Clustering With MyHeritage Labels

By Michelle Leonard M.A. PgCert

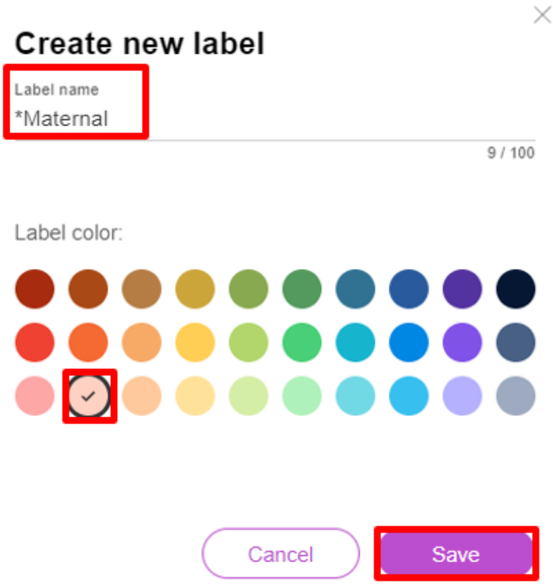
This webinar demonstrates how to use MyHeritage 'Labels' to organize your match list. It outlines different clustering methods that can be used in different situations and provides step-by-step guidance to help you on the path to a more organised MyHeritage match list.

What Is Clustering?

- Clustering is a technique that can be used to organise DNA matches into distinct groups that should generally represent different specific lines on your tree
- Clusters are creating by gathering together groups of shared matches and this means shared match lists are the key to clustering
- If a group of matches mostly all share DNA with each other as well as the test taker, then it's very likely the test taker and the matches in that group share a common ancestor or common ancestral couple
- The matches could be on many different generational levels along that line depending on the amount of DNA shared with each of them
- Some shared matches, however, will match the test taker and the comparison match in different ways so we always need to be on the lookout for red herrings that could corrupt our clusters
- There are several ways that your MyHeritage DNA match list can be clustered:
    - You can manually cluster matches on MyHeritage using the in-site 'Labels'
    - You can also create off-site custom spreadsheets
    - In recent years excellent developers have also created automated clustering tools such as 'AutoCluster' on MyHeritage
- This webinar will concentrate on how to use the in-site 'Labels' on MyHeritage to best effect for clustering purposes

Tips For Clustering with MyHeritage Labels

- MyHeritage provide 30 colour-coding 'Labels' you can use to create clusters and these work in a very similar way to the 'Groups' facility provided by Ancestry
- Since there is no SideView split on MyHeritage, the first task I recommend you undertake is to try to separate out the match list into high level 'Maternal' and 'Paternal' side clusters
- If you have a parent tested, always use that parent to separate out the match list as accurately as possible



Create new label

Label name: *Maternal

9 / 100

Label color:

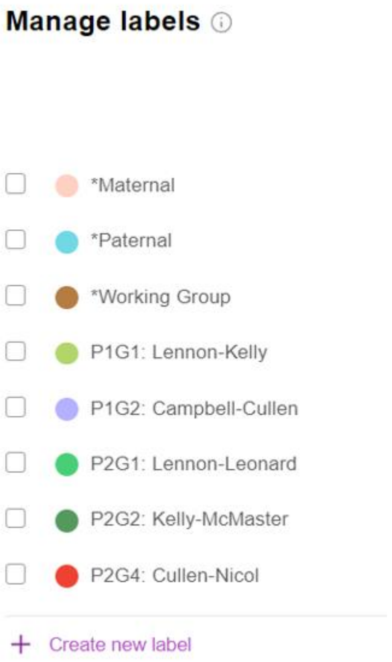
- If you don’t have a parent tested (or you are working on your parent’s match list or the match list of someone else without a parent tested) then use the closest identified matches that pertain to just one side whichever side that may be
- When naming a ‘Label’ on MyHeritage, remember the hierarchy of the naming organisation is symbols first followed by numbers then alphabetical – always try to organise your naming in a way that makes sense to you
- Be consistent with your group names and colours across all sites on which you undertake colour-coded clustering
- Be aware of your browser’s zoom ratio as you must access the ‘Show More DNA Matches’ button in order to add more than 10 matches at a time to your clusters
- Decrease the zoom ratio to dot more matches in a quicker time frame and reduce the need to scroll as much but remember that this works best on a larger screen
- The ‘Click All Checkboxes’ Google Chrome extension can’t be used on MyHeritage so remember that you have to do the dotting manually
- Bank your clicked matches after every few rounds of 10 clicks just in case something goes wrong (e.g. an inadvertent refresh after clicking several hundred matches would result in you having to do those clicks all over again)
- Add all shared matches down to at least the 20cM level to the label (maybe even go lower if using a parent)
- Be careful with both sides matches and deal with those separately
- Switch between the ‘Sort By’ options – sort by both shared DNA with the comparison match and shared DNA with the test taker to make sure you don’t miss adding any matches over 20cM for one or the other to the label
- I generally dot down to 20cM (using both versions of ‘Sort By’) so that all matches in the group share over 20cM with at least one of the test taker or the comparison match but there’s no hard and fast rule here (use your judgement!)

- Remember that MyHeritage does not have a 20cM shared match list cut-off like Ancestry so you could go lower with your clusters, but be aware that you will bring in more anomalies and false matches the lower you choose to go
- If you see a 'Triangulated Segment Icon' next to a shared match lower than 20cM, then it is more likely this is a match that should be added to the label
- Beware of duplicate dotting! If you have clicked on a match that is already in the 'Label' then the system won't let you click 'Apply' until you have clicked on the line next to that label and changed it to a check mark
- Always expand the side 'Label' with as many other identified matches as you can! You don't need to do this when you have a parent tested for that side but if you don't have a parent then this is an important step as different relatives on a particular side of your ancestry will have inherited different pieces of DNA from the same ancestors and will, therefore, have some different shared matches that need to be gathered into your 'Maternal' or 'Paternal' label

The DAG & ABCD Methods

- Once you have created and expanded your 'Side' labels, move on to the Direct Ancestors Group (DAG) Method and/or ABCD Method depending on what your particular situation or research goal requires
- The Direct Ancestor Groups Method uses identified matches (the matches with whom you know your common ancestors) to create labels for those shared common ancestors
- It is generally best to name your labels after ancestral couples (e.g. Kelly-McMaster)
- Add a label to a 'base match' who must be an 'identified' match that descends from the ancestral couple (or single ancestor) for whom you want to create a label (the DNA you share with the base match must track back exactly to that couple)
- You need a match at the 2nd cousin level for a great grandparent label, for instance, and a match at the 3rd cousin level for a 2nd great grandparent label
- I generally recommend starting to create labels from the great grandparent or 2nd great grandparent couple level so you have more colours to play with
- It is a good idea to create a 'Working Group' label for matches you wish to investigate when you have time and an 'Uncategorized' label for matches you can't currently assign to either one side or the other
- As your 'Labels' start to take shape, try not to use up all of your colours – it's important to hold back a number to be used on mystery groups and specific research questions
- It is generally best not to create a 'Label' if you can only find 1-3 matches that would belong in it as that's a waste of a colour (instead add good notes to these matches)
- Once you have exhausted your identified matches with the DAG Method (or if you don't have many identified matches to use that method) move to the ABCD Method

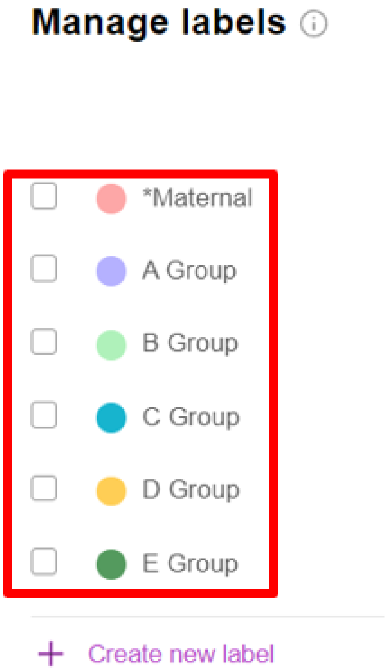
- The ABCD clustering method separates unidentified matches into genetic labels using an alphabetical system



Manage labels

☐ ● *Maternal
☐ ● *Paternal
☐ ● *Working Group
☐ ● P1G1: Lennon-Kelly
☐ ● P1G2: Campbell-Cullen
☐ ● P2G1: Lennon-Leonard
☐ ● P2G2: Kelly-McMaster
☐ ● P2G4: Cullen-Nicol

+ Create new label



Manage labels

☐ ● *Maternal
☐ ● A Group
☐ ● B Group
☐ ● C Group
☐ ● D Group
☐ ● E Group

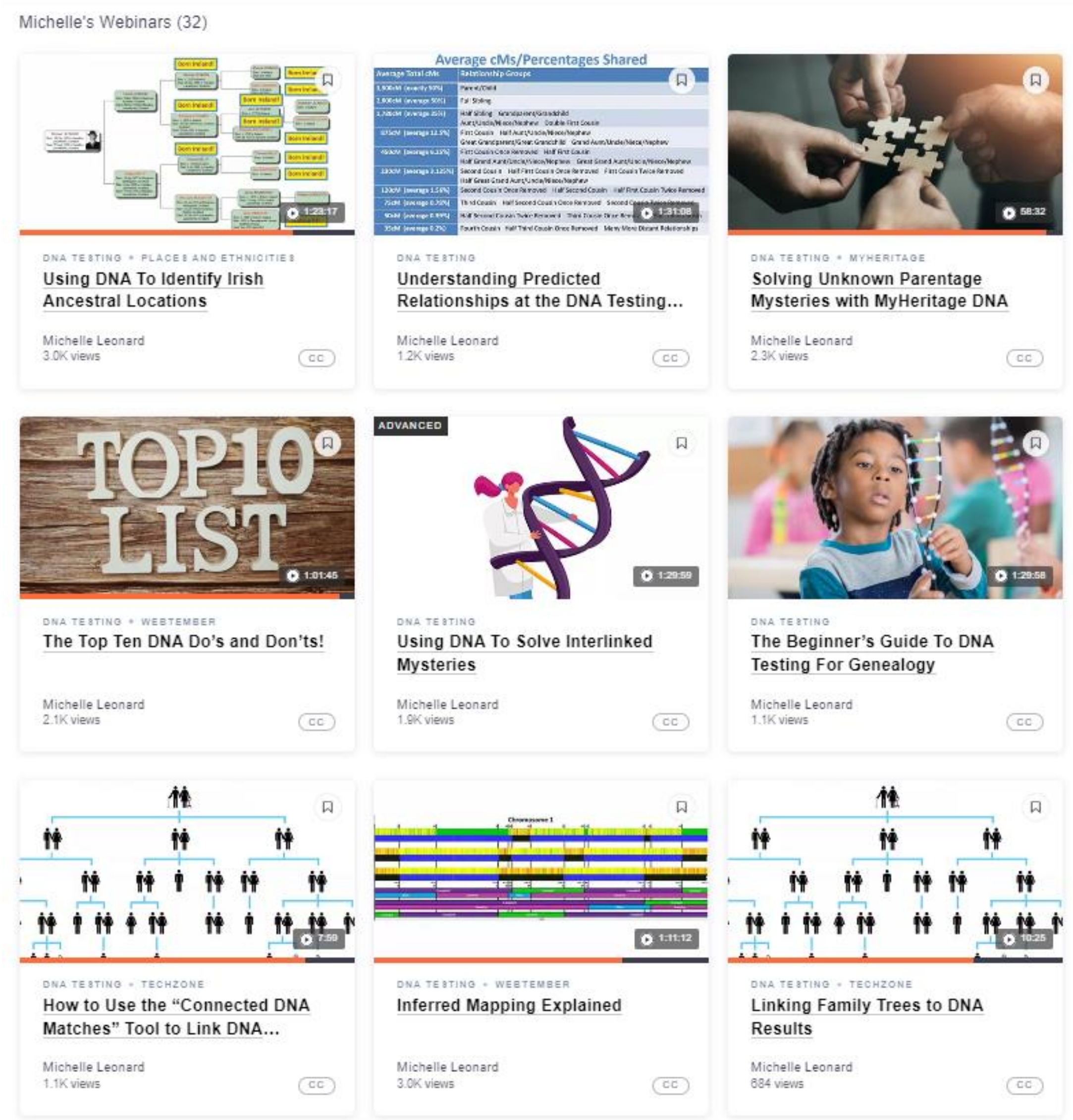
+ Create new label

- It’s important to choose a ‘base match’ that doesn’t share too much DNA with the tester to avoid merging together labels that should be separate
- I tend to start with a match who shares in the 100-250cM region
- You don’t need to know anything about how that match relates to you – just give them a label colour, dot their shared matches and name the label ‘A Group’
- Expand the label with other shared matches who end up in it, then repeat the process with new matches who didn’t receive the first dot colour (B Group, C Group etc)
- Simply scroll down your match list until you find another undotted unidentified match and use them as the ‘base match’ for B Group and so on
- Don’t be scared of making mistakes – if you feel that one of your ‘Labels’ has got corrupted in some way (it will happen!) then you can always easily delete it and start again being more careful to avoid any red herring matches

Summing Up

- Clustering is a powerful technique that can provide crucial insights into how groups of matches may relate to you
- Clustering with the in-site ‘Labels’ is the best way to organise your MyHeritage match list and create a reference for narrowing your new matches as they appear on the list
- It can be used to narrow connections down and identify red herrings
- Working through both manual and automated clusters can help you spot connections you may otherwise miss

If you enjoyed my presentation and would like to hear more, I now have over 30 long-form and TechZone presentations in the Legacy Family Tree Webinars library. You can check them all out here:



Michelle's Webinars (32)

DNA TESTING • PLACES AND ETHNICITIES
Using DNA To Identify Irish Ancestral Locations
Michelle Leonard
3.0K views

DNA TESTING
Understanding Predicted Relationships at the DNA Testing...
Michelle Leonard
1.2K views

DNA TESTING • MYHERITAGE
Solving Unknown Parentage Mysteries with MyHeritage DNA
Michelle Leonard
2.3K views

DNA TESTING • WEBTEMBER
The Top Ten DNA Do's and Don'ts!
Michelle Leonard
2.1K views

DNA TESTING
Using DNA To Solve Interlinked Mysteries
Michelle Leonard
1.9K views

DNA TESTING
The Beginner's Guide To DNA Testing For Genealogy
Michelle Leonard
1.1K views

DNA TESTING • TECHZONE
How to Use the "Connected DNA Matches" Tool to Link DNA...
Michelle Leonard
1.1K views

DNA TESTING • WEBTEMBER
Inferred Mapping Explained
Michelle Leonard
3.0K views

DNA TESTING • TECHZONE
Linking Family Trees to DNA Results
Michelle Leonard
684 views

https://familytreewebinars.com/michelleleonard

Contact & Social Media Links:

Email: michelle@genesandgenealogy.com
Facebook: www.facebook.com/genesandgenealogy
X (Formerly Twitter): www.twitter.com/genealogylass
LinkedIn: https://www.linkedin.com/in/michelleleonardgenealogist
APG: https://www.apgen.org/profiles/michelle-leonard

Genes & Genealogy