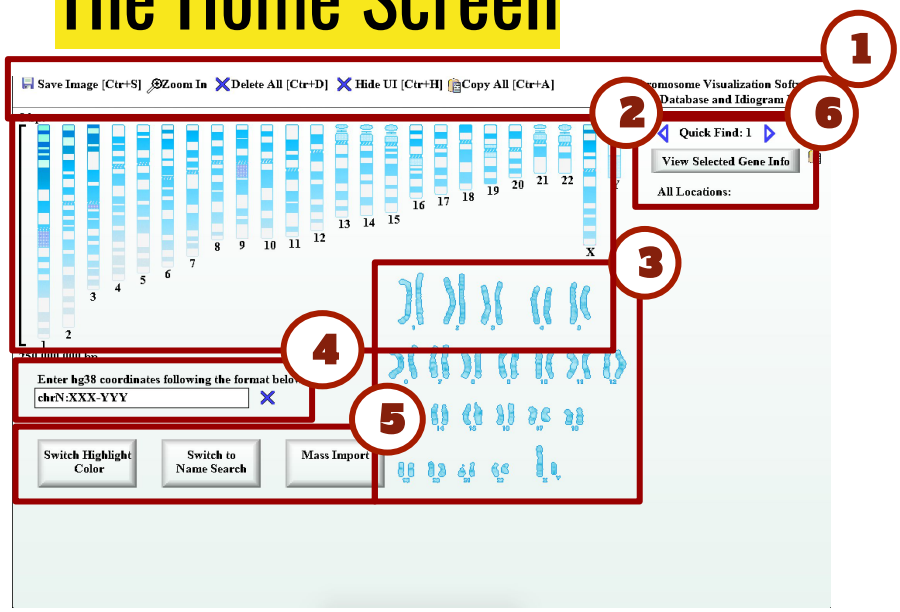
Windows

1. Download ChromosomeVisualizationSoftware.zip
2. Extract ChromosomeVisualizationSoftware.exe from the .zip file
3. Run the .exe file.

Acknowledgements

Programmer, Graphics, and Finalization: Ariadne Dulchinos
Concepts & Prototyping: Ariadne Dulchinos, James Weng
Research Mentors: Dr. Ece Uzun, Dr. Diana Siliezar

The Home Screen



Pictured to the left is the 'UI', the user interface of the program.

Observe: the UI is labeled numerically with common terms that will be referred to throughout the manual.

Throughout the manual, the person utilizing the program will be referred to as 'the user'.

1. The 'top bar' contains useful functions of the program, including downloading images, deleting all loci, and hiding the UI. If clicking a button does not work after multiple tries, use the keyboard shortcut to perform the desired action.
2. The 'idiogram' is a to-scale representation of each chromosome, including the X and Y chromosomes, in idiogram format.
3. The 'search bar' is the main area for user prompted data entry of a string of genomic coordinates, gene names, phenotypes, and disorders, simply by typing using the keyboard.
4. The 'buttons', labeled as Switch Highlight Color, Switch to X Search, and Mass Import, control various aspects related to searching and data entry.
5. The 'karyotype' is a simultaneous depiction of each chromosome and its identical pair in karyotype format.
6. The 'quick find bar' enables users to quickly find genes they have previously inputted as a real-time list of all genes currently on screen. The user can also copy the list of genes to the clipboard.

Icons & Keyboard Shortcuts

Note: Although this is the mac version, use CONTROL, not Command for keyboard shortcuts.

Save image

Delete or clear

Zoom

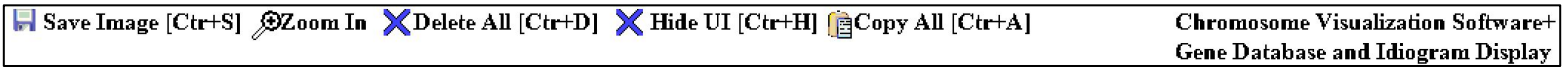
Search left or right

Copy to clipboard

Ctr + S = Save image
Ctr + D = Delete all loci
Ctr + A = Copy all genetic information to clipboard, helpful when exporting to text files.

Ctr + C = Copy currently selected item
Ctr + V = Cut currently selected item
Ctr + H = Hide or unhide the UI

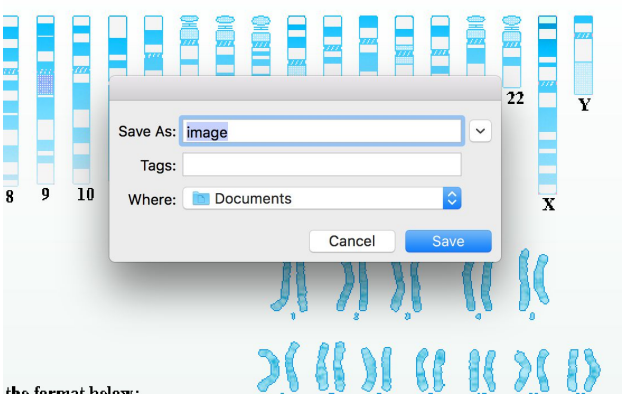
The Top Bar



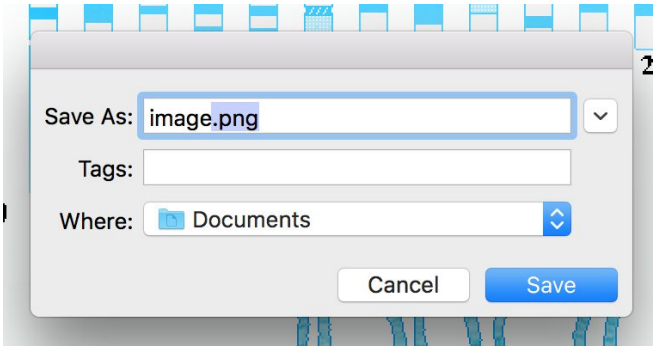
Save Image [Ctr+S]

The purpose of this button and its designated keyboard shortcut ‘Ctr+S’ is to allow the user to save whatever is currently displayed on screen as a **.png** (Portable Network Graphic) image file. This is very helpful to create diagrams for data analysis or records of genes.

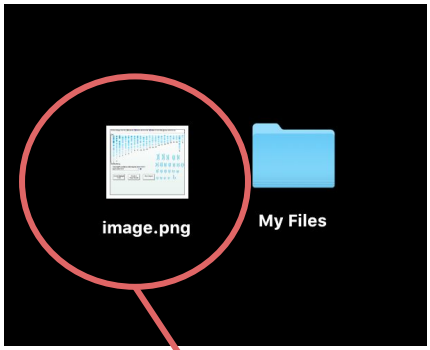
When this button is clicked, the following menu prompt appears:



The menu prompt for saving.

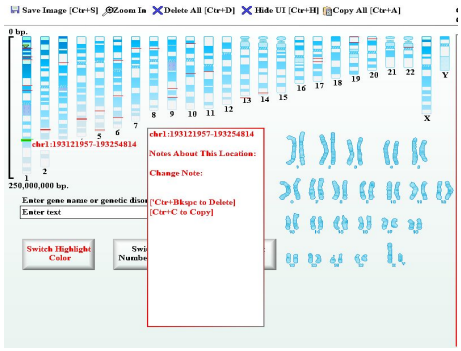


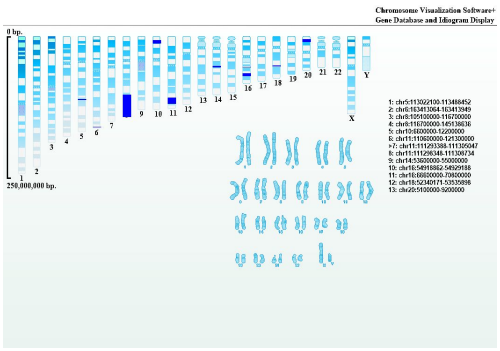
IMPORTANT: Remember to include **‘.png’** at the end of the file name for it to save as an image.

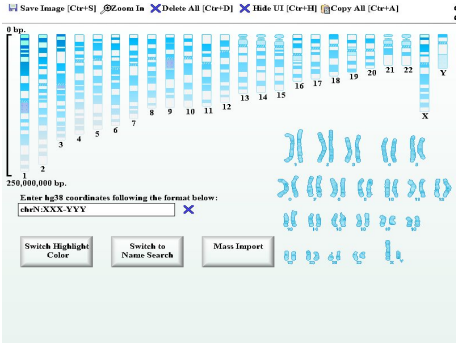


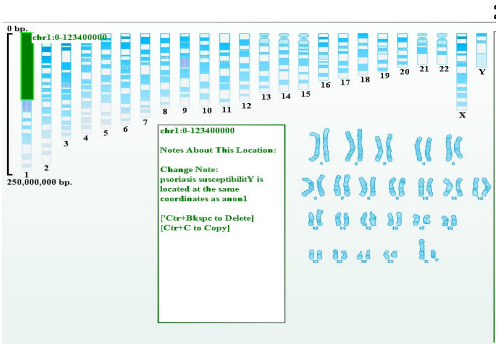
The screenshot is now visible at the chosen file path.

Example Screenshots

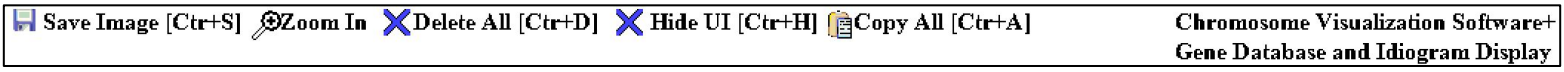






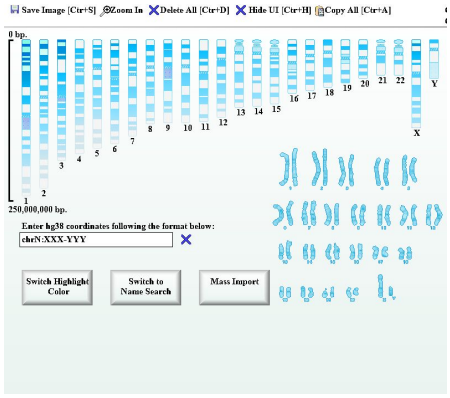


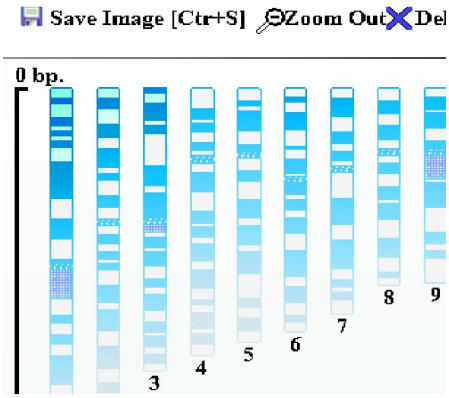
The Top Bar



The purpose of this button is to allow the user to zoom in and out. In zoom in mode, the user's cursor becomes a magnifying glass. When the button is clicked again, the magnifying glass returns to normal.

Below is an example of screen perception when zooming in vs. when zoomed out.





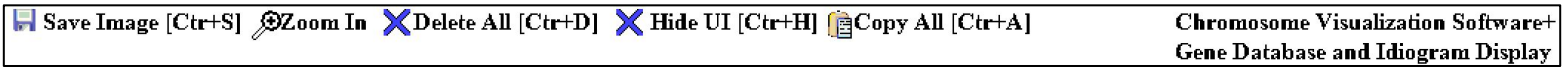
Zoomed out.

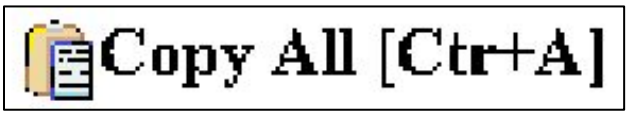
Zoomed in. Screenshots can still be taken in zoomed in mode.

This button deletes all loci currently on screen, reverting the UI to its default state. This is helpful when the user has finished a dataset and needs a clean canvas to work with in preparation for another set. WARNING, this action is not undoable; once the button is pressed, the loci are gone.

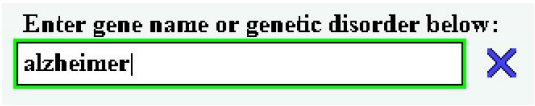
This button hides the user interface and all buttons on the top bar, which is helpful for removing distracting elements while saving an image.

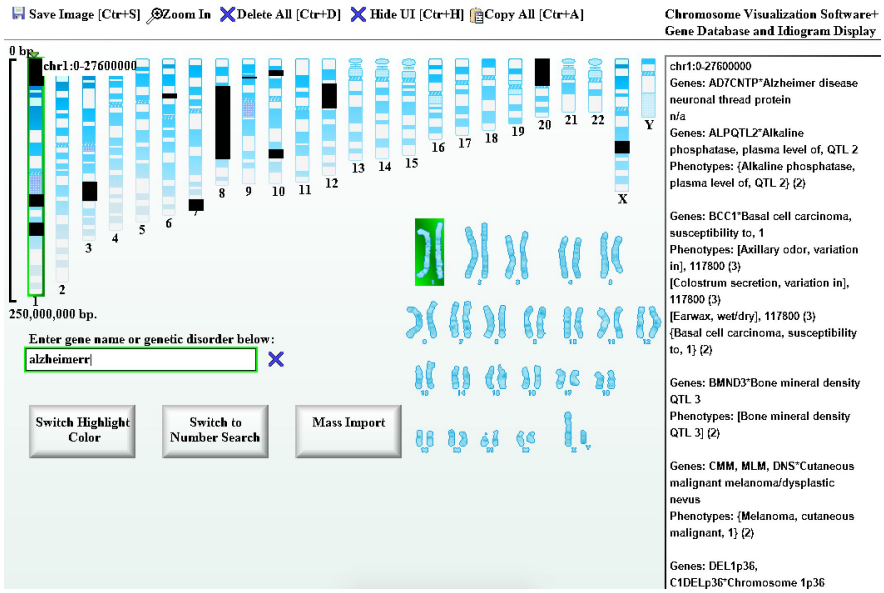
The Top Bar





The purpose of this button is to allow the user to copy a complete list of all genes on screen. For example, say the user is looking up all known genes associated with Alzheimer's. The user types in 'alzheimer' into the program to ensure no apostrophes or capitals are discounted in the search.





This is what the user gets. That's a lot of genes, and as can be seen from the highlighted locus, there are multiple genes and phenotypes associated with that locus.

To facilitate the export of this data to a text file, the user hits 'Ctr+A'. Now all of the text is copied to the user's clipboard.

Now if the user goes to a text processing software like Google Docs or MS Word, they will see their data, separated by # symbols.

If this is too messy and the user desires cleaner formatting, they can go to http://www.unit-conversion.info/texttools/add-line-breaks/#data and under 'input data' hit Ctr+V to paste the information. Then, set the next two parameters to below:

Upon clicking submit, the user will see a cleaner output with line breaks separating the data, which they can paste into a text document.

"What does this mean?"

In a dataset, this is the number assigned to this specific datapoint.

The chromosome where the locus is found.

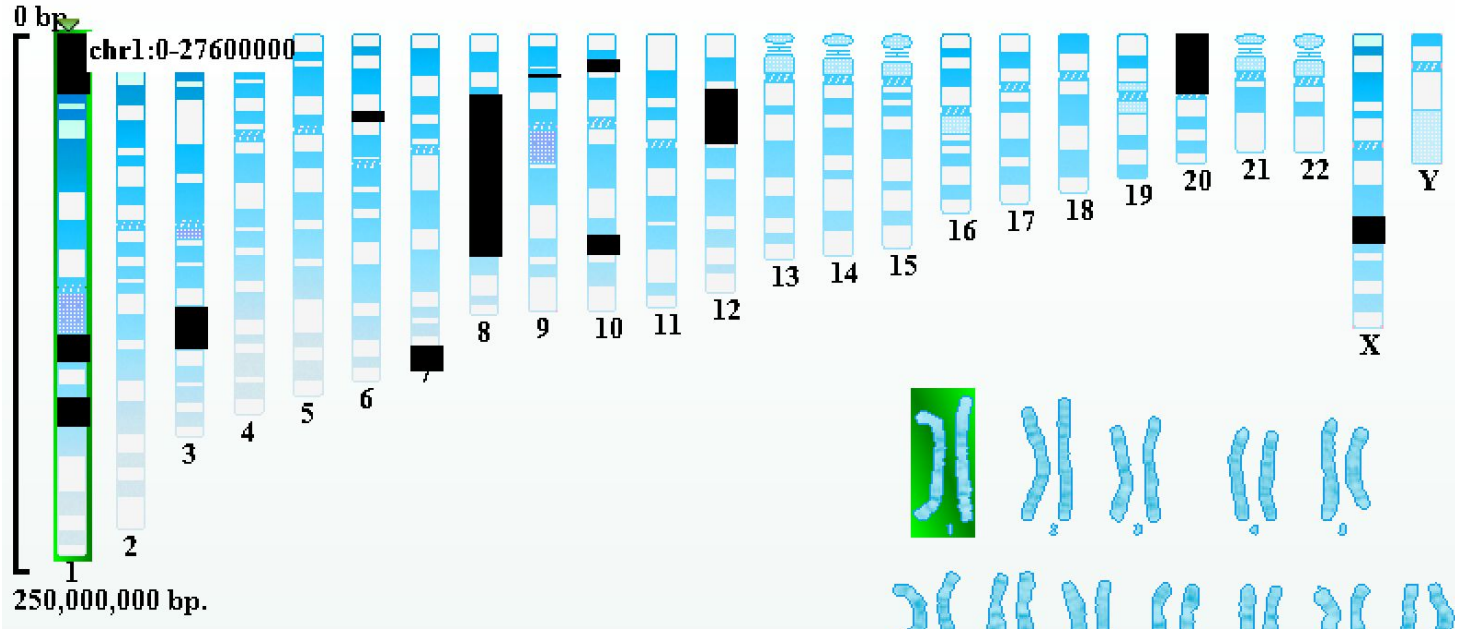
The start/end basepairs of the locus, determining where on the chromosome it is and its size.

9: chr10 : 12200000- 17300000
chr10 : 12200000- 17300000
Genes: AD7 * Alzheimer disease 7
Phenotypes: Alzheimer disease-7 (2)
{Colorectal cancer, susceptibility to, 3}, 612229 (3)

What genes (abbreviated and full name) are known to be associated with this locus, there may be multiple.

What phenotypes (diseases and traits) are associated with the gene above. Some genes require further study to identify phenotypes not currently listed.

The Idiogram



This is an idiogram employed by the program's UI, a visual representation of all chromosomes simultaneously. At the left, there is a scale,which is a precise translation of basepairs to pixel (the screen resolution).

Idiograms are an efficient means of visualizing the relative locations of genes on all chromosomes at once. The blue stripes on the chromosomes represent cytobands, which are highly concentrated regions of base pairs that can be visualized using certain stains. Cytobands also contain sub-bands. The black bars indicate loci which the user has inputted. The diagonally striped blue indentations towards the center of each chromosome is the centromere. The two ends of the chromosome are telomeres.

This hierarchical viewing of chromosomes is key for accurately locating genes and their proximity to others. In this way, the program can visualize characteristics such as synteny of genes, and contribute to data analysis regarding the probability that two loci will cross over together and display correlated phenotypes.

The Karyotype

The karyotype is another method employed by the program to visualize chromosomes. Instead of showing each cytoband, it shows the trademark x shape with a well depicted centromere. It also shows the chromosomal pair.

In the program, then the user hovers over a locus on the idiogram, it shows up on the karyotype, delineated by a green box outline.

The Search Bar

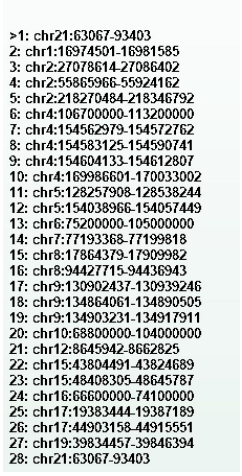
How to use the search bar to search by number (genomic coordinate):

1. Click it, until a green outline bar shows up like this:

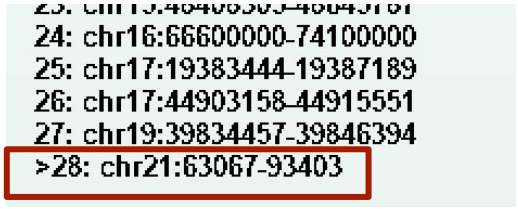
2. Enter or paste genomic coordinates in this specific format ‘chrN:123-456, where N = the chromosome number, and 123 is the starting base pair, and 456 is the ending base pair. Do not include any capital letters or leading zeroes. If this format is not adhered to, the program may not work.

3. If the program was able to plot the locus, a message box will show up stating “Possible Match Found.” Click this to exit.

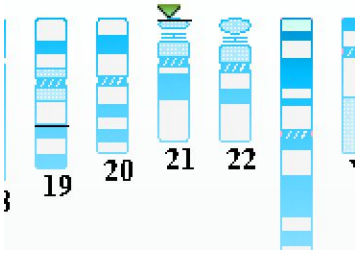
    a. If this does not occur, please consult the troubleshooting page later in the manual.



Having trouble finding a locus? Use the quick find menu to navigate through all these genes. Tap the right arrow until the desired locus is reached.



A ‘>’ before the coordinate indicates that it is the currently selected locus.



Once the locus is selected, there will be a green upside-down triangle above it.

Clicking this button will bring up gene and phenotype information for this locus if it exists.

4. The locus should now be visible on screen. If the user has difficulty finding their locus (ie: too small) they can check if their locus was plotted successfully by using the Quick Find Bar, and tapping the right arrow until they reach the desired locus. When this locus is reached, there will be a green upside-down triangle above the locus.

5. To view gene information, the user can either left-click the locus (which may be difficult depending on its size) or click the ‘View Selected Gene Info’ button.

6. To copy a list of all genes (but not their information, just coordinates) click the clipboard icon next to the View Selected Gene Info button.

The Search Bar

How to use the search bar to search by name:

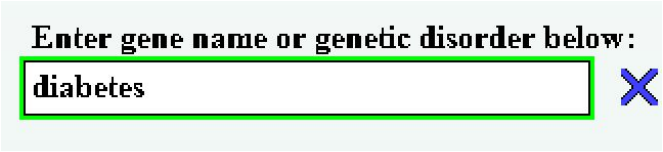
Searching by name entails a plethora of parameters useful for scientists, researchers and educators. It is possible to search by:

- Gene name (ie: neurexin)
- Abbreviated gene name (ie: NRXN)
- Phenotypes (ie: skeletal dysplasia)
- Symptoms (ie: asthma)
- Diseases (ie: diabetes)

Important: For best results, only use capital letters for abbreviated gene names. Also, do not enter characters besides letters and numbers into the search bar.

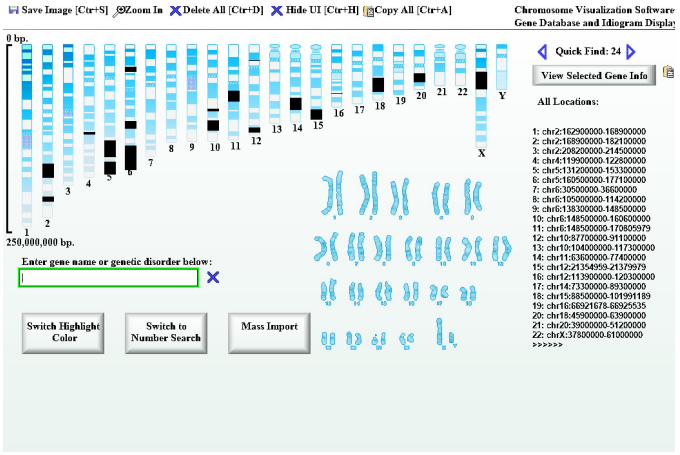
1. Click the button below the search bar which says ‘Switch to Name Search’. It can be switched back to number search at any time.

2. Click the search bar so that a green border appears around it, and type in the desired search term. Press enter.



3. Much like searching by number, if a possible match is found, a message box will pop up. Click it to exit the prompt.

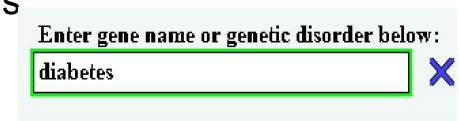
4. The genes should now be visible on the chromosomes. Please note that when searching by name there is an increased likelihood that more than one gene will pop up as a match, because there can be multiple genes associated with a specific phenotype/disorder/etc.



Troubleshooting

Having trouble getting genes to pop up as a possible match? Follow the checklist below to maximize results.

1. **Use the correct mode:** it is crucial not to search a coordinate in the name box, and vice versa. This is an easy fix, simply click the button 'Switch to X Search' below the bar.
2. **Use the correct format:** coordinates must be typed as chrN:123-456, and names may not include special characters If searching for a plural or possessive such as 'Alzheimer's Disease', instead try 'alzheimer' to get results.
3. **The end base must be greater than the start base**.
4. **The end base must be less than the total number of bases on a particular chromosome**.



What about near matches, and non-phenotype attributed genes?

Regardless of what coordinate is typed in on a chromosome, as long as it is between 0 and the total number of base pairs, the locus can be visually represented. However, sometimes, there is no gene associated with a locus (aka, neither OMIM nor UCSC genome browser has reported an existing gene at that locus). If this is the case, the algorithm is programmed to try and find a near match, which is a genetically valid locus on the same chromosome but has slightly different coordinates, to ensure that a gene and/or phenotype exists. If no near match exists, the locus will still be plotted, but under gene information it will show 'n/a' for one or more categories.

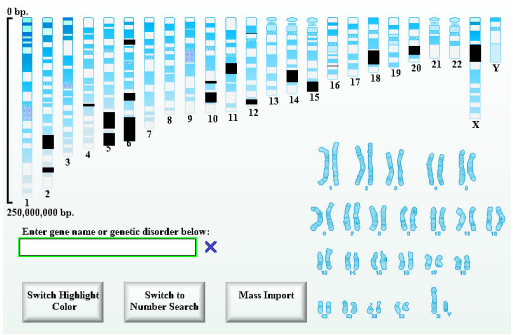
Likewise, if the coordinates of a locus are entered that are WITHIN a larger locus that does contain a gene/phenotype, the algorithm will output the inputted coordinate plus the genes and phenotypes comprised by the larger locus. This ensures that if a user mistypes one or two digits but still has a general locus, the algorithm will catch this with precision.

The Buttons

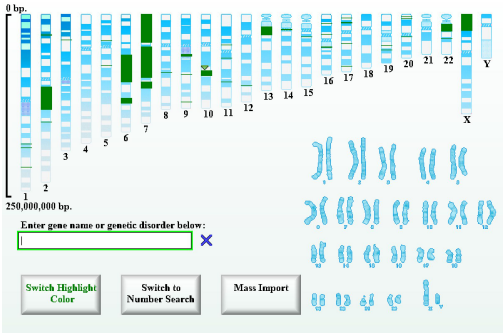
Switch to X Search: The function of this button has been covered extensively in previous sections. It enables the user to search by genomic coordinate, gene name, phenotype, and more.

Mass Import: When this button is clicked, a menu pops up prompting the user to enter or paste multiple coordinates at once (this button does not support name searching) Ensure all coordinates are valid and separated with commas, and hit ‘Enter’ to simultaneously plot every coordinate at once. Note: sometimes excludes the last coordinate entered.

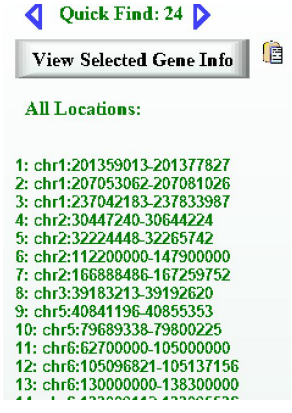
Switch Highlight Color: When clicked, this button changes color, toggling between the six colors of the rainbow and black. The color of the button indicates what color the bars of genes will show up as when plotted. This is helpful for organizational purposes, especially if many groups of genes are visible on screen. See examples below:

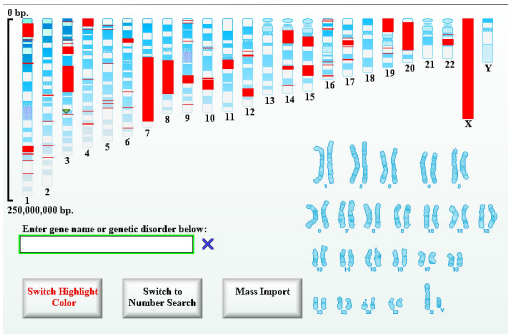


Genes when the highlight color is black (default)



Genes highlighted green, note that the quick find menu also changes color





Genes when the highlight color is red

The Quick Find Bar

The function of the quick find bar is to enable the user to navigate through all of the genes they have plotted. Although the number begins to go off screen after some time, especially for quite large data sets, each gene is still there and able to be navigated.

To navigate the Quick Find Bar, the user can click the blue left and right arrows.

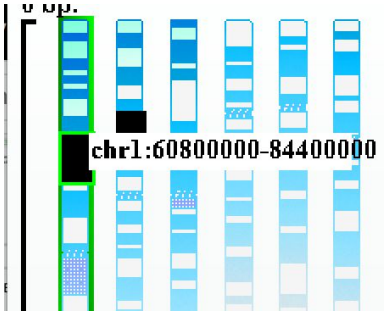
Between the arrows shows the currently selected locus, which is delineated by a '>' next to the coordinate name. If the coordinate is off screen, this functionality still works, and the user will still be able to see the green upside down triangle corresponding to their gene.

Below the arrows is a 'View Selected Gene Info' button which enables the user to view gene and phenotype info for the locus they have selected by clicking the button.

To the right of this button is the copy feature, which copies the list of genes, not the information but merely coordinate lists, which is helpful for mass importing later with a bit of reformatting. If the user desires to copy all gene information they can use Ctr+A.

The Loci

Once a loci is highlighted, either by left clicking it or selecting View Selected Gene Info, there are a variety of functions that can be performed.

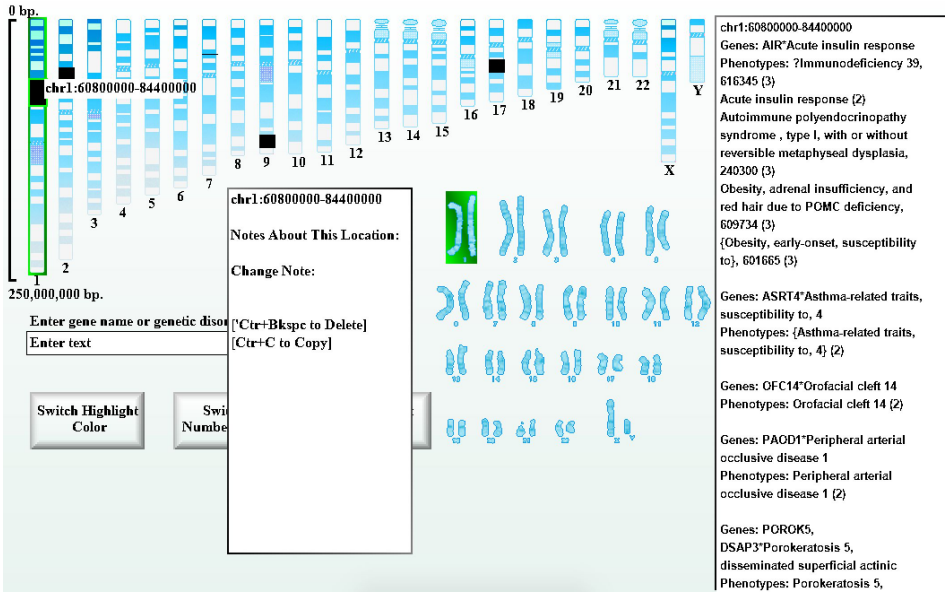


The locus should have a green square around it, like this:

The white square with the coordinate written inside it is the genomic locus. The black bar is the gene itself.

Once the gene is highlighted, these two white menus should pop up:

The center menu, when clicked, allows the user to type in a note to the gene. It can be any string of text. Its purpose is organizational, such as if a biologists wants to mark a certain gene for review.



The purpose of the right menu is to display all gene and phenotype information known about the locus. It can be copied to the clipboard by pressing Ctr+C.

A gene can be deleted by either right clicking it or hitting Ctr+Backspace.