Chromosome Mapping your DNA Clusters with DNA Painter

Chip Mansure
cmansure@gmail.com

Chromosome Mapping your DNA Clusters with DNA Painter

At the dashboard select the Chromosome Maps tab and you will see the link for painting your clusters automatically. The Cluster Auto Painter process automatically creates a new chromosome map for you. You can merge your cluster chromosome map with existing segment data later. Merging is easy compared to trying to unmerge. Everything I will be showing you can be done with the free DNA painter account. The one issues with the free account is you cannot have more than one chromosome map. Thus, mapping your clusters cannot be done with a free account if you already have another chromosome map. The Cluster Auto Painter link brings up the following page.

My goal today is to give you an overview of how to paint clusters. There are too many options for painting clusters to go over all the details, instead I'll get you started and point you to where to find more information. I will share with you an example of how painting 23&Me clusters has helped find leads to follow-up with the traditional paperwork and contacting matches. Thus, painting my clusters has helped prioritize my work. As you can see DNA Painter provides 4 ways to paint clusters. I'll talk about the first three. I have not tired using DNAGedcom. You can import clusters directly from MyHeritage and Gedmatch. To import data from 23&Me and FTDNA you must use Genetic Affairs or DNAGedcom.

CLUSTER AUTO PAINTER FAQ

How do I edit a cluster to make it maternal or paternal?
I've written an example covering how to do this, including other tips on annotating your clusters:
- Annotating a Cluster Auto Painter chromosome map

Why are some of the matches and/or clusters missing?
There are a few reasons why this can happen:
- A match may have opted out of sharing segments with you.
- A match may not have any segments over the cM threshold (the default is set to 7cM)
- If this applies to all matches in a cluster, the cluster itself will be missing
- You might accidentally have used a segments file for just one match, or a segments file for another family member

Why do some of my clusters contain only one match?
Please see above. Also:
- Gedmatch appears to use thresholds very differently between its cluster tool and the segment search tool.
- Depending on the threshold you use, you may find your clusters include many matches that don't appear in your Gedmatch segments file.
- You may be able to work around this by either adjusting the thresholds, or by increasing the number of matches included in your segment search report at Gedmatch.

How do I add my painted clusters to my existing DNA Painter chromosome map?
At present CAP generates a new chromosome map. To import this into an existing profile:
- In the cluster profile, click on the settings cog above chromosome 1 and click 'All segment data'.
- Optionally filter this table to include just certain clusters
- Click 'CSV file' at the top left
- Within the profile you want to import the clusters, click the settings cog, and click 'Import segment data'. Browse for your downloaded CSV file from the clusters profile, adjust the threshold settings if desired (the default is 10cM), and click 'Import this file'

MORE ABOUT CLUSTERING

For more background information on how clusters are generated, please take a look at the following:

- https://www.danaleeds.com/the-leeds-method/
- https://www.geneticaffairs.com/media.html
- https://blog.kittycooper.com/2018/12/automatic-clustering-from-genetic-affairs/
- https://dna-explained.com/2018/12/04/autoclustering-by-genetic-affairs/
- https://www.danaleeds.com/dnagedcom-clients-new-collins-leeds-method-3d/

Lower down on the same page you will find helpful information. This is where you should go for help plus the help tab on the home page has a place to search for help.
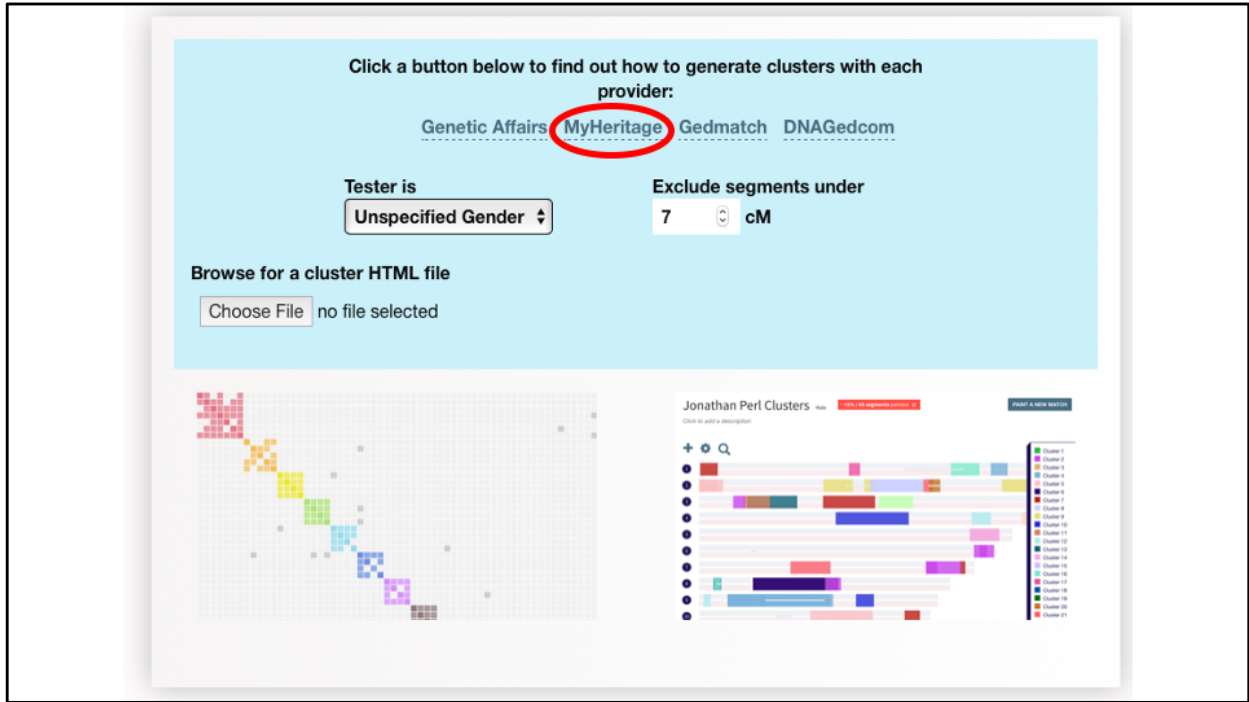
To understand what is required to paint clusters let's consider MyHeritage.
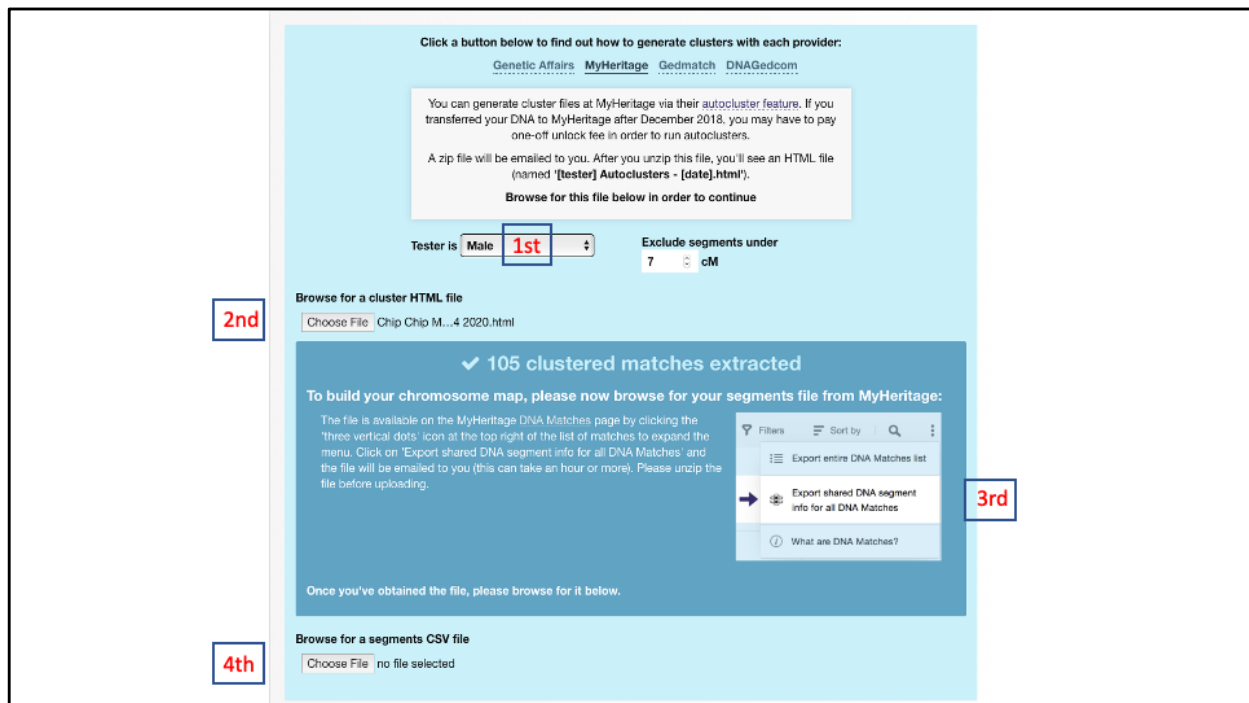
MyHeritage is easiest way to get started painting your clusters. First you choose whether DNA Painter will map one or two X-chromosomes. Second you choose the html cluster file you downloaded from MyHeritage. Third, you download the segment data for your matches, a csv file. DNA Painter provides you directions of where to find your segment data at MyHeritage. Finally, you select the segment data file you just downloaded.
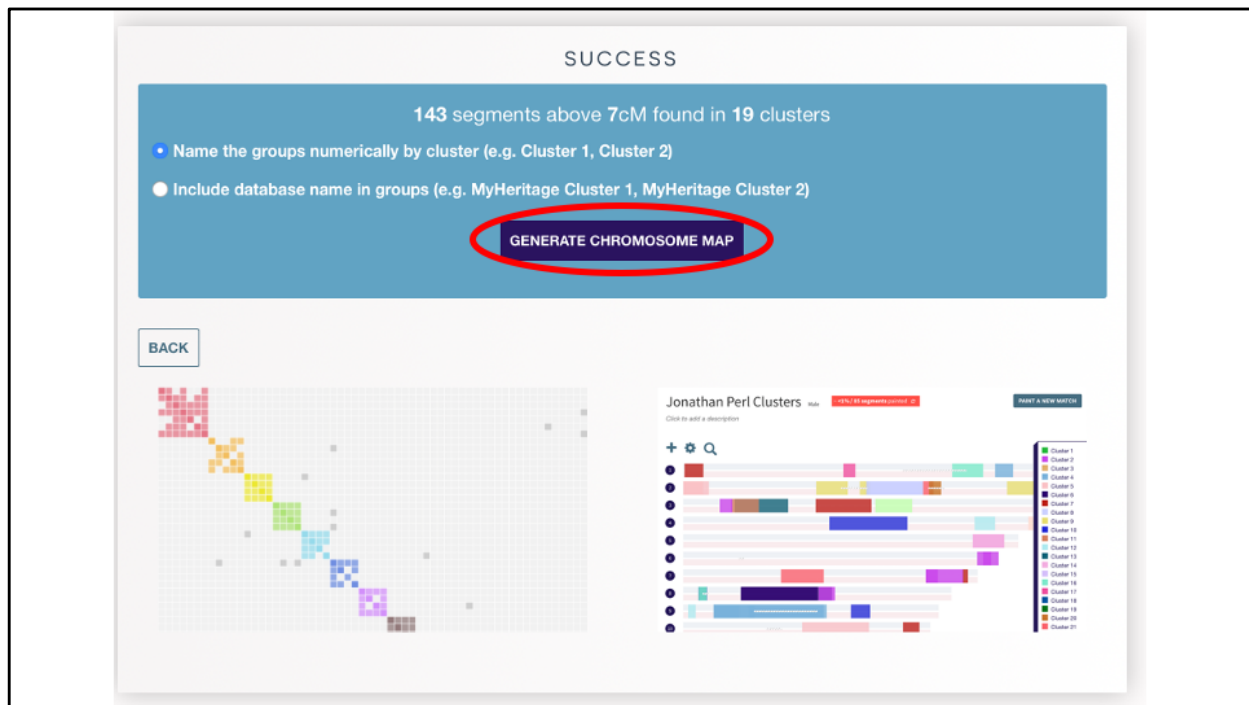
Now DNA Painter is ready to generate your cluster's chromosome map.
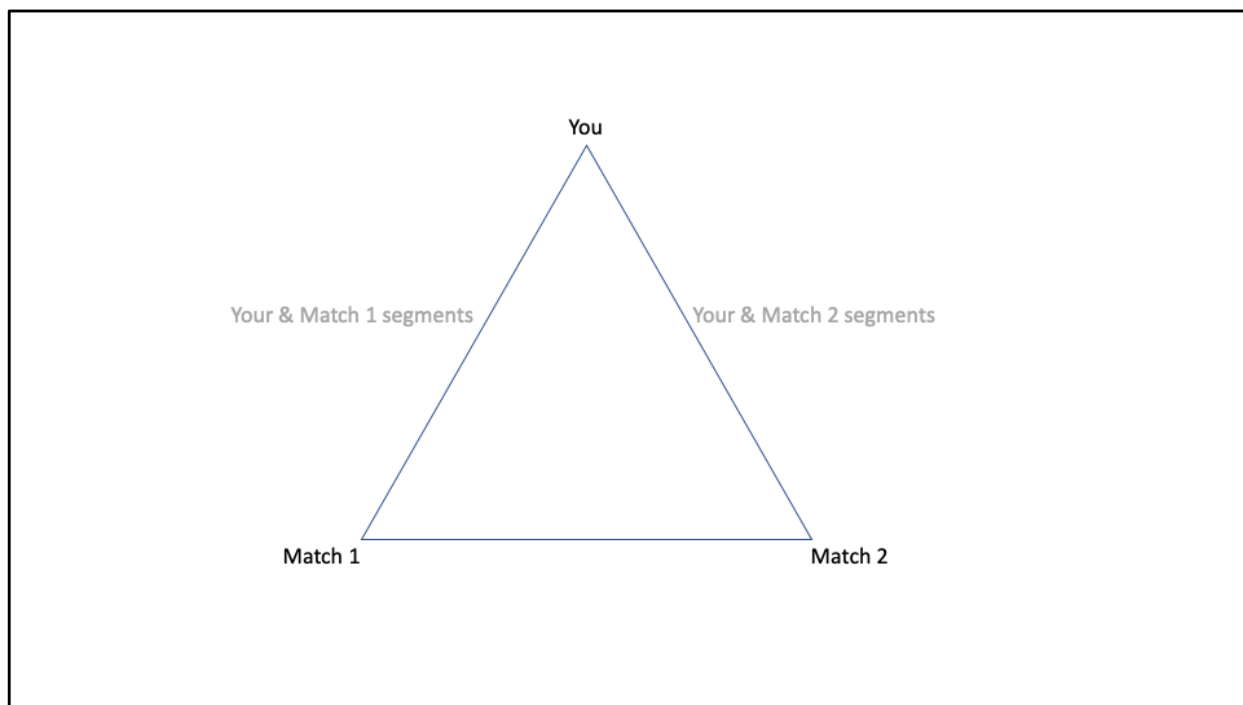
Why do you have to upload two files to DNA Painter to paint your MyHeritage clusters?   The tradition clusters diagram displays information on your matches and who you have in common with your matches. For example, you match person 1 and person 2 and that they match each other. Segment data may have been used to create the clusters, but it is not displayed on a traditional clustering diagram. Without segment data there is nothing to display using DNA Painter. Your segment information provides a quantitative link between you and your matches. This is the first reason to use DNA Painter: to examine the segments of your clusters. DNA painter provides a convenient way to examine the individual segments your share with your matches.
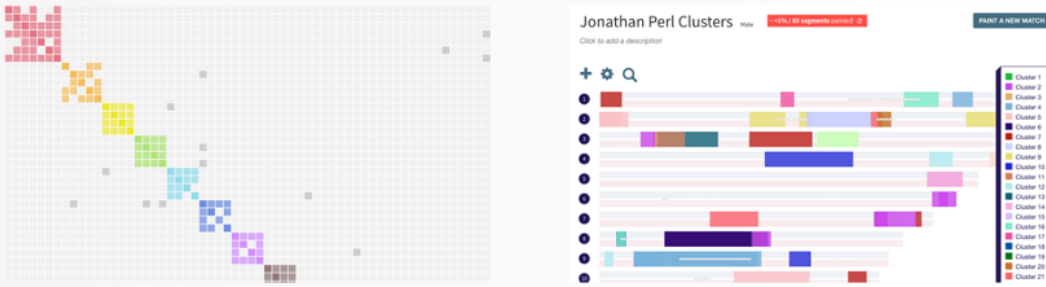
If you choose to use clusters created by Genetic Affairs, segment data is imbedded in the html file, so you don't have to upload a separate segment file.

When you choose Genetic Affairs instead of MyHeritage, this is what is displayed. AncestryDNA should be removed from the first paragraph. What does the second paragraph mean? AncestryDNA does not provide segment data so you cannot map AncestryDNA segment data using DNA Painter.

If you don't want to give Genetic Affairs your passwords, you can create your own tables of matches and segments and have it processed into clusters by Genetic Affairs. There is a small fee to use Genetic affairs. I recommend using the "*Purchase credits using a single charge*" option. For me that has been much cheaper than a subscription.

Settings used for this AutoCluster analysis

As requested, cM thresholds of 250 cM and 50 cM were used. In addition, shared matches should at least share 10 cM with each other. A total number of 21 matches were identified that were used for an AutoCluster analysis. There should be one ZIP file attached to this email. The unzipped files contain an HTML file that holds a visual representation of the AutoCluster analysis if enough matches were present for the clustering analysis. The unzipped Excel file holds the matches from the clustering as well as the compelete matches in several worksheets. Please note that some files might be displayed incorrectly when directly opened from this email. Instead, save the zip files to your local drive, unzip them and then open the files.

Note: The following 7 matches met the inclusion parameters but were placed in an cluster without other members and so are not included in the chart:

These are my 23&Me clusters created using Gedmatch for matches >50 cM & <250 cM requiring the matches share at least 10 cM. One reason I like Genetic Affairs is that you can change these parameters. Lowering the minimum cM results in more clusters which, unless they contain matches with known MRCAs, adds segments to you DNA Painter chromosome map that you can't parse.
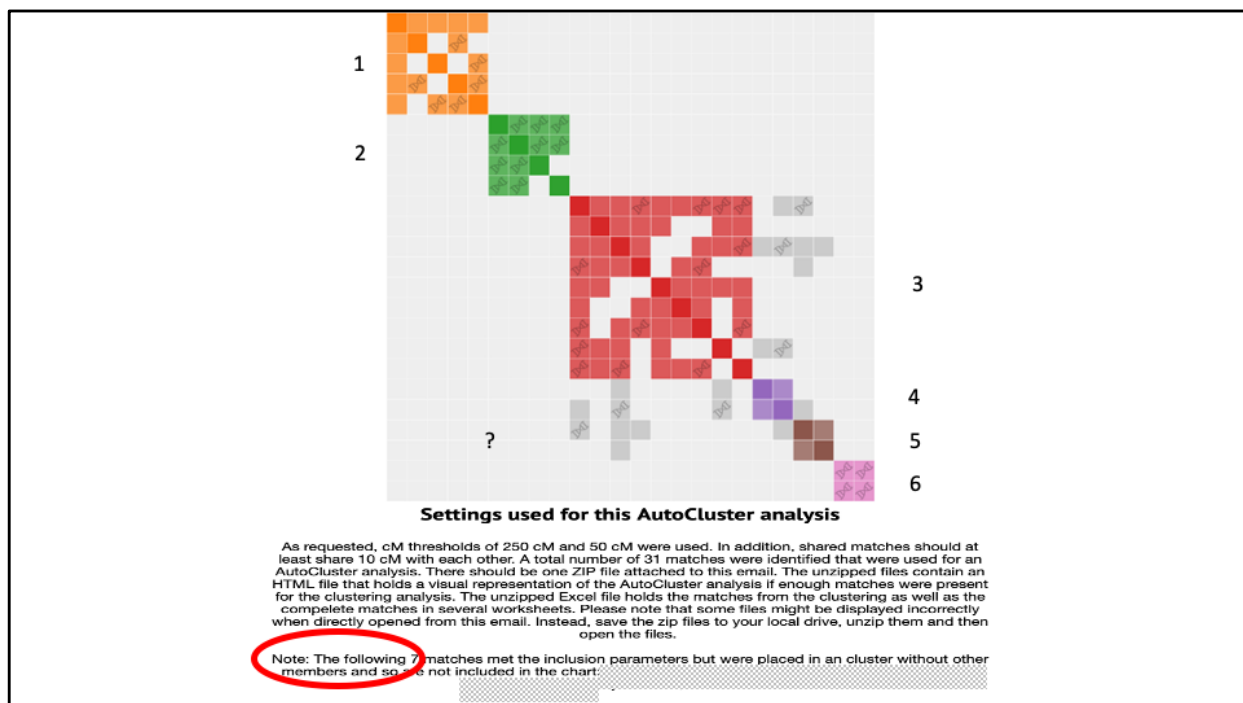
Looking at the list at the bottom of the page, I saw there are two matches with good papers trails to our MRCAs that were not included in the cluster. They were not included because they did not meet the minim 50 cM cutoff. That suggests I may want to import those segments into my cluster chromosome map later. I can also do the same for matches from MyHeritage, FTDNA, and Gedmatch.

Lest you think my presentation is a plug for Genetic Affairs and clustering, I tried both some months ago and thought they weren't helpful. Now that I am using them to feed DNA Painter, I have found both Genetic Affairs and clustering useful.
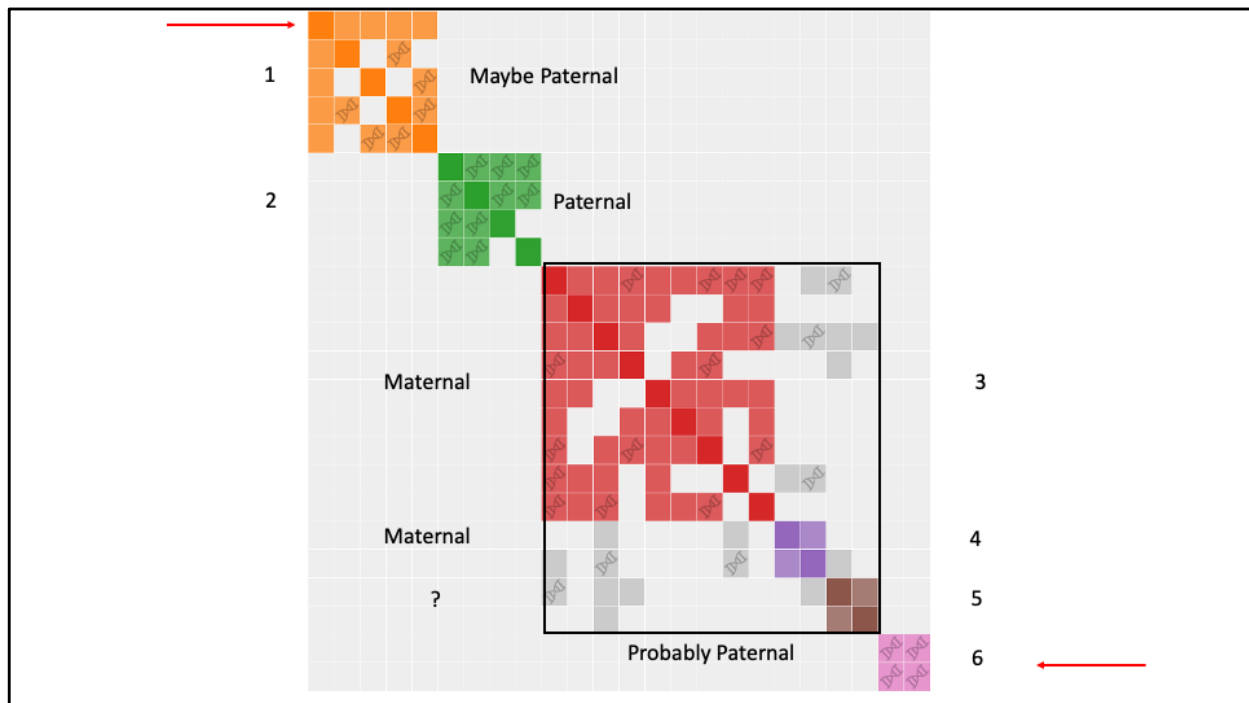
I have labeled clusters 2, 3, and 4 paternal or maternal based on matches in those clusters for whom I have good paper trails to our MCRAs. Cluster 1, 23&Me places on my paternal side. Painting that cluster, I noticed that the person of the top line didn't show up in my DNA painter chromosome map. That is because even though he is in this cluster, he has not shared his segment data. If and when, I am able to determine the MCRAs for the cluster and place the other people in my tree, this person would be a candidate for the WATO tool. Cluster 6 is equally interesting. I believe it should be paternal, but 23&Me switched it from paternal to maternal when my brother joined 23&Me. I hope by examining segment date with DNA Painter I'll identify leads to solve this dilemma.

Note that clusters 1, 2, and 6 have no matches with any other cluster. Whereas clusters 3, 4, and 5 have crossmatches suggesting they are a super cluster. Thus, I may have only four distinct MCRAs in this diagram.
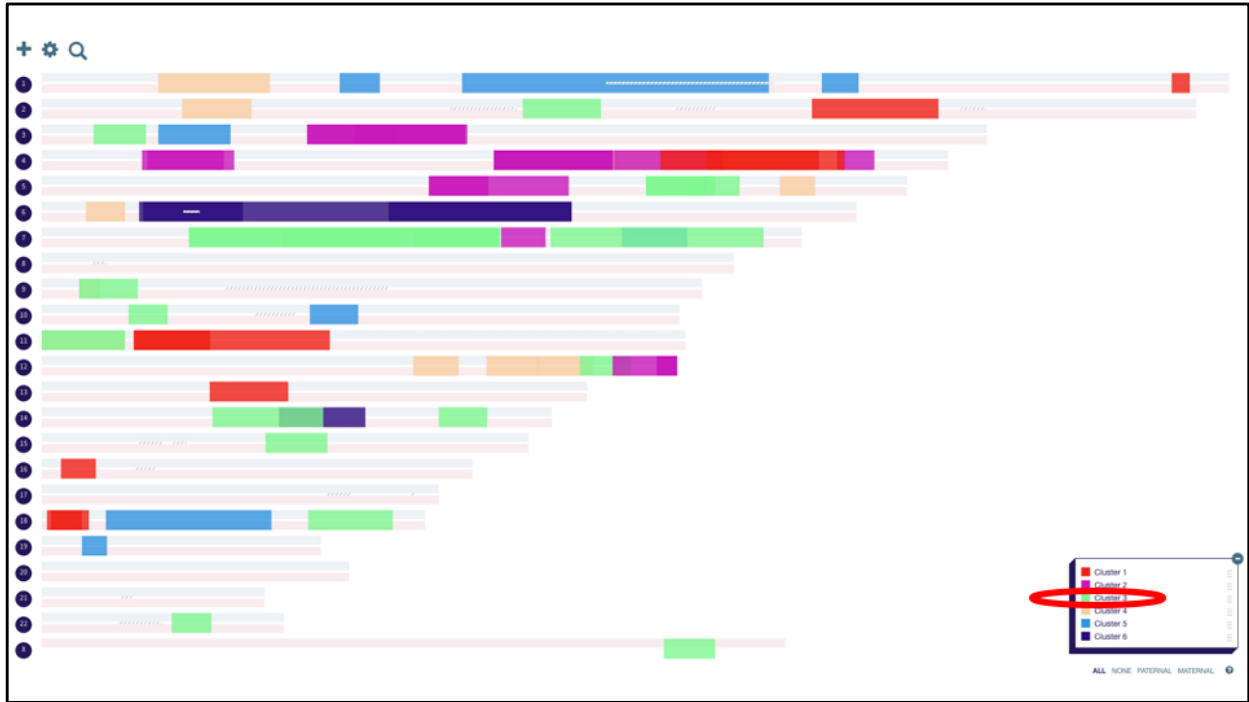
When my clusters were imported into DNA Painter this is what I saw. The data spans both the paternal and maternal horizontal bars. One needs knowledge of matches with MRCAs to parse the data and update the labels. To do this click on the appropriate label. Let's start with cluster three.
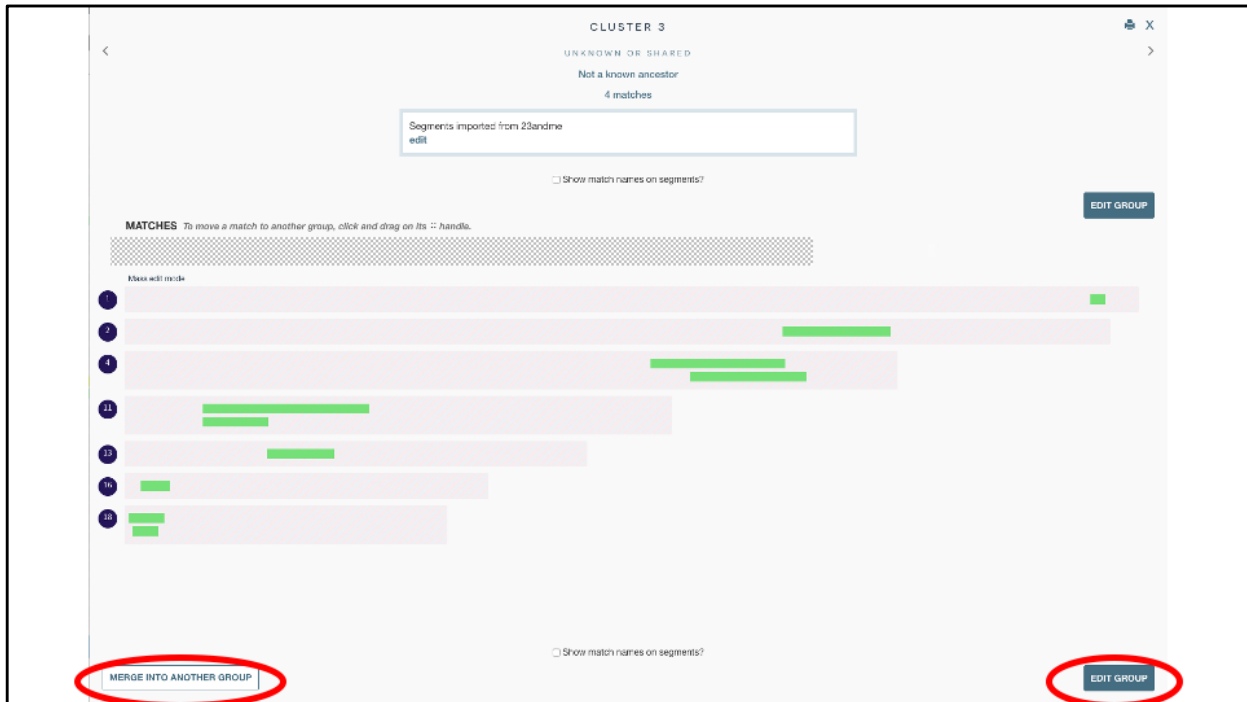
When you click on the label it brings up this dialogue. At the bottom of the page there are two options: Merge into Another Group and Edit Group. To parse cluster 3 into my maternal side and change the label to reflect my MRCAs for this group of matches we need to choose Edit Group.

EDIT GROUP   X

Group Name

3: Jedediah Allen & Sarah Jane Cope

☐ **Known ancestor(s)?**

☐ **This is a group of ancestry composition segments**

Side

Maternal ⬍

Colour

🟩 ▾

☐ **Set a confidence level for the group**

Optional: indicate how certain you are that you've identified these segments correctly for the entire group.

The level shown above is a placeholder and does not reflect the current certainty of segments in this group. If you tick this box, the level you choose will overwrite any individual match settings.

☐ **Add a dividing line below this group in the key**

Notes

Text only please! Links starting with http:// and https:// will be converted to hyperlinks when this text is displayed.

Segments imported from 23andme

DELETE ENTIRE GROUP   VIEW GROUP   SAVE GROUP

Clicking on Edit Group pops up this dialogue. I have typed into the Group Name my MRCAs and have selected maternal for the side.

Note that the Green group includes a significant X-chromosome segment. A check that I parsed the group correctly. Clicking on the circle labeling a particular chromosome opens that chromosome. You can now see that the Green cluster is maternal whereas the other clusters are shared or both. For the chromosomes that are not opened you can see the green as maternal and the other groups spanning both paternal and maternal.

A feature of DNA Painter I like is that when you click on a particular segment (in this case the green segment of chromosome 11), this navy colored dialogue pops up. It tells me the base pair span of the segment I clicked on, the number of cM, and SNPs. If I were to click on beige colored segment just above the green segment currently selected, I would find exactly the same base pair span. This makes sense because looking back to see who these two individuals are, I my paper trail work implies they are brother and sister. That could be further checked by entering their shared cM in the Shared cM Tool. Knowing this I'll merge cluster 4 into cluster 3. I'll also parse cluster #2 since I know the MRCA for members of that group. Next, I'll import the two known paternal matches that Genetic Affairs did not included in the cluster diagram. I add those matches by exporting the segments using the All Segment Data feature in my primary DNA Painter chromosome map and the using the import segment data feature to incorporate them here. Before doing that, just to be safe, I made a copy of my work so far.
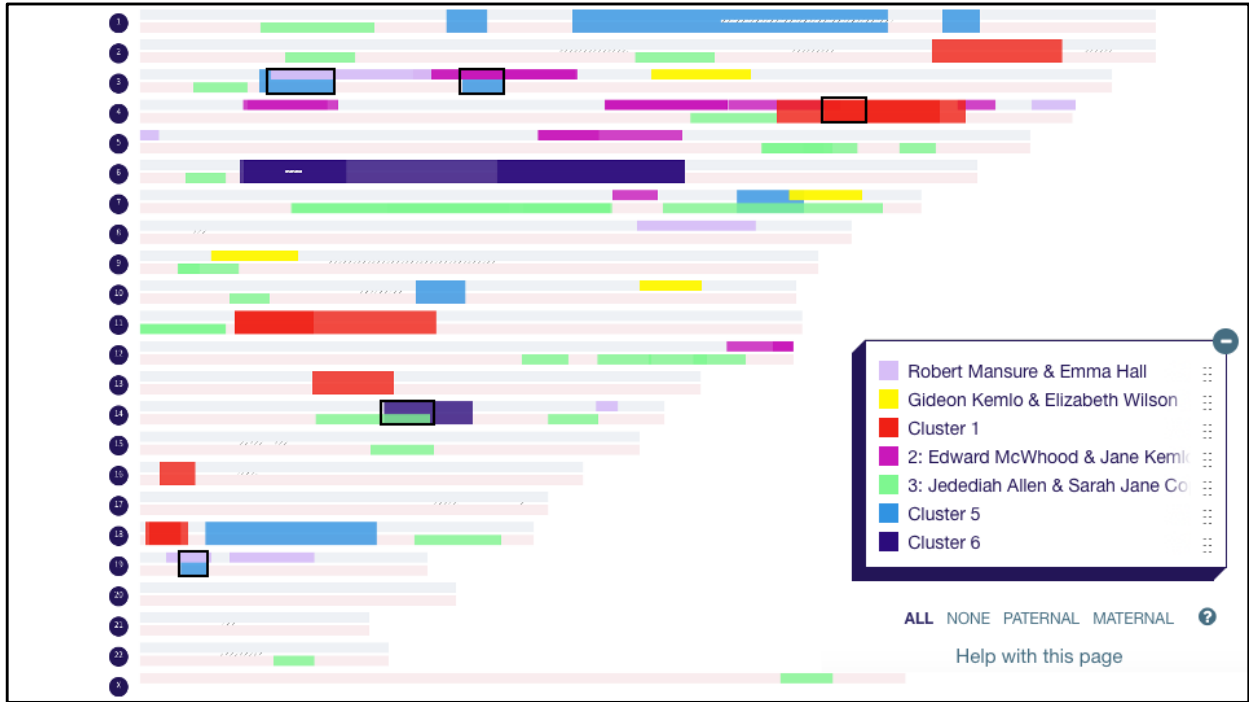
At this point I have noted with boxes overlaps that should be examined.

Looking at chromosome 3. MS and WS have significant overlap. Thus, I should check if they are a match. Given that I have a good paper trail for WS, a match may provide clues for finding my MRCAs for MS. WS and BF also overlap. They are supposed to be on distinct parental lines, however. The overlap is small enough that it could be just by chance. In my wife's DNA chromosome map, I find large overlaps with conflicting MRCAs with fairly good paper trails. That is not surprising because that part of her tree is endogamous. The action in such cases is to look for additional consistent common ancestors.

Examining the results of chromosome mapping my clusters has already led to number of leads. In conjunction with my 23&Me tree I have identified cases where I can say this person should be a descendent for either of two couples. I thus can message them: I think we are third cousins and descent from couple A or B. Hopefully, that is enough to entice them to reply.
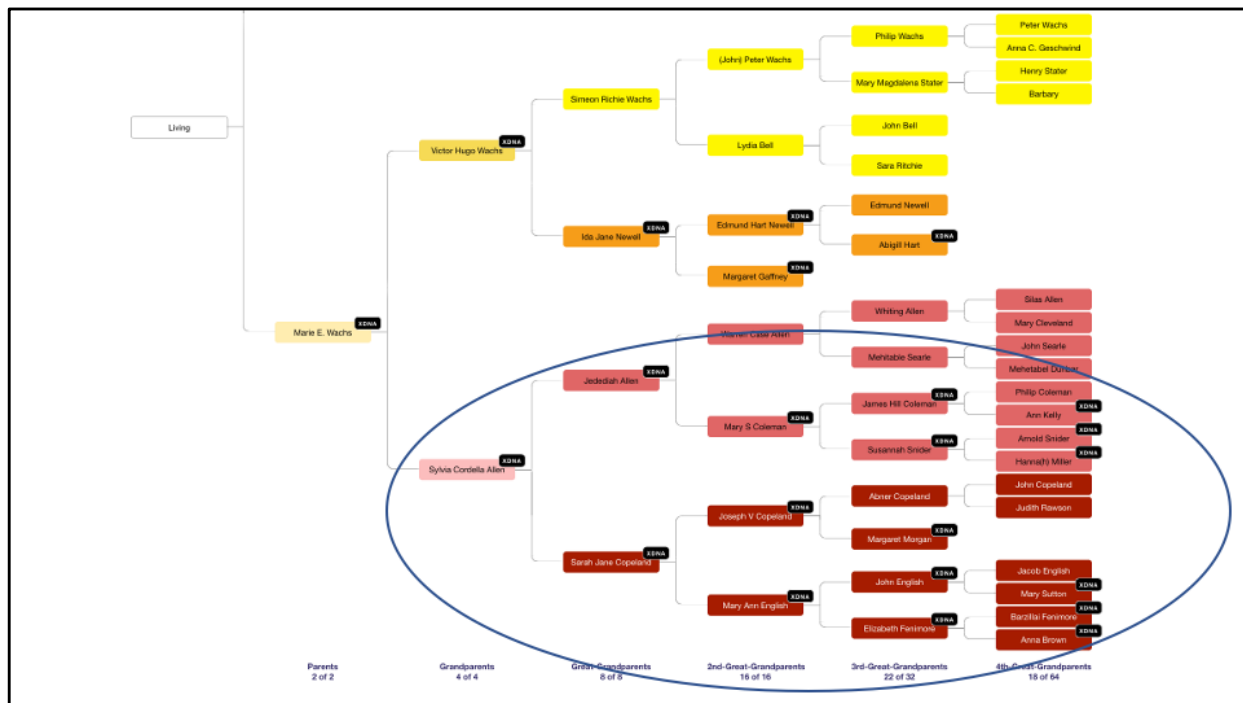
Before painting my Jedidiah Allen & Sarah Jane Copeland cluster, I had not noticed the X-DNA match. Using the X-DNA filter on my tree at DNA Painter I can see potential ancestors that I have in common with this X-DNA match. Further knowing the MRCAs of the cluster that my x-match belongs to, I can limit the possible common ancestors to my X matches in this oval. With that information I contacted the match suggesting several possible MRCAs based on the cM we share and am hoping for a reply. Thus, DNA Painter has given me leads to follow and help prioritize my work.