AutoSegment tutorial

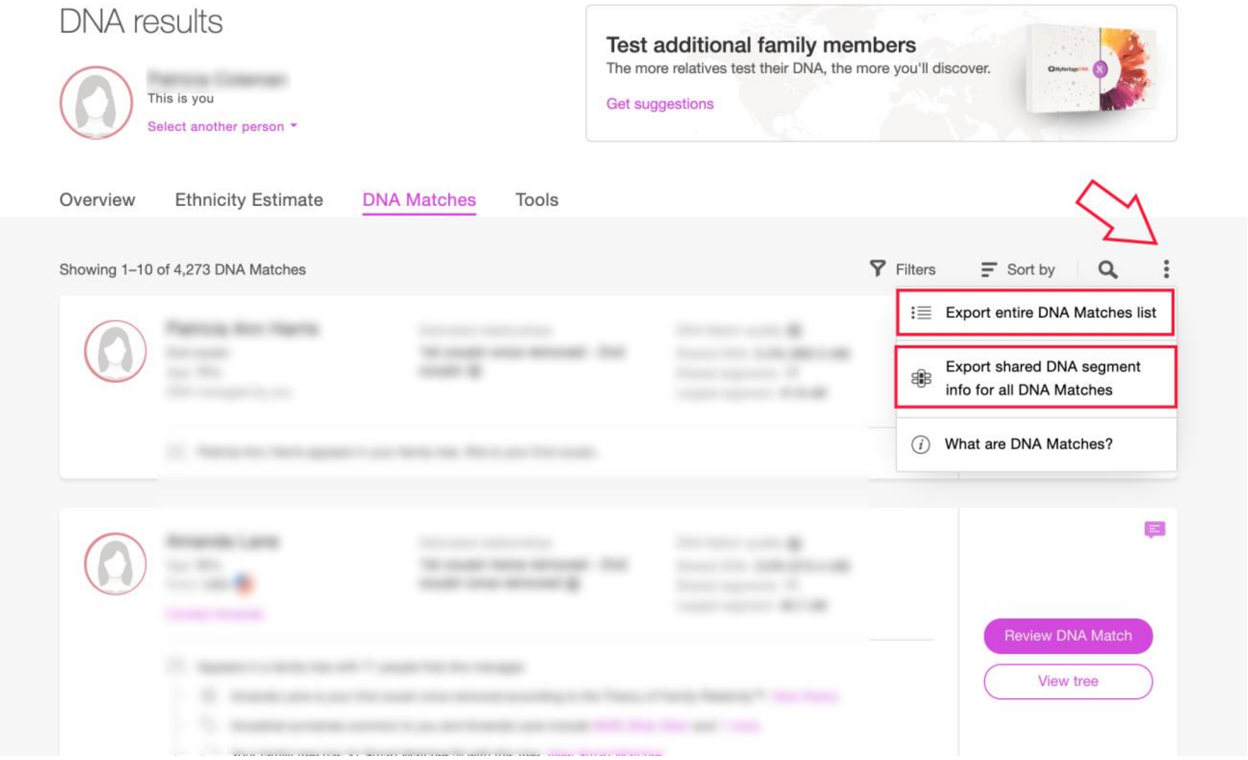
Retrieve offline data for MyHeritage/23andme/FamilyTreeDNA and GEDmatch

AutoSegment is a DNA segment-based clustering method that is available for segment data from MyHeritage, 23andme, FamilyTreeDNA and GEDmatch. Requires segment data downloaded from testing companies and it does not require credential information. An AutoSegment analysis can be started using this link: https://members.geneticaffairs.com/autosegment

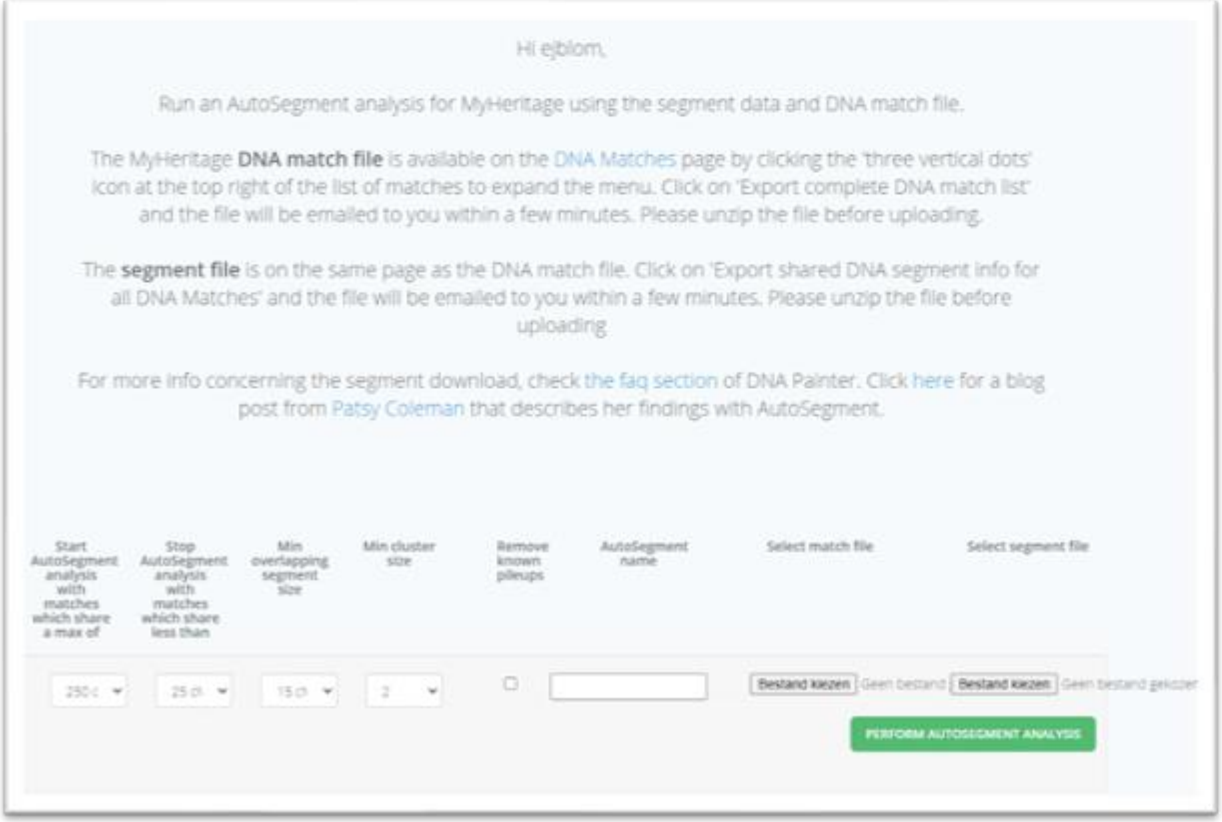
Retrieving segment data for MyHeritage

Login to MyHeritage and visit the DNA matches page: https://www.myheritage.com/dna/matches

Click on the three vertical dots to expand the menu. Select the first option to retrieve the DNA match list, the second option will allow the retrieval of the segment data. Both files will be mailed.



Save the attachments from the e-mails and go to the Genetic Affairs AutoSegment page for MyHeritage: https://members.geneticaffairs.com/autosegment/addWebsite/MyHeritage



Hi ejblom,

Run an AutoSegment analysis for MyHeritage using the segment data and DNA match file.

The MyHeritage **DNA match file** is available on the DNA Matches page by clicking the 'three vertical dots' icon at the top right of the list of matches to expand the menu. Click on 'Export complete DNA match list' and the file will be emailed to you within a few minutes. Please unzip the file before uploading.

The **segment file** is on the same page as the DNA match file. Click on 'Export shared DNA segment info for all DNA Matches' and the file will be emailed to you within a few minutes. Please unzip the file before uploading.

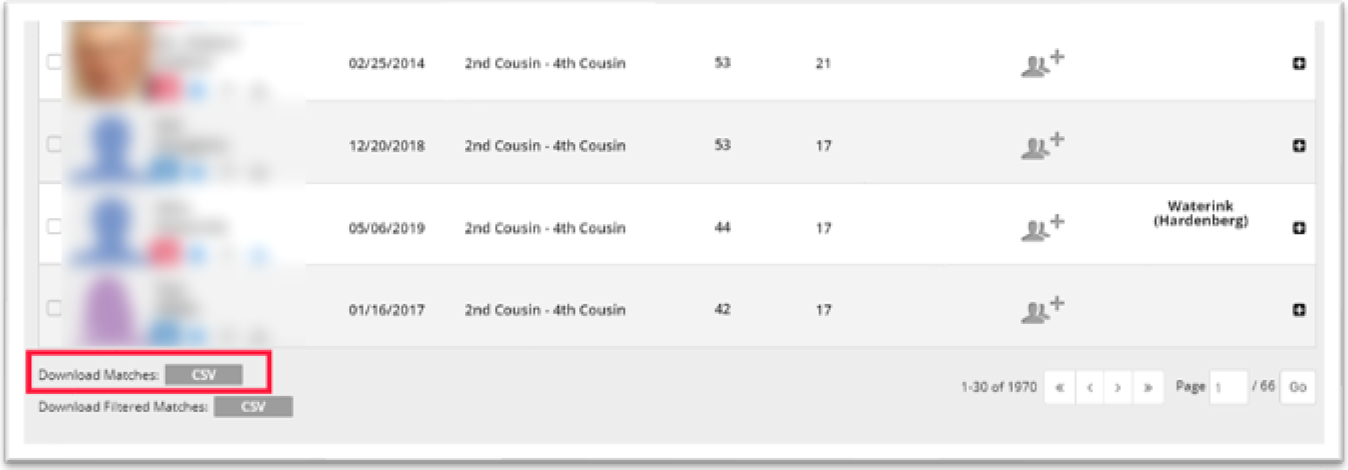
For more info concerning the segment download, check the faq section of DNA Painter. Click here for a blog post from Patsy Coleman that describes her findings with AutoSegment.

Start AutoSegment analysis with matches which share a max of	Stop AutoSegment analysis with matches which share less than	Min overlapping segment size	Min cluster size	Remove known pileups	AutoSegment name	Select match file	Select segment file
250 c	25 c	15 c	2	☐		Bestand kiezen Geen bestand	Bestand kiezen Geen bestand gekozen

Adjust the search parameters, fill in the name field and select the segment file and match file. After you start the analysis, results will appear within 15 min in your mailbox. If no results appear, it is possible that too low cM values were employed. Try raising the minimum cM settings and retry the analysis. If it still doesn't provide results, contact us via info@geneticaffairs.com

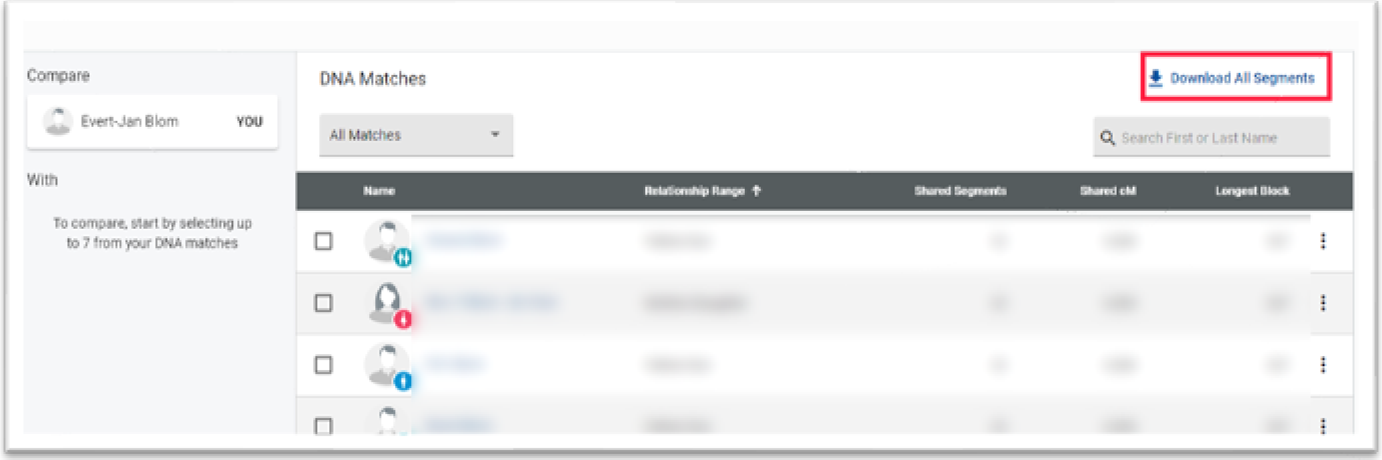
Retrieving segment data for FamilyTreeDNA

Login into FamilyTreeDNA and visit your Family Finder page: https://www.familytreedna.com/my/familyfinder. Go to the bottom of the page and click on "Download Matches: CSV"



02/25/2014	2nd Cousin - 4th Cousin	53	21	
12/20/2018	2nd Cousin - 4th Cousin	53	17	
05/06/2019	2nd Cousin - 4th Cousin	44	17	Waterink (Hardenberg)
01/16/2017	2nd Cousin - 4th Cousin	42	17	

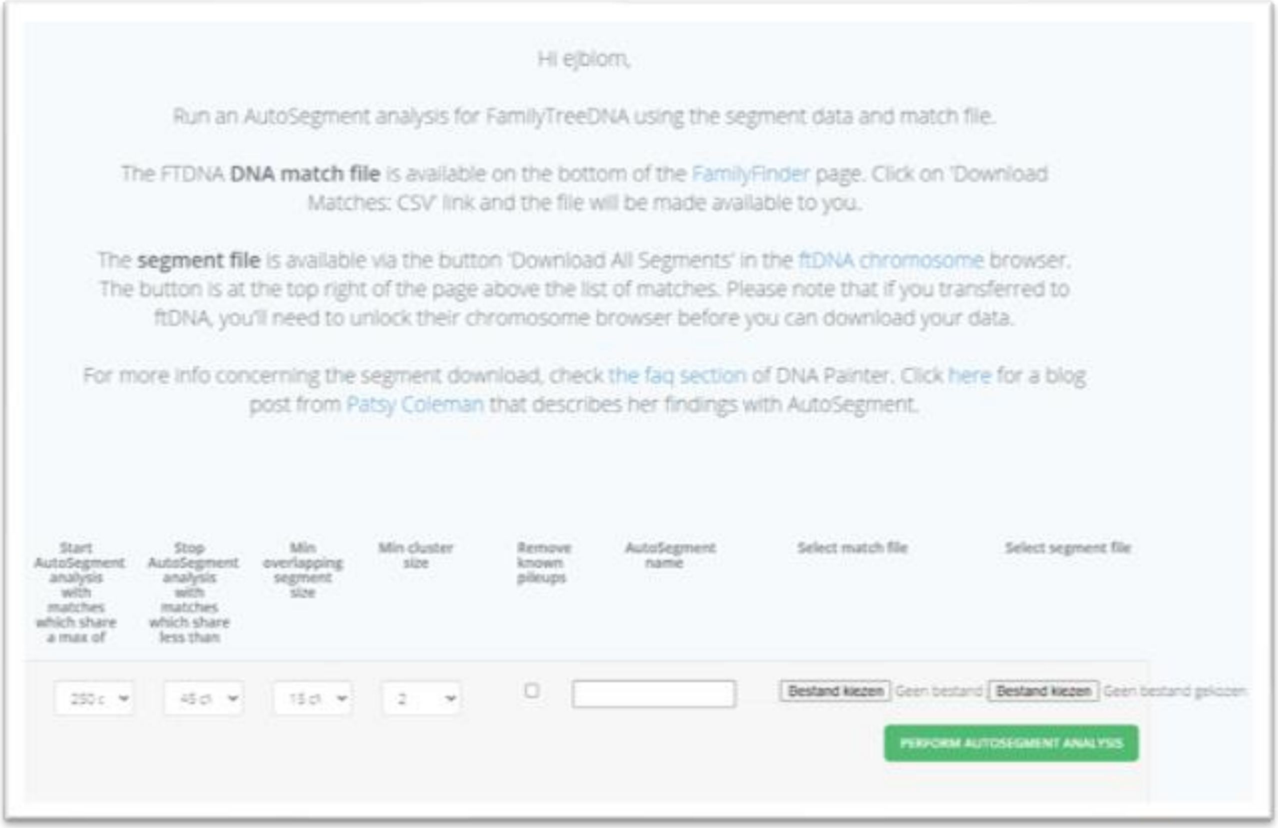
Save the matches file to your local drive. Next, visit the chromosome browser page for the segment data: https://www.familytreedna.com/my/family-finder/chromosome-browser



To compare, start by selecting up to 7 from your DNA matches

Click on "Download all segments" to download the segment data.

Next, go to the Genetic Affairs AutoSegment page for FamilyTreeDNA:
https://members.geneticaffairs.com/autosegment/addWebsite/FamilyTreeDNA



Hi ejblom,

Run an AutoSegment analysis for FamilyTreeDNA using the segment data and match file.

The FTDNA **DNA match file** is available on the bottom of the FamilyFinder page. Click on 'Download Matches: CSV' link and the file will be made available to you.

The **segment file** is available via the button 'Download All Segments' in the ftDNA chromosome browser. The button is at the top right of the page above the list of matches. Please note that if you transferred to ftDNA, you'll need to unlock their chromosome browser before you can download your data.

For more info concerning the segment download, check the faq section of DNA Painter. Click here for a blog post from Patsy Coleman that describes her findings with AutoSegment.

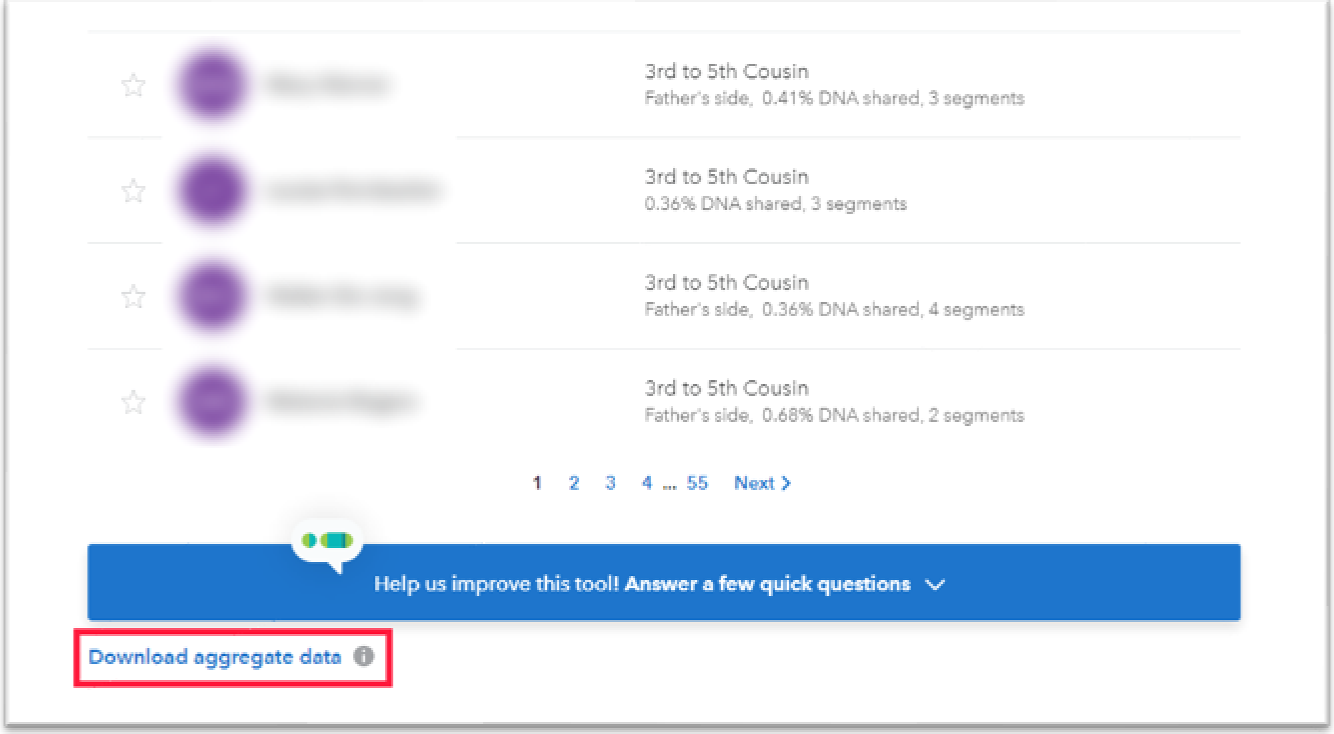
Start AutoSegment analysis with matches which share a max of	Stop AutoSegment analysis with matches which share less than	Min overlapping segment size	Min cluster size	Remove known pileups	AutoSegment name	Select match file	Select segment file
250 c	45 c	15 c	2	☐		Bestand kiezen Geen bestand	Bestand kiezen Geen bestand gekozen

Adjust the search parameters, fill in the name field and select the segment file and match file. After you start the analysis, results will appear within 15 min in your mailbox. If no results appear, it is possible that too low cM values were employed. Try raising the minimum cM settings and retry the analysis. If it still doesn't provide results, contact us via info@geneticaffairs.com

Retrieve data for 23andme

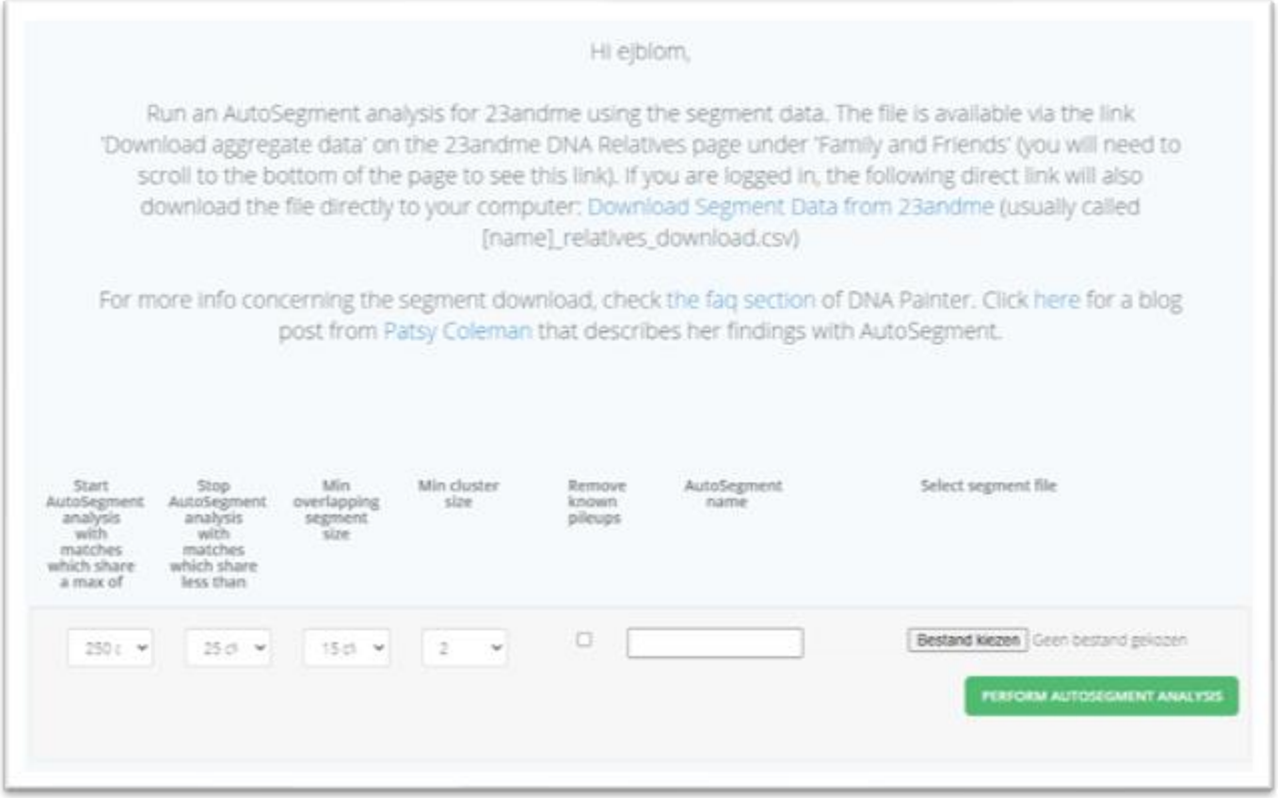
The AutoSegment analysis for 23andme is the only analysis that only requires one file. This is because 23andme has merged the DNA match data into the segment file. To obtain this file, first go to https://www.23andme.com and login. Next, go to the DNA relatives page (https://you.23andme.com/tools/relatives/):

On the DNA relatives page, scroll to the bottom and click on "Download aggregate data"



This will allow you to download all segment/match data to your local drive. You can also obtain this file by using to this direct link: https://you.23andme.com/tools/relatives/download/

Next, go to the Genetic Affairs AutoSegment page for 23andme:
https://members.geneticaffairs.com/autosegment/addWebsite/23andme



Hi ejblom,

Run an AutoSegment analysis for 23andme using the segment data. The file is available via the link 'Download aggregate data' on the 23andme DNA Relatives page under 'Family and Friends' (you will need to scroll to the bottom of the page to see this link). If you are logged in, the following direct link will also download the file directly to your computer: Download Segment Data from 23andme (usually called [name]_relatives_download.csv)

For more info concerning the segment download, check the faq section of DNA Painter. Click here for a blog post from Patsy Coleman that describes her findings with AutoSegment.

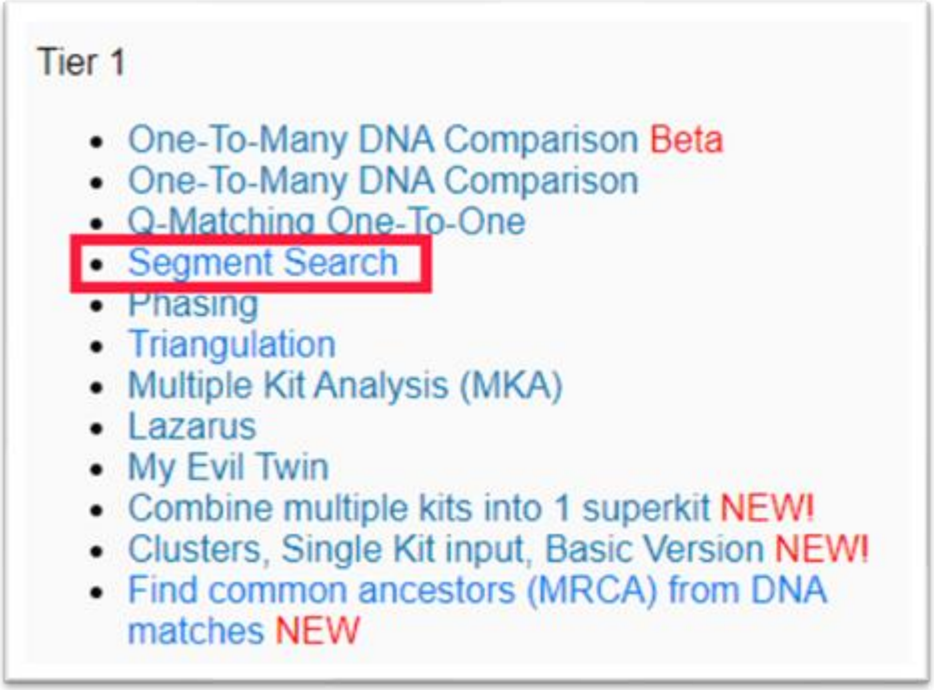
Start AutoSegment analysis with matches which share a max of	Stop AutoSegment analysis with matches which share less than	Min overlapping segment size	Min cluster size	Remove known pileups	AutoSegment name	Select segment file
250 c	25 cl	15 cl	2	☐		Bestand kiezen Geen bestand gekozen

Adjust the search parameters, fill in the name field and select the segment/match file. After you start the analysis, results will appear within 15 min in your mailbox. If no results appear, it is possible that too low cM values were employed. Try raising the minimum cM settings and retry the analysis. If it still doesn't provide results, contact us via info@geneticaffairs.com

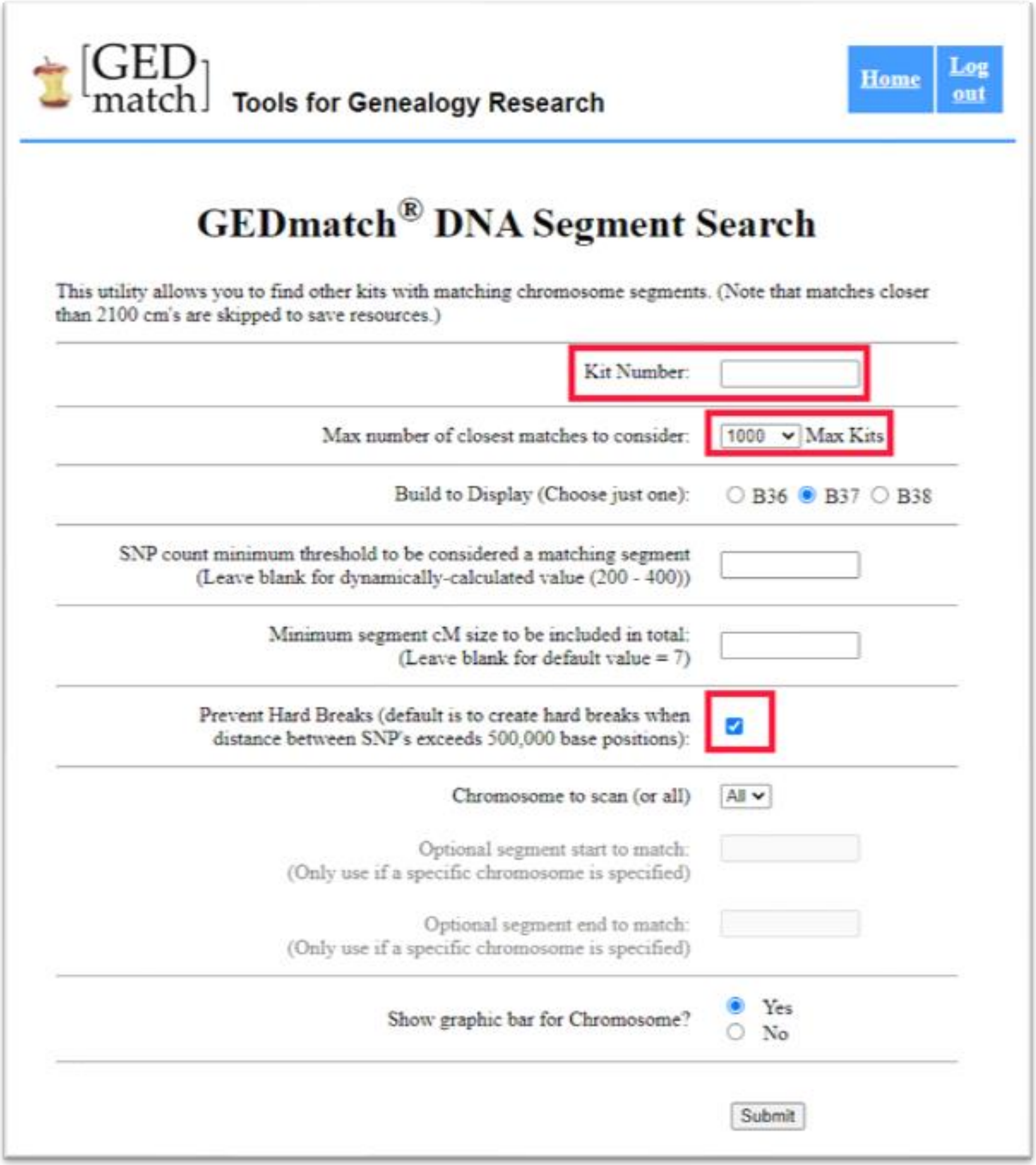
Retrieve data for GEDmatch

GEDmatch is the only company that provides triangulated segments in addition to regular DNA segment data. AutoSegment employs both files since the regular DNA segment file contains information concerning the DNA matches that is not available in the triangulated segments file.

Log into GEDmatch and select the Tier 1 – Segment Search (https://www.gedmatch.com/segment_search.php):



Fill in the concerned Kit Number, select the max number of closest matches to consider (for instance 5000) and enable the “Prevent Hard Breaks” option:



GEDmatch® DNA Segment Search

This utility allows you to find other kits with matching chromosome segments. (Note that matches closer than 2100 cm's are skipped to save resources.)

Kit Number:

Max number of closest matches to consider: 1000 Max Kits

Build to Display (Choose just one): ○ B36 ● B37 ○ B38

SNP count minimum threshold to be considered a matching segment (Leave blank for dynamically-calculated value (200 - 400)):

Minimum segment cM size to be included in total: (Leave blank for default value = 7)

Prevent Hard Breaks (default is to create hard breaks when distance between SNP's exceeds 500,000 base positions): ☑

Chromosome to scan (or all) All

Optional segment start to match: (Only use if a specific chromosome is specified)

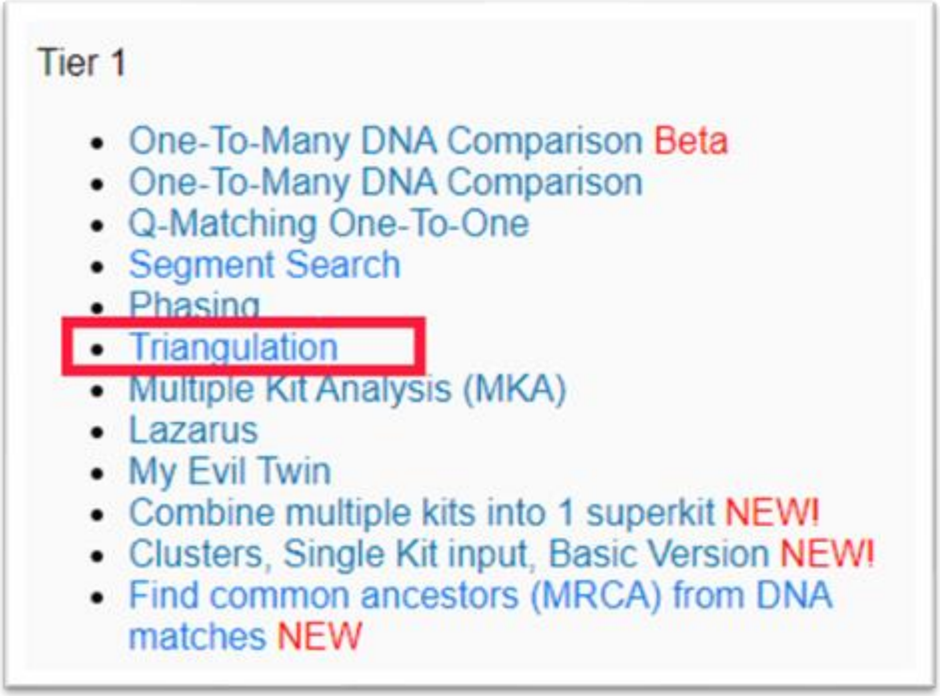
Optional segment end to match: (Only use if a specific chromosome is specified)

Show graphic bar for Chromosome? ● Yes ○ No

Submit

Click on submit and let the tool analyze the data. After it is finished, click on the “here” button in the top of the screen and save the csv file to your local drive.

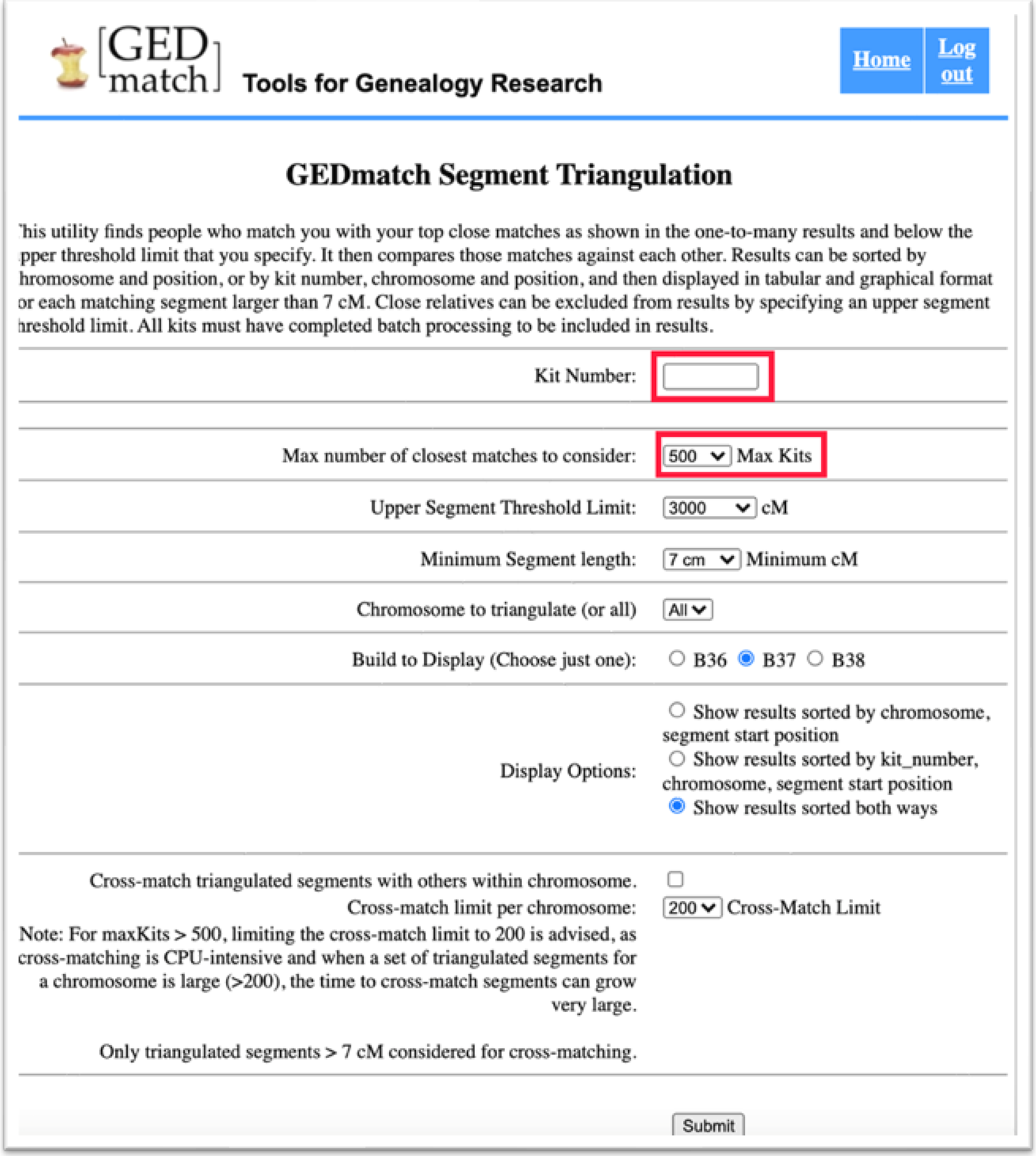
Next, we will download the triangulated data from GEDmatch. From the main page, select the triangulation option:



Tier 1

- One-To-Many DNA Comparison Beta
- One-To-Many DNA Comparison
- Q-Matching One-To-One
- Segment Search
- Phasing
- Triangulation
- Multiple Kit Analysis (MKA)
- Lazarus
- My Evil Twin
- Combine multiple kits into 1 superkit NEW!
- Clusters, Single Kit input, Basic Version NEW!
- Find common ancestors (MRCA) from DNA matches NEW

Fill in the concerned Kit Number, select the max number of closest matches to consider (for instance 5000) and start the analysis by clicking on the submit button:



GEDmatch Segment Triangulation

his utility finds people who match you with your top close matches as shown in the one-to-many results and below the pper threshold limit that you specify. It then compares those matches against each other. Results can be sorted by hromosome and position, or by kit number, chromosome and position, and then displayed in tabular and graphical format or each matching segment larger than 7 cM. Close relatives can be excluded from results by specifying an upper segment hreshold limit. All kits must have completed batch processing to be included in results.

Kit Number:

Max number of closest matches to consider: 500 Max Kits

Upper Segment Threshold Limit: 3000 cM

Minimum Segment length: 7 cm Minimum cM

Chromosome to triangulate (or all) All

Build to Display (Choose just one): ○ B36 ● B37 ○ B38

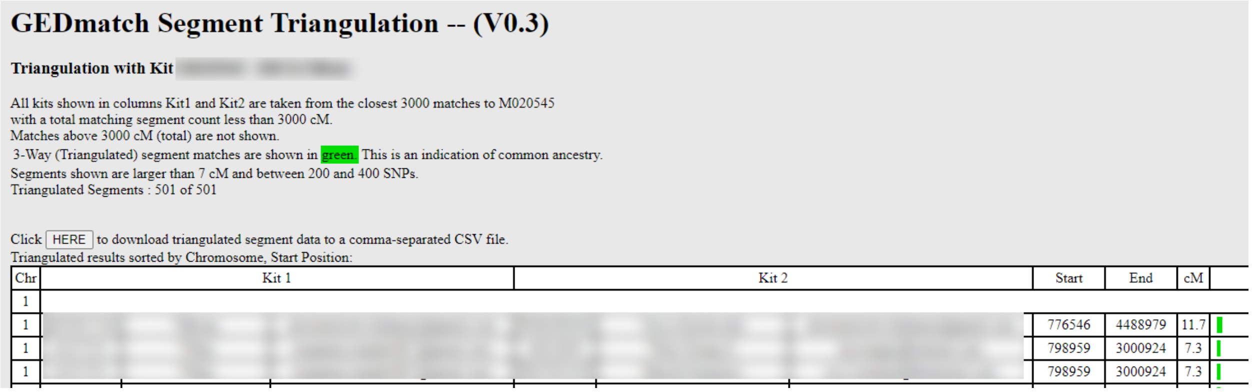
Display Options:
○ Show results sorted by chromosome, segment start position
○ Show results sorted by kit_number, chromosome, segment start position
● Show results sorted both ways

Cross-match triangulated segments with others within chromosome. ☐
Cross-match limit per chromosome: 200 Cross-Match Limit

Note: For maxKits > 500, limiting the cross-match limit to 200 is advised, as cross-matching is CPU-intensive and when a set of triangulated segments for a chromosome is large (>200), the time to cross-match segments can grow very large.

Only triangulated segments > 7 cM considered for cross-matching.

Browse to the end of the page and locate the download link:



GEDmatch Segment Triangulation -- (V0.3)

Triangulation with Kit [redacted]

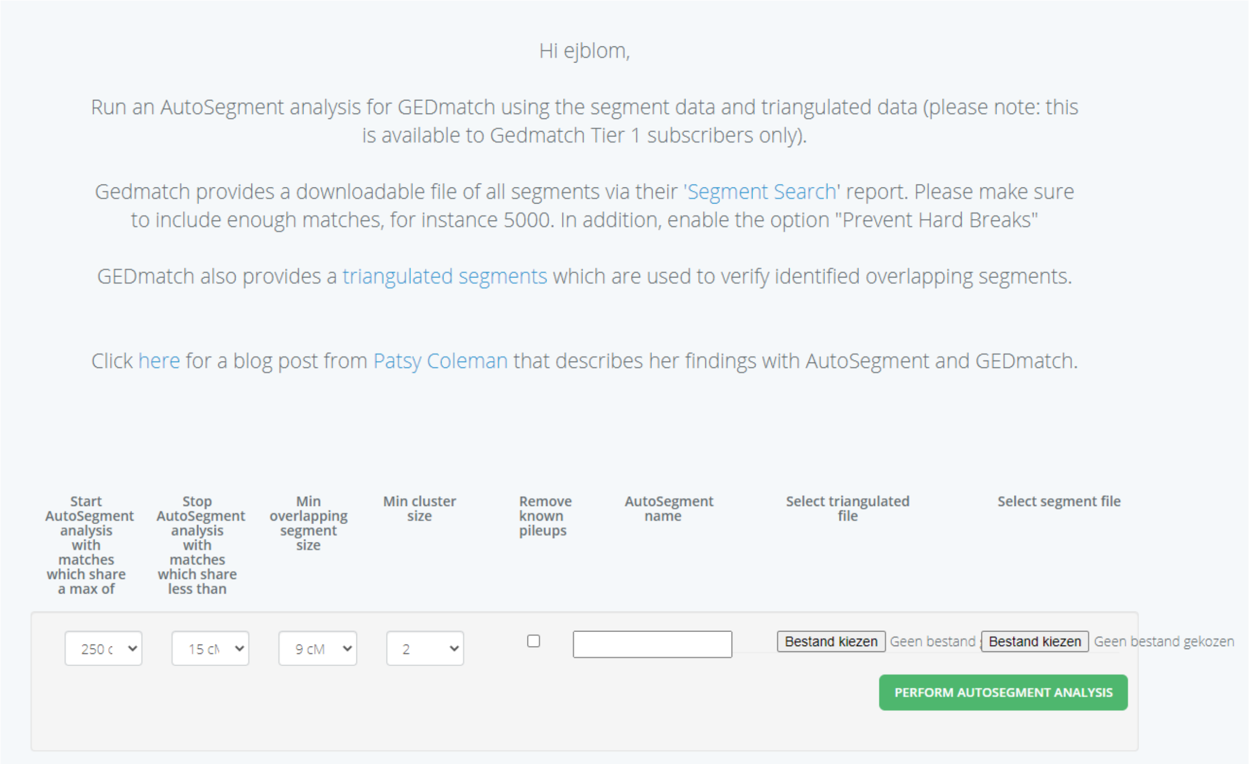
All kits shown in columns Kit1 and Kit2 are taken from the closest 3000 matches to M020545
with a total matching segment count less than 3000 cM.
Matches above 3000 cM (total) are not shown.
3-Way (Triangulated) segment matches are shown in green. This is an indication of common ancestry.
Segments shown are larger than 7 cM and between 200 and 400 SNPs.
Triangulated Segments : 501 of 501

Click HERE to download triangulated segment data to a comma-separated CSV file.
Triangulated results sorted by Chromosome, Start Position:

Chr	Kit 1	Kit 2	Start	End	cM	
1						
1	[redacted]	[redacted]	776546	4488979	11.7	🟩
1	[redacted]	[redacted]	798959	3000924	7.3	🟩
1	[redacted]	[redacted]	798959	3000924	7.3	🟩

Download the triangulated data by clicking on the 'here' link in the bottom of the screen. This will allow you to save a csv file to your local drive.

Next, go to the Genetic Affairs AutoSegment page for GEDmatch: https://members.geneticaffairs.com/autosegment/addWebsite/GEDmatch



Hi ejblom,

Run an AutoSegment analysis for GEDmatch using the segment data and triangulated data (please note: this is available to Gedmatch Tier 1 subscribers only).

Gedmatch provides a downloadable file of all segments via their 'Segment Search' report. Please make sure to include enough matches, for instance 5000. In addition, enable the option "Prevent Hard Breaks"

GEDmatch also provides a triangulated segments which are used to verify identified overlapping segments.

Click here for a blog post from Patsy Coleman that describes her findings with AutoSegment and GEDmatch.

Start AutoSegment analysis with matches which share a max of	Stop AutoSegment analysis with matches which share less than	Min overlapping segment size	Min cluster size	Remove known pileups	AutoSegment name	Select triangulated file	Select segment file
250 c	15 cM	9 cM	2	☐		Bestand kiezen Geen bestand	Bestand kiezen Geen bestand gekozen

Adjust the search parameters, fill in the name field and select the triangulated and segment file. After you start the analysis, results will appear within 15 min in your mailbox. If no results appear, it is possible that too low cM values were employed. Try raising the minimum cM settings and retry the analysis. If it still doesn't provide results, contact us via info@geneticaffairs.com