Ancestry.com Matches by Cluster Tool

GSSCC DNA Special Interest Group - - handout compiled by Gail Burk
September 2, 2025

Ancestry.com's clustering tool, called "Matches by Cluster," organizes your DNA matches into colored blocks, each representing a shared family line. A Pro Tools subscription is required in order to view AncestryDNA clusters. At the top of your DNA match list, click on the tab for "By Clusters." This feature helps users, especially beginners, to better understand their matches by providing a visual organization of potential family connections without requiring manual grouping. Matches by Cluster is being rolled out in phases; some Ancestry DNA members will not be able to access this feature until December, 2025. Clustered matches don't necessarily match on the same segments, which is what defines triangulation. Matches by Cluster is not a triangulation tool. Matches by Cluster is currently in Beta mode; future changes to the clustering tool are likely! Ancestry has announced that soon you'll be able to create custom clusters with specific matches and cM ranges.

The Matches by Cluster tool organizes matches into color-coded blocks that represent distinct family lines. Clusters help identify ancestors by showing which clusters your matches belong to, enabling you to more easily identify unknown common ancestors. Ancestry's Clusters use "grey cells" to indicate intersections. Grey cells in the visual display show connections between matches who are assigned to more than one cluster, helping researchers to trace relationships across multiple family lines.

The tool may be less effective for users whose DNA results indicate endogamy, that is, those DNA testers whose family members are part of isolated communities with significant intermarriage. Endogamy leads to large, overlapping clusters which are less helpful for confirming DNA match relationships.

Matches by Cluster charts matches between shared DNA levels of 65 cM and 1300 cM. By focusing on matches between first and third cousins (people with whom you share between 65cM and 1,300cM of DNA), Ancestry can form clusters that are smaller and more informative. These clusters often contain matches who share an ancestor with you at the grandparent or great-grandparent level.

Not everyone in a cluster matches everyone else in that cluster, but everyone matches some of the other cluster members. Ancestry includes some qualifying people in clusters, but not all people who qualify. It is unclear why, or how, that decision was made, but not every match who meets the qualifications is included in a cluster. Stated simply, people who all match you above 65 cM, AND match each other at 20 cM or greater, after application of Ancestry's Timber algorithm, are candidates to form a cluster.

Ancestry's clustering tool organizes your DNA matches based on their relationships to each other. They start by mapping the connections between your matches and all of their matches. This tracks how you're all connected and how much DNA you share. Ancestry removes your closest relatives (like parents, siblings, half-siblings, aunts/uncles, and nieces/nephews) and removes your matches who are more distant than third cousins.

Adding too many distant matches can make it harder to identify meaningful clusters. Some distant matches may be a result of people descending from a common population rather than a common ancestor. This can disrupt the clustering in a way that isn't informative about your family history. After lots of experimentation, Ancestry believes it has found a good balance by leaving out matches who are too distantly related, those who share less than 20cM of DNA with each other.

Ancestry's clustering algorithms find sets of matches who are more closely related to each other than to your other matches. Your entire group of clusters is shown together initially, either in the grid format, which Ancestry named "chart format" if 100 or fewer, and in a list accessible via dropdowns for everyone. If you have more than 100 clustered matches, you will not see a grid view; but you will still see the clusters in the list view.

As mentioned above, not everyone included in a cluster, however, will match every other person in that cluster. When viewing your clusters, the colored cells indicate matches, meaning those people also match each other at 20 cM or greater. The non-colored or "blank" cells in each colored cluster block indicate that those two intersecting people don't match each other at 20 cM or more. It does not mean they don't match each other at all, just not above 20 cM, which is the lowest amount of shared DNA between your matches that you can see using Pro Tools.

In forming clusters, Ancestry uses its SideView™ technology to separate the map of your matches into two sides, representing matches from each side of your family, maternal and paternal. This helps build more accurate clusters. Everyone who can should utilize SideView. Ancestry does its best to assign clusters to one parent or the other, even if you don't (or can't) designate parental "sides," meaning which parent is which. Matches on your list who are "unassigned" aren't placed into any cluster.

Matches by Cluster will generally work for most people who have taken an AncestryDNA test. However, if you don't have many matches sharing 65 cM of DNA or more with you, or your matches were left "Unassigned" by SideView, then you may not see a Matches by Cluster result. Some factors complicating clustering include: individuals of recent origin from East Asia, South Asia, and South America who will likely not have enough matches in the database to generate many clusters. Individuals of Ashkenazi, French Canadian,

Puerto Rican, Pacific Island, or other populations where there was a lot of intermarriage, will likely have large, overlapping clusters. That makes it difficult to identify a single common ancestor.

Every time you sign in to AncestryDNA, or click on "By cluster," Ancestry recalculates your clusters, so the clusters may be slightly different from one time to the next. This could be a function of Clusters being in Beta. New matches may also affect reclustering. Basically, Ancestry uses between first and third cousins to define clusters.

There are many different clustering methods, each with its own strengths and weaknesses. Ancestry uses several and compares results. This approach is called an ensemble strategy. Ancestry also applies additional rules about how closely people are related when performing the clustering in order to handle complex relationships created by endogamy or pedigree collapse, where people are genetically connected in more than one way.

Once your match clusters are ready, you can explore each one to discover which common ancestors those matches share. Click on any spot in a cluster to view information about the match, or scroll down the page to look at the list of members in each cluster. With Matches by Cluster, you can research the members of the cluster by looking at their family trees and the amount of DNA they share with each other to help identify common ancestors. You can then compare this information with your own family tree and add the DNA matches who are descendants of your common ancestors to your tree. Clusters can help to validate, confirm, or refine your genealogical research. If there is a cluster you cannot place in your family tree, that cluster might help you break down one of your genealogical brick walls.

Clusters are particularly valuable for adoptees and others with recent unknown parentage. You can use them to identify close relatives and biological ancestral lines. Because clusters are displayed directly on the Ancestry site, you can easily add notes and colored dots to group matches as you analyze each cluster. For DNA matches without trees, matches by cluster will still help you identify which branch of your family tree the person belongs to.

In working with Matches by Cluster, go to the top of the cluster page, where you'll see tabs for "All," "Parent 1," and "Parent 2." The default view is "All," so clusters from both parental sides are included in this display, assuming you have clusters on both sides. By clicking on either Maternal or Paternal at the top, you see a view of multiple maternal or paternal clusters, so long as the total number of cluster members isn't over 100.

Hovering over any colored cell tells you how closely these two individuals are related to each other. Scrolling down below the cluster chart displays your match information to each cluster member, including whether they have a tree and their estimated relationship to you. It's important to keep in mind that clusters may contain people with generational differences. Matches in one cluster may also be in other clusters, as represented by the grey cells on the grid. So, for any one person, to see how many clusters in which they are included, and who they match in every cluster, just follow their row straight across, left to right. Or straight down, top to bottom.

Note that when viewing cluster members in the cluster dropdown boxes, if someone is a member of two or more clusters, he or she is only listed as a member of one cluster, not both. Therefore, any individual person is only listed once, not in each cluster of which they are a member. This makes it unnecessarily difficult to identify match connections and associate them with ancestors. Every cluster has its own unique history and it is your job to figure out which ancestral couple each cluster represents.

Roberta Estes suggests this workaround: Create a spreadsheet or chart and manually add the people associated with two or more clusters. Just follow each person's row across left to right and use the grey squares in the cluster chart. Your analysis will reveal why they are members of multiple clusters. This approach works so long as you don't have more than 100 people on either your maternal or paternal side, respectively. If the page of clusters is larger than 100, you can't see the multiple cluster chart image, so you're out of luck tracking matches in multiple clusters because you need to see those grey cells!

Roberta also recommends creating what she calls "Cluster Groups." Use Ancestry's colored dots to tag a group for each of your clusters. You can either select an existing group, or create a new one. You can assign this group identifier to everyone in this cluster, or just some people by checking their box (or not). Remember, your matches are only listed in one cluster, so you'll need to add people into multiple cluster groups manually. After creating cluster groups, when you return to your primary DNA match list, it will be easy to see who is and is not included in a cluster, or multiple clusters, based on the group tags.

There are some frustrations with Ancestry's Matches by Cluster, to be sure. To quote another genetic genealogist, many qualifying matches who clearly meet the cluster criteria "have been left on the cutting room floor." Several of my own close cousins are missing from my clusters. In spite of having identified dozens of close cousins, Ancestry's Matches by Clusters only shows four clusters for me (two paternal and two maternal), and a total of a mere 43 DNA matches in those four clusters.

Known cousins are key to identifying DNA matches. Adding to the confusion, many people who match you more distantly, and match your excluded cousins, may be included in clusters. Even more confusing is that a missing cousin may be found in Ancestry's ThruLines, but not in your cluster where that cousin belongs. ThruLines makes it even easier to identify the cluster, so one would think that matches with ThruLines would be a priority to include in clusters.

AncestryDNA testers who have already done a lot of grouping and analysis may be underwhelmed by the current version of Matches by Cluster. I share that disappointment. Randomly, of my six paternal first cousins who have tested, half of them are assigned to a cluster tied to our grandmother's Scottish father, and the other half are assigned to the cluster for our grandfather's Nova Scotian mother. No clusters at all for DNA matches tied to our grandmother's English mother or our grandfather's Pennsylvania biological father, despite my having matches for these omitted groups who meet ancestry's cM criteria. I, for one, would like to be able to set a lower maximum cM threshold so that my first cousins aren't complicating the clusters. I am awaiting the day when we can select cM parameters that will make sense in meeting our personal research goals.

Another issue with Matches by Cluster may be Ancestry's Timber algorithm (which attempts to help offset excess shared DNA caused by endogamy). Ancestry's minimum threshold of 65 cM of shared DNA would often be met except for Timber's interference. But until Ancestry tweaks Matches by Cluster to enable users to select the unweighted amount of DNA, rather than the Timber-adjusted quantity of DNA, Timber will inevitably eliminate cluster matches that otherwise could be included. Many experienced genealogists are working on brick walls more distant than a grandparent or great-grandparent. Until Ancestry makes it possible to adjust the minimum amount of shared DNA in forming clusters, Matches by Clusters will fail to show the helpful matches that are found below 65 cM.

One reason that someone might not be clustered is that two matches are too closely related to each other, such as parent and child. In cases of close cousins, without chromosome mapping, there is no way to determine why Ancestry would break up close cousins who logically would be in the same cluster as each other. As you work down your own match list, you may find that there's little consistency between who is and is not a cluster member. Roberta Estes uses her spreadsheet as a starting point to flesh out these incomplete clusters. While Ancestry's clusters may not cluster everyone that they logically should, you can use the clusters that have formed, and combine that information with shared matches to other cluster members in order to further identify, or at least find, hints pointing to common ancestors.

Ancestry explains that a sister may end up in a different cluster from her sister, her own son, and two granddaughters, due to the concept of "generations" in family trees. In this context, individuals in the same generation are the same number of degrees away from a common ancestor. For example, the sister and her sister are in the same generation as they share the same parents. However, the sister's son and granddaughters belong to different generations, as they are one and two additional steps away from the common ancestor (the sister herself), respectively. This generational difference can lead to them being categorized into different clusters when analyzing family relationships. While this is an explanation, it doesn't seem especially helpful.

There are some ways to use Matches by Cluster to further your DNA analysis. Identify cluster members by different methods:

Because the tester is known to you
Because a match has a common ancestor in their tree
Because you extended their tree to find a common ancestor
Because a ThruLine has been formed

Use the clusters as a starting point, and continue working down your match list. Use shared matches with cluster members to continue to associate your matches with genetic groups, even if Ancestry doesn't assign them to clusters. Your spreadsheet is your friend, and so are notes and grouping tags.

The answer to how clusters can help you depends, in part, on your goals. If you're searching for unknown parents, grandparents, or even great-grandparents, and assuming enough other people have tested, clusters should work well for you. Clusters are essential for people seeking unknown, relatively closely related family members. Note that clusters always identify a couple, not an individual.

You may only be able to identify a cluster match to a generation closer in time. Remember that Ancestry's clustering is not triangulation, so your matches may not match on the same segment. You could match person A due to one set of ancestors, person B due to another set of ancestors, and A and B could match each other due to a third set of ancestors. This, in part, is why clustering is useful, though, as it reduces although does not eliminate, the possibility of that happening. With clustering, you're dealing with groups of people, not just three triangulated people.

When pedigree collapse or endogamy is involved, multiple people may match due to different ancestors that they can't identify because the group of matches shares multiple ancestors. Think of either first cousins marrying each other a couple of generations ago, which is pedigree collapse, or endogamous groups like Acadians or Jewish people, any isolated cultural groups who intermarried for generations.

Triangulation, which clusters are not, further reduces ambiguity because the same segment of DNA is being measured and compared. Ancestry does not offer triangulation, but both FamilyTreeDNA and MyHeritage offer triangulation analysis.

If you're looking to break through a recent brick wall, you may be able to do that with Matches by Cluster. In part, it's a roll of the dice depending on who has tested, the size of the testing pool where your ancestors are from, combined with internal Ancestry algorithms. However, if you descend from ancestors in an under-tested part of the world, you may have fewer or no clusters.

If you're an experienced genealogist trying to break through a distant brick wall, Ancestry's clusters, as they are today, probably aren't going to help you much. If you're hunting for the identity of a 4th great-grandparent, pay close attention to the common ancestors of the people in your closest cluster to that unknown ancestor in your tree. Work on each cluster. If you find a cluster you can't attribute to one of your ancestors, compare the ancestors in the trees of each cluster member, looking for commonality. Ancestors shared between them and not you may point to your brick wall.

Try additional cluster resources at other sites. Each vendor has a different pool of people who have tested. Other vendors and third-party tools provide cluster resources and various types of automated tree-building.

MyHeritage – licenses Genetic Affairs for custom clustering and is optimized for endogamy
GEDmatch – licenses Genetic Affairs for custom clustering and includes unique features
Genetic Affairs – Provides advanced tools and clusters results from FamilyTreeDNA and 23andMe. Also has an AutoTree and AutoKinship feature that compares trees of your matches.
DNAGedcom – Clusters results from various vendors

Sources

Ancestry.com support - - article "Science of Matches by Cluster," by Ross E Curtis, Principal Scientist at AncestryDNA®
https://support.ancestry.com/s/article/Science-of-Matches-by-Cluster?language=en_US

Ancestry.com blog post - - post by Angie Bush, Senior Research Manager, AncestryProGenealogists®
https://www.ancestry.com/c/ancestry-blog/dna/dna-matches-by-cluster

DNAeXplained blog post
https://dna-explained.com/2025/07/10/how-to-use-ancestrys-new-match-clusters-and-what-they-mean/
including comments by Wallace Fullerton and Gary Bookhammer