DNA Segments

Basic Information

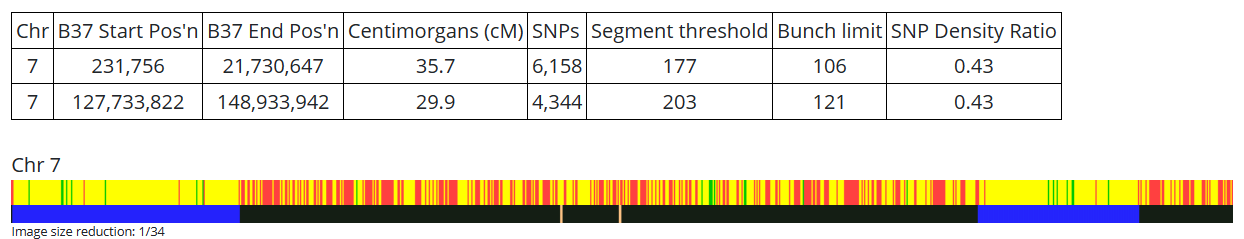
A DNA segment is defined by three data points:

1. Chromosome number.
2. Start position base pair number.
3. End position base pair number.

It is very important to remember we all have two versions of each chromosome, one from our mother and one from our father. Therefore, for any given start and end position on a specific chromosome we will have two different DNA segments.

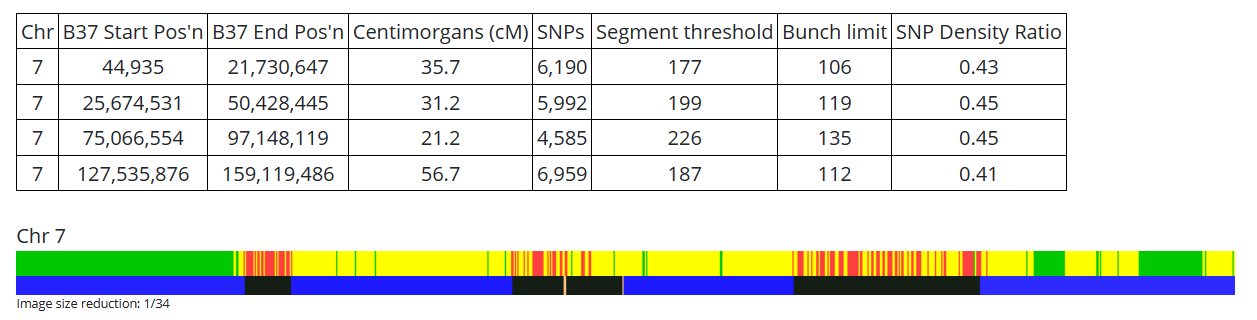
A DNA segment can be measured in different ways:

1. The number of base pairs (bp) or Mega base pairs (Mbp) it contains.
   1. 1 Mbp = 1,000,000 bp.
2. The number of centi Morgans (cMs) it contains. This is the most useful.
3. The number of Single Nucleotide Polymorphisms (SNPs) it contains.



Example of two segments from GEDmatch. Start and end positions are in bp.

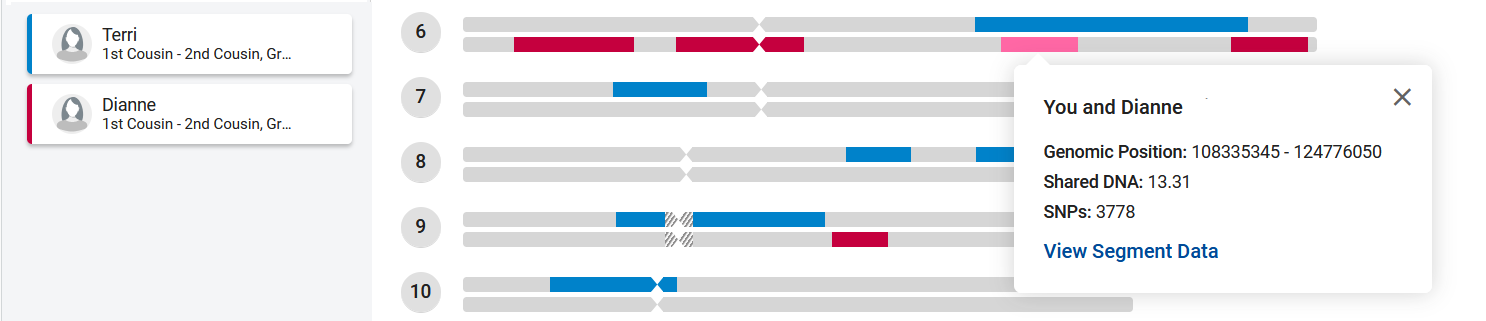
Matches can be half or full. A full match generally means a match on both the maternal and paternal chromosomes and are usually found in sibling matches.



This GEDmatch example shows where my brother and I match on chromosome 7.

The blue areas show common segments, green shows full matches, yellow shows half matches.

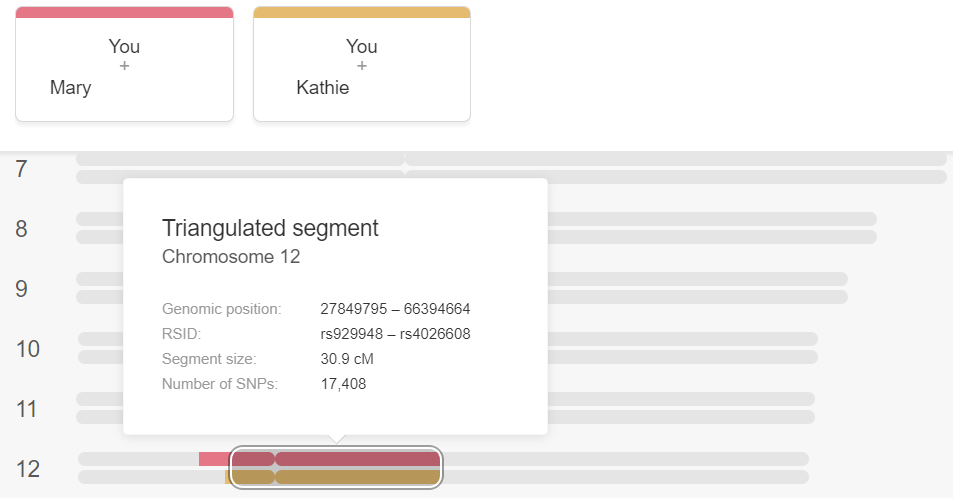
A shared DNA segment includes segment data but also specifies *who* shares the DNA segment. Most companies only show segments shared by 2 people. Triangulated segments shared by 3 are available on My Heritage and GEDmatch.



This Family Tree DNA (FTDNA) example shows matches I have with Terri and Dianne.

FTDNA does not triangulate so I cannot infer Terri matches Dianne on this segment.

In fact, Terri is my maternal cousin and Dianne my paternal cousin, they share no DNA.



In this My Heritage example Mary and Kathie also match each other on this segment.

If they did not match each other then this would not be a triangulated segment.

Mary, Kathie and I all inherited the circled segment from a common ancestor.

If two or more people share the exact same DNA segment of sufficient length, then all of them will have inherited that segment from a common ancestor who would also have had that segment as would each individual in direct line of descent from the common ancestor to each of the people who have that segment. In cases where the common ancestor is not known genealogical research is required!

Sorting Segments

I downloaded all the segment matches of my top 20 matches (excluding my siblings and my niece) from GEDmatch. Here are all the matches on chromosome 16 up to about 6 Mbp (note each line is a segment shared by two people):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **cM** | **Person 1** | **Person 2** |
| 16 | 101,263 | 5,467,795 | 11.7 | \*Dan | Donald |
| 16 | 103,423 | 5,908,224 | 13.6 | Dianne | Donald |
| 16 | 103,423 | 21,243,950 | 45.0 | \*Dan | Dianne |
| 16 | 105,444 | 16,291,983 | 36.4 | Marlyn | Sherri |
| 16 | 115,657 | 5,441,411 | 11.6 | \*AT | \*J. O. |
| 16 | 580,124 | 21,222,941 | 43.1 | \*L. M. | Donald |
|  |  |  |  |  |  |

How many different DNA segments are represented in this table? In order to answer this question, the matches have to be sorted into Triangulation Groups (TGs). It helps to color code the names:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **cM** | **Person 1** | **Person 2** |
| 16 | 101,263 | 5,467,795 | 11.7 | \*Dan | Donald |
| 16 | 103,423 | 5,908,224 | 13.6 | Dianne | Donald |
| 16 | 103,423 | 21,243,950 | 45.0 | \*Dan | Dianne |
| 16 | 105,444 | 16,291,983 | 36.4 | Marlyn | Sherri |
| 16 | 115,657 | 5,441,411 | 11.6 | \*AT | \*J. O. |
| 16 | 580,124 | 21,222,941 | 43.1 | \*L. M. | Donald |
|  |  |  |  |  |  |

The first thing to notice is Marlyn & Sherri do not match anyone else nor do \*AT & \*J. O. so these are two distinct segments. The second thing to note is I match three different people who do not all match each other. \*L. M. doesn’t match either \*Dan or Dianne while these last two do match each other. Therefore, my match with \*L. M. is a different segment than my match with \*Dan and Dianne. Since these two segments overlap or are in the same area of chromosome 16 one must be on my maternal chromosome and the other on my paternal chromosome. We can now label the different segments:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **cM** | **Person 1** | **Person 2** | **Segment** |
| 16 | 101,263 | 5,467,795 | 11.7 | \*Dan | Donald | A |
| 16 | 103,423 | 5,908,224 | 13.6 | Dianne | Donald | A |
| 16 | 103,423 | 21,243,950 | 45.0 | \*Dan | Dianne | A |
| 16 | 105,444 | 16,291,983 | 36.4 | Marlyn | Sherri | B |
| 16 | 115,657 | 5,441,411 | 11.6 | \*AT | \*J. O. | C |
| 16 | 580,124 | 21,222,941 | 43.1 | \*L. M. | Donald | D |

So, there are four different segments represented in the table. What does my genealogy tell me about these four segments? First, \*L. M. is my maternal first cousin while Dianne is my paternal half first cousin and this fits with the conclusion drawn above. Segment A is on my paternal chromosome, D on my maternal. Dianne and I have the same grandmother, Lenora, so she must also have had segment A. \*Dan’s grandmother was Lenora’s sister so she must also have had segment A. Our grandmothers inherited segment A from one of their parents. We don’t have enough information to know which parent it was inherited from.

Segment A is the only TG within the table since it is the only segment shared by at least three people. It is a confirmed TG since all three people match each other on the same segment. It is important to note that this TG is based on three different matching segments that overlap but do not do so perfectly. Most of the start and end positions are different. The triangulated segment does not exactly match any of the individual matched segments, it consists of only the portion of DNA shared by all three. This is found by taking the highest start position and the lowest end position as highlighted below. The size of the TG segment can be estimated using [DNA Painter’s CM Estimator tool](https://dnapainter.com/tools/cme). In this case it is 10.7 cM.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **cM** | **Person 1** | **Person 2** | **Segment** |
| 16 | 101,263 | 5,467,795 | 11.7 | \*Dan | Donald | A |
| 16 | 103,423 | 5,908,224 | 13.6 | Dianne | Donald | A |
| 16 | 103,423 | 21,243,950 | 45.0 | \*Dan | Dianne | A |
|  |  |  |  |  |  |  |

\*AT is my maternal 1C1R, we are related through my maternal grandfather, John and through his mother, Minnie. \*AT and \*J. O. are 1C1R so segment C must have been inherited from one of their Most Recent Common Ancestors (MRCAs), either William Martz or his wife Mary Ellen Mohney (Minnie’s parents). They are also common ancestors for \*L. M. and me but we did not inherit segment D from them. If we did inherit segment D from our grandfather John, it must be from his father rather than Minnie. Segment D could also have come from our grandmother, Creo. There is not enough information to say more about segment D.

Finally, Marlyn and Sherri are close relatives, either siblings or parent and child. Our MRCAs are William Martz’s grandparents, John Shirey and Polly Shaffer and our connection is through Sherri’s mother. Unfortunately, we don’t know enough about segment B to let us draw any further conclusions since we don’t even know at this point if this segment is even on Sherri’s maternal chromosome.

Practice #1

Here are a couple lists of matches I took from GEDmatch. Everyone in these lists has at least one common ancestor. The easiest way to sort these segments is to copy them into a spreadsheet. During the presentation on May 9 I will show the sorted segments and discuss what conclusions can be drawn from the sorting. In the meantime you can try sorting the segments yourself.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **cM** | **Person 1** | **Person 2** |
| 7 | 61,994,170 | 74,489,889 | 8.6 | \*L. M. S. | \*A. M. S. |
| 7 | 61,994,170 | 74,489,889 | 8.6 | \*W. S. | \*LGS |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*BASS | \*NJSD |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*P.B.F. | Sylvia |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*P.B.F. | \*Adarsh |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*SMiller | \*HSmith |
| 7 | 62,035,570 | 74,489,889 | 8.6 | Amber | \*Olivia P |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*J.E. | \*G.E. |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*J.E. | \*4.E. |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*G.E. | \*4.E. |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*CS | \*Adarsh |
| 7 | 62,035,570 | 74,489,889 | 8.6 | \*CAL | Terri |
| 7 | 62,035,570 | 74,494,207 | 8.6 | Margaret | Michael D |
| 7 | 62,035,570 | 74,489,889 | 8.6 | Sylvia | \*Adarsh |
| 7 | 62,047,108 | 74,489,889 | 8.6 | \*CLE | \*J.E. |
| 7 | 62,047,108 | 74,489,889 | 8.6 | \*CLE | \*G.E. |
| 7 | 62,047,108 | 74,489,889 | 8.6 | \*CLE | \*4.E. |
| 7 | 62,047,108 | 74,489,889 | 8.6 | \*HSmith | Kristy S |
| 7 | 62,047,108 | 74,489,889 | 8.6 | George E | \*grc51 |
| 7 | 63,298,188 | 74,489,889 | 8.5 | \*V. L. | \*E.G. |
| 7 | 66,213,108 | 74,489,889 | 7.4 | Janice M | \*TJM |

Practice #2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **cM** | **Person 1** | **Person 2** |
| 5 | 49,470,083 | 68,828,296 | 14.6 | \*ram0610 | Marjorie |
| 5 | 49,470,083 | 68,828,296 | 14.6 | \*W. S. | Donald\_S |
| 5 | 49,565,100 | 67,534,039 | 13.4 | \*E.G. | \*TKW |
| 5 | 49,625,550 | 68,828,296 | 14.6 | E B | Sylvia W |
| 5 | 49,636,753 | 68,828,296 | 14.6 | \*HSmith | \*SMiller |
| 5 | 49,636,753 | 68,828,296 | 14.6 | \*G.E. | \*J.E. |
| 5 | 49,636,753 | 68,828,296 | 14.6 | \*4.E. | \*J.E. |
| 5 | 49,636,753 | 68,828,296 | 14.6 | \*4.E. | \*G.E. |
| 5 | 49,636,753 | 68,828,296 | 14.6 | \*Adarsh | \*CS |
| 5 | 49,636,753 | 65,538,791 | 10.9 | \*Shelly | \*TT's F |
| 5 | 49,636,753 | 68,828,296 | 14.6 | Margaret | Michael |
| 5 | 49,651,872 | 59,995,742 | 8.2 | \*PD | \*T. E. |
| 5 | 49,874,330 | 68,828,296 | 14.6 | \*CLE | \*G.E. |
| 5 | 49,874,330 | 68,828,296 | 14.6 | \*4.E. | \*CLE |
| 5 | 49,874,330 | 68,828,296 | 14.6 | \*HSmith | Kristy S |
| 5 | 49,874,330 | 68,828,296 | 14.6 | \*J. O. | \*T. E. |
| 5 | 49,874,330 | 61,977,629 | 9.2 | \*E. F. R | Katie J |
| 5 | 49,874,330 | 68,828,296 | 14.6 | \*grc51 | George E |
| 5 | 49,874,330 | 68,200,642 | 14.0 | Carol D | Jean L |
| 5 | 50,114,852 | 68,820,348 | 14.5 | \*TJM | Janice M |
| 5 | 50,162,499 | 68,820,348 | 14.5 | \*O. A. | Amber B |
| 5 | 50,205,230 | 65,895,193 | 11.3 | \*J.E. | Richard |
| 5 | 52,069,314 | 65,603,639 | 10.1 | \*Shelly | John R |
| 5 | 52,178,056 | 65,247,893 | 9.7 | \*TT's F | John R |
| 5 | 55,445,129 | 68,427,072 | 10.2 | \*J. O. | Jennifer |
| 5 | 55,601,387 | 68,396,580 | 10.0 | \*T. E. | Jennifer |
| 5 | 58,520,735 | 68,828,296 | 7.2 | \*Sandy | Katie J |