

Sainaltanna / Features

Our feature articles present a report on how the Y-DNA results were processed, and an account of the recent National Clan Gathering in Nenagh.



Y-DNA PROJECT

Since this is the first report, we will keep the explanation as simple as possible to help everyone get up to speed. We cannot entirely avoid the jargon, so we shall explain it as we go along. Basically, we can split it up into five easy steps.

Step one: Calculate the genetic distance using the test results

Each test result is a string of 25 numbers. Comparing my numbers with your numbers, we might see a few differences here and there. Adding up all the differences between my results and your results gives us a value for the **genetic distance** between us. For example, let's assume that all but two of our numbers matched perfectly; my first number was 12 but yours was 13 (a difference of 1), and my second number was 25 but yours was 24 (a difference of 1 again). The genetic distance between us would be 2.

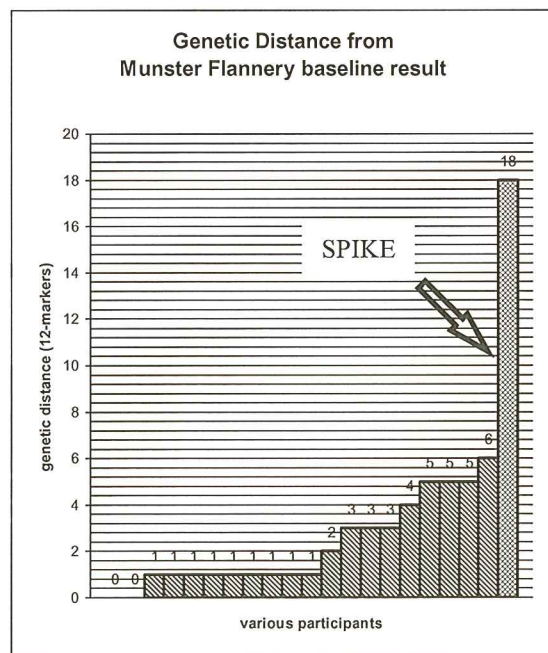
Step two: Graph the genetic distances to identify groups and wildcards

If you like numbers, you can arrange all of the genetic distances in a table and the unusual results might jump out at you. If you want to make the comparisons even easier, you can just graph all the genetic distances and the unusual results show up as spikes on the graph. The easiest results to spot are these "wildcard" results. The experts reckon that, on average, 1 in 20

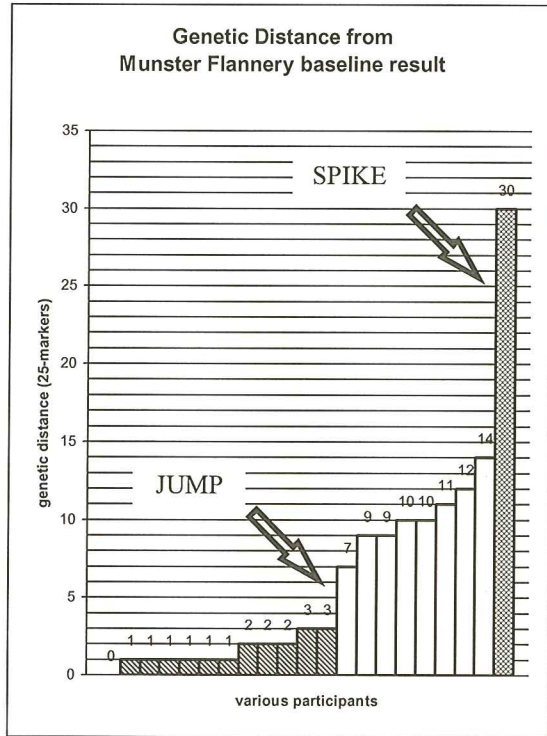
results will be very different from the others. The second easiest results to spot are the perfect matches, because that means that the genetic distance = 0 and they just sit on the bottom of the graph. Obviously, all perfect matches belong to the same group.

The tricky bit is deciding what to do with the slightly different results. How close a match is a proper match? Is there a magic number for genetic distance that confirms a relationship? The graph helps...

If we graph the genetic distances calculated using just the first 12 numbers of the test, and arrange them in ascending order of genetic distance, the flat-bottomed perfect matches appear on the left, then we get a gradual sloping line as the results gently increase, and then finally we get spiked on the wildcard results on the right. The flat-bottom and spikes are easy to group, but the sloping line is too gradual for us to pick a value where the group stops. Here is a typical example:



However, if we plot the same graph using genetic distances calculated using all 25 numbers of the test, then the extra markers greatly accentuate the results. The steeper sloping line suddenly jumps and indicates a quantum leap that defines the end of the group. Here is the revised graph:



So we can split up the results into genetically related groups using the jump in the graph.

Step three: Calculate the baseline value for each group

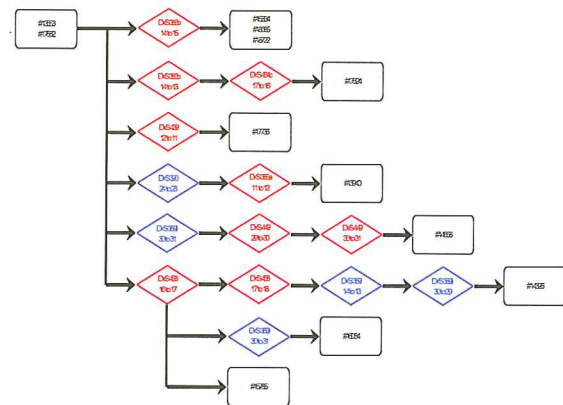
Once we have grouped all the results, we can then have a look at each group in turn and see if they are all the same or slightly different. In all cases, most of the results are the same. Where numbers differ, one guy will have one number slightly off, and another guy will have a different number slightly off. Picking out the numbers that are most common amongst all members of the group, it is possible to draw up the ideal result for the group. This is called the **baseline value**. It represents the original Y-DNA profile of the common ancestor. Initially, it is a theoretical value. As more test results roll in, eventually

someone actually comes along and scores the baseline value. That means that their Y-DNA signature has remained unchanged over the generations whilst others in the group have undergone one or more tiny changes over a long period of time.

Step four: Draw a cladogram for each group using the baseline value as the origin

A **cladogram** is a diagram that helps to illustrate how each result is related to the others in the group. It starts with the baseline value, and then sprouts branches (similar to a family tree) as successive results adopt tiny genetic variations.

Here is the cladogram for the Munster Flannerys:



Step five: Where known, highlight documented ancestral locations on the cladogram

Once the cladogram has been drawn, it only remains to highlight ancestral locations wherever they are known. Ultimately, the common origin of the group and the migrations of its members become apparent as more and more results are added.

In the next report, we shall focus in detail on how the Appalachian Flannerys fit into the group of Munster Flannerys.

