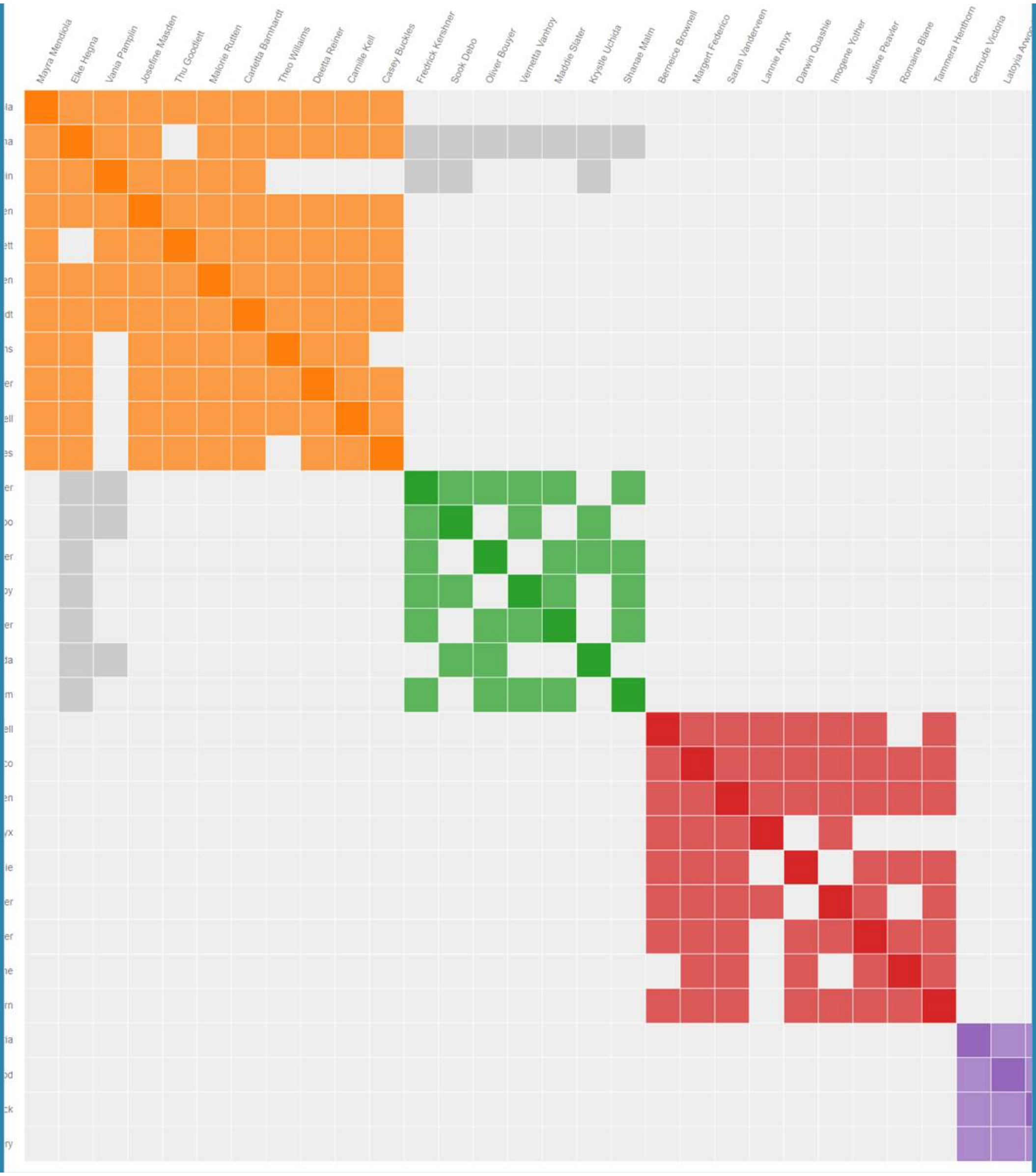


Unraveling your genealogy with

Genetic Affairs

Evert-Jan Blom PhD

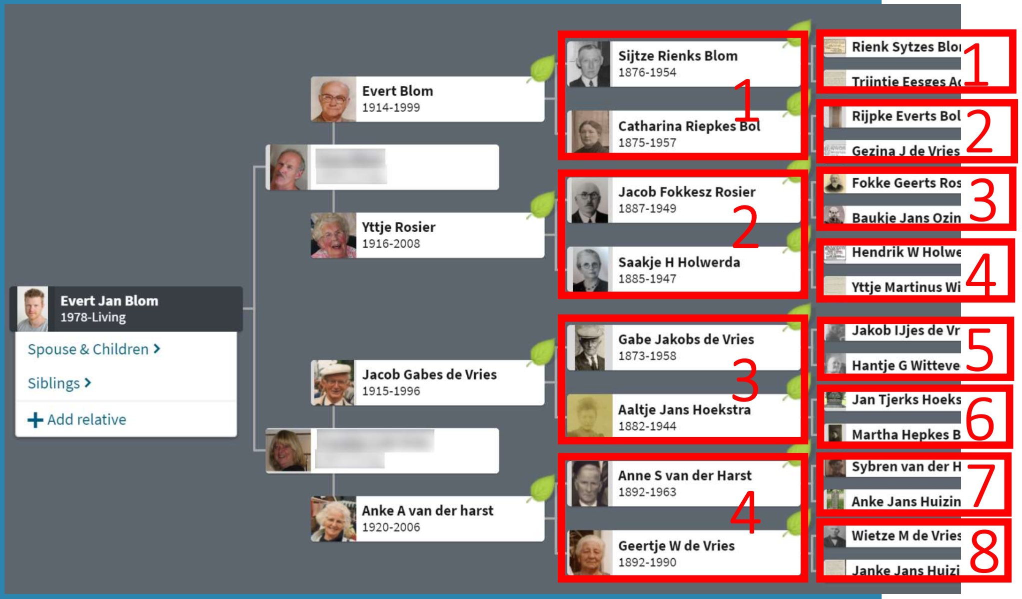
Outline of presentation

- AutoCluster – grouping your matches
- AutoSegment – finding common DNA segments
- AutoTree – mining trees for common ancestors
- AutoKinship – create trees based on shared DNA
- AutoKinship on GEDmatch

Outline AutoCluster

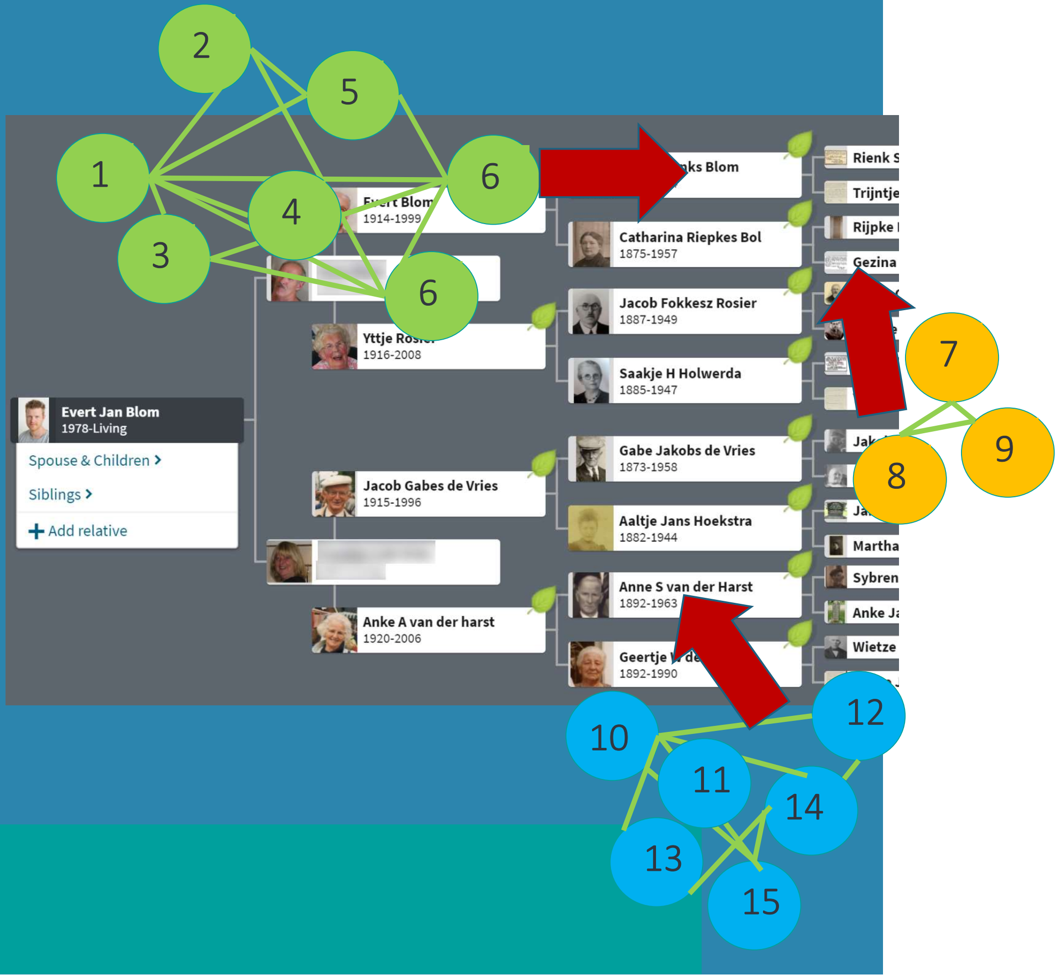
- Why clustering?
- How does AutoCluster work
- AutoCluster visualization
- AutoCluster concepts
- Ordering of clusters
- Start a clustering
  - FTDNA, 23andme & Ancestry
  - MyHeritage & GEDmatch
- Conclusion

Why clustering?



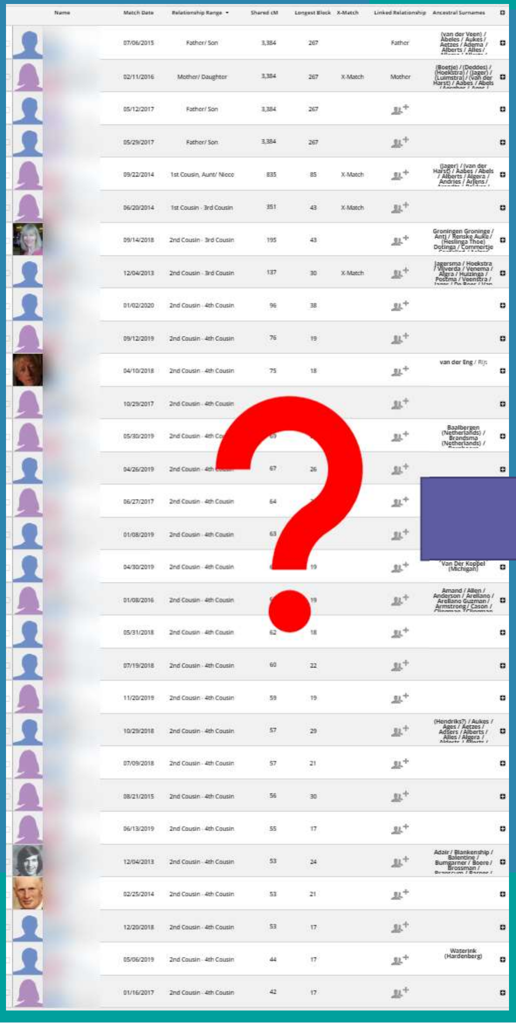
- Break up your DNA match list into groups
- 4 groups that represent your great-gp
- 8 groups that represent your 2g-gp

Why clustering?

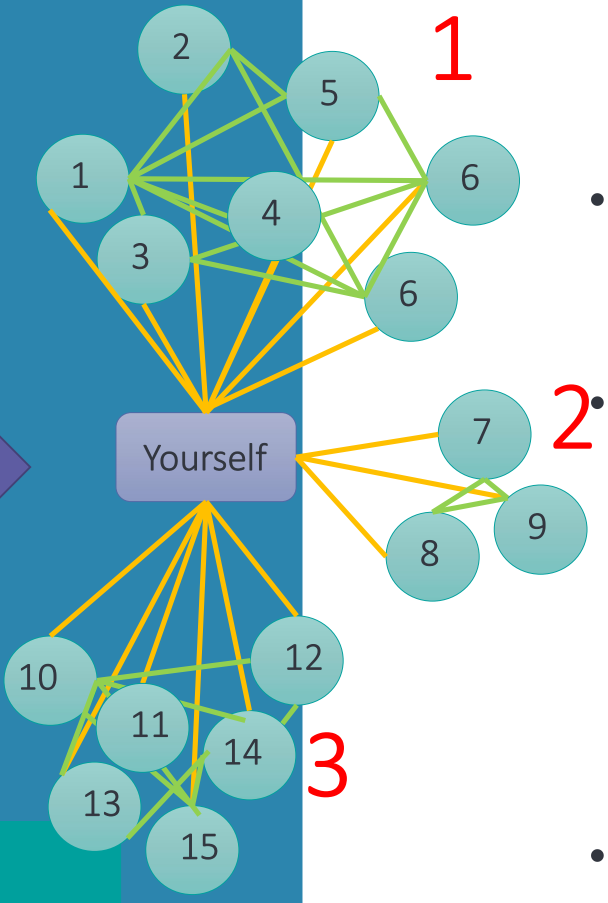


- Provides hypotheses:
- Members of clusters are likely to share the same common ancestor and represent a branch of your family.
- Analyzing clusters:
  - Interpret trees of matches from a cluster
    - Identify one or more ancestors in common
  - Common surnames/locations
  - Contact matches, build Q&D trees
  - Find shared segments

How does it work?



Name	Match Date	Relationship Range	Shared cM	Longest Block	X-Match	Linked Relationship	Ancestral Surnames
[redacted]	07/06/2015	Father/ Son	3,384	267		Father	(van der Veen) / Abeles / Aukes / Aetzes / Adema / Alberts / Alles /
[redacted]	02/11/2016	Mother/ Daughter	3,384	267	X-Match	Mother	(Boetje) / (Deddes) / (Hoekstra) / (Jager) / (Luimstra) / (van der Harst) / Aabes / Abels
[redacted]	05/12/2017	Father/ Son	3,384	267			
[redacted]	05/29/2017	Father/ Son	3,384	267			
[redacted]	09/22/2014	1st Cousin, Aunt/ Niece	835	85	X-Match		(Jager) / (van der Harst) / Aabes / Abels / Alberts / Algera / Andries / Arjens /
[redacted]	06/20/2014	1st Cousin - 3rd Cousin	351	43	X-Match		
[redacted]	09/14/2018	2nd Cousin - 3rd Cousin	195	43			Groningen Groninge / Antj / Renske Auke / (Heslinga Thoe) Dotinga / Commertje
[redacted]	12/04/2013	2nd Cousin - 3rd Cousin	137	30	X-Match		Jagersma / Hoekstra / Viverda / Venema / Algra / Huizinga / Postma / Veenstra /
[redacted]	01/02/2020	2nd Cousin - 4th Cousin	96	38			
[redacted]	09/12/2019	2nd Cousin - 4th Cousin	76	19			
[redacted]	04/10/2018	2nd Cousin - 4th Cousin	75	18			van der Eng / Rijs
[redacted]	10/29/2017	2nd Cousin - 4th Cousin	[illegible]	[illegible]			
[redacted]	05/30/2019	2nd Cousin - 4th Cousin	69	[illegible]			Baalbergen (Netherlands) / Brandsma (Netherlands) /
[redacted]	04/26/2019	2nd Cousin - 4th Cousin	67	26			
[redacted]	06/27/2017	2nd Cousin - 4th Cousin	64	[illegible]			[illegible]
[redacted]	01/08/2019	2nd Cousin - 4th Cousin	63	[illegible]			[illegible]
[redacted]	04/30/2019	2nd Cousin - 4th Cousin	[illegible]	19			Van Der Koppel (Michigan)
[redacted]	01/08/2016	2nd Cousin - 4th Cousin	[illegible]	19			Amand / Allen / Anderson / Arellano / Arellano Guzman / Armstrong / Cason /
[redacted]	05/31/2018	2nd Cousin - 4th Cousin	62	18			
[redacted]	07/19/2018	2nd Cousin - 4th Cousin	60	22			
[redacted]	11/20/2019	2nd Cousin - 4th Cousin	59	19			
[redacted]	10/29/2018	2nd Cousin - 4th Cousin	57	29			(Hendriks?) / Aukes / Ages / Aetzes / Adsers / Alberts / Alles / Algera /
[redacted]	07/09/2018	2nd Cousin - 4th Cousin	57	21			
[redacted]	08/21/2015	2nd Cousin - 4th Cousin	56	30			
[redacted]	06/13/2019	2nd Cousin - 4th Cousin	55	17			
[redacted]	12/04/2013	2nd Cousin - 4th Cousin	53	24			Adair / Blankenship / Balentine / Bumgarner / Boere / Brossman /
[redacted]	02/25/2014	2nd Cousin - 4th Cousin	53	21			
[redacted]	12/20/2018	2nd Cousin - 4th Cousin	53	17			
[redacted]	05/06/2019	2nd Cousin - 4th Cousin	44	17			Waterink (Hardenberg)
[redacted]	01/16/2017	2nd Cousin - 4th Cousin	42	17			



- Organize unorganized lists to groups of related matches

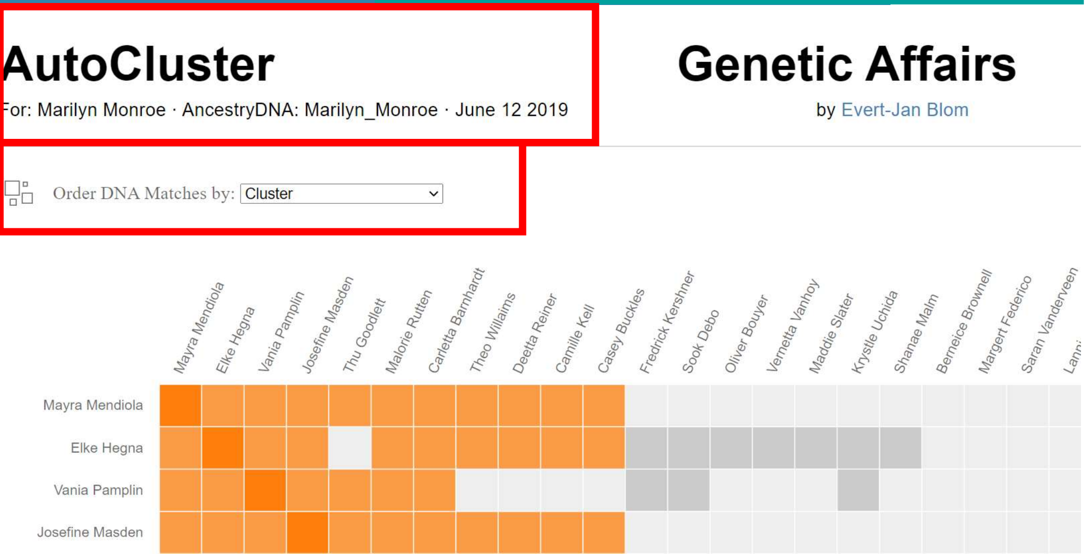
- Take your match list
  - For each match:
  - Take the shared (or “in common with”) matches and connect linked matches
  - Create a (genetic) network

- Identify groups (clusters) of linked matches within the network

AutoCluster visualization

- AutoCluster automatically organizes your matches into shared match clusters .
- Provides a visual chart of your clusters.
- The chart starts out unsorted and organizes itself into the colored clusters

AutoCluster concepts



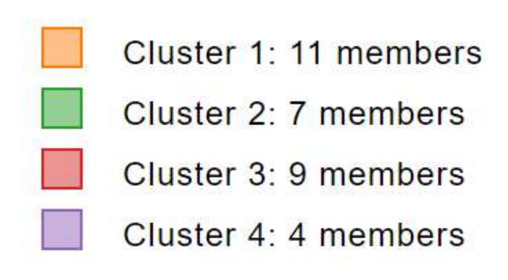
AutoCluster

For: Marilyn Monroe · AncestryDNA: Marilyn_Monroe · June 12 2019

Genetic Affairs

by Evert-Jan Blom

Order DNA Matches by: Cluster



🟧 Cluster 1: 11 members

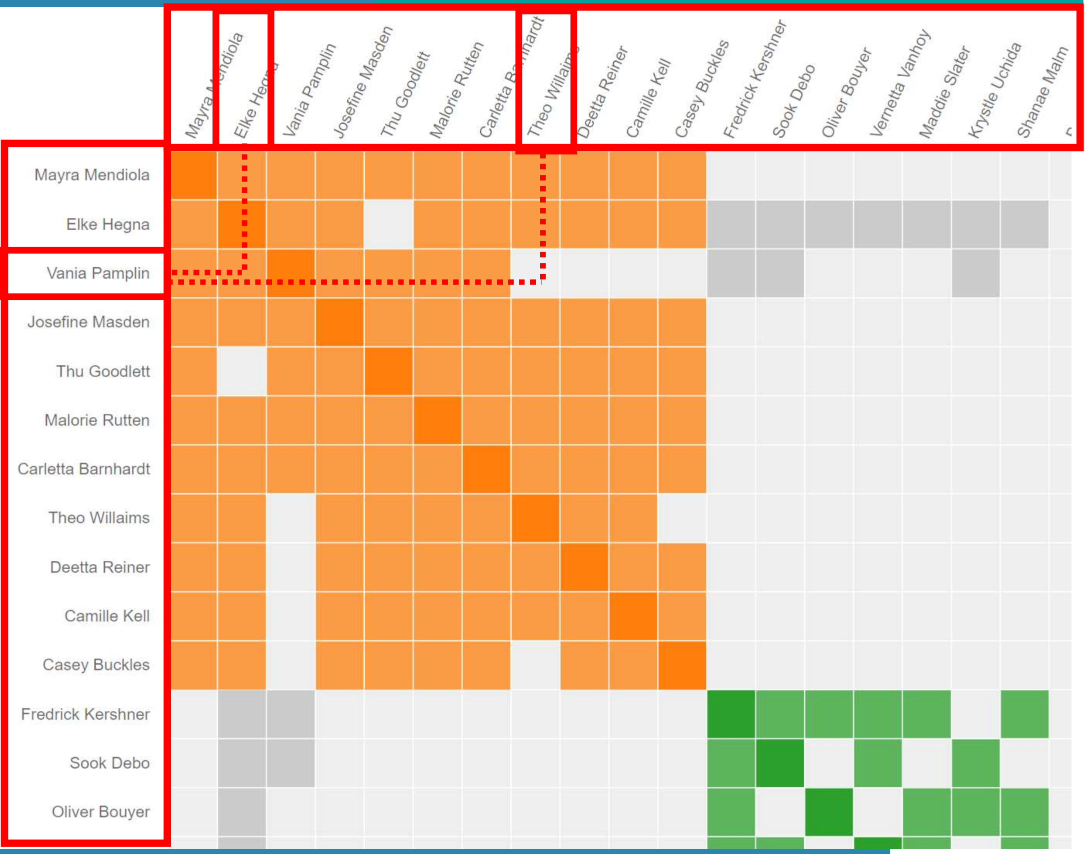
🟩 Cluster 2: 7 members

🟥 Cluster 3: 9 members

🟪 Cluster 4: 4 members

- Top section contains information concerning the analysis
  - Person tested
  - DNA testing company
  - Date of analysis
- Order DNA matches
- Cluster information

AutoCluster concepts



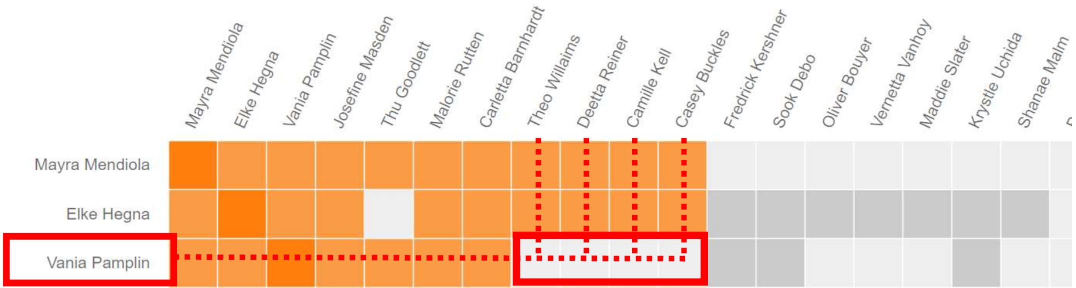
- your DNA matches on X and Y-axis
- You match Vania Pamplin & Elke Hegna
- Vania Pamplin & Elke Hegna also match!
- You also match Theo Willaims
- Vania Pamplin does not match Theo Willaims

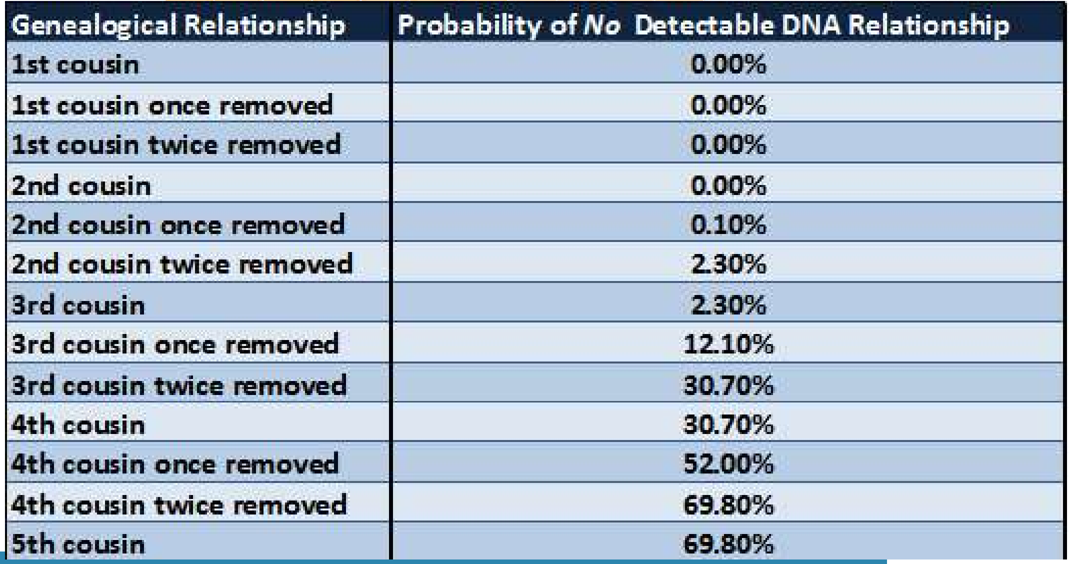
AutoCluster grey cells

- Grey cells can link clusters
- Indicate a relationship between clusters

- Elke Hegna could be a high cM match linking two more distant common ancestors
- Elke is a member of cluster 1 because she matches the most matches
- E.g., Elke Hegna is a maternal 2nd cousin
    - Orange cluster MMP
    - Green cluster MMM

AutoCluster empty cells





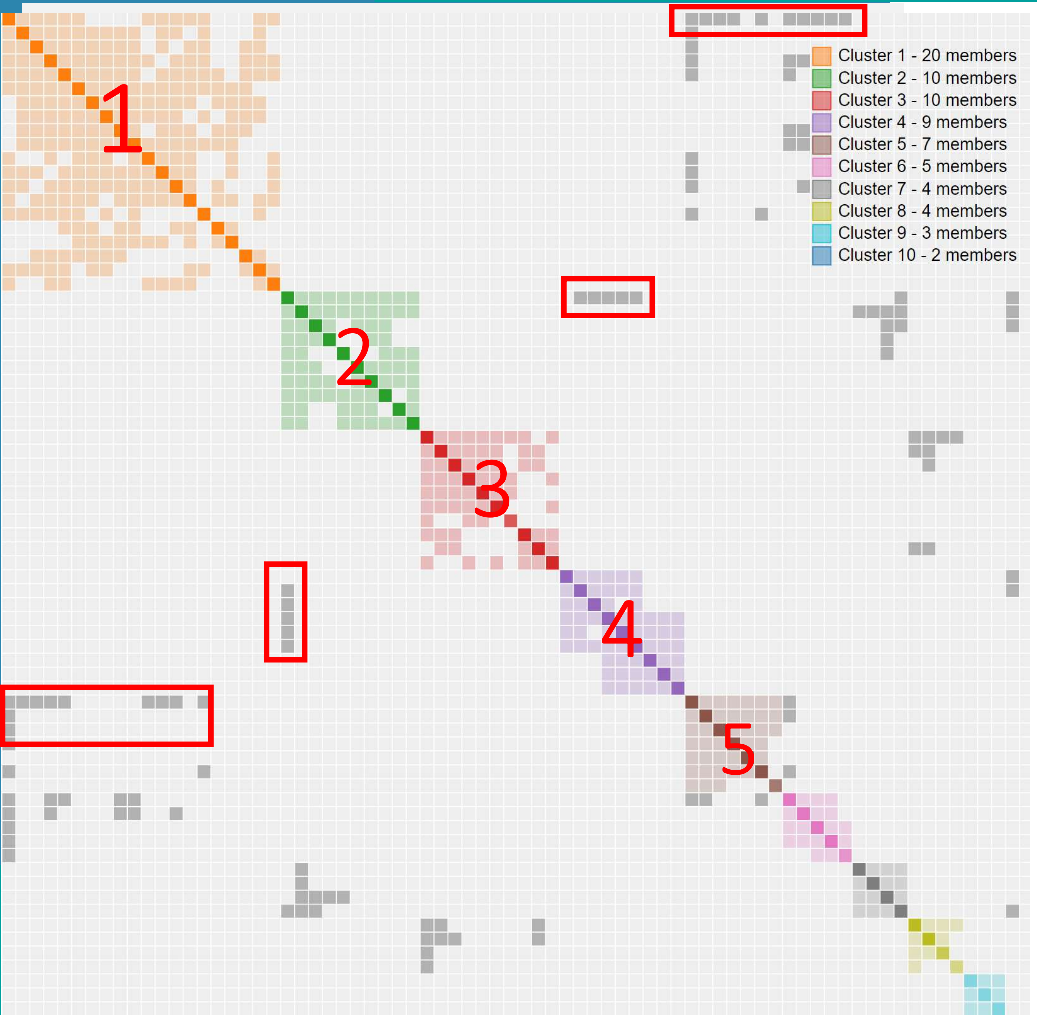
Genealogical Relationship	Probability of *No* Detectable DNA Relationship
1st cousin	0.00%
1st cousin once removed	0.00%
1st cousin twice removed	0.00%
2nd cousin	0.00%
2nd cousin once removed	0.10%
2nd cousin twice removed	2.30%
3rd cousin	2.30%
3rd cousin once removed	12.10%
3rd cousin twice removed	30.70%
4th cousin	30.70%
4th cousin once removed	52.00%
4th cousin twice removed	69.80%
5th cousin	69.80%

https://isogg.org/wiki/Cousin_statistics

- Vania does not match 4 matches in the cluster
- Based on chance, you sometimes do not match with a genealogical match
- 31% chance no DNA with 4C

- Shared cM threshold could be too high
  - On Ancestry, default is 20 cM

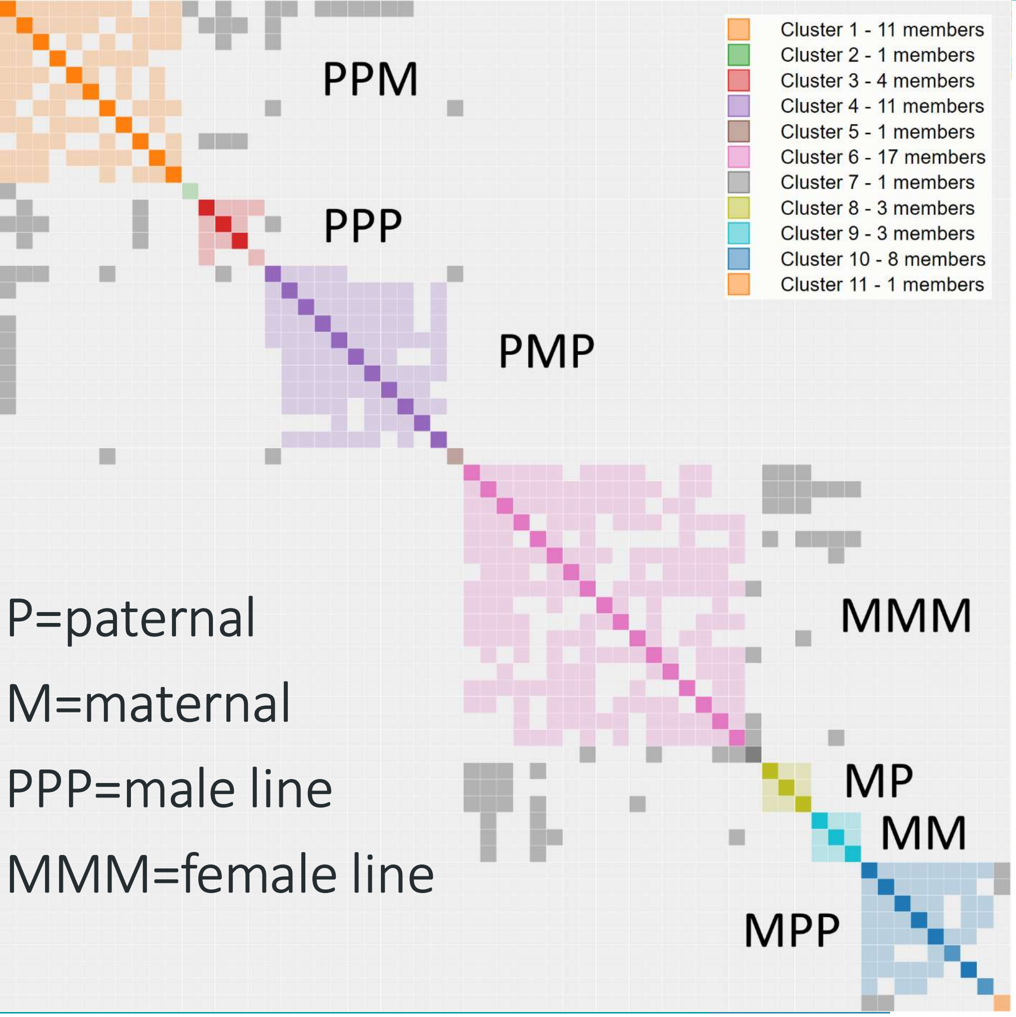
Ordering of clusters



- Initial sorting of clusters was based on cluster size
- More intuitive sorting; based on interactions between clusters

- Link cluster 5 with cluster 1
- Link cluster 4 with cluster 2

Ordering of clusters



P=paternal

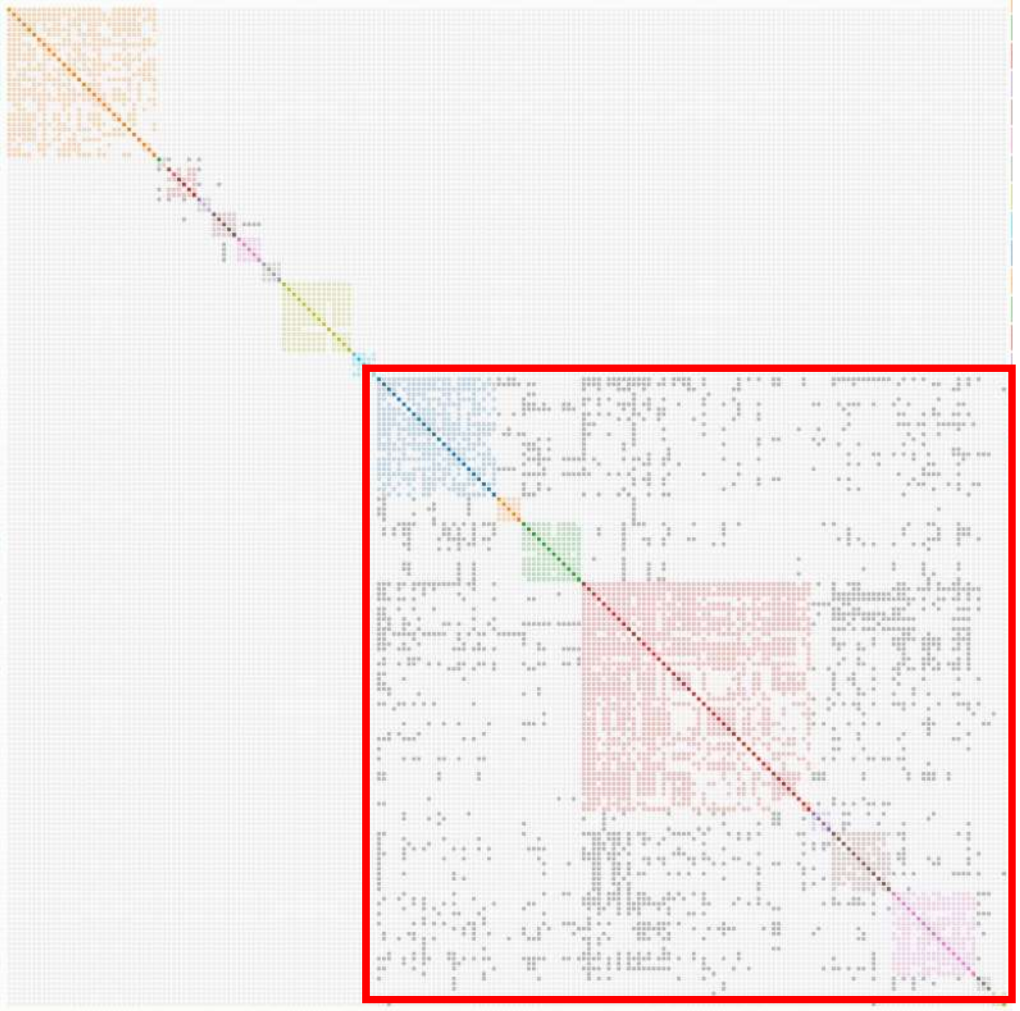
M=maternal

PPP=male line

MMM=female line

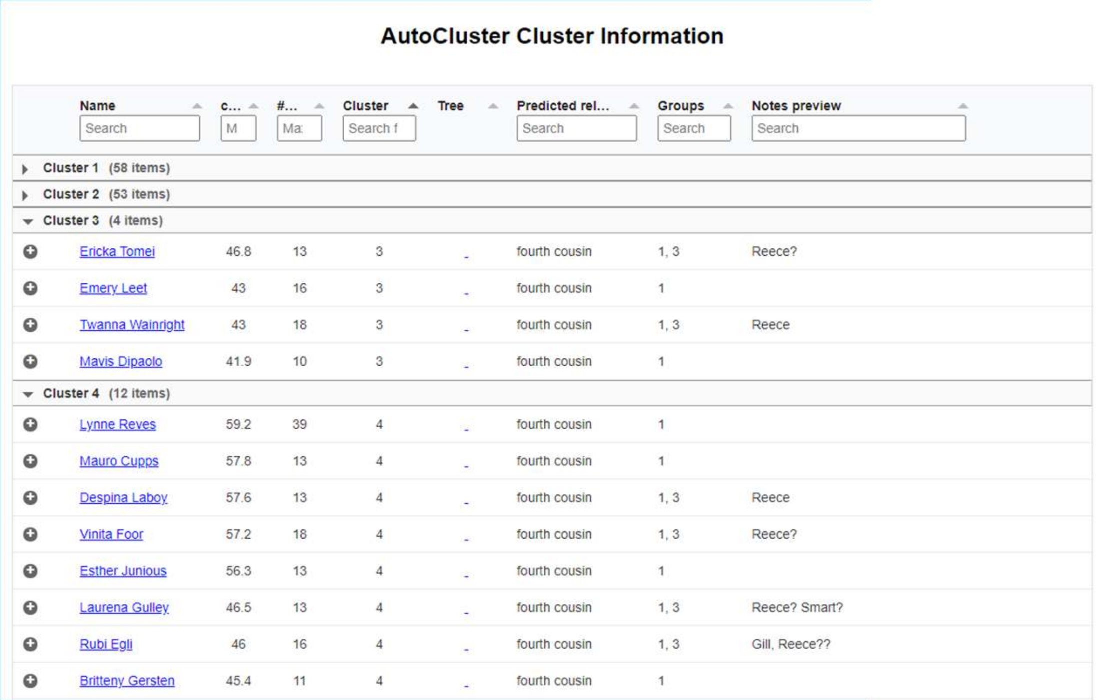
- Connected clusters often in concordance with ancestry tree
- High cM matches can be used to “link” clusters together

Ordering of clusters



- Connected clusters often in concordance with ancestry tree
- High cM matches can be used to “link” clusters together
- A group of clusters that are closely related to each other => supercluster

AutoCluster table

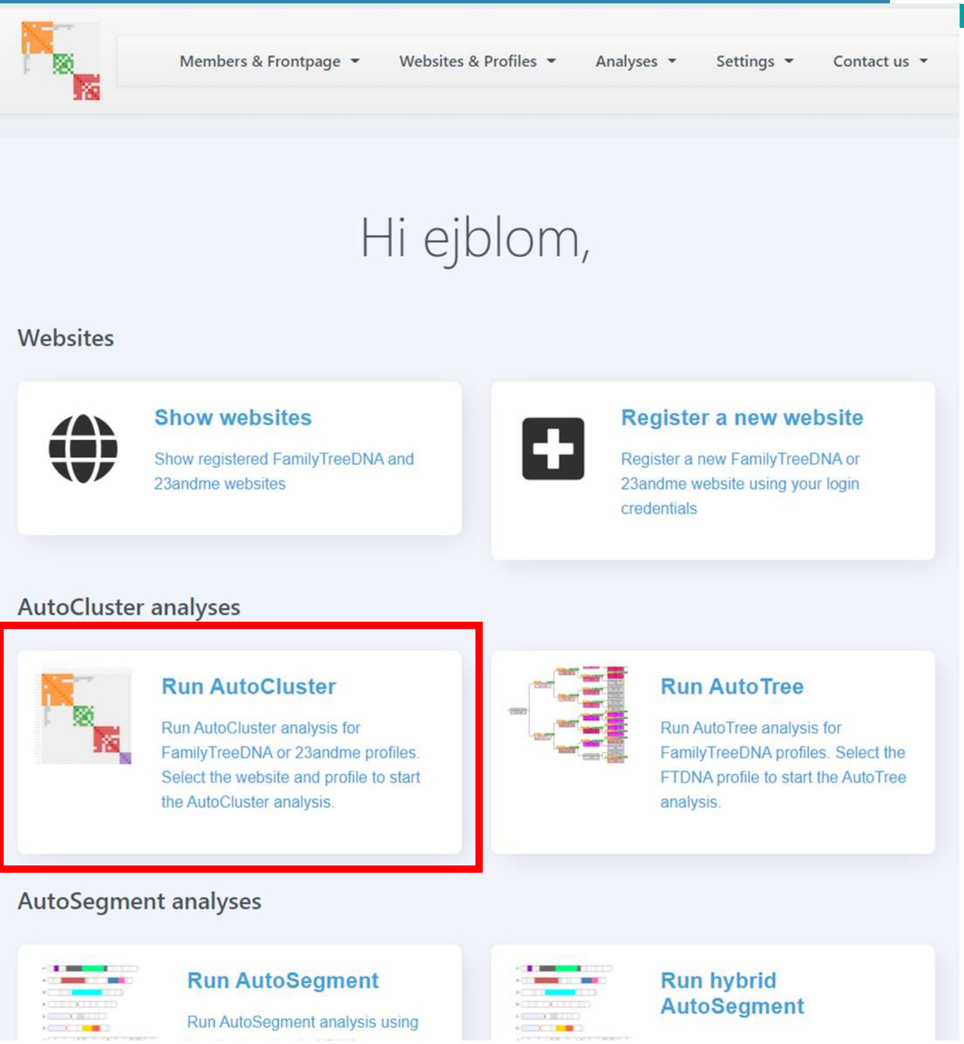


AutoCluster Cluster Information

Name	c...	#...	Cluster	Tree	Predicted rel...	Groups	Notes preview
Cluster 1 (58 items)							
Cluster 2 (53 items)							
Cluster 3 (4 items)							
Ericka Tomei	46.8	13	3	_	fourth cousin	1, 3	Reece?
Emery Leet	43	16	3	_	fourth cousin	1	
Twanna Wainright	43	18	3	_	fourth cousin	1, 3	Reece
Mavis Dipaolo	41.9	10	3	_	fourth cousin	1	
Cluster 4 (12 items)							
Lynne Reves	59.2	39	4	_	fourth cousin	1	
Mauro Cupps	57.8	13	4	_	fourth cousin	1	
Despina Laboy	57.6	13	4	_	fourth cousin	1, 3	Reece
Vinita Foor	57.2	18	4	_	fourth cousin	1, 3	Reece?
Esther Junious	56.3	13	4	_	fourth cousin	1	
Laurena Gulley	46.5	13	4	_	fourth cousin	1, 3	Reece? Smart?
Rubi Egli	46	16	4	_	fourth cousin	1, 3	Gill, Reece??
Britteny Gersten	45.4	11	4	_	fourth cousin	1	

- Matches per cluster available in table underneath chart
- Link to DNA match, shared cM and notes

Start clustering



Hi ejblom,

Websites

Show websites

Show registered FamilyTreeDNA and 23andme websites

Register a new website

Register a new FamilyTreeDNA or 23andme website using your login credentials

AutoCluster analyses

Run AutoCluster

Run AutoCluster analysis for FamilyTreeDNA or 23andme profiles. Select the website and profile to start the AutoCluster analysis.

Run AutoTree

Run AutoTree analysis for FamilyTreeDNA profiles. Select the FTDNA profile to start the AutoTree analysis.

AutoSegment analyses

Run AutoSegment

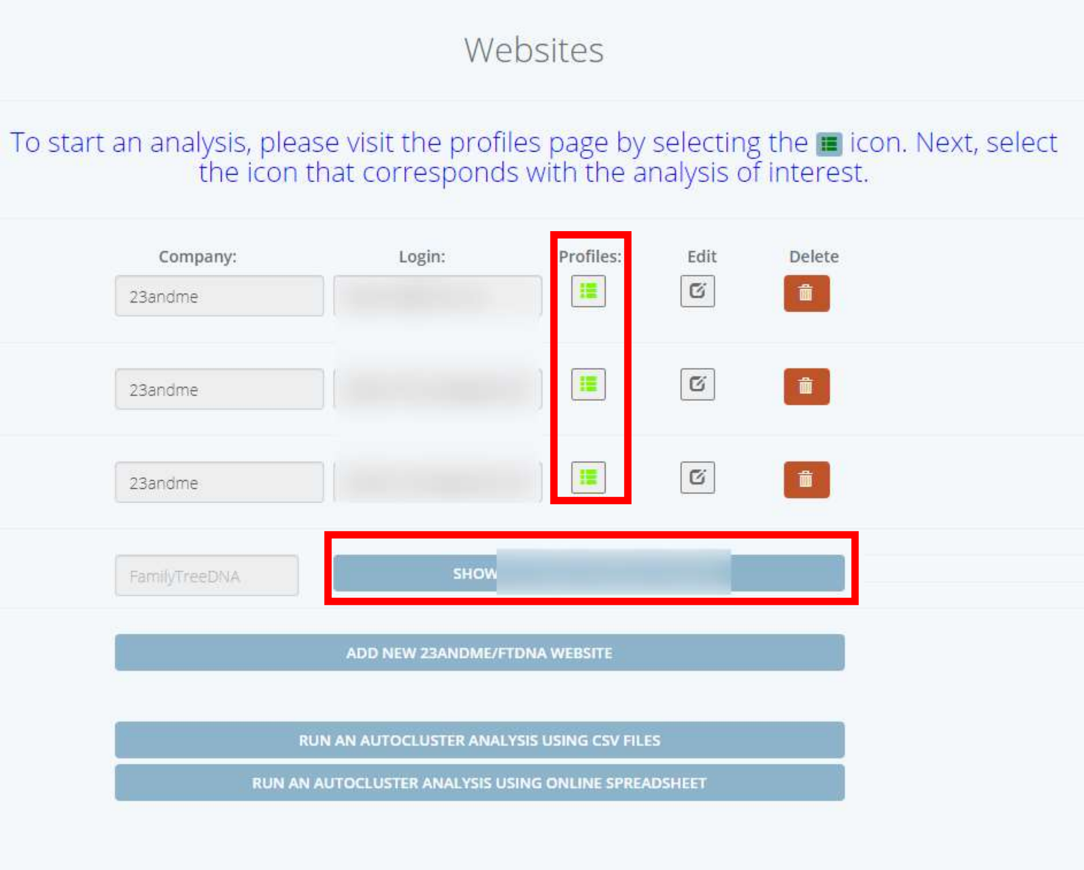
Run AutoSegment analysis using

Run hybrid AutoSegment

- Genetic Affairs members page: https://members.geneticaffairs.com/

- Free registration and 200 credits
- Shows registered websites
- Register new websites (23andme/FTDNA)
- Start analyses

FTDNA & 23andme



To start an analysis, please visit the profiles page by selecting the 🟩 icon. Next, select the icon that corresponds with the analysis of interest.

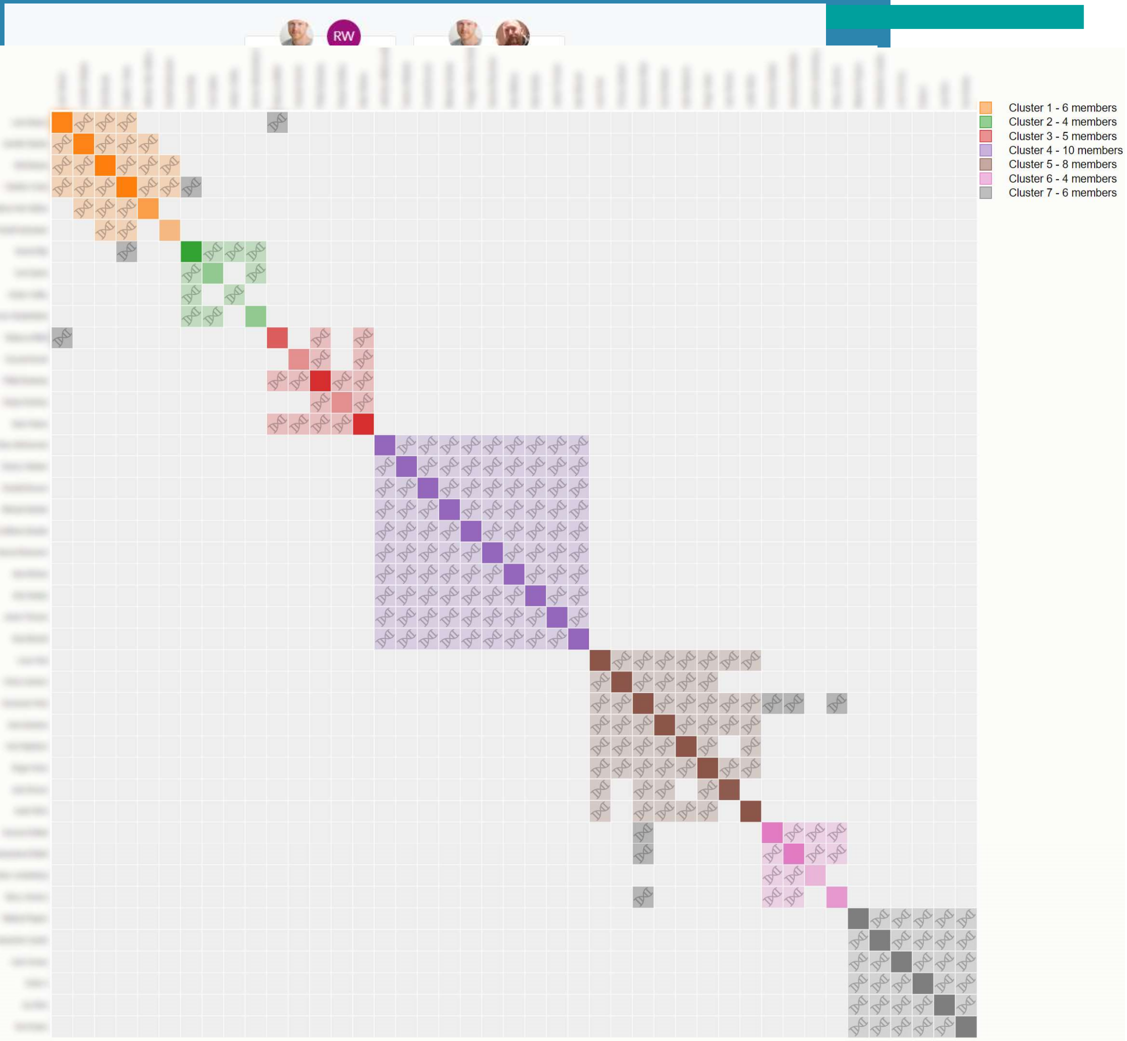
Company:	Login:	Profiles:	Edit	Delete
23andme	[redacted]	🟩	✎	🗑
23andme	[redacted]	🟩	✎	🗑
23andme	[redacted]	🟩	✎	🗑
FamilyTreeDNA				

- Register FTDNA or 23andme account

- Select registered FTDNA/23andme account
- Select profile of interest
- Set parameters:
  - max & min cM (e.g., between 400 -50 cM)
  - lower cM limit = more matches/clusters
  - min shared segment size
  - For 23andme: min shared DNA between matches

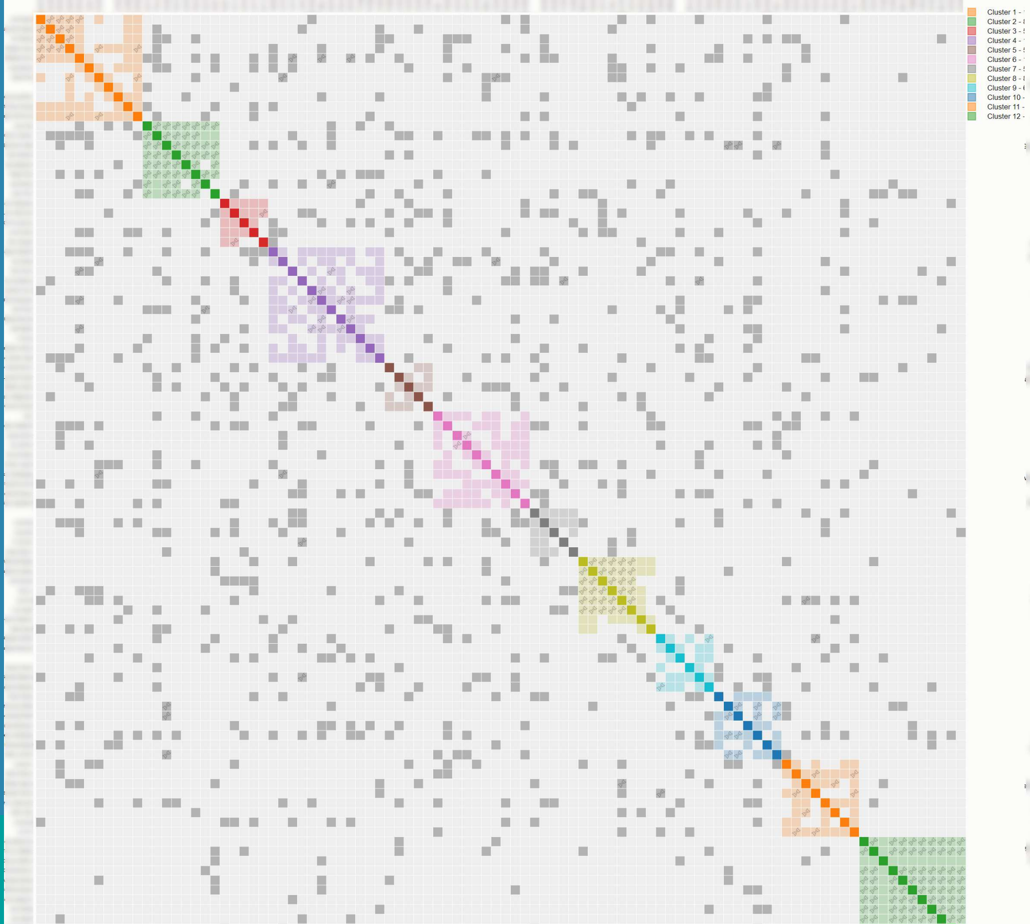
- Start analysis

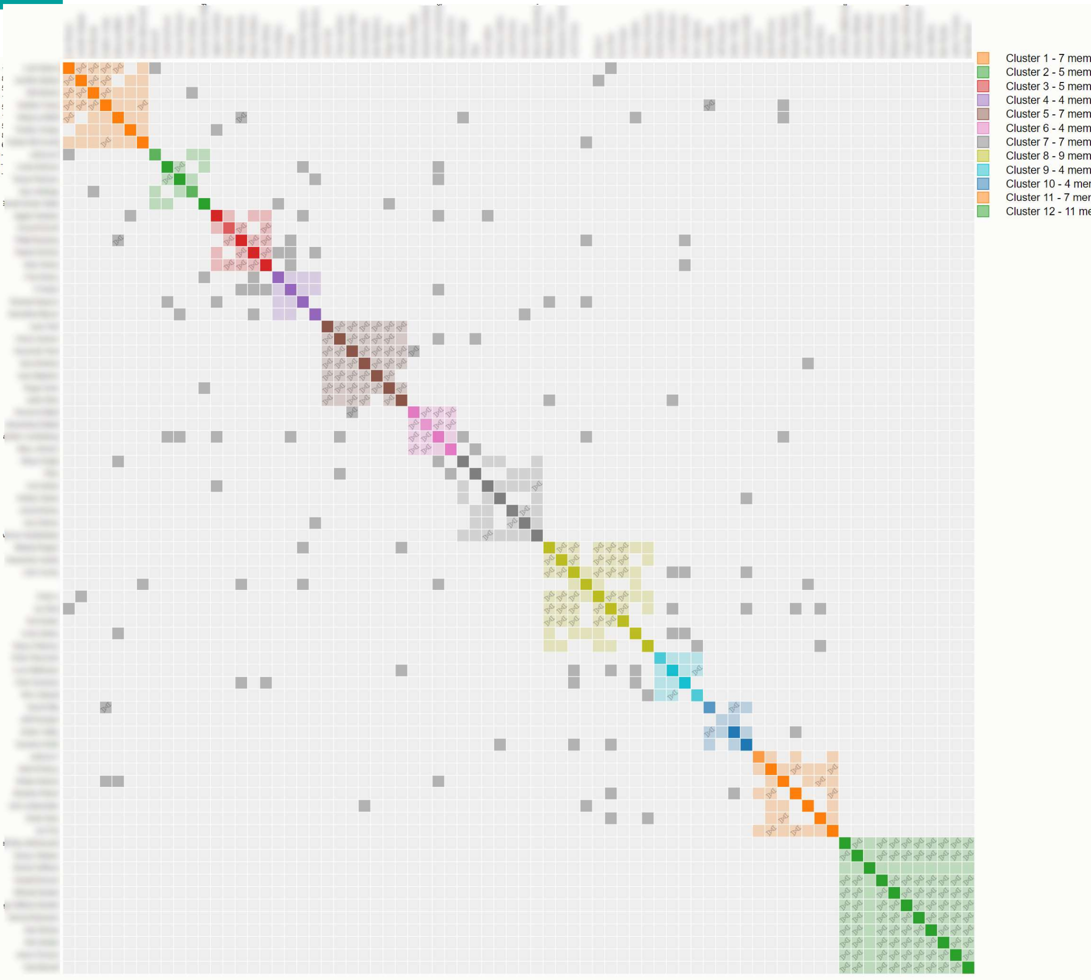
23andme shared DNA



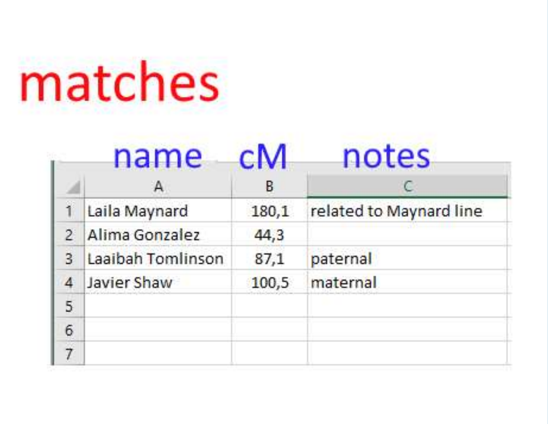
- 23andme displays “shared DNA” for relatives in common
- Click on the “Yes” : segments in chromosome browser with overlap
- Helix symbols represent these overlapping (and triangulated) segments
- Possible to run clustering only using overlapping segments (AutoSegment)

Shared DNA: 5 vs 10 cM



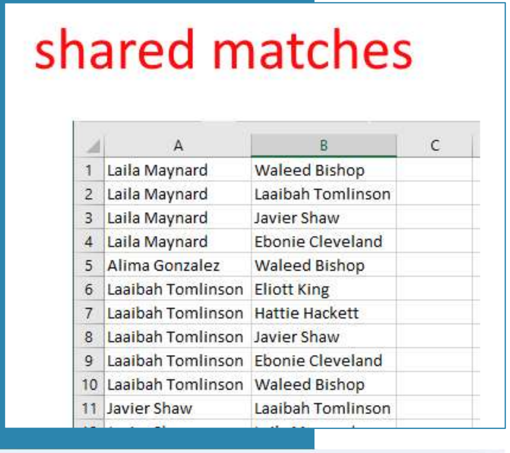


AutoCluster for Ancestry



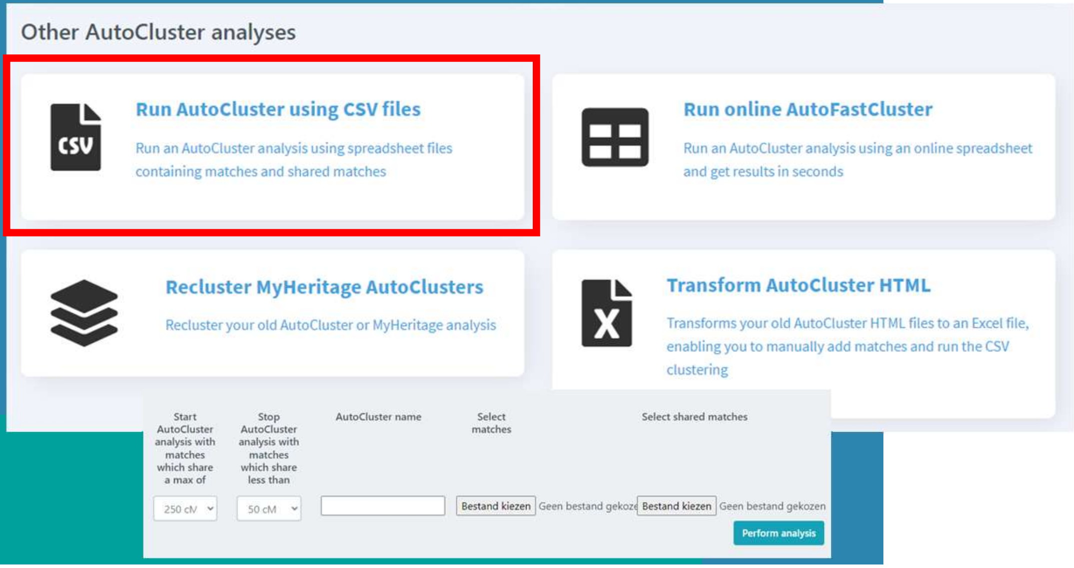
matches

	name A	cM B	notes C
1	Laila Maynard	180,1	related to Maynard line
2	Alima Gonzalez	44,3	
3	Laaibah Tomlinson	87,1	paternal
4	Javier Shaw	100,5	maternal
5			
6			
7			



shared matches

	A	B	C
1	Laila Maynard	Waleed Bishop	
2	Laila Maynard	Laaibah Tomlinson	
3	Laila Maynard	Javier Shaw	
4	Laila Maynard	Ebonie Cleveland	
5	Alima Gonzalez	Waleed Bishop	
6	Laaibah Tomlinson	Eliott King	
7	Laaibah Tomlinson	Hattie Hackett	
8	Laaibah Tomlinson	Javier Shaw	
9	Laaibah Tomlinson	Ebonie Cleveland	
10	Laaibah Tomlinson	Waleed Bishop	
11	Javier Shaw	Laaibah Tomlinson	



Other AutoCluster analyses

Run AutoCluster using CSV files
Run an AutoCluster analysis using spreadsheet files containing matches and shared matches

Run online AutoFastCluster
Run an AutoCluster analysis using an online spreadsheet and get results in seconds

Recluster MyHeritage AutoClusters
Recluster your old AutoCluster or MyHeritage analysis

Transform AutoCluster HTML
Transforms your old AutoCluster HTML files to an Excel file, enabling you to manually add matches and run the CSV clustering

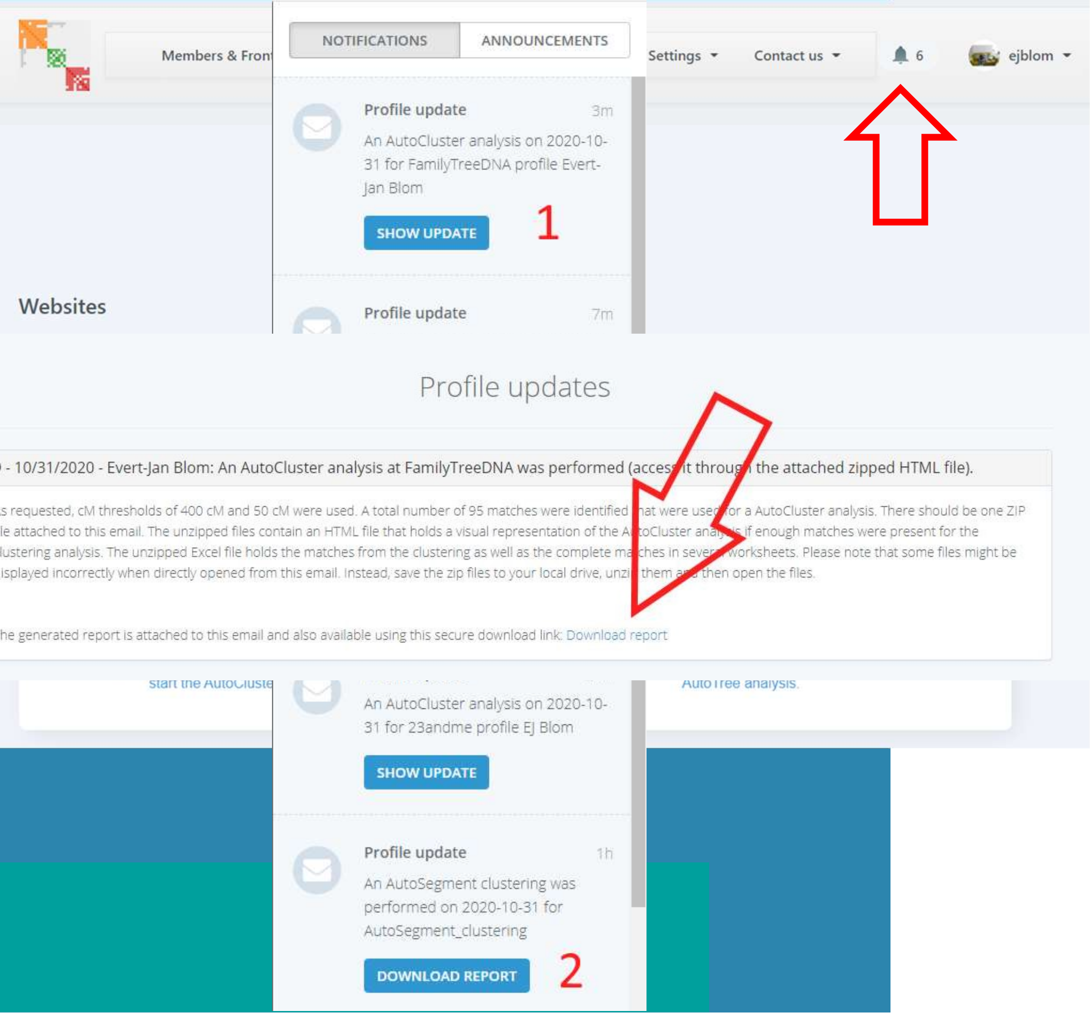
Start AutoCluster analysis with matches which share a max of: 250 cM
Stop AutoCluster analysis with matches which share less than: 50 cM
AutoCluster name:
Select matches: Bestand kiezen Geen bestand gekoze
Select shared matches: Bestand kiezen Geen bestand gekozen

- Manually obtain matches & shared matches
- Using a spreadsheet, save 2 CSV files

- DNA matches : 3 columns
  - name, cM & notes
- shared matches: 2 columns
  - match, shared match

- Run AutoCluster using CSV files

Receive results



Profile update 3m

An AutoCluster analysis on 2020-10-31 for FamilyTreeDNA profile Evert-Jan Blom

Profile update 7m

Profile updates

) - 10/31/2020 - Evert-Jan Blom: An AutoCluster analysis at FamilyTreeDNA was performed (access it through the attached zipped HTML file).

s requested, cM thresholds of 400 cM and 50 cM were used. A total number of 95 matches were identified that were used for a AutoCluster analysis. There should be one ZIP le attached to this email. The unzipped files contain an HTML file that holds a visual representation of the AutoCluster analysis if enough matches were present for the lustering analysis. The unzipped Excel file holds the matches from the clustering as well as the complete matches in several worksheets. Please note that some files might be isplayed incorrectly when directly opened from this email. Instead, save the zip files to your local drive, unzip them and then open the files.

he generated report is attached to this email and also available using this secure download link: Download report

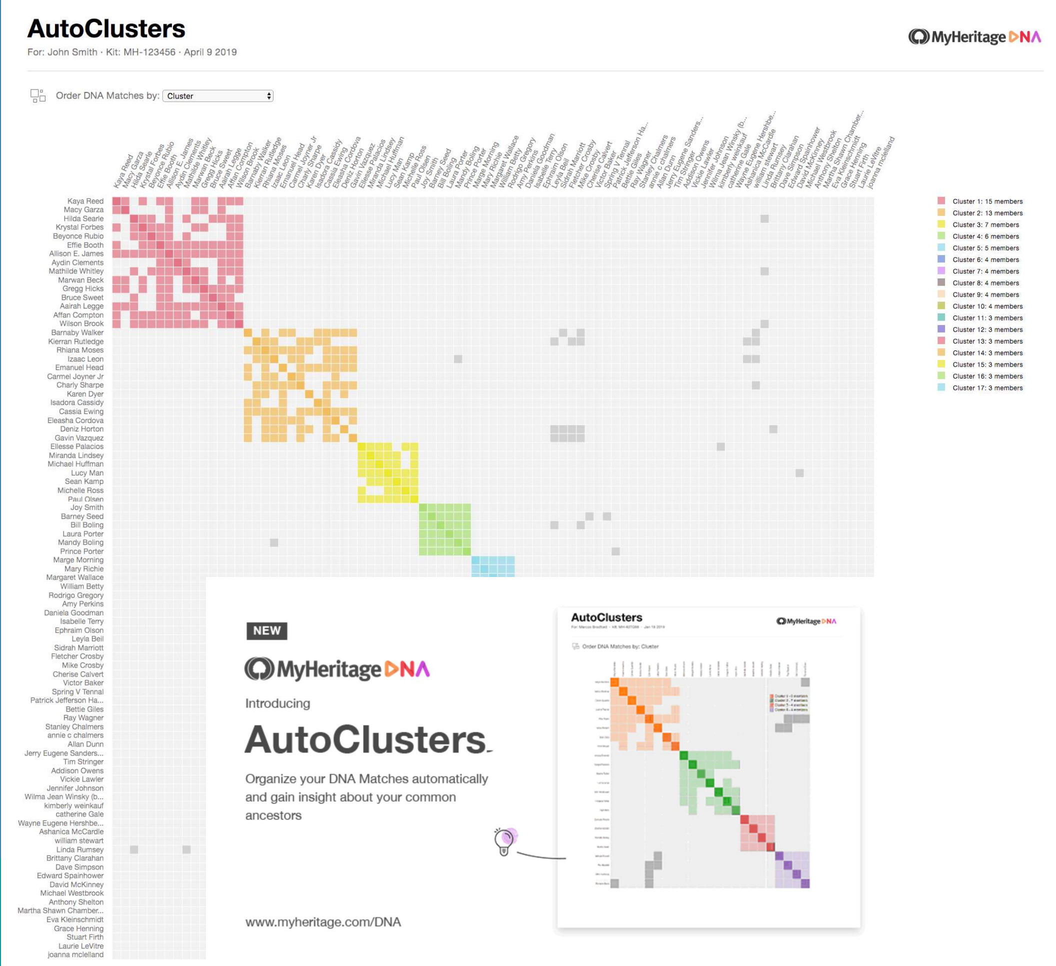
An AutoCluster analysis on 2020-10-31 for 23andme profile EJ Blom

Profile update 1h

An AutoSegment clustering was performed on 2020-10-31 for AutoSegment_clustering

- Results are emailed (with attachment)
- or
- Download analysis from our site
- Save ZIP file to local drive
- Unzip ZIP file and open HTML

AutoCluster on MyHeritage



AutoClusters

For: John Smith · Kit: MH-123456 · April 9 2019

Order DNA Matches by: Cluster

Cluster 1: 15 members
Cluster 2: 13 members
Cluster 3: 7 members
Cluster 4: 6 members
Cluster 5: 5 members
Cluster 6: 4 members
Cluster 7: 4 members
Cluster 8: 4 members
Cluster 9: 4 members
Cluster 10: 4 members
Cluster 11: 3 members
Cluster 12: 3 members
Cluster 13: 3 members
Cluster 14: 3 members
Cluster 15: 3 members
Cluster 16: 3 members
Cluster 17: 3 members

NEW

MyHeritage DNA

Introducing

AutoClusters

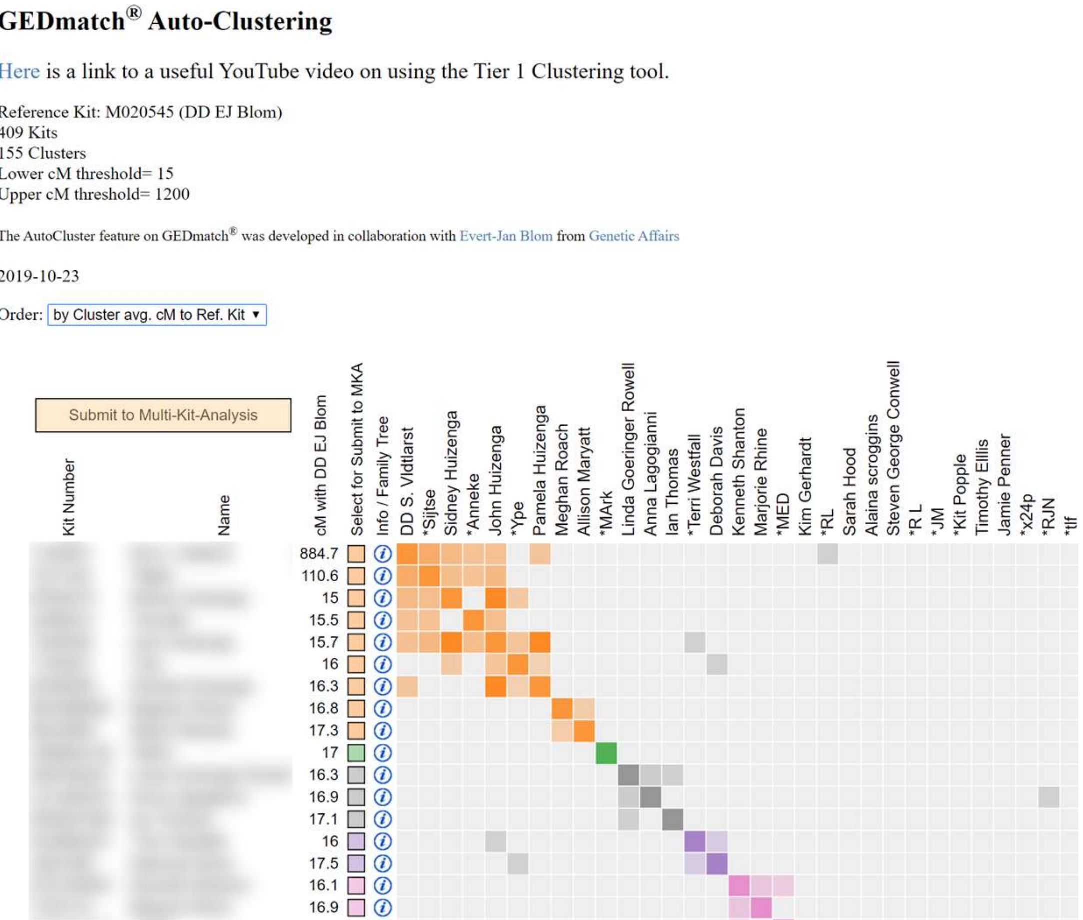
Organize your DNA Matches automatically and gain insight about your common ancestors

www.myheritage.com/DNA

Introducing AutoClusters for DNA Matches

- Implementation of AutoCluster algorithm running at MyHeritage
- Special initial analysis:
  - For top 500 matches (max 400 cM)
  - Test many cM ranges (min max) as well as shared cM between matches
  - Take clustering result with around 100 matches
- Increased shared cM improves clusters for endogamous populations

AutoCluster for GEDmatch



GEDmatch® Auto-Clustering

Here is a link to a useful YouTube video on using the Tier 1 Clustering tool.

Reference Kit: M020545 (DD EJ Blom)
409 Kits
155 Clusters
Lower cM threshold= 15
Upper cM threshold= 1200

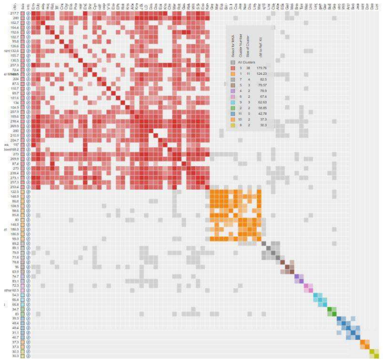
The AutoCluster feature on GEDmatch® was developed in collaboration with Evert-Jan Blom from Genetic Affairs

2019-10-23

Order: by Cluster avg. cM to Ref. Kit ▼

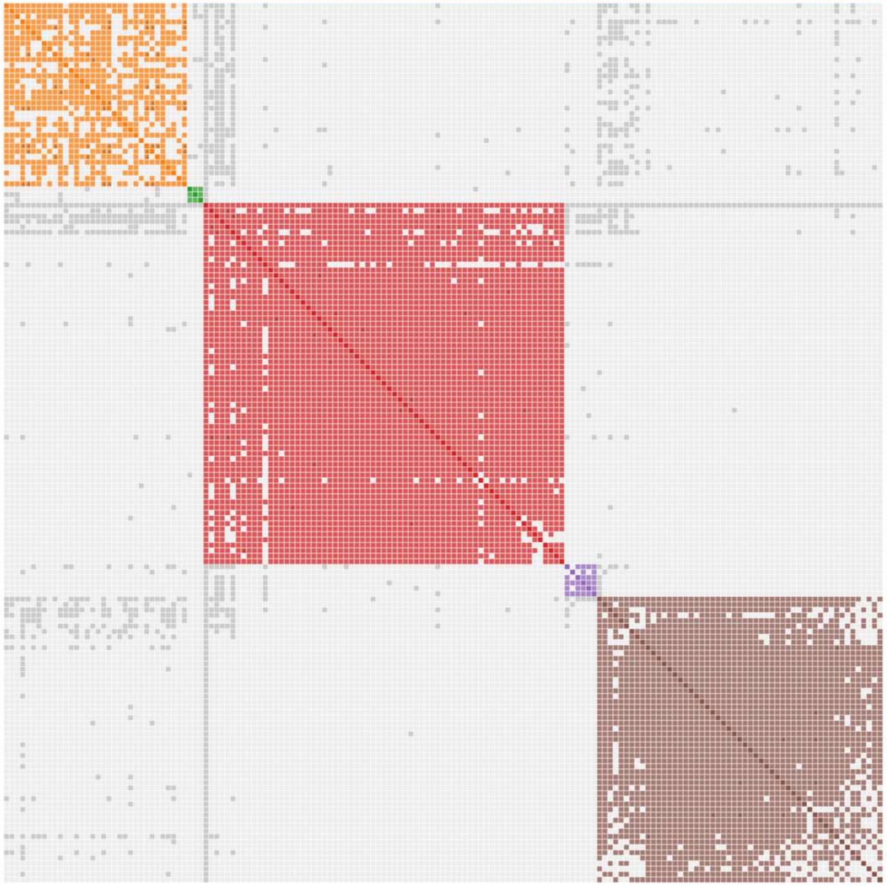
Kit Number	Name	cM with DD EJ Blom	Select for Submit to MKA	Info / Family Tree	DD S. Vldtlarst	*Sijtse	Sidney Huizenga	*Anneke	John Huizenga	*Ype	Pamela Huizenga	Meghan Roach	Allison Maryatt	*MArk	Linda Goeringer Rowell	Anna Lagogianni	Ian Thomas	*Terri Westfall	Deborah Davis	Kenneth Shanton	Marjorie Rhine	*MED	Kim Gerhardt	*RL	Sarah Hood	Alaina scroggins	Steven George Conwell	*R L	*JM	*Kit Popple	Timothy Ellis	Jamie Penner	*x24p	*RJN	*tlf
[redacted]	[redacted]	884.7	☐	ⓘ																															
[redacted]	[redacted]	110.6	☐	ⓘ																															
[redacted]	[redacted]	15	☐	ⓘ																															
[redacted]	[redacted]	15.5	☐	ⓘ																															
[redacted]	[redacted]	15.7	☐	ⓘ																															
[redacted]	[redacted]	16	☐	ⓘ																															
[redacted]	[redacted]	16.3	☐	ⓘ																															
[redacted]	[redacted]	16.8	☐	ⓘ																															
[redacted]	[redacted]	17.3	☐	ⓘ																															
[redacted]	[redacted]	17	☐	ⓘ																															
[redacted]	[redacted]	16.3	☐	ⓘ																															
[redacted]	[redacted]	16.9	☐	ⓘ																															
[redacted]	[redacted]	17.1	☐	ⓘ																															
[redacted]	[redacted]	16	☐	ⓘ																															
[redacted]	[redacted]	17.5	☐	ⓘ																															
[redacted]	[redacted]	16.1	☐	ⓘ																															
[redacted]	[redacted]	16.9	☐	ⓘ																															

- Integration for Tier 1 users



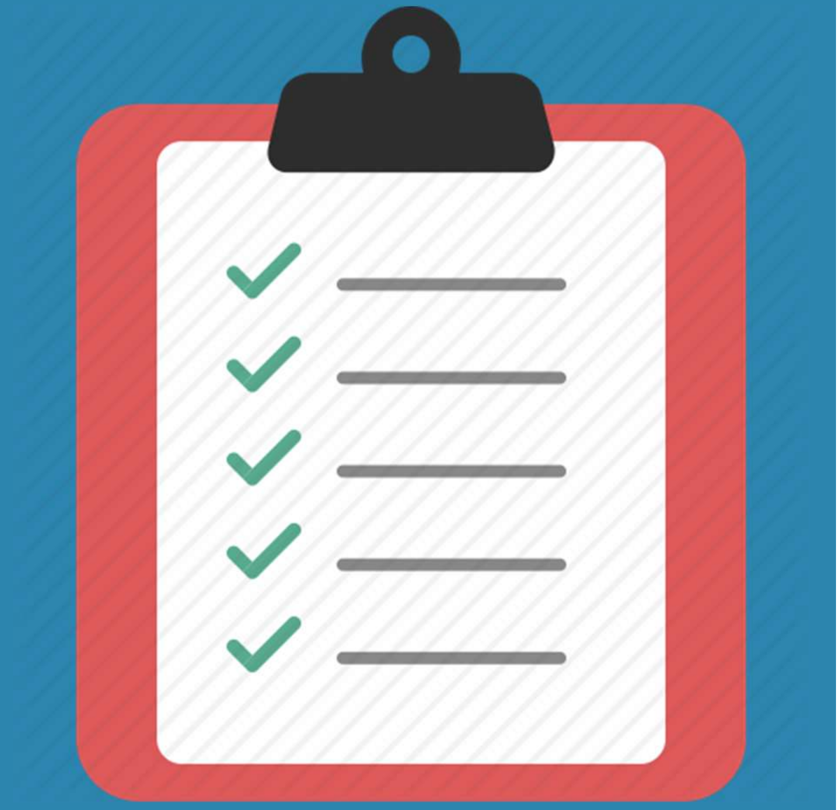
- Interactive chart
- Ability to employ the Multi-Kit-Analysis for downstream analyses

Conclusion AutoCluster



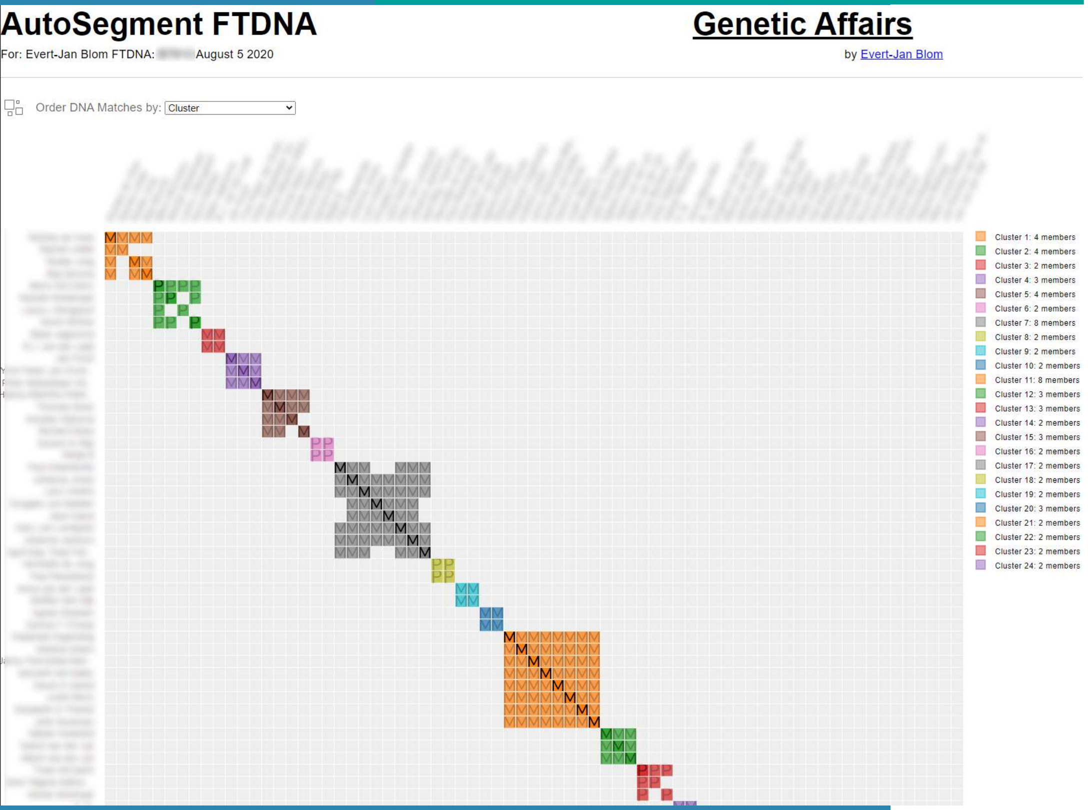
- AutoCluster groups DNA matches that likely share a common ancestor
- Clustering can be performed automatically for:
    - FamilyTreeDNA
    - 23andme
- Using CSV files for Ancestry/LivingDNA
- On GEDmatch for tier 1 users

Outline AutoSegment



- Clustering using segments
- AutoCluster ICW clusters
- Example
- ICW vs shared segments
- Concepts
- Visualization
- Caveats
- Start AutoSegment
  - FTDNA, 23andme & GEDmatch
  - Hybrid AutoSegment
- Conclusion

AutoSegment



AutoSegment FTDNA

For: Evert-Jan Blom FTDNA: [redacted] August 5 2020

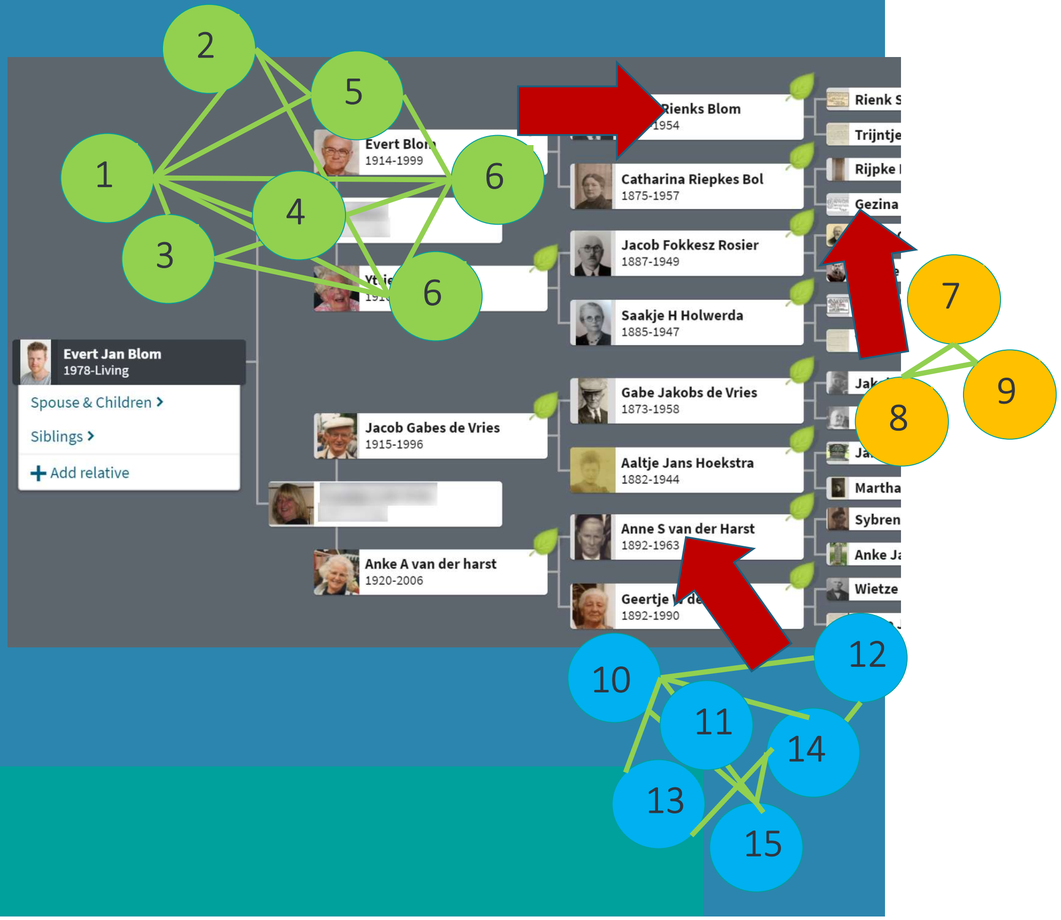
Genetic Affairs

by Evert-Jan Blom

Order DNA Matches by: Cluster

- AutoSegment automatically organizes your matches into clusters based on overlapping segments or triangulating segments.
- Works for MyHeritage, 23andme, FTDNA and GEDmatch segments
- DISCLAIMER: in some scenarios it requires follow up analyses

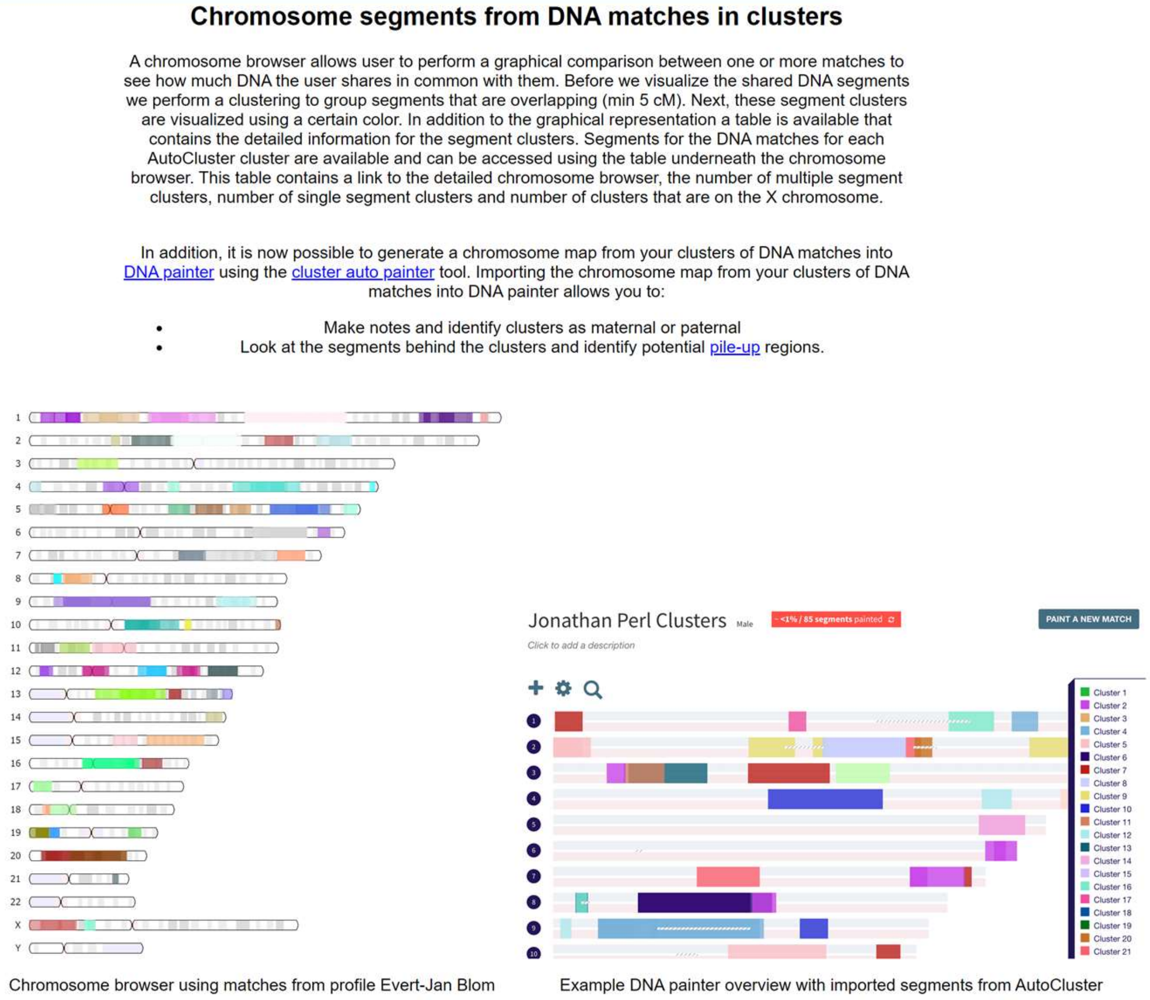
Clustering using segments



- Members of clusters with the same segment share a common ancestor and represent a branch of your family.

- Analyzing clusters:
  - Interpret trees of matches from a cluster
  - Contact matches, build Q&D trees
  - Painting of DNA segments to help classify future DNA matches

Clustering using segments



Chromosome segments from DNA matches in clusters

A chromosome browser allows user to perform a graphical comparison between one or more matches to see how much DNA the user shares in common with them. Before we visualize the shared DNA segments we perform a clustering to group segments that are overlapping (min 5 cM). Next, these segment clusters are visualized using a certain color. In addition to the graphical representation a table is available that contains the detailed information for the segment clusters. Segments for the DNA matches for each AutoCluster cluster are available and can be accessed using the table underneath the chromosome browser. This table contains a link to the detailed chromosome browser, the number of multiple segment clusters, number of single segment clusters and number of clusters that are on the X chromosome.

In addition, it is now possible to generate a chromosome map from your clusters of DNA matches into DNA painter using the cluster auto painter tool. Importing the chromosome map from your clusters of DNA matches into DNA painter allows you to:

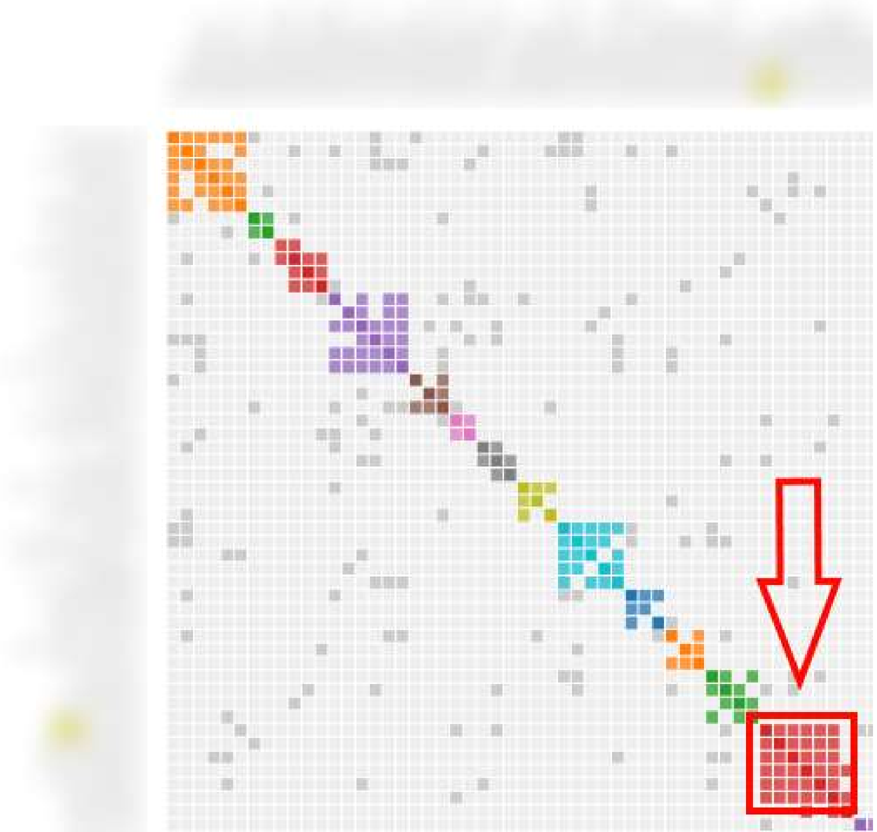
- Make notes and identify clusters as maternal or paternal
- Look at the segments behind the clusters and identify potential pile-up regions.

Chromosome browser using matches from profile Evert-Jan Blom

Example DNA painter overview with imported segments from AutoCluster

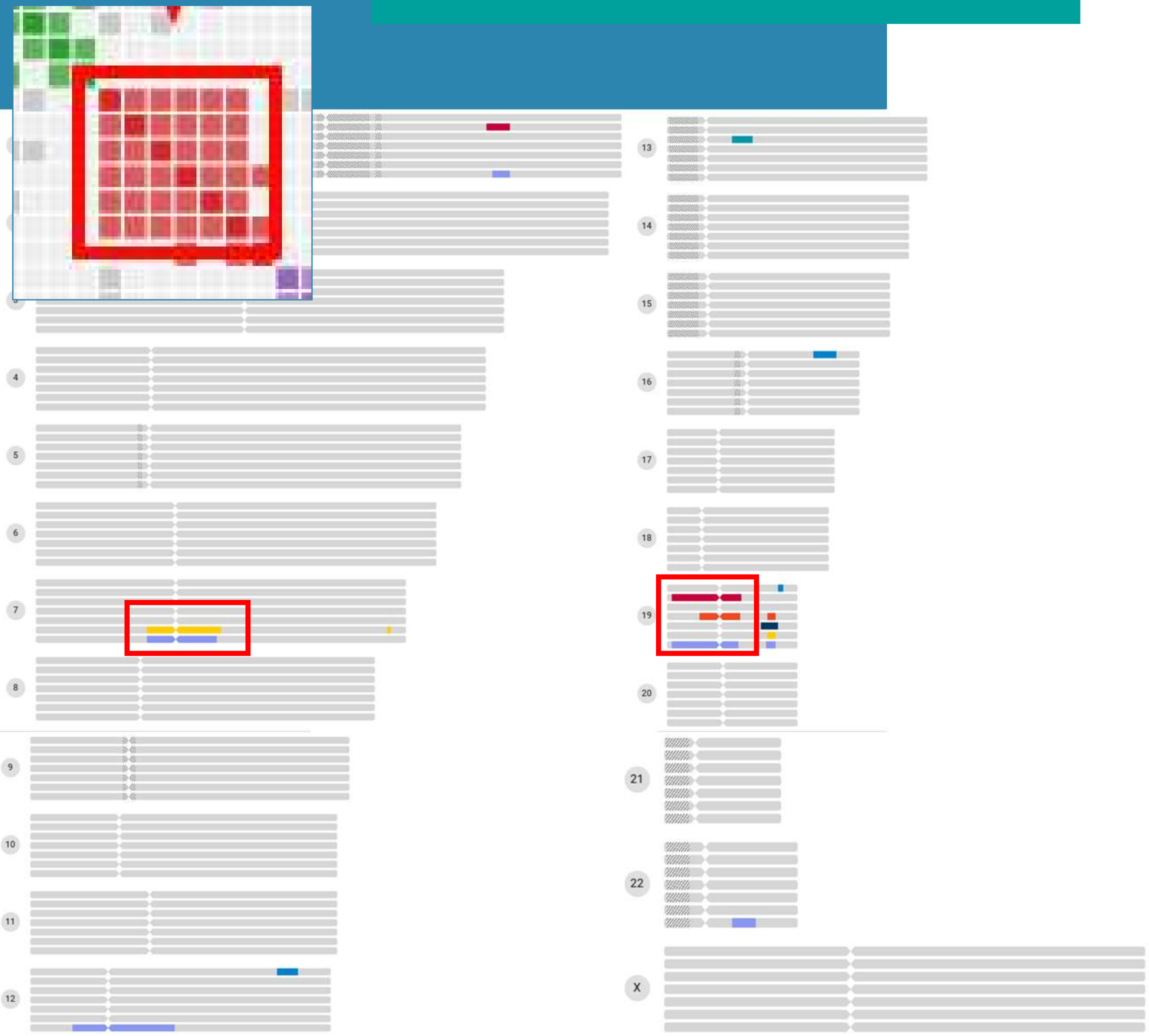
- Segment data can be imported into DNA Painter
- Cluster Auto Painter
- Generate a chromosome map from your clusters of DNA matches
- Provide hypotheses for future matches
- See https://www.dnapainter.com/tools/cap

AutoCluster clustering



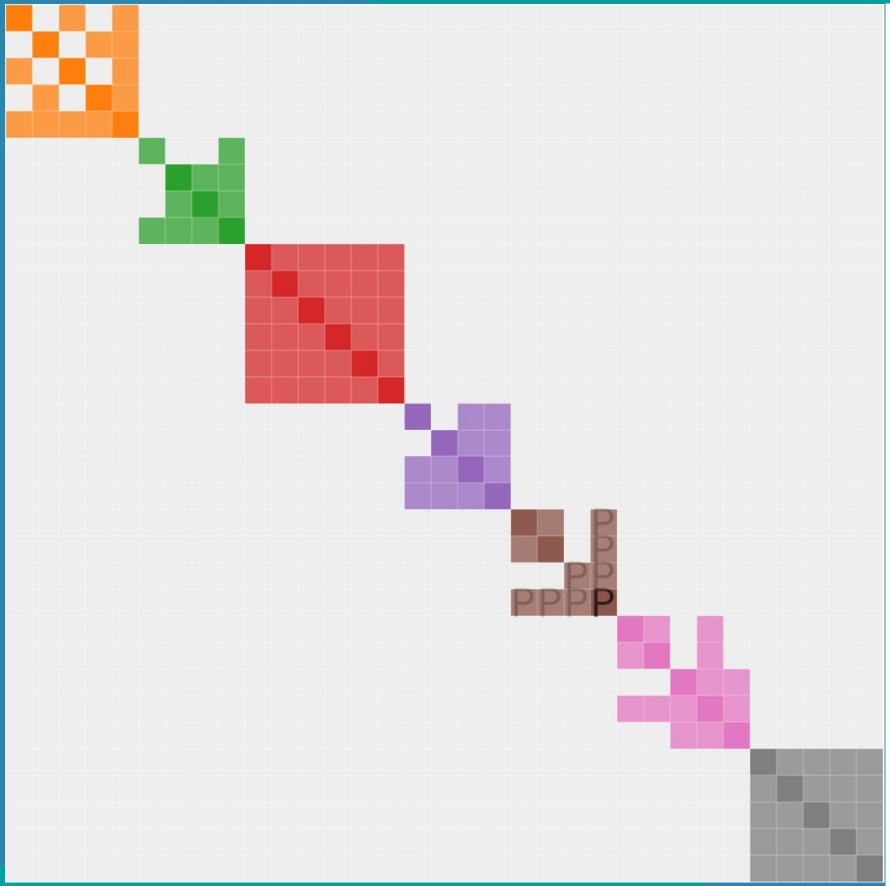
- Output from a “regular” AutoCluster clustering
- Based on shared matches – In Common With
- Red FTDNA cluster – all matches are shared between each other
    - Check shared segments for matches from red cluster using FTDNA chromosome browser

AutoCluster ICW clusters



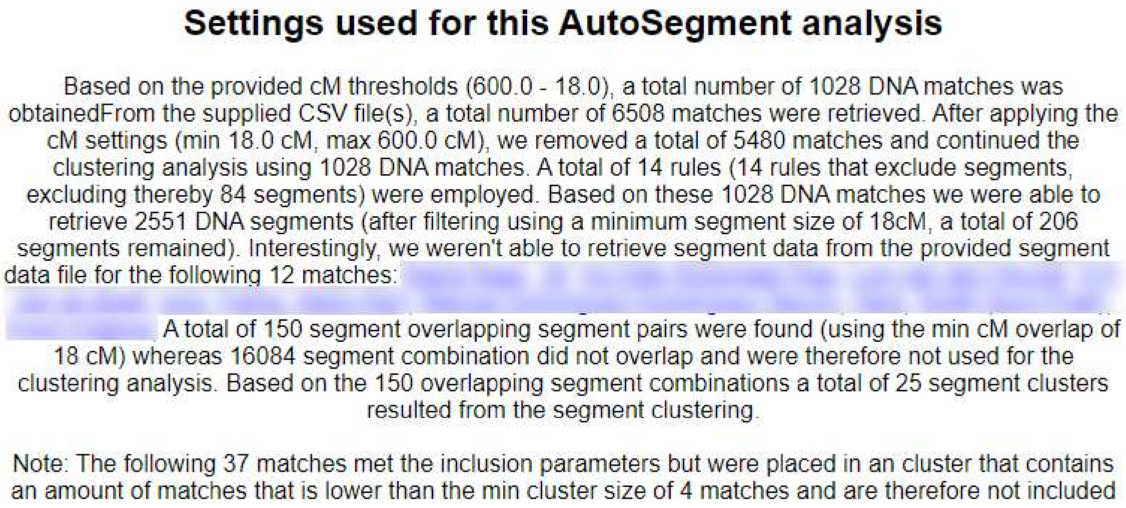
- Shared segments from the members of the red cluster
- Only some matches have overlapping segments (on chr 7 and chr 19)
- AutoSegment will only identify these overlapping segments

AutoSegment example



MyHeritage: 600-18 cM – 18 cM overlap

- MyHeritage segment and match data



Settings used for this AutoSegment analysis

Based on the provided cM thresholds (600.0 - 18.0), a total number of 1028 DNA matches was obtainedFrom the supplied CSV file(s), a total number of 6508 matches were retrieved. After applying the cM settings (min 18.0 cM, max 600.0 cM), we removed a total of 5480 matches and continued the clustering analysis using 1028 DNA matches. A total of 14 rules (14 rules that exclude segments, excluding thereby 84 segments) were employed. Based on these 1028 DNA matches we were able to retrieve 2551 DNA segments (after filtering using a minimum segment size of 18cM, a total of 206 segments remained). Interestingly, we weren't able to retrieve segment data from the provided segment data file for the following 12 matches: [redacted] A total of 150 segment overlapping segment pairs were found (using the min cM overlap of 18 cM) whereas 16084 segment combination did not overlap and were therefore not used for the clustering analysis. Based on the 150 overlapping segment combinations a total of 25 segment clusters resulted from the segment clustering.

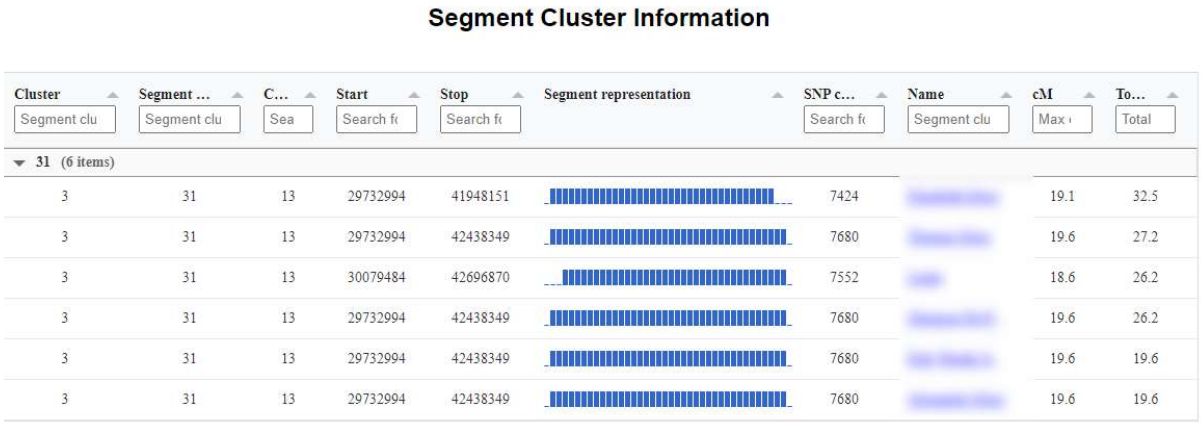
Note: The following 37 matches met the inclusion parameters but were placed in an cluster that contains an amount of matches that is lower than the min cluster size of 4 matches and are therefore not included

- Let’s analyze the third red cluster

AutoSegment Example

Chromosome browser

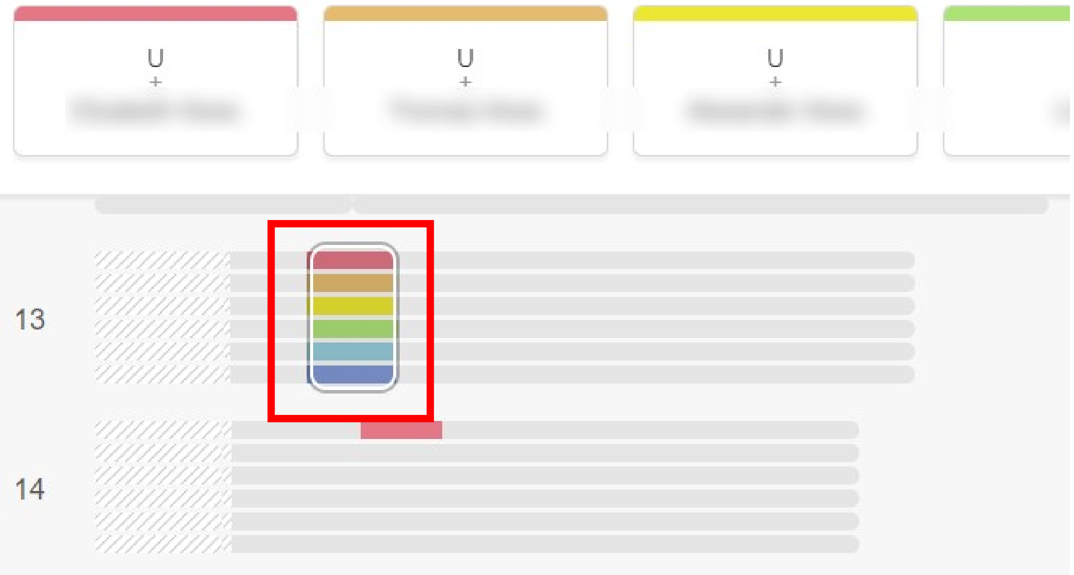
- Single well-defined segment on chr 13
- Segment table shows 6 matches with this segment & good overlap between segments



Segment Cluster Information

Cluster	Segment ...	C...	Start	Stop	Segment representation	SNP c...	Name	cM	To...
Segment clu	Segment clu	Sea	Search fc	Search fc		Search fc	Segment clu	Max	Total
31 (6 items)									
3	31	13	29732994	41948151		7424	[redacted]	19.1	32.5
3	31	13	29732994	42438349		7680	[redacted]	19.6	27.2
3	31	13	30079484	42696870		7552	[redacted]	18.6	26.2
3	31	13	29732994	42438349		7680	[redacted]	19.6	26.2
3	31	13	29732994	42438349		7680	[redacted]	19.6	19.6
3	31	13	29732994	42438349		7680	[redacted]	19.6	19.6

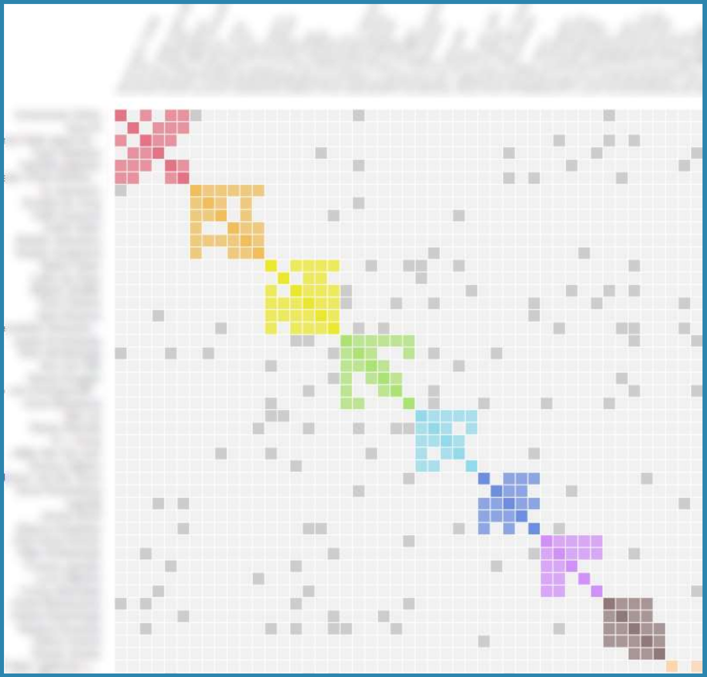
AutoSegment example

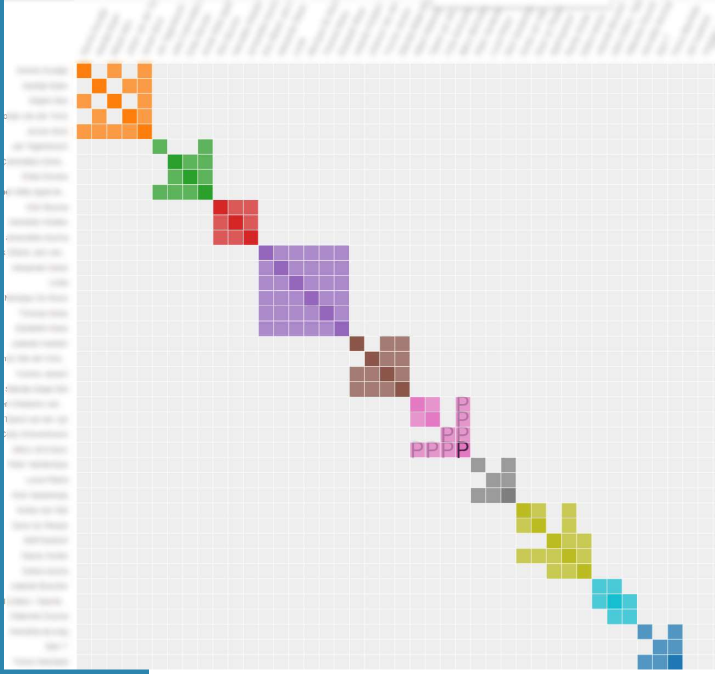


Chromosome browser

- Use MyHeritage chromosome browser to check if they are on the same side
- Segments triangulate!

ICW vs shared segments



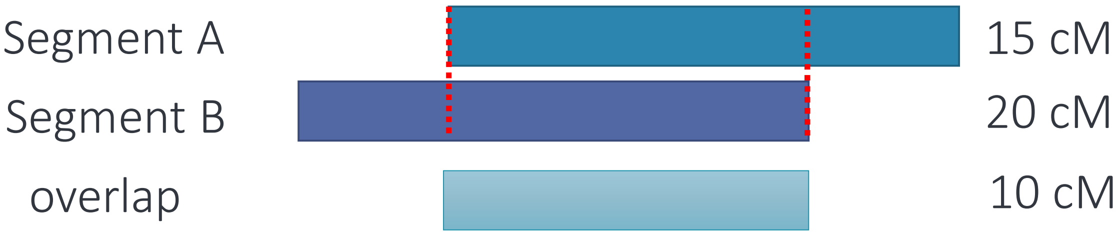


- Example of MyHeritage AutoCluster (ICW) vs AutoSegment result.
- AutoCluster DNA matches in a cluster do not always share the same segment
- AutoSegment DNA matches in a cluster based on local segment files (can) share one or more segment(s)

AutoSegment concepts

1. Import matches
2. Import segments
3. Compare segments

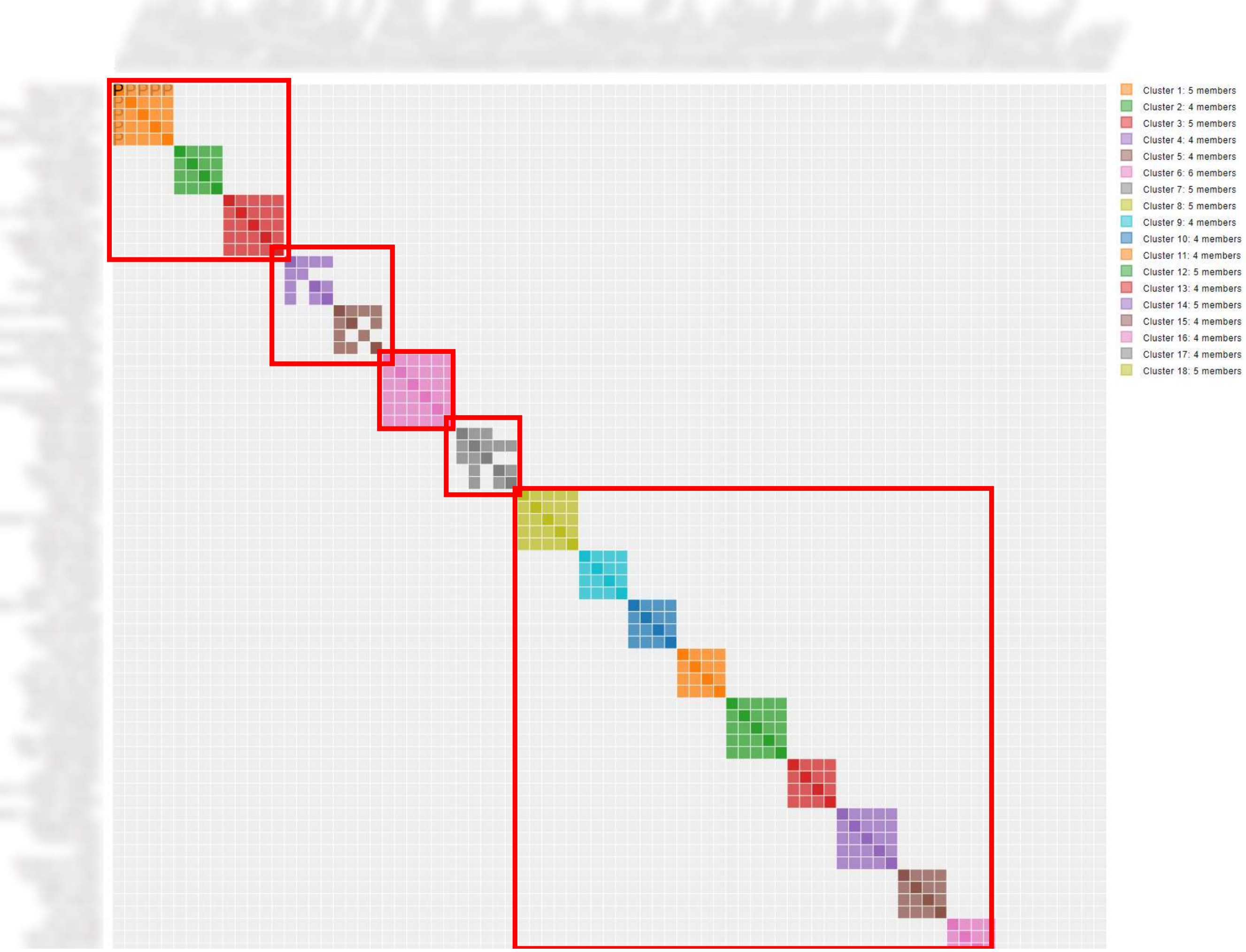
- User defines the cM range and the min segment overlap (e.g., matches between 400 cM – 20 cM and segments with at least 10 cM overlap)
- Discard matches outside range
- Discard segments smaller than min segment overlap



Segment A	15 cM
Segment B	20 cM
overlap	10 cM

- calculate cM of the overlap of two segments based on human genetic map (build37)

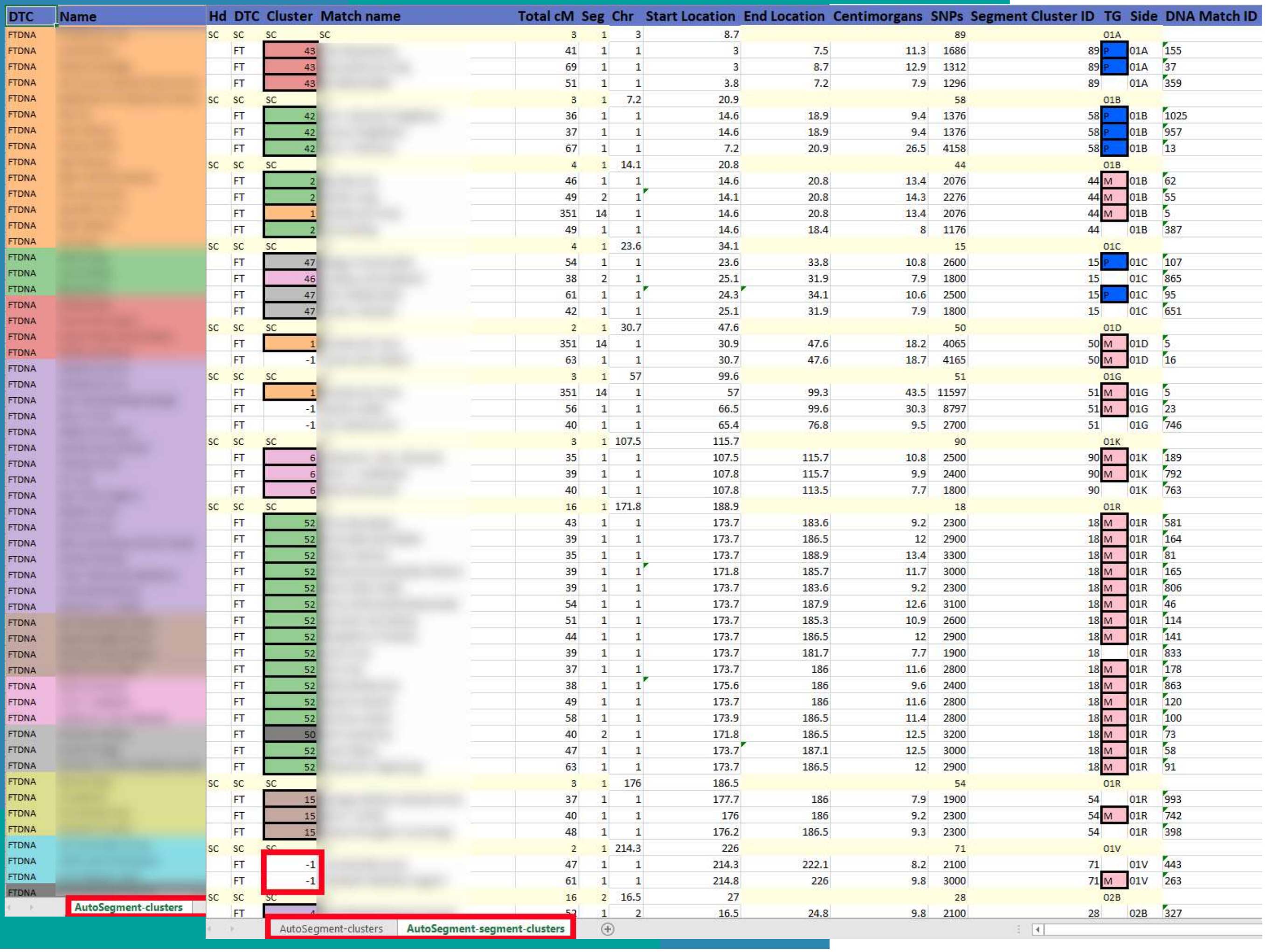
AutoSegment main chart



- Fully connected clusters usually represent a single shared segment

- Not fully connected clusters usually represent multiple shared segments

Excel representation



DTC	Name
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]
FTDNA	[redacted]

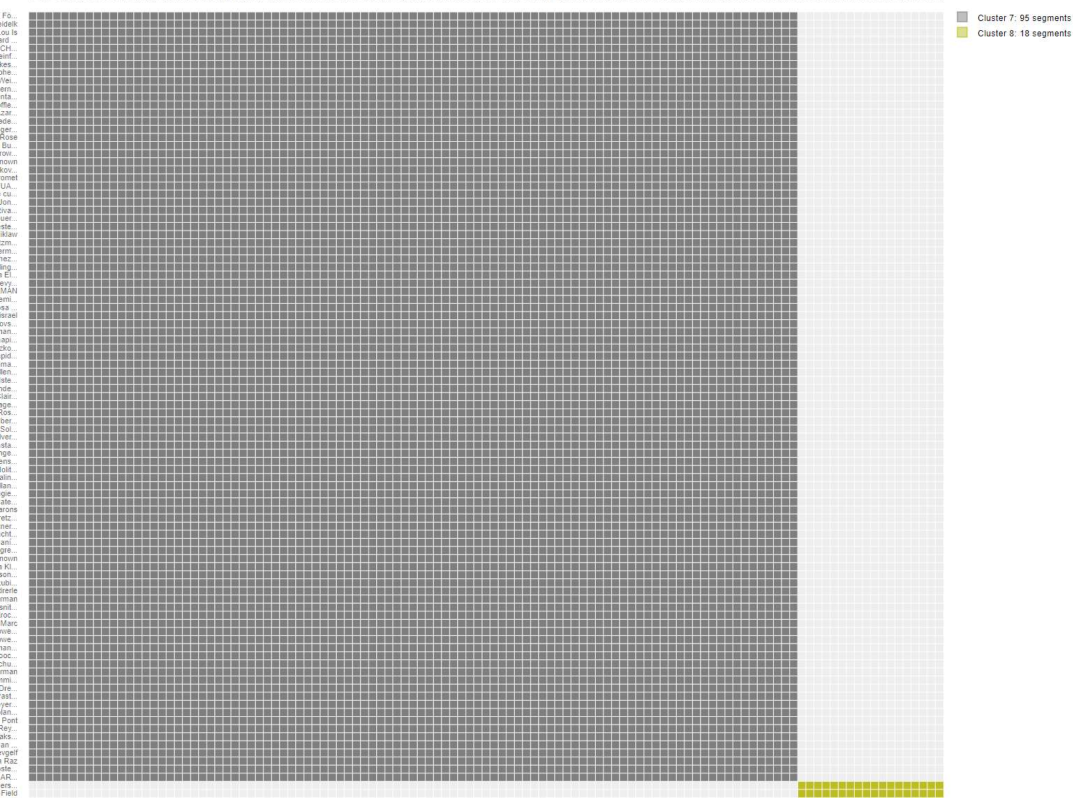
AutoSegment-clusters

Hd	DTC	Cluster	Match name	Total cM	Seg	Chr	Start Location	End Location	Centimorgans	SNPs	Segment Cluster ID	TG	Side	DNA Match ID
SC	SC	SC	SC	3	1	3	8.7			89		01A		
	FT	43	[redacted]	41	1	1	3	7.5	11.3	1686	89	P	01A	155
	FT	43	[redacted]	69	1	1	3	8.7	12.9	1312	89	P	01A	37
	FT	43	[redacted]	51	1	1	3.8	7.2	7.9	1296	89		01A	359
SC	SC	SC		3	1	7.2	20.9			58		01B		
	FT	42	[redacted]	36	1	1	14.6	18.9	9.4	1376	58	P	01B	1025
	FT	42	[redacted]	37	1	1	14.6	18.9	9.4	1376	58	P	01B	957
	FT	42	[redacted]	67	1	1	7.2	20.9	26.5	4158	58	P	01B	13
SC	SC	SC		4	1	14.1	20.8			44		01B		
	FT	2	[redacted]	46	1	1	14.6	20.8	13.4	2076	44	M	01B	62
	FT	2	[redacted]	49	2	1	14.1	20.8	14.3	2276	44	M	01B	55
	FT	1	[redacted]	351	14	1	14.6	20.8	13.4	2076	44	M	01B	5
	FT	2	[redacted]	49	1	1	14.6	18.4	8	1176	44		01B	387
SC	SC	SC		4	1	23.6	34.1			15		01C		
	FT	47	[redacted]	54	1	1	23.6	33.8	10.8	2600	15	P	01C	107
	FT	46	[redacted]	38	2	1	25.1	31.9	7.9	1800	15		01C	865
	FT	47	[redacted]	61	1	1	24.3	34.1	10.6	2500	15	P	01C	95
	FT	47	[redacted]	42	1	1	25.1	31.9	7.9	1800	15		01C	651
SC	SC	SC		2	1	30.7	47.6			50		01D		
	FT	1	[redacted]	351	14	1	30.9	47.6	18.2	4065	50	M	01D	5
	FT	-1	[redacted]	63	1	1	30.7	47.6	18.7	4165	50	M	01D	16
SC	SC	SC		3	1	57	99.6			51		01G		
	FT	1	[redacted]	351	14	1	57	99.3	43.5	11597	51	M	01G	5
	FT	-1	[redacted]	56	1	1	66.5	99.6	30.3	8797	51	M	01G	23
	FT	-1	[redacted]	40	1	1	65.4	76.8	9.5	2700	51		01G	746
SC	SC	SC		3	1	107.5	115.7			90		01K		
	FT	6	[redacted]	35	1	1	107.5	115.7	10.8	2500	90	M	01K	189
	FT	6	[redacted]	39	1	1	107.8	115.7	9.9	2400	90	M	01K	792
	FT	6	[redacted]	40	1	1	107.8	113.5	7.7	1800	90		01K	763
SC	SC	SC		16	1	171.8	188.9			18		01R		
	FT	52	[redacted]	43	1	1	173.7	183.6	9.2	2300	18	M	01R	581
	FT	52	[redacted]	39	1	1	173.7	186.5	12	2900	18	M	01R	164
	FT	52	[redacted]	35	1	1	173.7	188.9	13.4	3300	18	M	01R	81
	FT	52	[redacted]	39	1	1	171.8	185.7	11.7	3000	18	M	01R	165
	FT	52	[redacted]	39	1	1	173.7	183.6	9.2	2300	18	M	01R	806
	FT	52	[redacted]	54	1	1	173.7	187.9	12.6	3100	18	M	01R	46
	FT	52	[redacted]	51	1	1	173.7	185.3	10.9	2600	18	M	01R	114
	FT	52	[redacted]	44	1	1	173.7	186.5	12	2900	18	M	01R	141
	FT	52	[redacted]	39	1	1	173.7	181.7	7.7	1900	18		01R	833
	FT	52	[redacted]	37	1	1	173.7	186	11.6	2800	18	M	01R	178
	FT	52	[redacted]	38	1	1	175.6	186	9.6	2400	18	M	01R	863
	FT	52	[redacted]	49	1	1	173.7	186	11.6	2800	18	M	01R	120
	FT	52	[redacted]	58	1	1	173.9	186.5	11.4	2800	18	M	01R	100
	FT	50	[redacted]	40	2	1	171.8	186.5	12.5	3200	18	M	01R	73
	FT	52	[redacted]	47	1	1	173.7	187.1	12.5	3000	18	M	01R	58
	FT	52	[redacted]	63	1	1	173.7	186.5	12	2900	18	M	01R	91
SC	SC	SC		3	1	176	186.5			54		01R		
	FT	15	[redacted]	37	1	1	177.7	186	7.9	1900	54		01R	993
	FT	15	[redacted]	40	1	1	176	186	9.2	2300	54	M	01R	742
	FT	15	[redacted]	48	1	1	176.2	186.5	9.3	2300	54		01R	398
SC	SC	SC		2	1	214.3	226			71		01V		
	FT	-1	[redacted]	47	1	1	214.3	222.1	8.2	2100	71		01V	443
	FT	-1	[redacted]	61	1	1	214.8	226	9.8	3000	71	M	01V	263
SC	SC	SC		16	2	16.5	27			28		02B		
	FT	4	[redacted]	52	1	2	16.5	24.8	9.8	2100	28		02B	327

AutoSegment-clusters | AutoSegment-segment-clusters

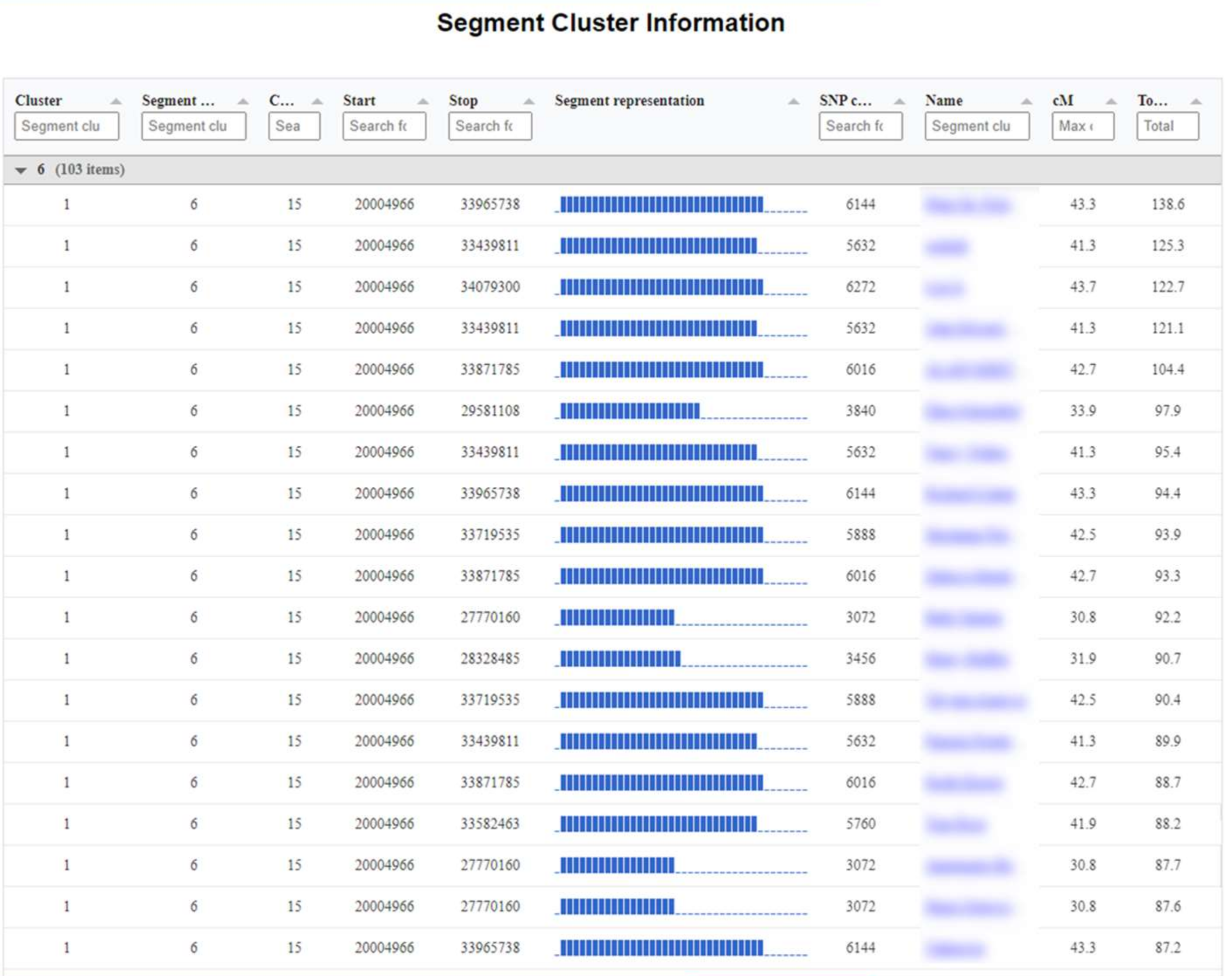
- Large HTML charts difficult to open in browsers
- Use the Excel file to view the match clusters and segment clusters
- Cluster colors are represented in the first columns
- Segment cluster sheet format inspired by segmentology blog: https://segmentology.org/

Caveats of AutoSegment



- Easy to pick up pile-ups
- Example: segment cluster 7 has 95 segments

Caveats of AutoSegment

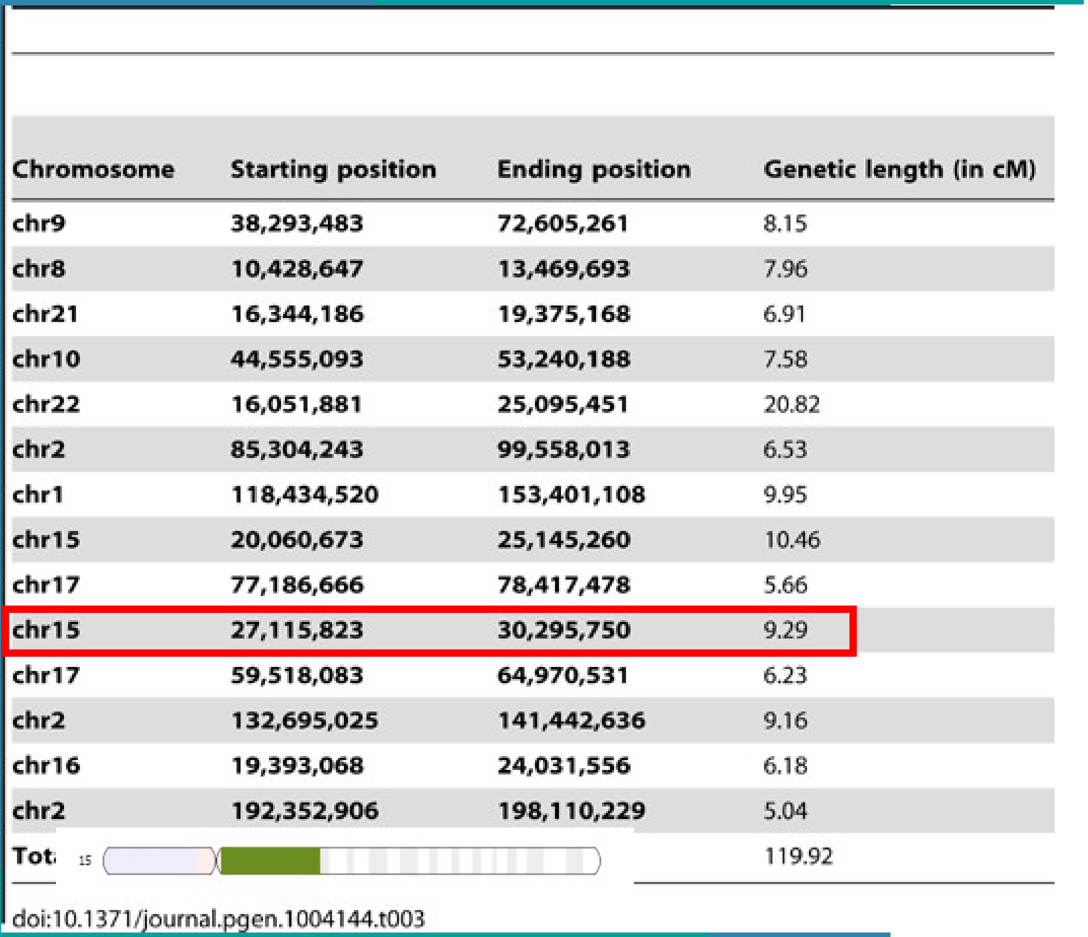


Segment Cluster Information

Cluster	Segment ...	C...	Start	Stop	Segment representation	SNP c...	Name	cM	To...
Segment clu	Segment clu	Sea	Search fc	Search fc		Search fc	Segment clu	Max (	Total
6 (103 items)									
1	6	15	20004966	33965738		6144	[redacted]	43.3	138.6
1	6	15	20004966	33439811		5632	[redacted]	41.3	125.3
1	6	15	20004966	34079300		6272	[redacted]	43.7	122.7
1	6	15	20004966	33439811		5632	[redacted]	41.3	121.1
1	6	15	20004966	33871785		6016	[redacted]	42.7	104.4
1	6	15	20004966	29581108		3840	[redacted]	33.9	97.9
1	6	15	20004966	33439811		5632	[redacted]	41.3	95.4
1	6	15	20004966	33965738		6144	[redacted]	43.3	94.4
1	6	15	20004966	33719535		5888	[redacted]	42.5	93.9
1	6	15	20004966	33871785		6016	[redacted]	42.7	93.3
1	6	15	20004966	27770160		3072	[redacted]	30.8	92.2
1	6	15	20004966	28328485		3456	[redacted]	31.9	90.7
1	6	15	20004966	33719535		5888	[redacted]	42.5	90.4
1	6	15	20004966	33439811		5632	[redacted]	41.3	89.9
1	6	15	20004966	33871785		6016	[redacted]	42.7	88.7
1	6	15	20004966	33582463		5760	[redacted]	41.9	88.2
1	6	15	20004966	27770160		3072	[redacted]	30.8	87.7
1	6	15	20004966	27770160		3072	[redacted]	30.8	87.6
1	6	15	20004966	33965738		6144	[redacted]	43.3	87.2

- Easy to pick up pile-ups
- Example: segment cluster 7 has 95 segments
- All on chromosome 15

Caveats of AutoSegment



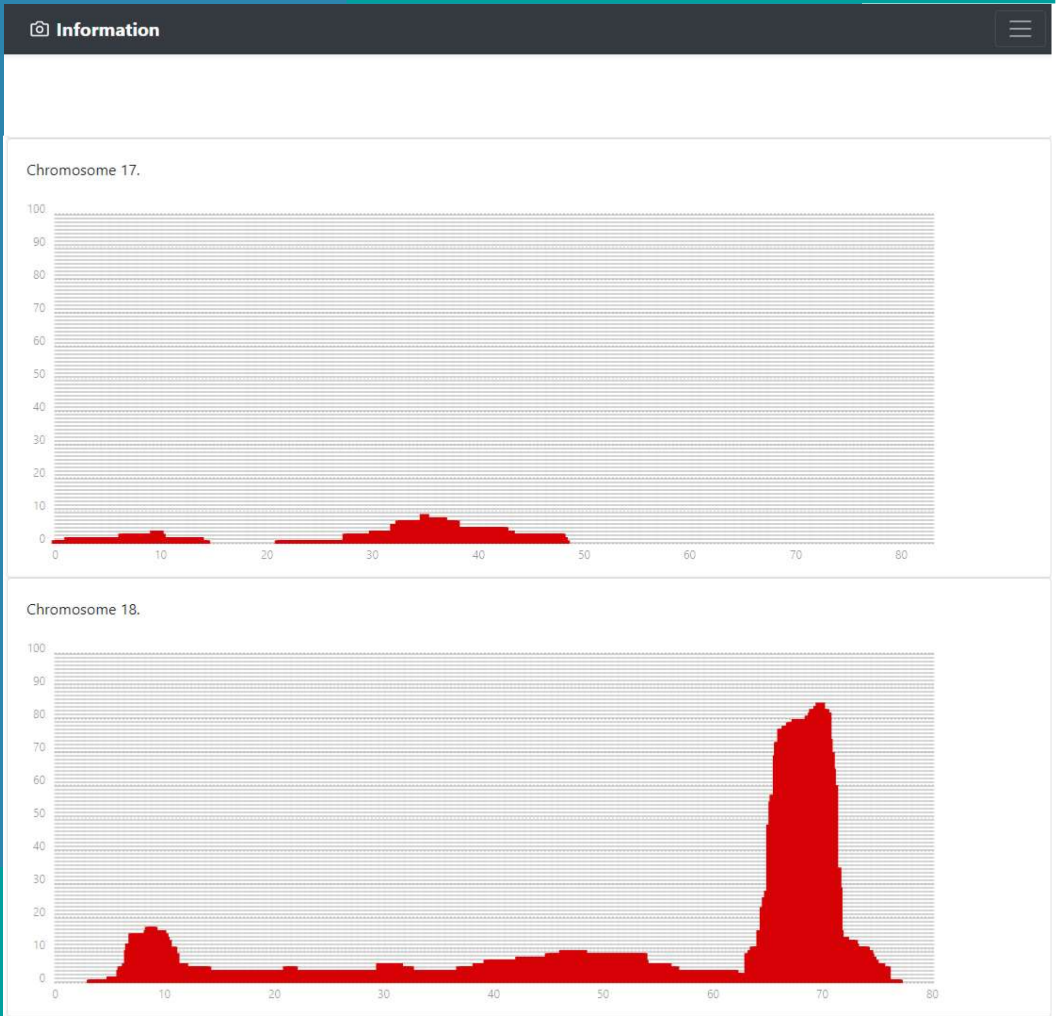
Chromosome	Starting position	Ending position	Genetic length (in cM)
chr9	38,293,483	72,605,261	8.15
chr8	10,428,647	13,469,693	7.96
chr21	16,344,186	19,375,168	6.91
chr10	44,555,093	53,240,188	7.58
chr22	16,051,881	25,095,451	20.82
chr2	85,304,243	99,558,013	6.53
chr1	118,434,520	153,401,108	9.95
chr15	20,060,673	25,145,260	10.46
chr17	77,186,666	78,417,478	5.66
chr15	27,115,823	30,295,750	9.29
chr17	59,518,083	64,970,531	6.23
chr2	132,695,025	141,442,636	9.16
chr16	19,393,068	24,031,556	6.18
chr2	192,352,906	198,110,229	5.04
Tot			119.92

doi:10.1371/journal.pgen.1004144.t003

Table from Li *et al* 2014

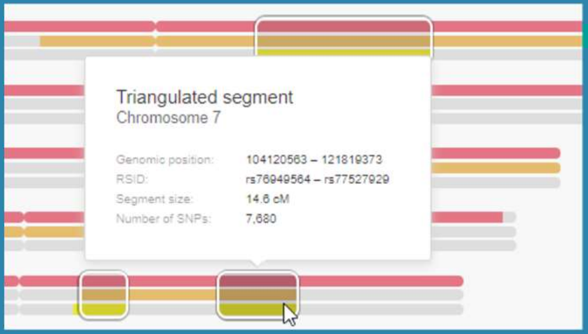
- Easy to pick up pile-ups
- Example: segment cluster 7 has 95 segments
- All on chromosome 15
- Overlaps with pile ups from Li *et al* 2014.

Caveats of AutoSegment



- In addition to known pile up regions
- Prone to picking up "personal pileups"
- Visualization of pile ups for analysis
  - Using the min segment size

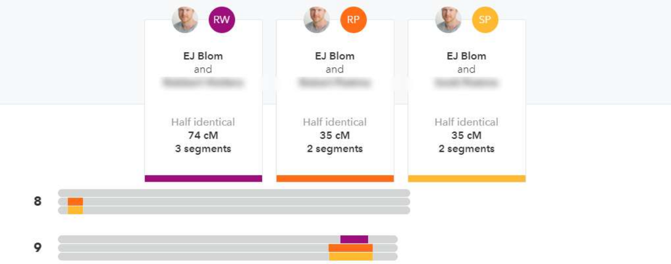
Caveats of AutoSegment



Triangulated segment
Chromosome 7

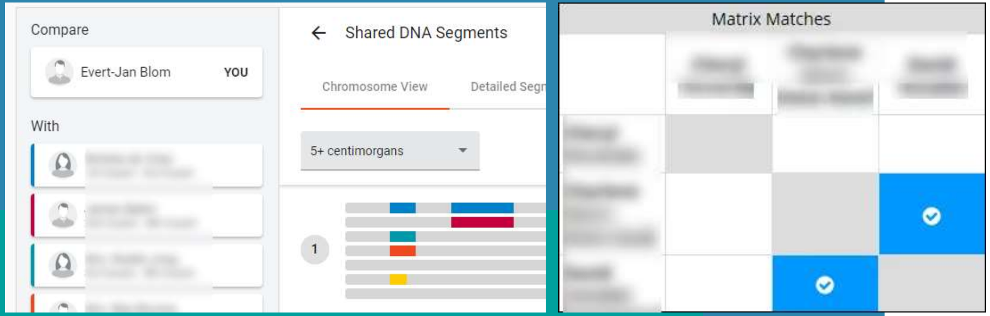
Genomic position: 104120563 – 121819373
RSID: rs76949564 – rs77527929
Segment size: 14.6 cM
Number of SNPs: 7,680

MyHeritage chromosome browser



EJ Blom and [redacted]	EJ Blom and [redacted]	EJ Blom and [redacted]
Half identical 74 cM 3 segments	Half identical 35 cM 2 segments	Half identical 35 cM 2 segments

23andme chromosome browser



Compare
Evert-Jan Blom YOU
With
[redacted]

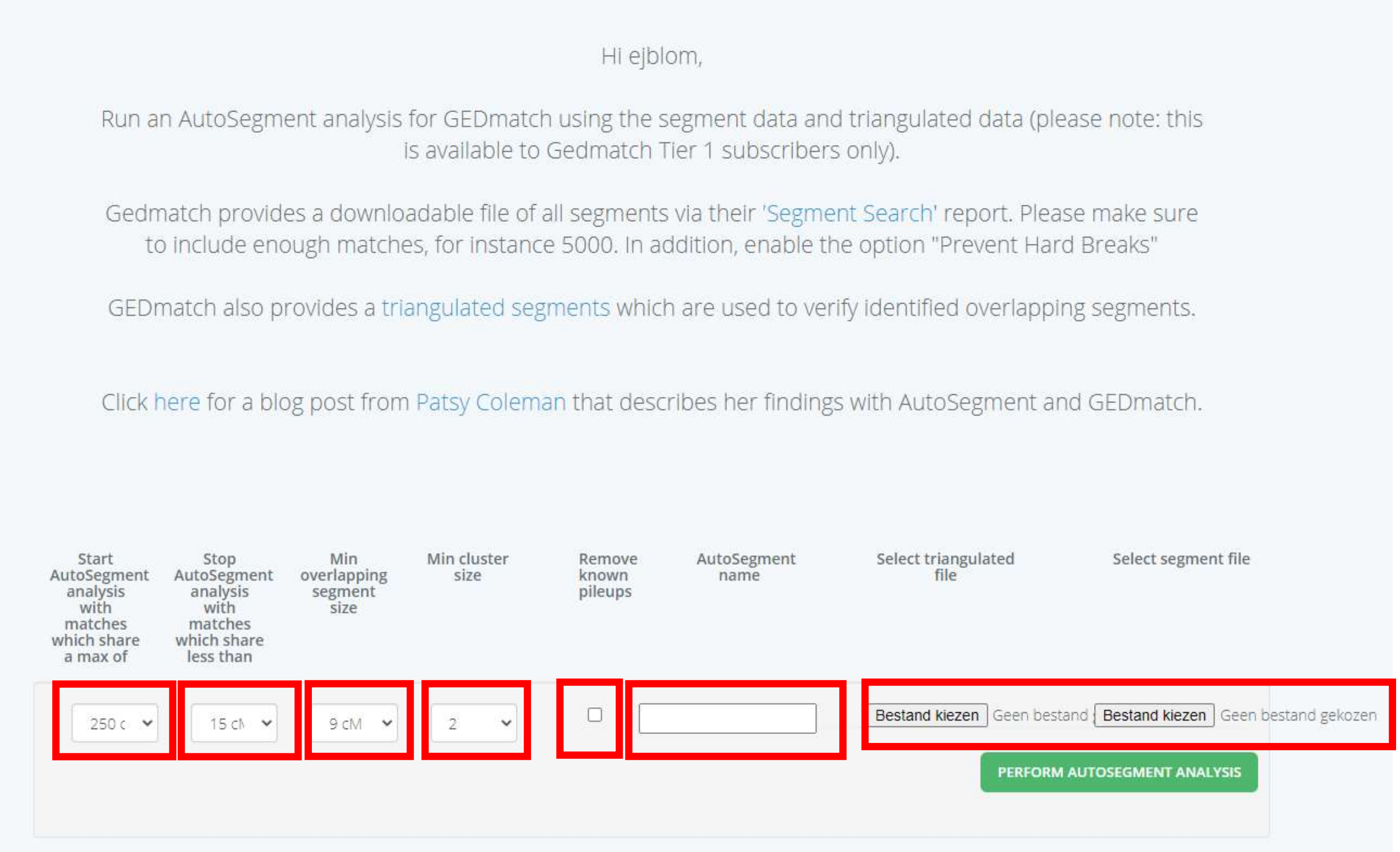
Shared DNA Segments
Chromosome View
Detailed Segm
5+ centimorgans

FTDNA chrom browser & autosomal matrix

- Segment clustering based on "flat" segment files
  - No method to infer triangulation based on flat files
- Important to check if overlapping segments from clusters triangulate!
- GEDmatch is the exception
  - Triangulation files used

Start AutoSegment

https://members.geneticaffairs.com/autosegment



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"

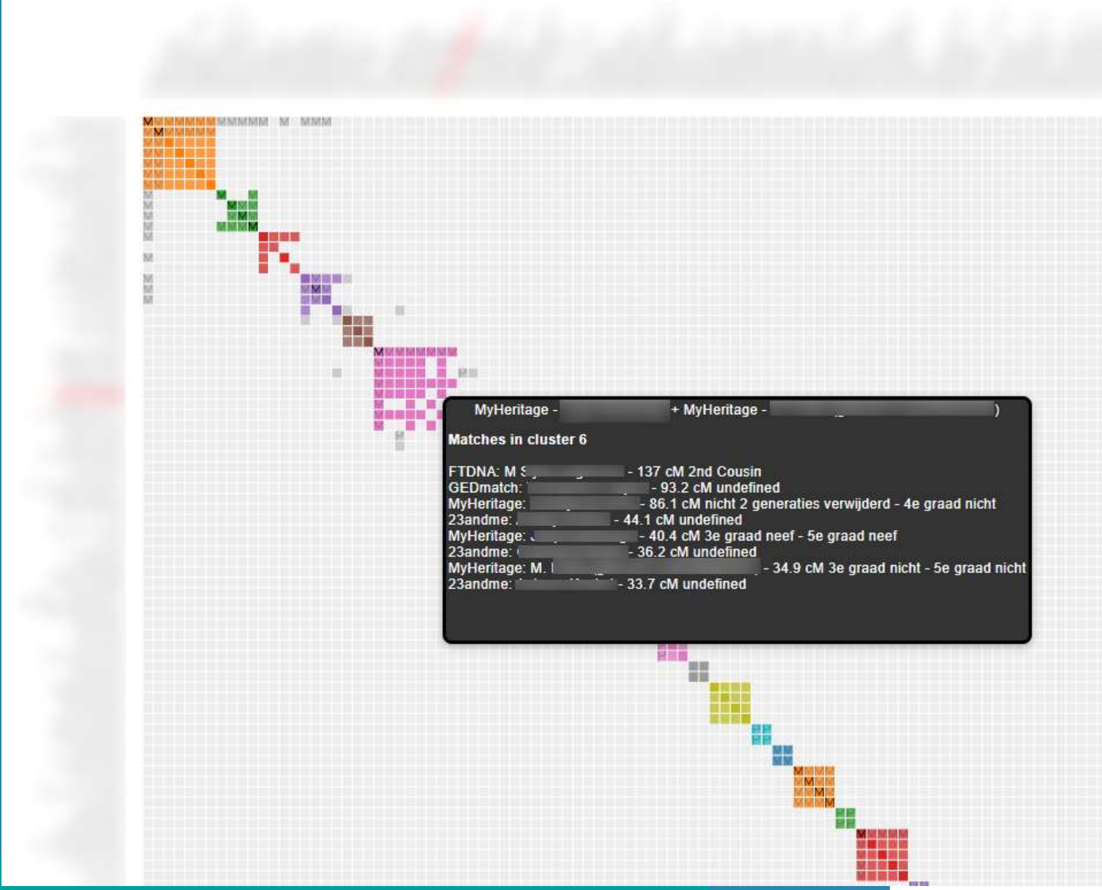
GEDmatch also provides a triangulated segments which are used to verify identified overlapping segments.

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 cN	9 cM	2	☐		Bestand kiezen Geen bestand	Bestand kiezen Geen bestand gekozen

- Max cM
- Min cM
- Min overlapping segment size
- Min cluster size
- Remove known pileups
- Name
- Select files

Hybrid AutoSegment



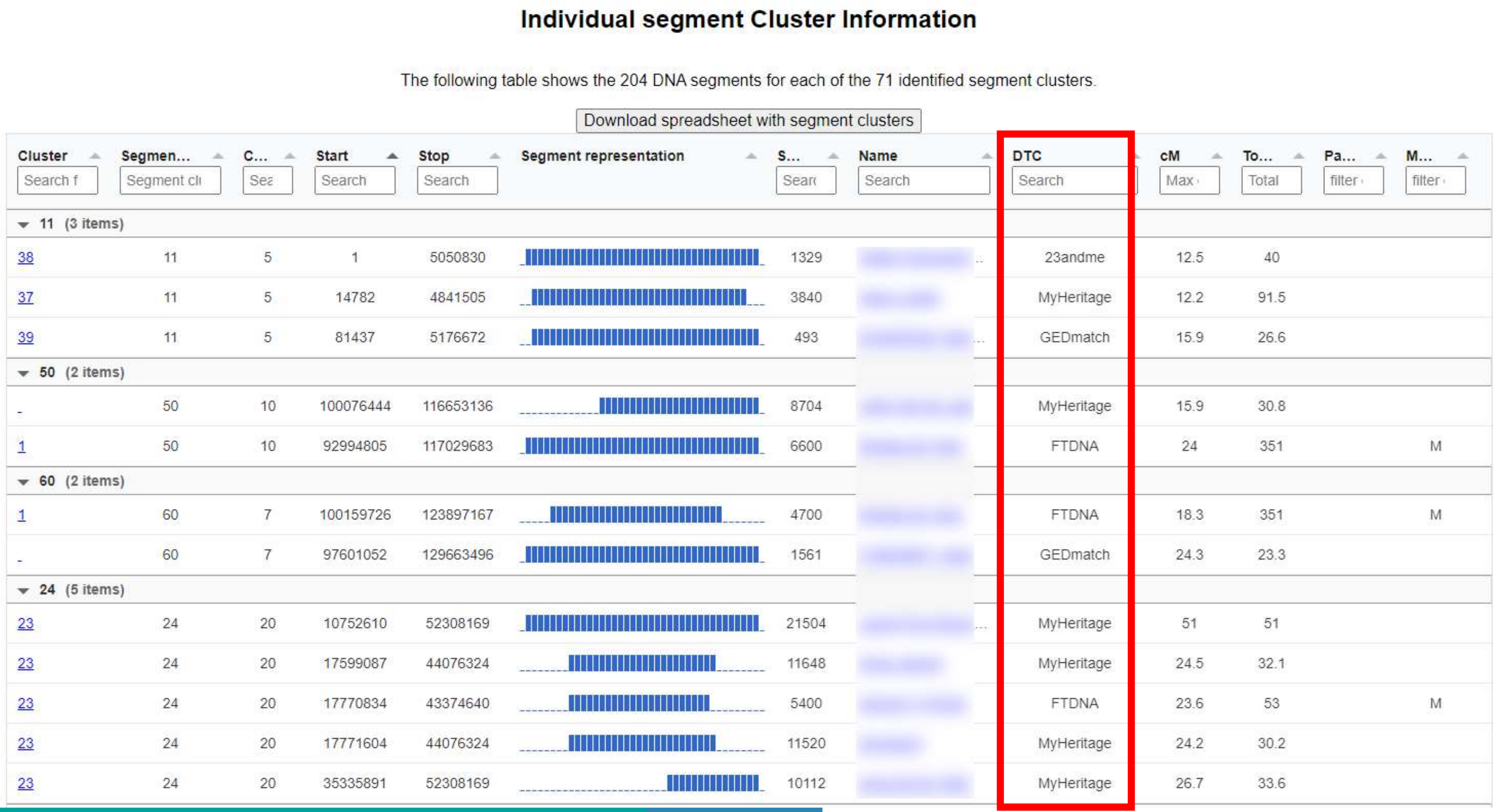
MyHeritage - [redacted] + MyHeritage - [redacted])

Matches in cluster 6

FTDNA: M [redacted] - 137 cM 2nd Cousin
GEDmatch: [redacted] - 93.2 cM undefined
MyHeritage: [redacted] - 86.1 cM nicht 2 generaties verwijderd - 4e graad nicht
23andme: [redacted] - 44.1 cM undefined
MyHeritage: [redacted] - 40.4 cM 3e graad neef - 5e graad neef
23andme: [redacted] - 36.2 cM undefined
MyHeritage: M. [redacted] - 34.9 cM 3e graad nicht - 5e graad nicht
23andme: [redacted] - 33.7 cM undefined

- Combines FTDNA/23andme/MyHeritage and GEDmatch into one analysis
- Overlapping segments from four different DNA match databases
- Overlapping GEDmatch segments should triangulate
- Same principles as AutoSegment

Hybrid AutoSegment



Individual segment Cluster Information

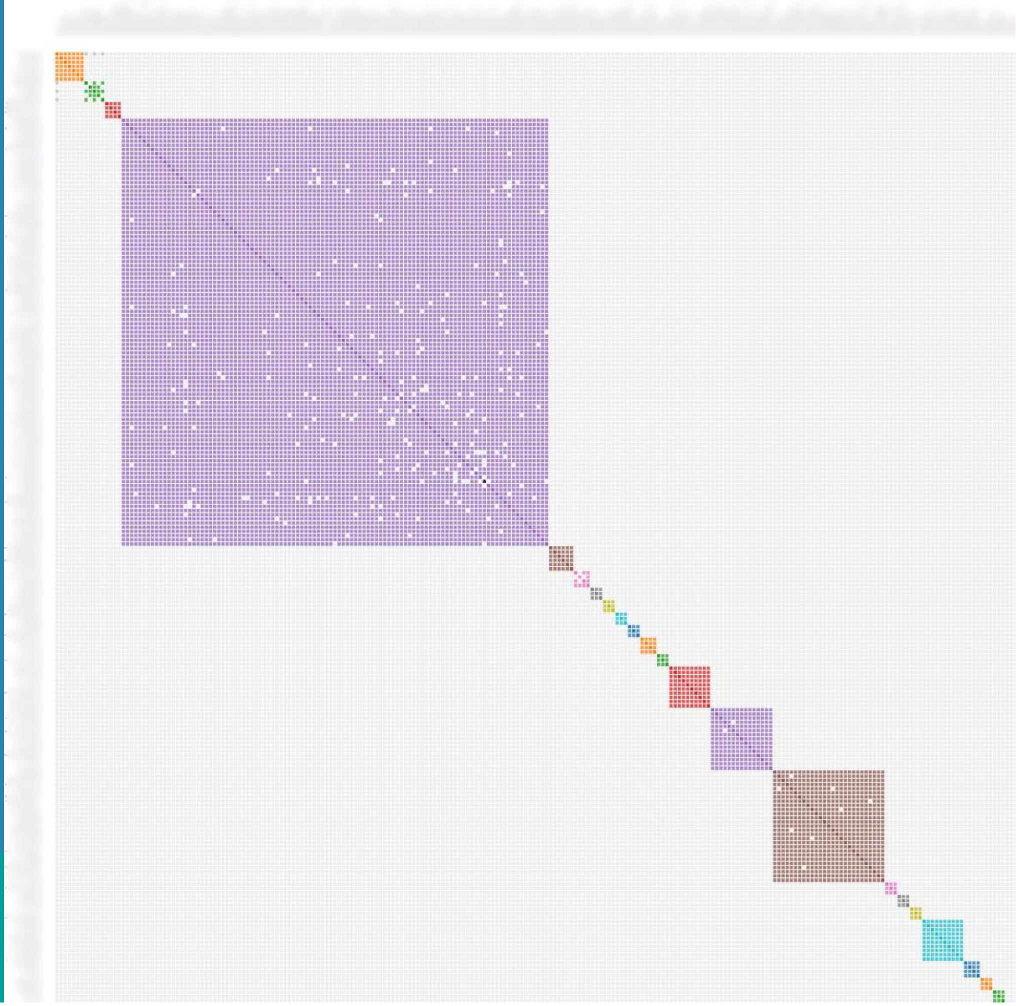
The following table shows the 204 DNA segments for each of the 71 identified segment clusters.

Download spreadsheet with segment clusters

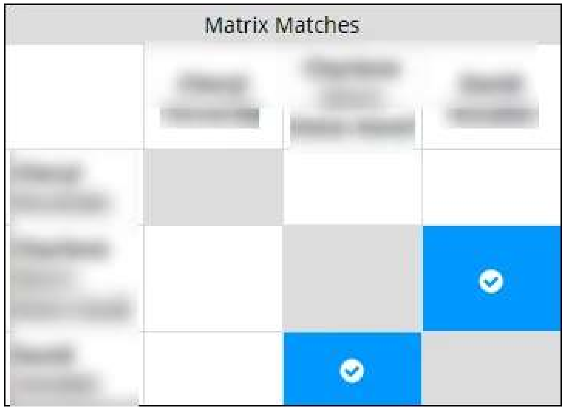
Cluster	Segmen...	C...	Start	Stop	Segment representation	S...	Name	DTC	cM	To...	Pa...	M...
Search f	Segment cl	Sea	Search	Search		Searc	Search	Search	Max	Total	filter	filter
11 (3 items)												
38	11	5	1	5050830		1329	[redacted]	23andme	12.5	40		
37	11	5	14782	4841505		3840	[redacted]	MyHeritage	12.2	91.5		
39	11	5	81437	5176672		493	[redacted]	GEDmatch	15.9	26.6		
50 (2 items)												
-	50	10	100076444	116653136		8704	[redacted]	MyHeritage	15.9	30.8		
1	50	10	92994805	117029683		6600	[redacted]	FTDNA	24	351		M
60 (2 items)												
1	60	7	100159726	123897167		4700	[redacted]	FTDNA	18.3	351		M
-	60	7	97601052	129663496		1561	[redacted]	GEDmatch	24.3	23.3		
24 (5 items)												
23	24	20	10752610	52308169		21504	[redacted]	MyHeritage	51	51		
23	24	20	17599087	44076324		11648	[redacted]	MyHeritage	24.5	32.1		
23	24	20	17770834	43374640		5400	[redacted]	FTDNA	23.6	53		M
23	24	20	17771604	44076324		11520	[redacted]	MyHeritage	24.2	30.2		
23	24	20	35335891	52308169		10112	[redacted]	MyHeritage	26.7	33.6		

- All segment clusters are listed in large table located in the main HTML file
- Some DNA matches are not in the chart
- Missing DNA matches could still be linked to segment clusters

FTDNA AutoSegment ICW

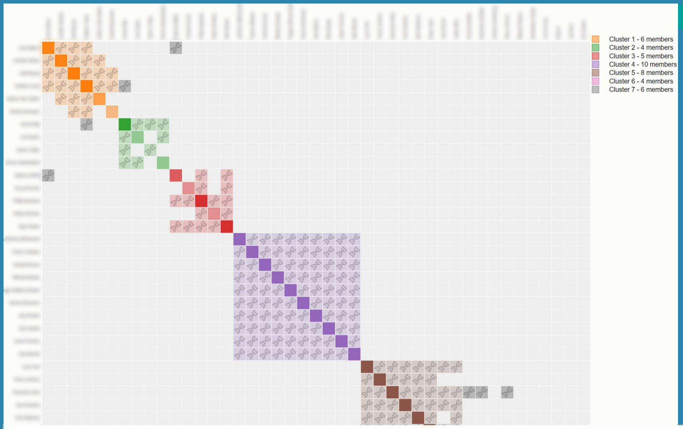


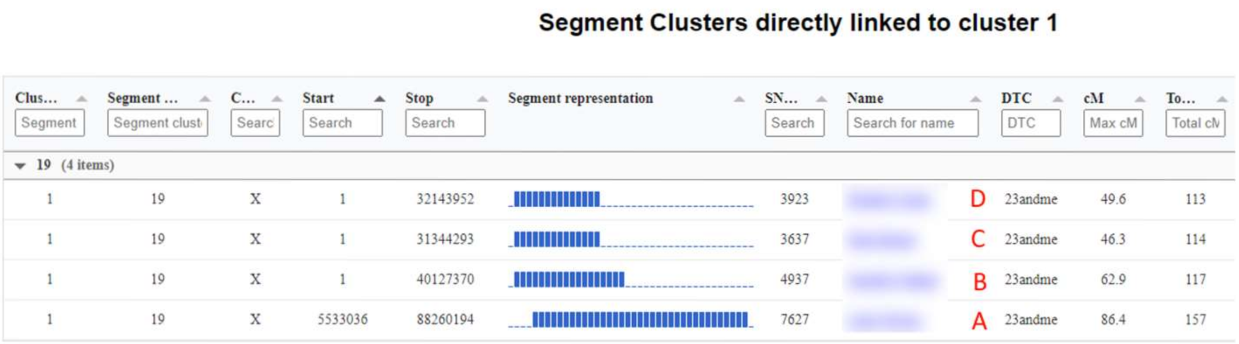
- Use shared match data to validate overlapping segments
- Only allow overlapping segments if the underlying DNA matches are a shared match



Matrix Matches

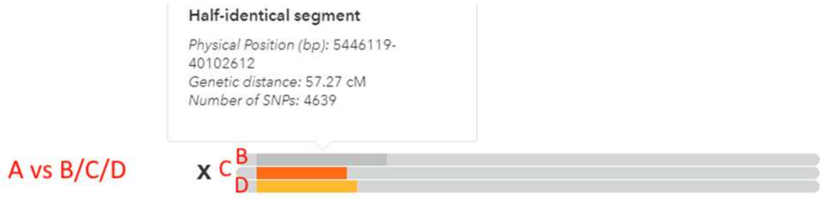
23andme AutoSegment





Segment Clusters directly linked to cluster 1

Clus…	Segment …	C…	Start	Stop	Segment representation	SN…	Name	DTC	cM	To…
Segment	Segment clust	Searc	Search	Search		Search	Search for name	DTC	Max cM	Total cN
▼ 19 (4 items)										
1	19	X	1	32143952		3923	D	23andme	49.6	113
1	19	X	1	31344293		3637	C	23andme	46.3	114
1	19	X	1	40127370		4937	B	23andme	62.9	117
1	19	X	5533036	88260194		7627	A	23andme	86.4	157



Half-identical segment

Physical Position (bp): 5446119-40102612
Genetic distance: 57.27 cM
Number of SNPs: 4639

A vs B/C/D

- Use DNA overlap feature on 23andme
- Only overlapping segments if DNA overlap is true

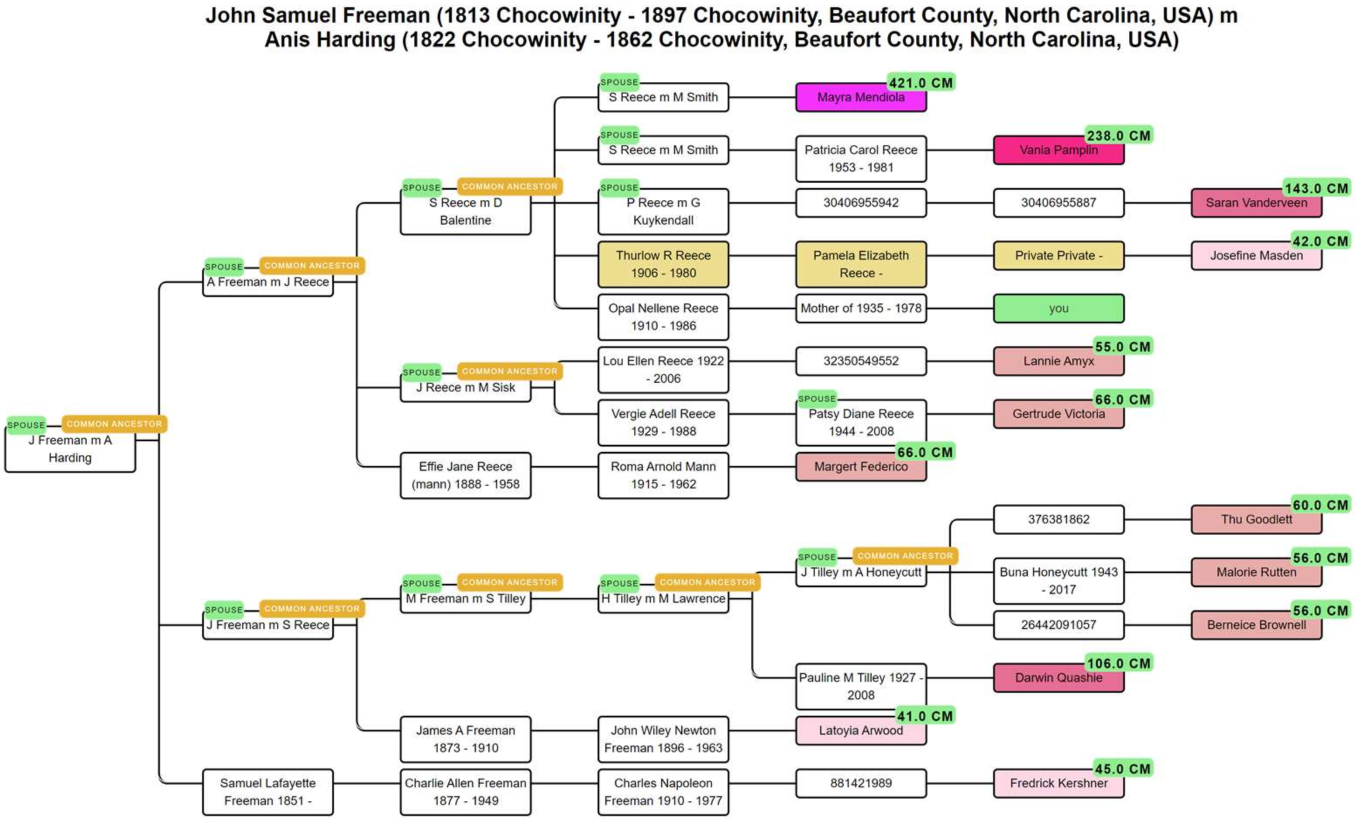
Conclusion AutoSegment

- AutoSegment groups DNA matches that share a common ancestor based on shared segments
- AutoSegment can be performed automatically for:
  - FamilyTreeDNA (using ICW data)
  - 23andme (using DNA overlap)
- AutoSegment can be performed using local files for:
  - FamilyTreeDNA, 23andme, MyHeritage and GEDmatch
- AutoSegment available for tier 1 users on GEDmatch
  - Ability to use the segments from a single chromosome

Outline AutoTree

- what is AutoTree
- AutoTree concepts
- visualization
- common ancestors
- Y-DNA and mtDNA
- starting AutoTree

AutoTree

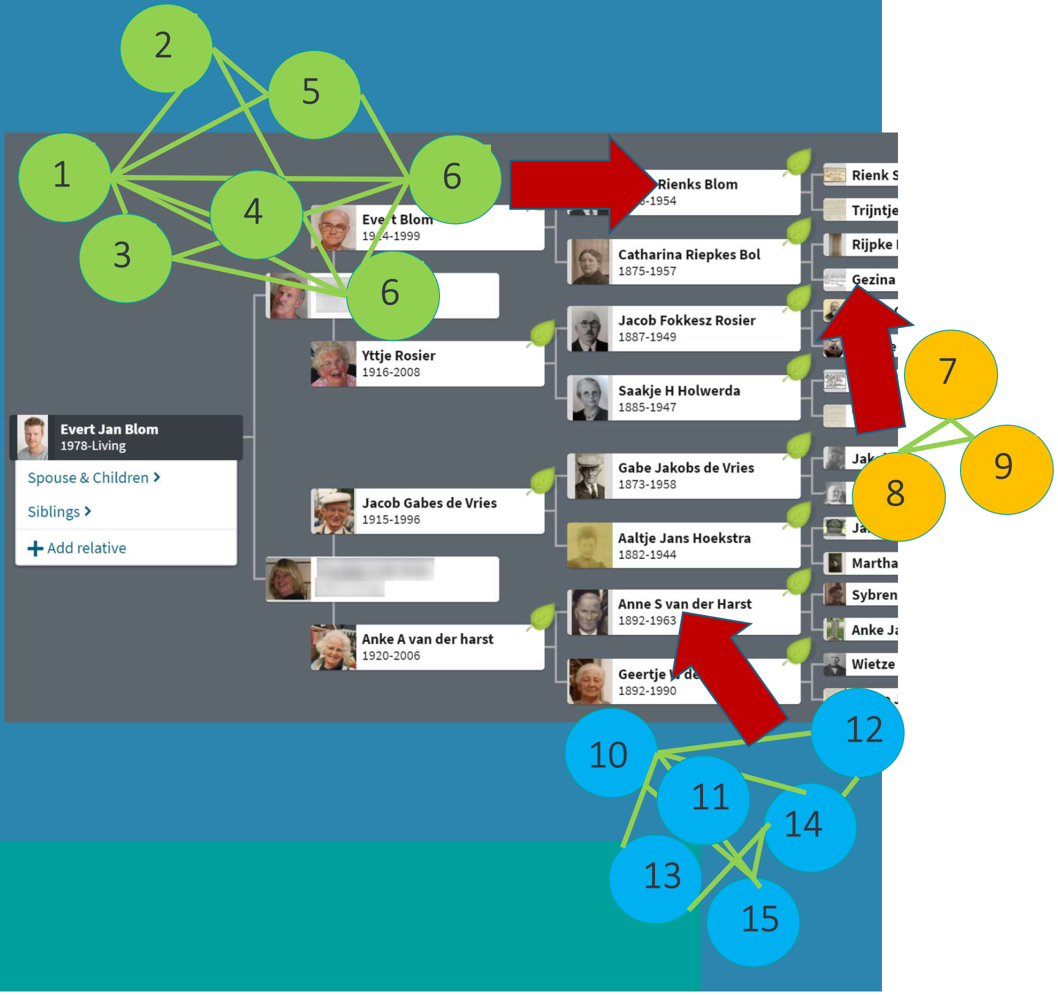

John Samuel Freeman (1813 Chocowinity - 1897 Chocowinity, Beaufort County, North Carolina, USA) m
Anis Harding (1822 Chocowinity - 1862 Chocowinity, Beaufort County, North Carolina, USA)

- AutoTree uses shared trees of DNA matches to find common ancestors (MRCAs) and can reconstruct a (partial) tree.

- Based on genealogical trees

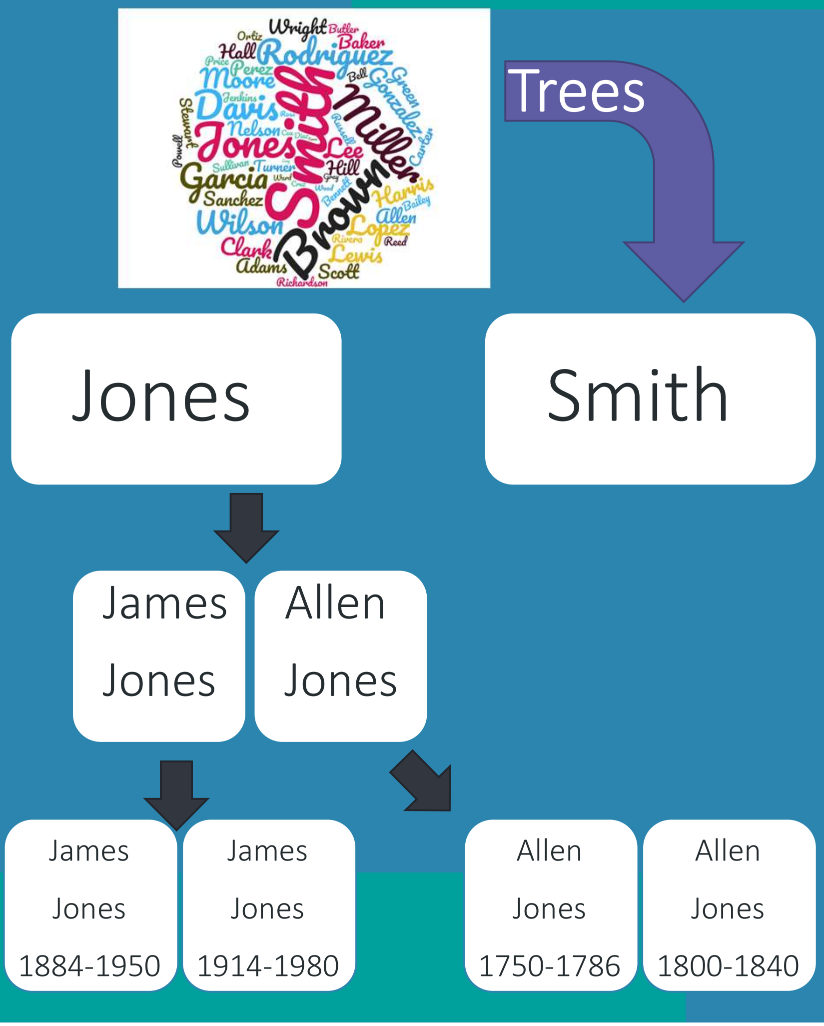
- AutoTree is available on Genetic Affairs for FTDNA profiles (and on GEDmatch for tier 1 users)

AutoTree



- AutoTree is performed using the trees of members of clusters obtained from an AutoCluster analysis.
- Reconstructed trees aid in the identification of the common ancestor of the remaining DNA matches in a cluster.
- Work for persons with unknown parentage to their birth families (for instance adoptees or donor-conceived persons) as well as people that have linked their DNA profile to a tree.

Concepts of AutoTree



Jones

Smith

James Jones

Allen Jones

James Jones 1884-1950

James Jones 1914-1980

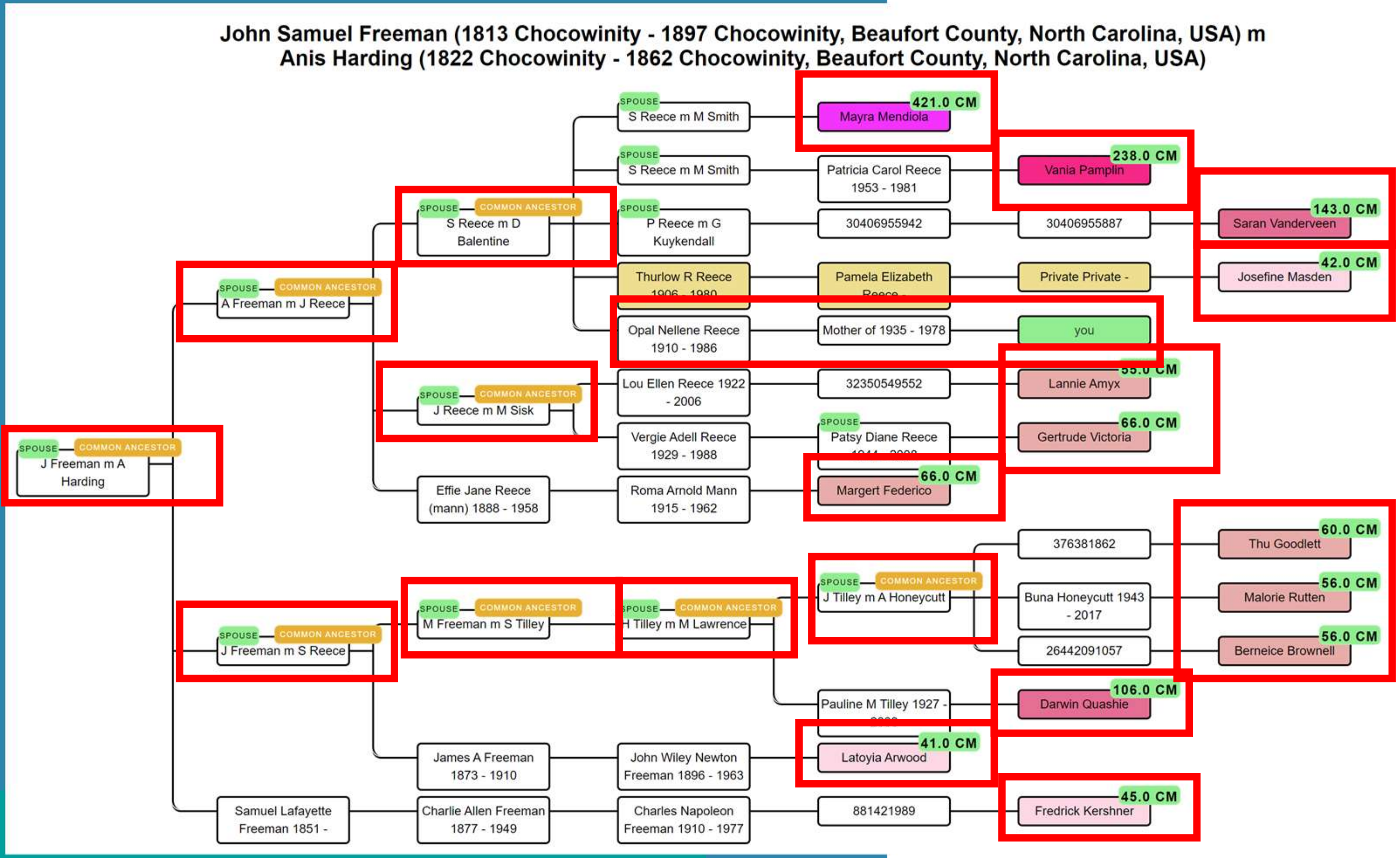
Allen Jones 1750-1786

Allen Jones 1800-1840

- Examines trees of matches to identify common ancestors and reconstruct a tree
- Employs three steps:

1. Surname clustering
2. First name clustering
3. Birth/Death year clustering

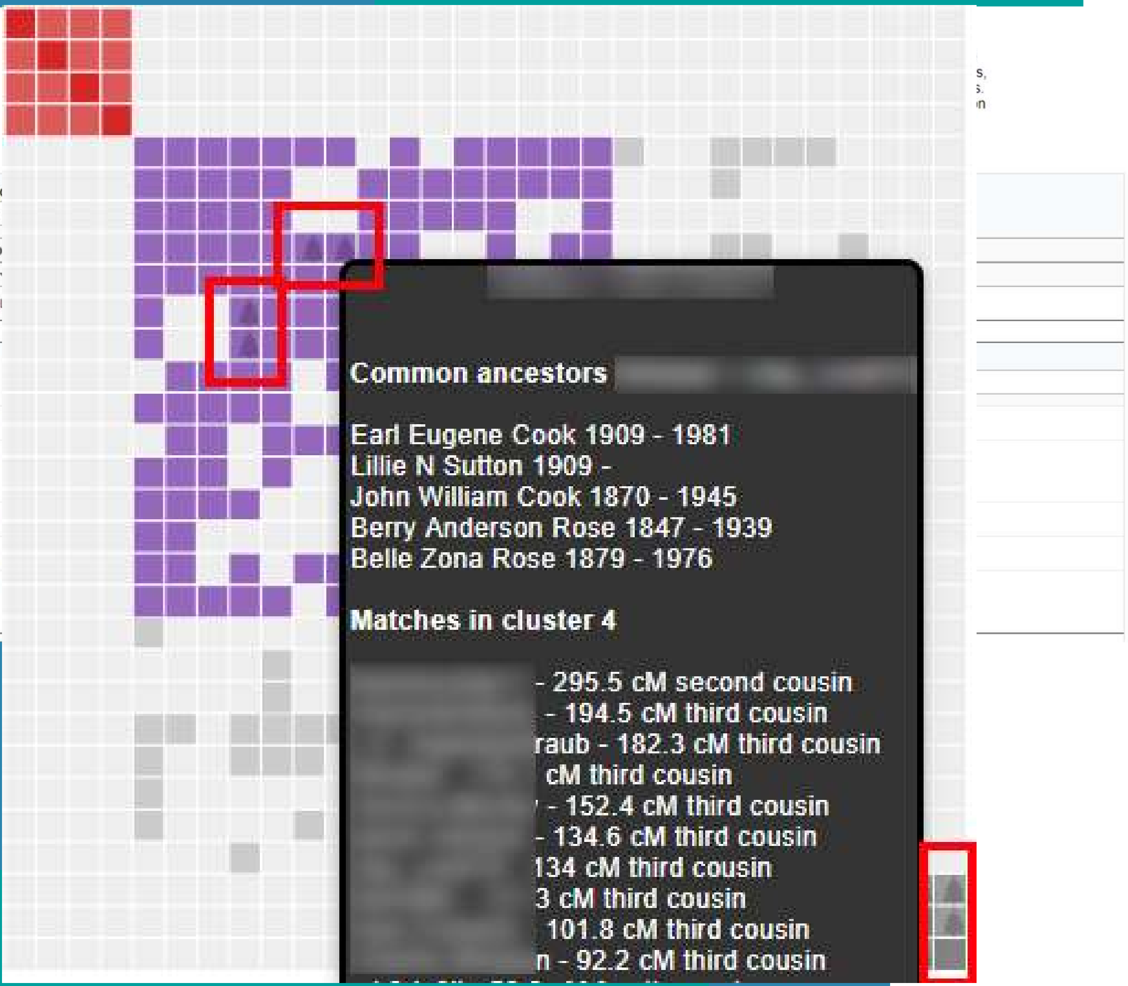
Visualization of trees



John Samuel Freeman (1813 Chocowinity - 1897 Chocowinity, Beaufort County, North Carolina, USA) m
Anis Harding (1822 Chocowinity - 1862 Chocowinity, Beaufort County, North Carolina, USA)

- Reconstructed trees are visualized for each common ancestor and linked DNA matches
- DNA matches visualized using a color gradient
- Tree persons from your own tree are also visualized (if available)

Visualization in chart



[redacted]

Common ancestors [redacted]

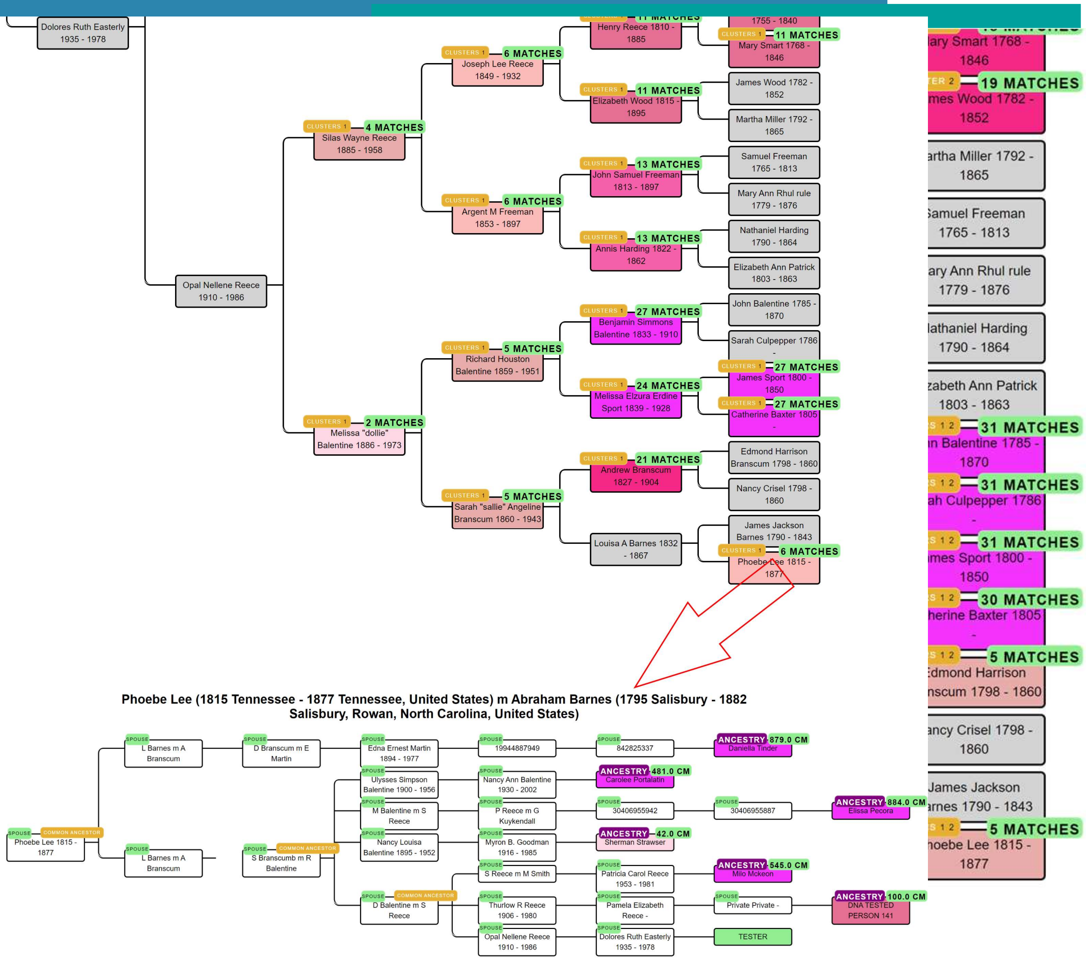
Earl Eugene Cook 1909 - 1981
Lillie N Sutton 1909 -
John William Cook 1870 - 1945
Berry Anderson Rose 1847 - 1939
Belle Zona Rose 1879 - 1976

Matches in cluster 4

[redacted]
- 295.5 cM second cousin
- 194.5 cM third cousin
raub - 182.3 cM third cousin
cM third cousin
- 152.4 cM third cousin
- 134.6 cM third cousin
134 cM third cousin
3 cM third cousin
101.8 cM third cousin
n - 92.2 cM third cousin

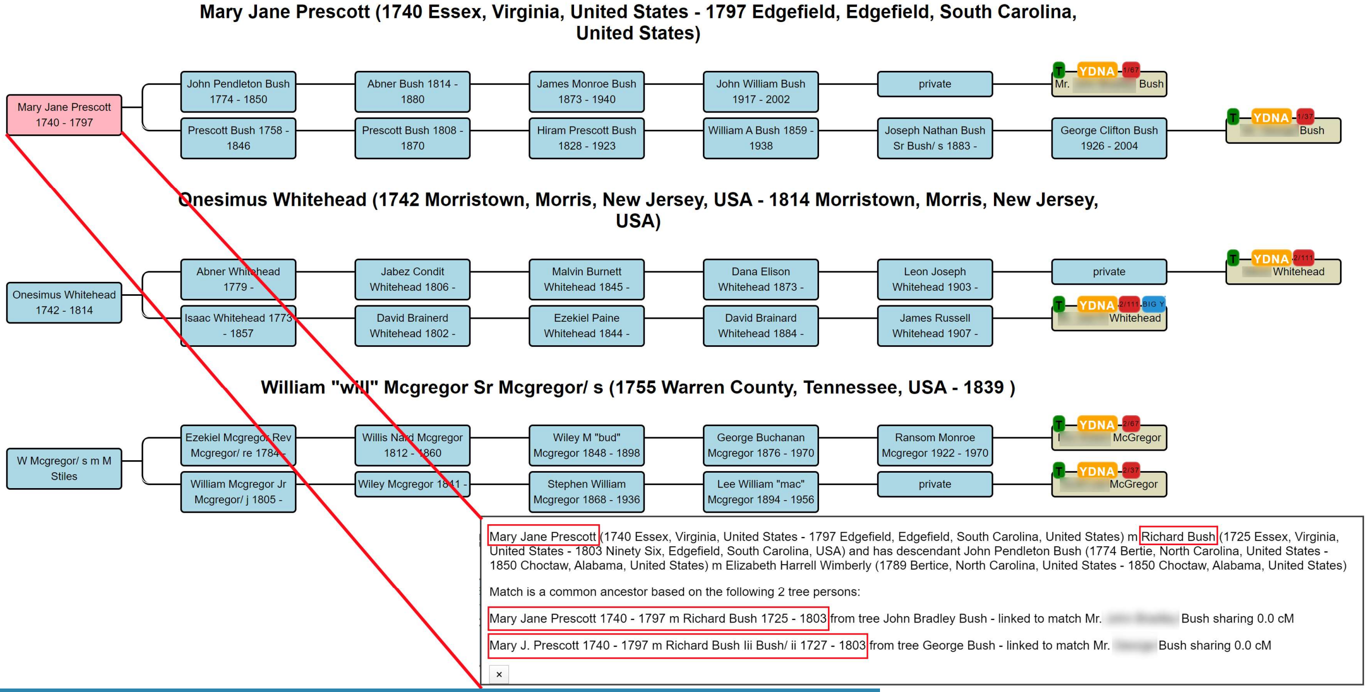
- Common locations using clustering of birth locations
- Integration within the AutoCluster chart (using Christmas trees and overlay information)

Common ancestors



- Overlay of common ancestors linked to DNA matches projected on your own tree
- Shows the clusters where the DNA matches originate from
- Clickable links for each common ancestor
- Show DNA matches that share the common ancestor

Y-DNA & mtDNA



Mary Jane Prescott (1740 Essex, Virginia, United States - 1797 Edgefield, Edgefield, South Carolina, United States)

Onesimus Whitehead (1742 Morristown, Morris, New Jersey, USA - 1814 Morristown, Morris, New Jersey, USA)

William "will" Mcgregor Sr Mcgregor/ s (1755 Warren County, Tennessee, USA - 1839)

Mary Jane Prescott (1740 Essex, Virginia, United States - 1797 Edgefield, Edgefield, South Carolina, United States) m Richard Bush (1725 Essex, Virginia, United States - 1803 Ninety Six, Edgefield, South Carolina, USA) and has descendant John Pendleton Bush (1774 Bertie, North Carolina, United States - 1850 Choctaw, Alabama, United States) m Elizabeth Harrell Wimberly (1789 Bertice, North Carolina, United States - 1850 Choctaw, Alabama, United States)

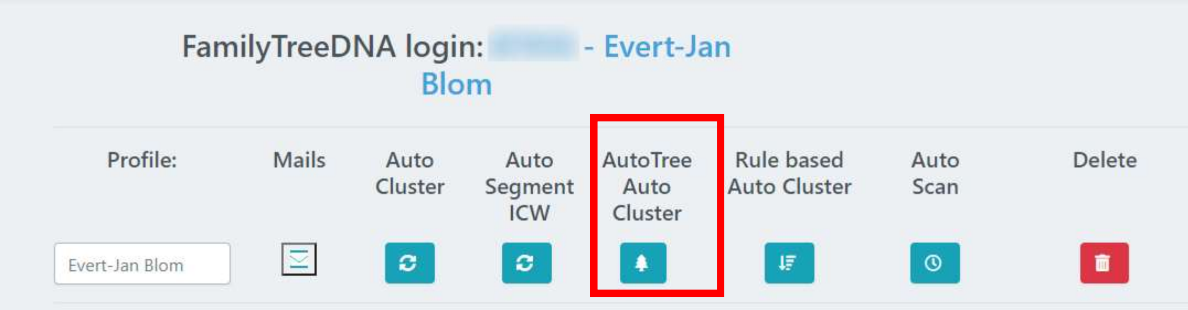
Match is a common ancestor based on the following 2 tree persons:

Mary Jane Prescott 1740 - 1797 m Richard Bush 1725 - 1803 from tree John Bradley Bush - linked to match Mr. [redacted] Bush sharing 0.0 cM

Mary J. Prescott 1740 - 1797 m Richard Bush Iii Bush/ ii 1727 - 1803 from tree George Bush - linked to match Mr. [redacted] Bush sharing 0.0 cM

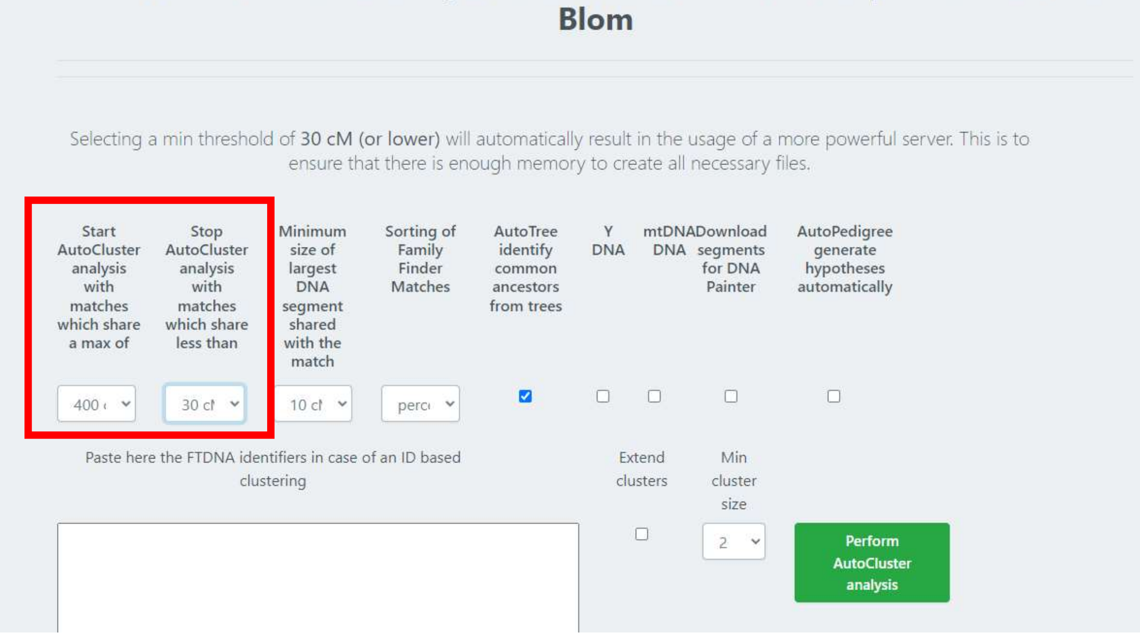
- Allows a quick analysis of shared common ancestors in trees of Y-DNA or mtDNA matches
- Comparison with trees from autosomal matches
- Male/female visualizations to improve analysis
- Available for FTDNA

Starting AutoTree



FamilyTreeDNA login: [redacted] - Evert-Jan Blom

Profile:	Mails	Auto Cluster	Auto Segment ICW	AutoTree Auto Cluster	Rule based Auto Cluster	Auto Scan	Delete
Evert-Jan Blom							



Perform AutoCluster analysis with AutoTree feature for profile **Evert-Jan Blom**

Selecting a min threshold of 30 cM (or lower) will automatically result in the usage of a more powerful server. This is to ensure that there is enough memory to create all necessary files.

Start AutoCluster analysis with matches which share a max of	Stop AutoCluster analysis with matches which share less than	Minimum size of largest DNA segment shared with the match	Sorting of Family Finder Matches	AutoTree identify common ancestors from trees	Y DNA	mtDNA DNA	Download segments for DNA Painter	AutoPedigree generate hypotheses automatically
400	30 c	10 c	perc	☑	☐	☐	☐	☐

Paste here the FTDNA identifiers in case of an ID based clustering

Extend clusters ☐

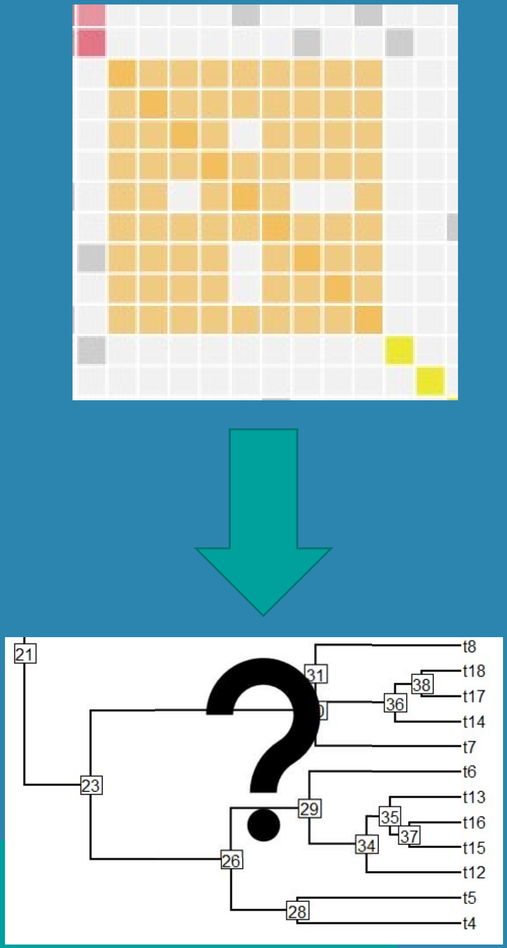
Min cluster size 2

- Available for FTDNA profiles
- Select FTDNA profile of interest
- Select AutoTree option
- Set parameters, and start analysis
- Retrieve results by e-mail or download results from site

Outline AutoKinship

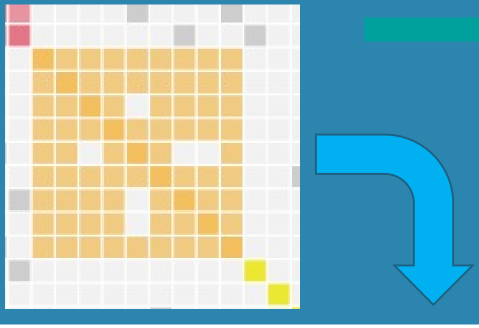
- What is AutoKinship
- AutoKinship concepts
- visualization
- AutoKinship & 23andme
- AutoKinship & MyHeritage
- Conclusion AutoTree & AutoKinship

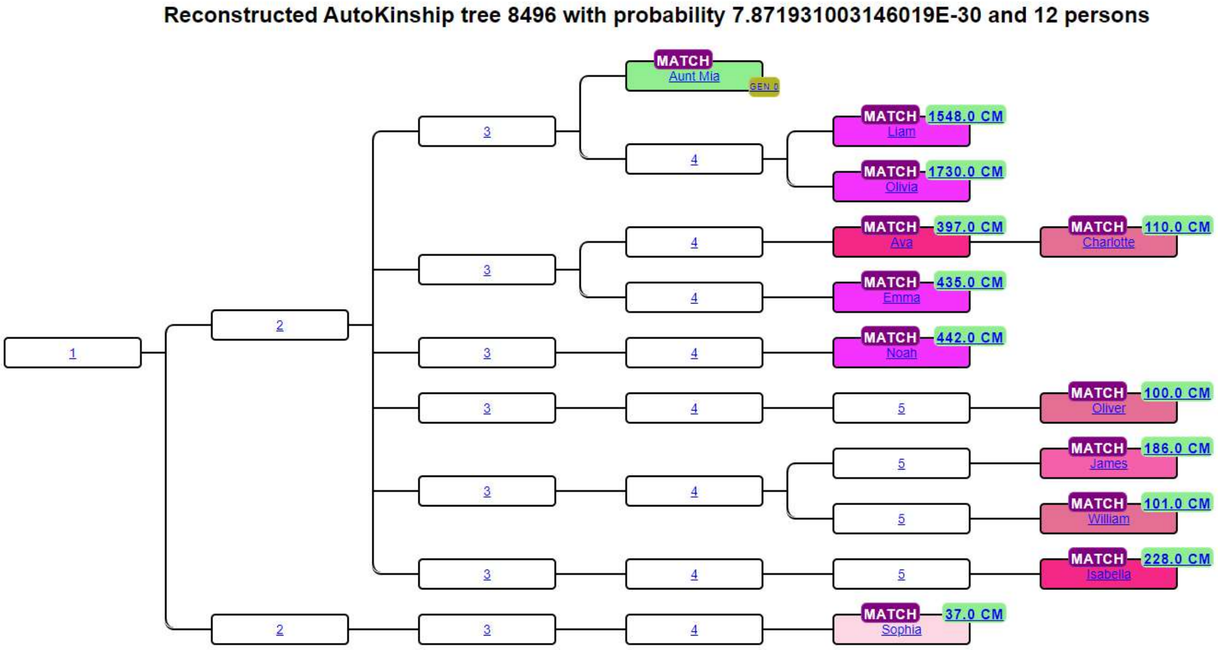
Based on AutoCluster



- AutoCluster provides a limited perspective regarding how matches are related
- Matches from a cluster likely share a common ancestor but how are they related?
- Are there different MRCAs between the matches?

AutoKinship

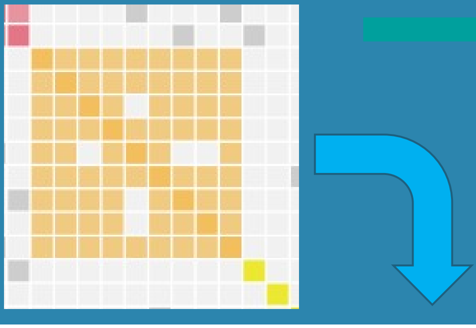


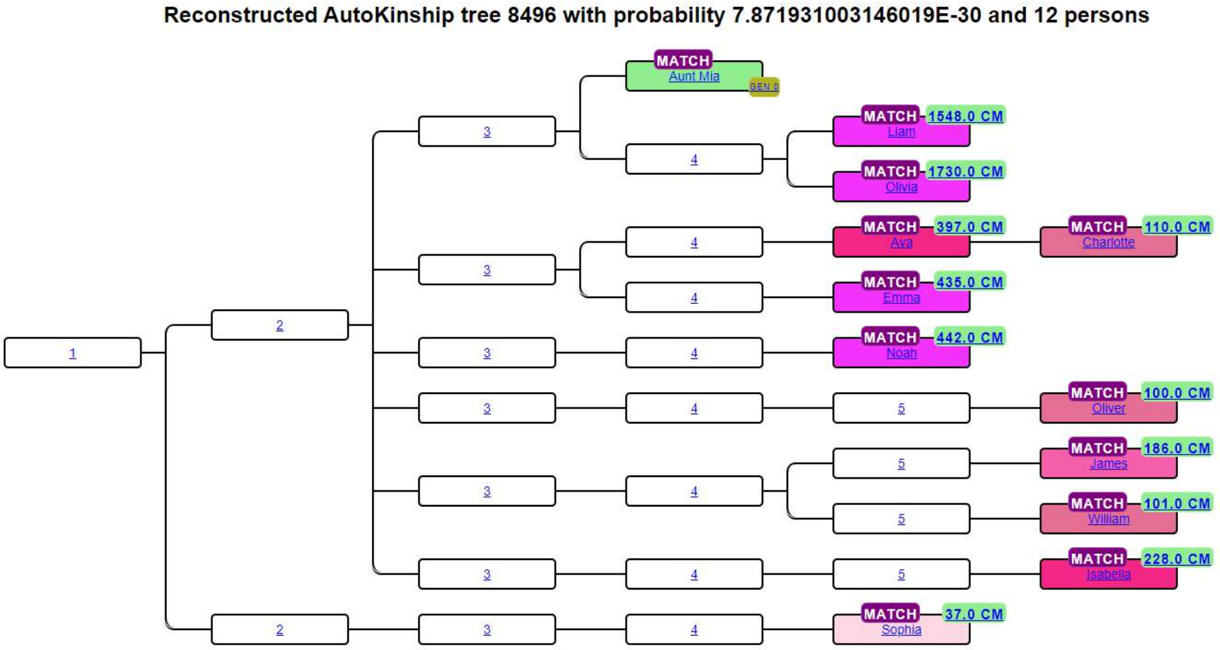


Reconstructed AutoKinship tree 8496 with probability 7.871931003146019E-30 and 12 persons

- AutoKinship is performed using members of clusters obtained from an AutoCluster analysis.
- No genealogical trees required
- Predicts family trees based on the amount of DNA your DNA matches share with you and each other.

AutoKinship

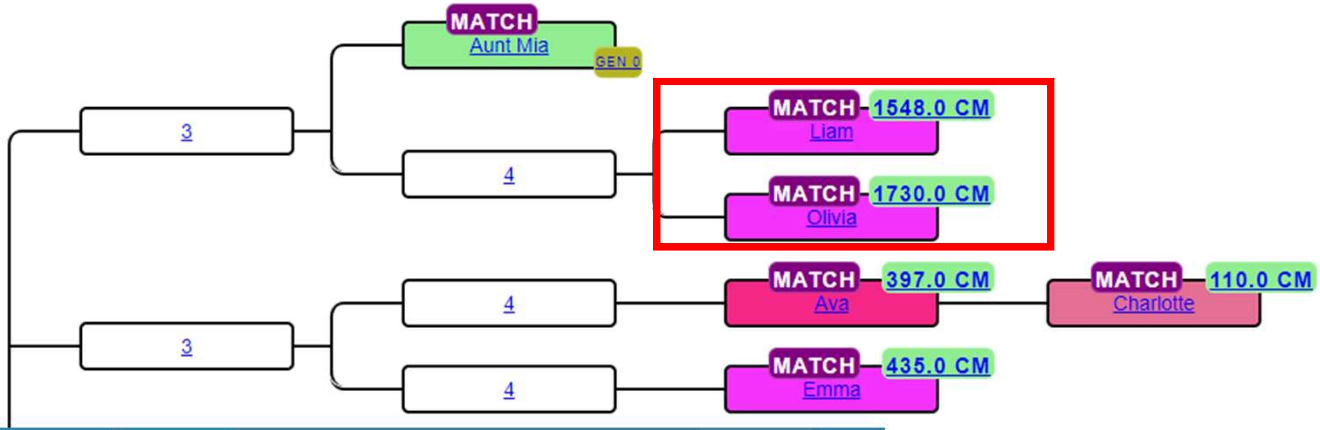


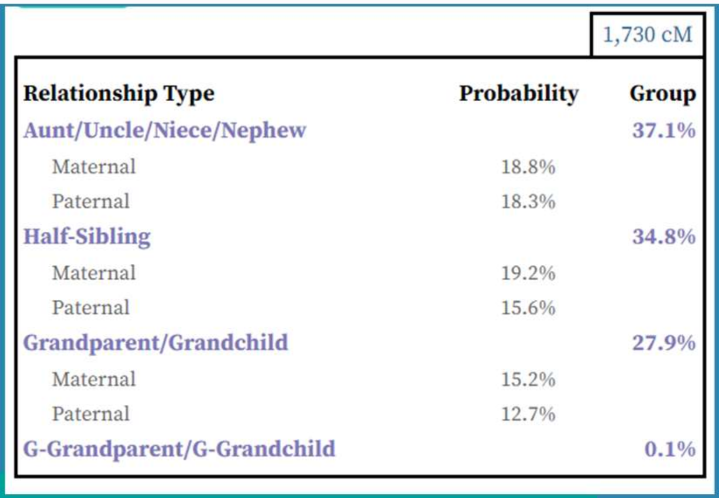


Reconstructed AutoKinship tree 8496 with probability 7.871931003146019E-30 and 12 persons

- Reconstructed trees provide hypotheses how matches are related to each other
- Work for persons with unknown parentage to their birth families (for instance adoptees or donor-conceived persons).
- Works for MyHeritage, 23andme and GEDmatch data.

AutoKinship example





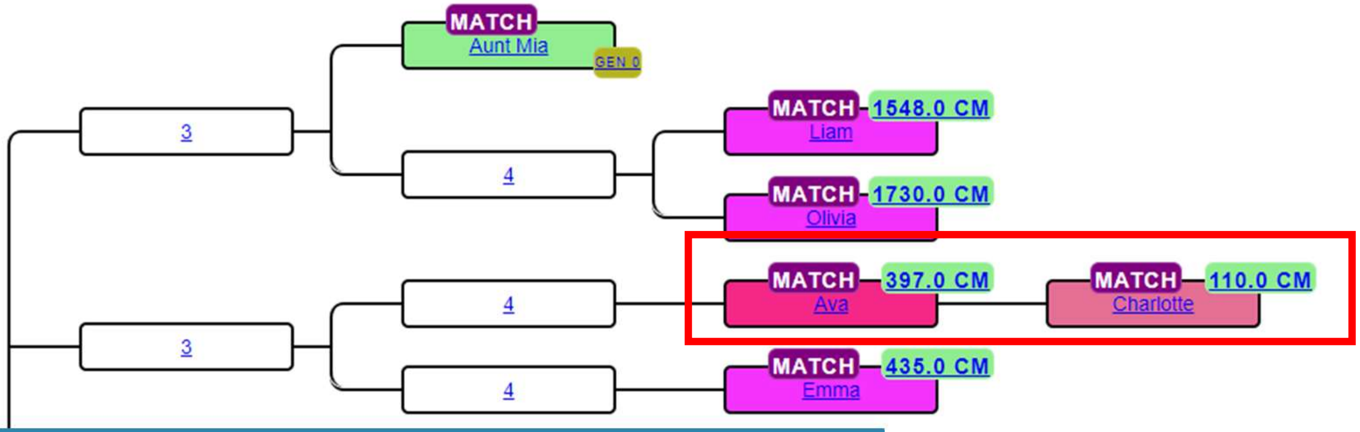
1,730 cM

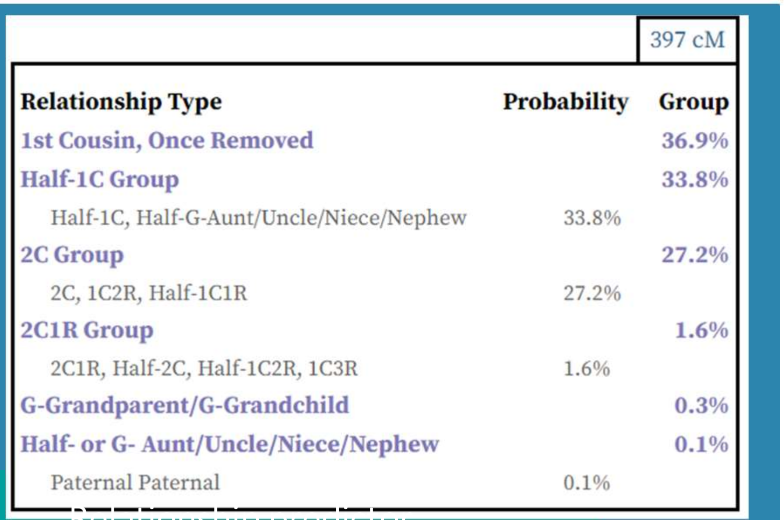
Relationship Type	Probability	Group
Aunt/Uncle/Niece/Nephew		37.1%
Maternal	18.8%	
Paternal	18.3%	
Half-Sibling		34.8%
Maternal	19.2%	
Paternal	15.6%	
Grandparent/Grandchild		27.9%
Maternal	15.2%	
Paternal	12.7%	
G-Grandparent/G-Grandchild		0.1%

Relationship predictor
https://dna-sci.com/tools/brit-cim/

- Aunt Mia is the tested person
- Liam and Olivia shared 2905 cM and are predicted siblings
- They share 1548/1730 cM with Aunt Mia

AutoKinship example





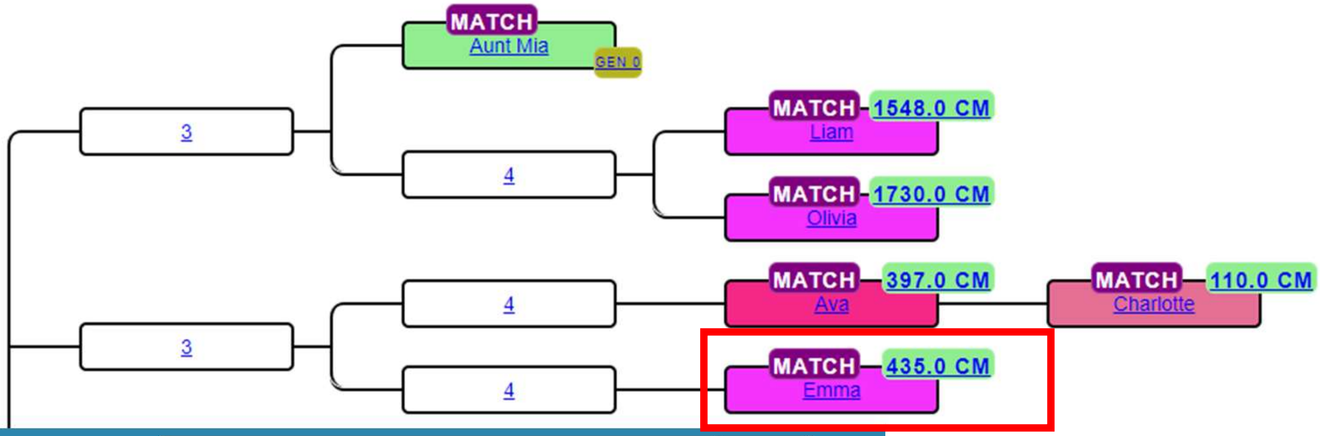
397 cM

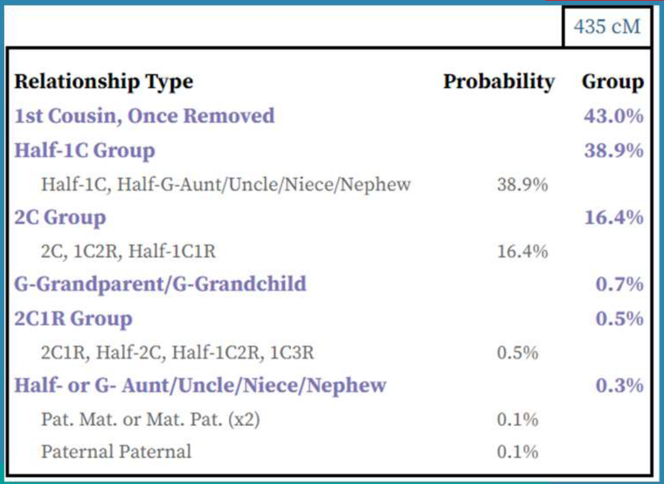
Relationship Type	Probability	Group
1st Cousin, Once Removed		**36.9%**
Half-1C Group		**33.8%**
Half-1C, Half-G-Aunt/Uncle/Niece/Nephew	33.8%	
2C Group		**27.2%**
2C, 1C2R, Half-1C1R	27.2%	
2C1R Group		**1.6%**
2C1R, Half-2C, Half-1C2R, 1C3R	1.6%	
G-Grandparent/G-Grandchild		**0.3%**
Half- or G- Aunt/Uncle/Niece/Nephew		**0.1%**
Paternal Paternal	0.1%	

Relationship predictor
https://dna-sci.com/tools/brit-cim/

- Ava and Charlotte share 3452 cM, parent child relationship. Who is the child?
- Aunt Mia shares 397 cM with Ava, and 110 cM with Charlotte
- Ava is the parent, most likely a 1C1R

AutoKinship example





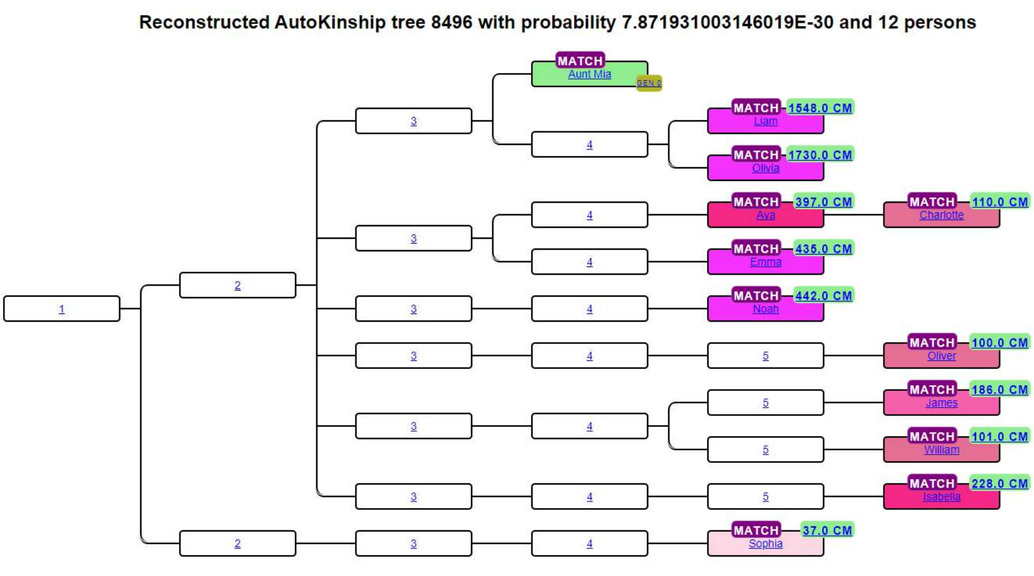
435 cM

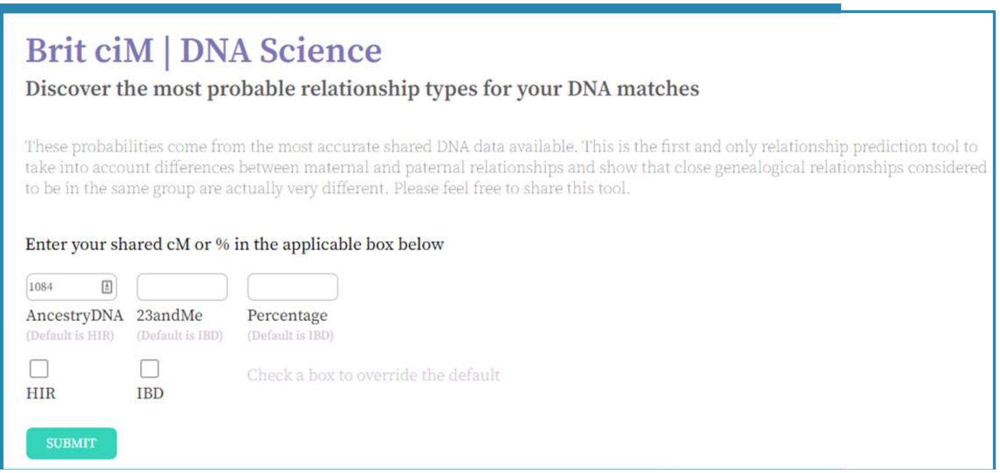
Relationship Type	Probability	Group
1st Cousin, Once Removed		**43.0%**
Half-1C Group		**38.9%**
Half-1C, Half-G-Aunt/Uncle/Niece/Nephew	38.9%	
2C Group		**16.4%**
2C, 1C2R, Half-1C1R	16.4%	
G-Grandparent/G-Grandchild		**0.7%**
2C1R Group		**0.5%**
2C1R, Half-2C, Half-1C2R, 1C3R	0.5%	
Half- or G- Aunt/Uncle/Niece/Nephew		**0.3%**
Pat. Mat. or Mat. Pat. (x2)	0.1%	
Paternal Paternal	0.1%	

Relationship predictor
https://dna-sci.com/tools/brit-cim/

- Aunt Mia and Emma share 435 cM, potential 1C1R?
- Emma and Ava share 1084 cM, so most likely a 1C
- The 1C prediction of Emma and Ava fits nicely with the 1C1R prediction of Emma with Aunt Mia.

AutoKinship concepts


Reconstructed AutoKinship tree 8496 with probability 7.871931003146019E-30 and 12 persons


Brit ciM | DNA Science

Discover the most probable relationship types for your DNA matches

These probabilities come from the most accurate shared DNA data available. This is the first and only relationship prediction tool to take into account differences between maternal and paternal relationships and show that close genealogical relationships considered to be in the same group are actually very different. Please feel free to share this tool.

Enter your shared cM or % in the applicable box below

AncestryDNA (Default is HIR): 1084

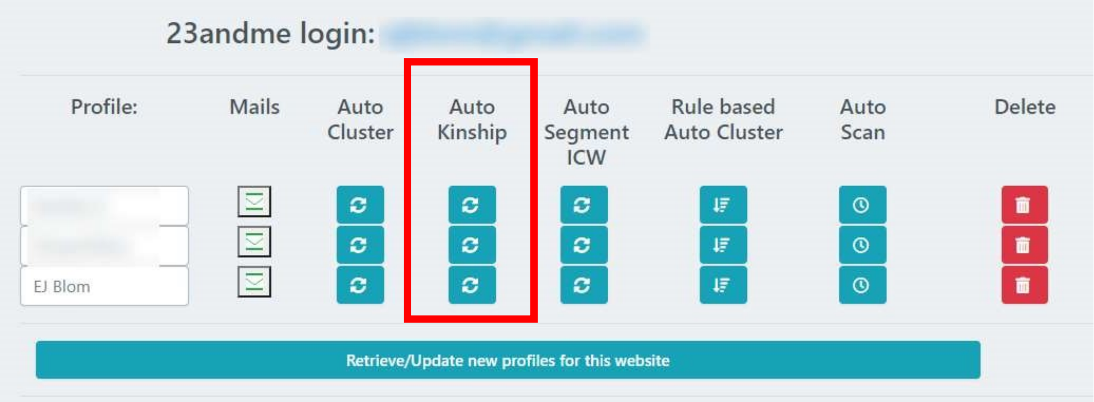
23andMe (Default is IBD):

Percentage (Default is IBD):

☐ HIR ☐ IBD Check a box to override the default

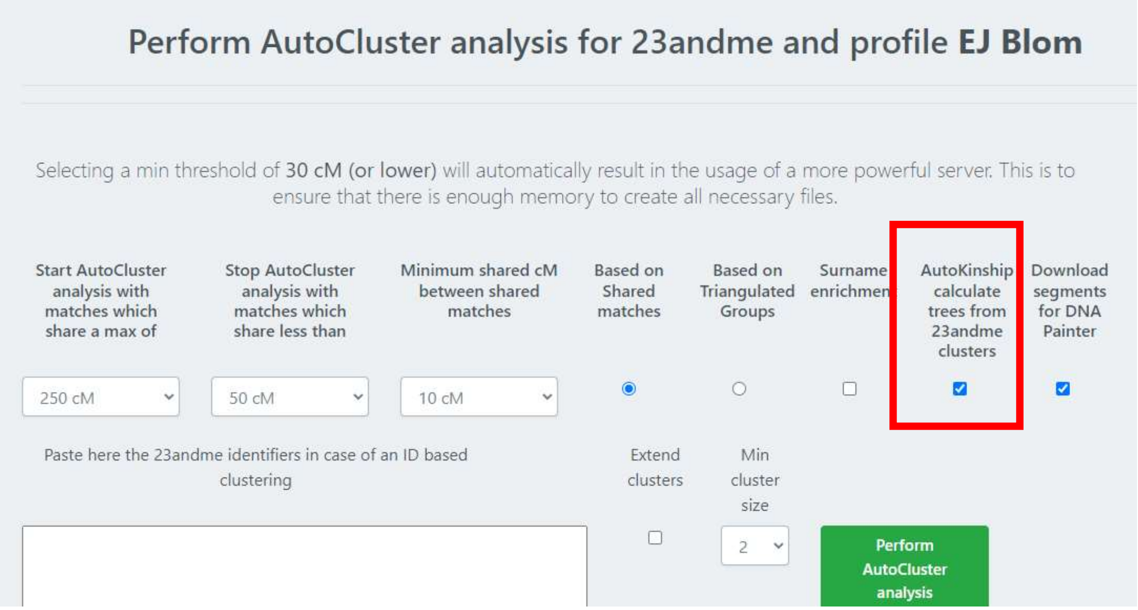
- Examines many possible relationships, similar to WATO but from the perspective of all matches
- Generated trees are sorted based on the combined probabilities
- Employs probabilities from https://dna-sci.com, created by Brit Nicholson

AutoKinship & 23andme



23andme login: [redacted]

Profile:	Mails	Auto Cluster	Auto Kinship	Auto Segment ICW	Rule based Auto Cluster	Auto Scan	Delete
[redacted]							
[redacted]							
EJ Blom							



Perform AutoCluster analysis for 23andme and profile **EJ Blom**

Selecting a min threshold of 30 cM (or lower) will automatically result in the usage of a more powerful server. This is to ensure that there is enough memory to create all necessary files.

Start AutoCluster analysis with matches which share a max of	Stop AutoCluster analysis with matches which share less than	Minimum shared cM between shared matches	Based on Shared matches	Based on Triangulated Groups	Surname enrichment	AutoKinship calculate trees from 23andme clusters	Download segments for DNA Painter
250 cM	50 cM	10 cM	●	○	☐	☑	☑

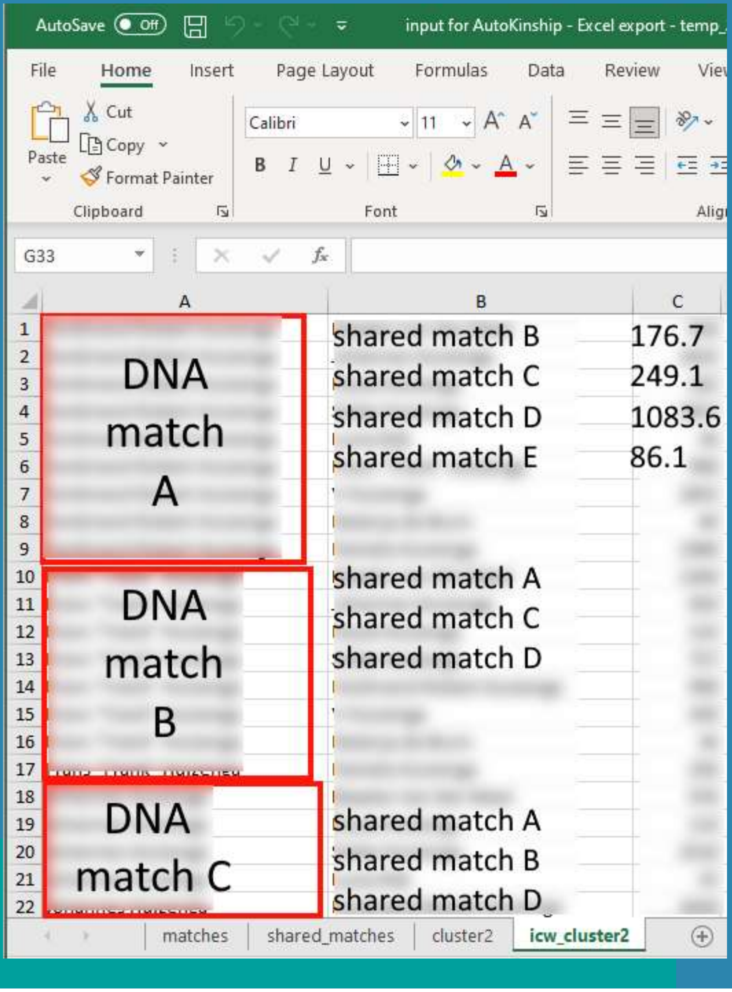
Paste here the 23andme identifiers in case of an ID based clustering

Extend clusters ☐

Min cluster size 2

- Available for 23andme profiles
- Select 23andme profile of interest
- Select AutoKinship option
- Set parameters, and start analysis
- Retrieve results by e-mail or download results from site

Running AutoKinship for MyHeritage



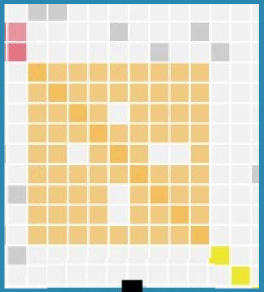
A	B	C
DNA match A	shared match B	176.7
	shared match C	249.1
	shared match D	1083.6
	shared match E	86.1
DNA match B	shared match A	
	shared match C	
	shared match D	
DNA match C	shared match A	
	shared match B	
	shared match D	

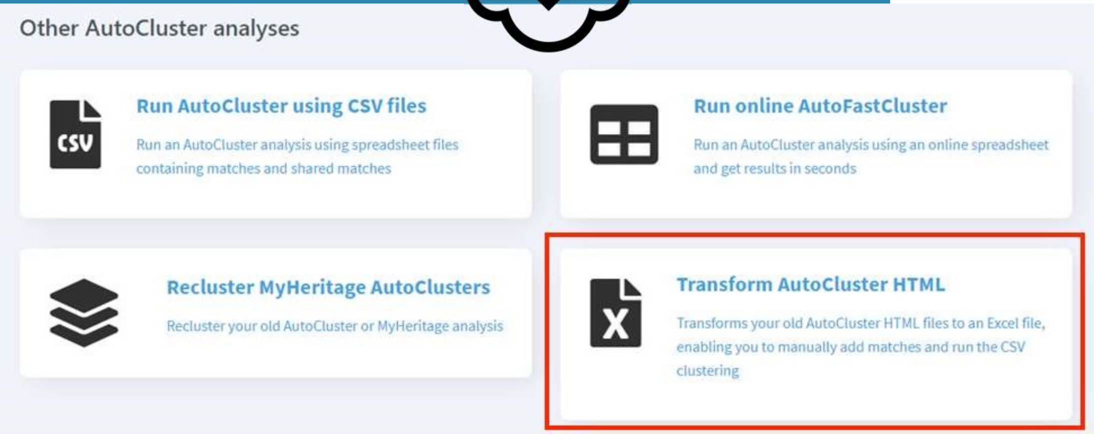
- (semi) manual data retrieval using 2 CSV files

- Requires matches file, 2 or 3 columns
  - 1st column: match name
  - 2nd column: amount of shared cM
  - 3rd column: generation (optional)

- Requires shared matches file, 3 columns
  - 1st column: match name
  - 2nd column: shared match name
  - 3rd column: amount of shared DNA between shared matches

Tutorial AutoKinship for MyHeritage





Other AutoCluster analyses

Run AutoCluster using CSV files
Run an AutoCluster analysis using spreadsheet files containing matches and shared matches

Run online AutoFastCluster
Run an AutoCluster analysis using an online spreadsheet and get results in seconds

Recluster MyHeritage AutoClusters
Recluster your old AutoCluster or MyHeritage analysis

Transform AutoCluster HTML
Transforms your old AutoCluster HTML files to an Excel file, enabling you to manually add matches and run the CSV clustering

- Run AutoClusters analysis on MyHeritage
- Save and unzip ZIP file from email
- Examine report and select cluster of interest
- Upload HTML to Genetic Affairs

- Retrieve Excel spreadsheet with matches & shared matches per cluster

Tutorial AutoKinship for MyHeritage

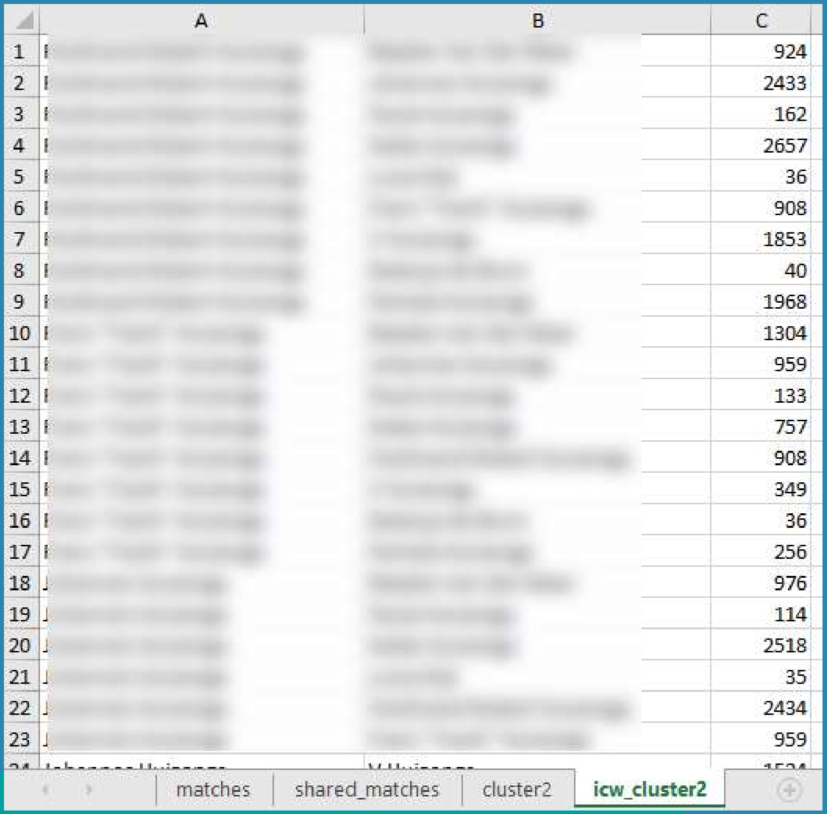
Gedeelde DNA Matches

en u delen de volgende 182 DNA Matches

Geschatte relatie tot **you**	Gedeelde Match		Geschatte relatie tot [redacted]
Moeder	**49,8%** (3.529,6 cM)	**3,3%** (236,6 cM)	Nicht 2 generaties verwijderd - 2e graad nicht 1 generatie verwijderd
Zoon	**49,9%** (3.538,9 cM)	**1,6%** (115,0 cM)	1e graad neef 2 generaties verwijderd - 2e graad neef 2 generaties verwijderd
Zoon	**50,0%** (3.541,1 cM)	**1,2%** (88,1 cM)	Neef 2 generaties verwijderd - 4e graad neef
Halfbroer, oom	**28,0%** (1.985,8 cM)	**3,8%** (268,1 cM)	Neef 1 generatie verwijderd - 2e graad neef

- Next, we need to obtain how much DNA our shared matches share.
- For each match
- There are a number of shared matches
- This is the amount of DNA I am sharing with the shared match
- My match is sharing this amount of DNA with each shared match
- The shared DNA information in this column needs to be copied into the Excel

Tutorial AutoKinship for MyHeritage

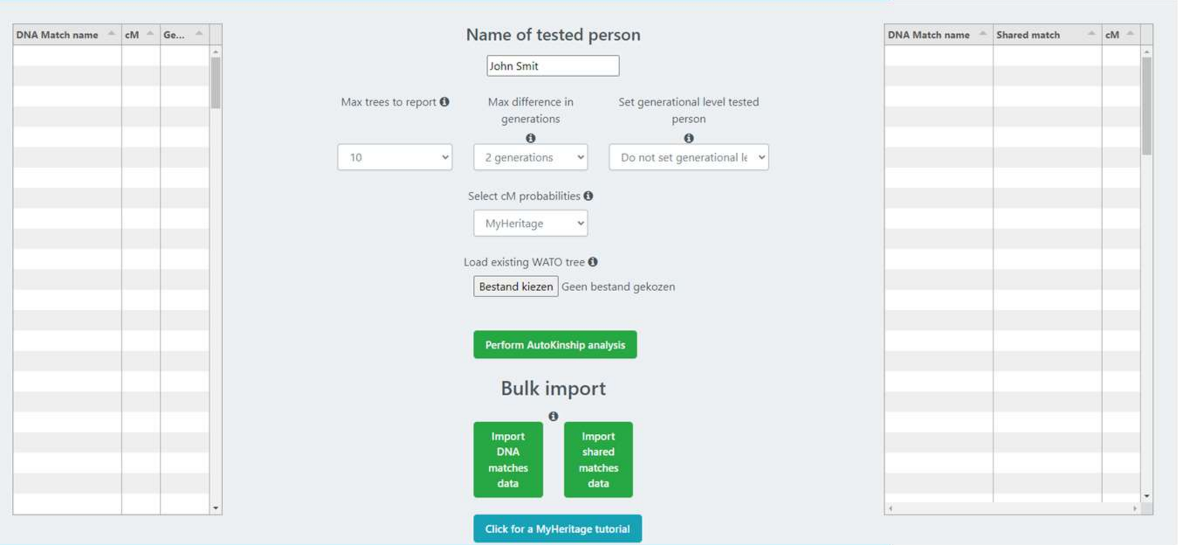


	A	B	C
1	[redacted]	[redacted]	924
2	[redacted]	[redacted]	2433
3	[redacted]	[redacted]	162
4	[redacted]	[redacted]	2657
5	[redacted]	[redacted]	36
6	[redacted]	[redacted]	908
7	[redacted]	[redacted]	1853
8	[redacted]	[redacted]	40
9	[redacted]	[redacted]	1968
10	[redacted]	[redacted]	1304
11	[redacted]	[redacted]	959
12	[redacted]	[redacted]	133
13	[redacted]	[redacted]	757
14	[redacted]	[redacted]	908
15	[redacted]	[redacted]	349
16	[redacted]	[redacted]	36
17	[redacted]	[redacted]	256
18	[redacted]	[redacted]	976
19	[redacted]	[redacted]	114
20	[redacted]	[redacted]	2518
21	[redacted]	[redacted]	35
22	[redacted]	[redacted]	2434
23	[redacted]	[redacted]	959

- For each match and shared match combination
- Find back how much DNA they share
- Fill it in the last column

- Quite time consuming!

Tutorial AutoKinship for MyHeritage



Name of tested person: John Smit

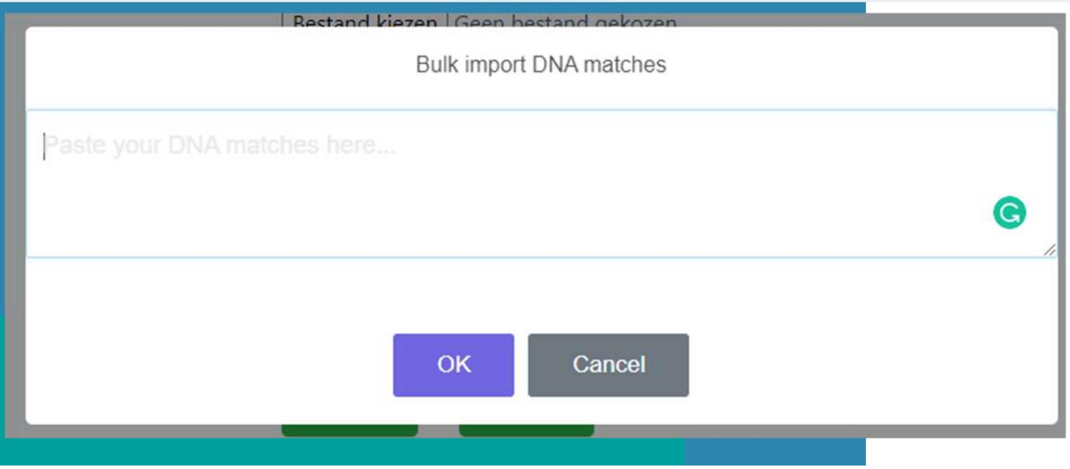
Max trees to report: 10

Max difference in generations: 2 generations

Set generational level tested person: Do not set generational le

Select cM probabilities: MyHeritage

Load existing WATO tree: Bestand kiezen Geen bestand gekozen



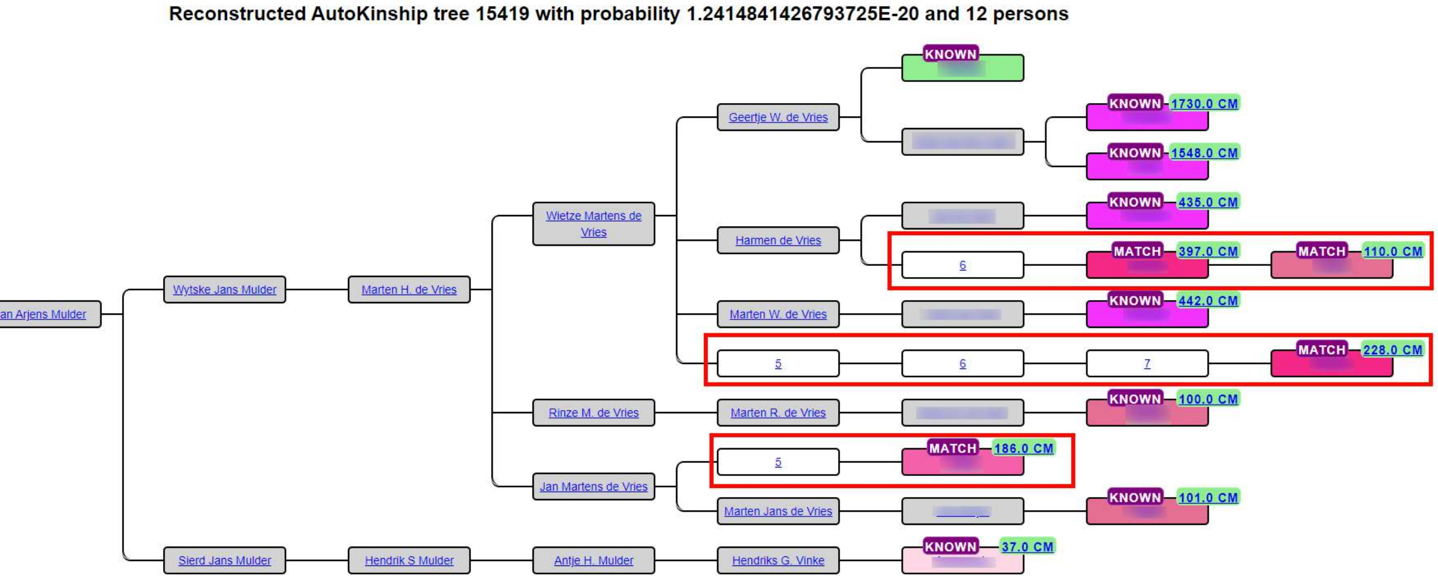
Bestand kiezen | Geen bestand gekozen

Bulk import DNA matches

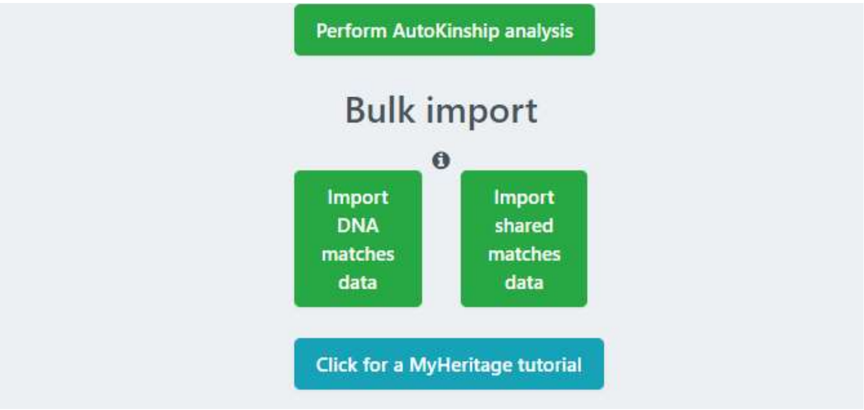
Paste your DNA matches here...

- Visit Genetic Affairs: https://members.geneticaffairs.com/autokinship
- Use bulk import to paste matches & shared matches
- (optional) set generational level of tested person
- Provide name, select cM probabilities and start analysis

Tutorial AutoKinship for MyHeritage



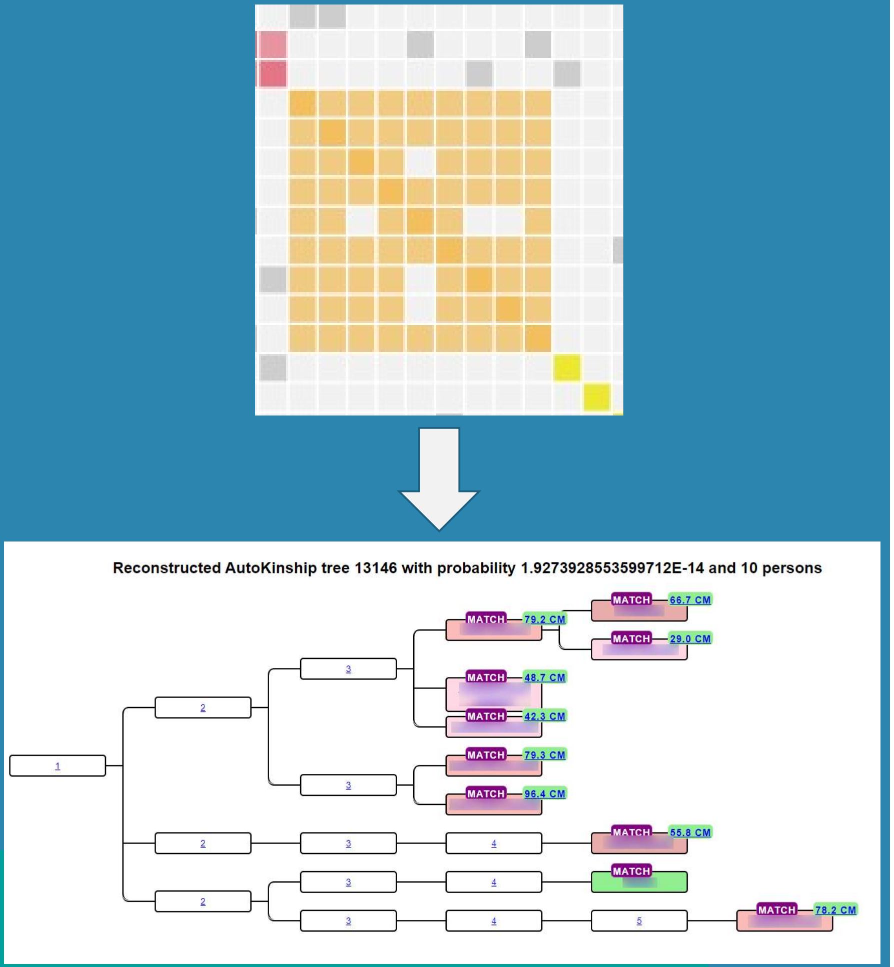
Reconstructed AutoKinship tree 15419 with probability 1.2414841426793725E-20 and 12 persons



- Parts of the predicted tree can be known

- Make a tree in WATO (https://dnapainter.com/tools/probability) and load it into AutoKinship

Conclusion AutoKinship & AutoTree



Reconstructed AutoKinship tree 13146 with probability 1.9273928553599712E-14 and 10 persons

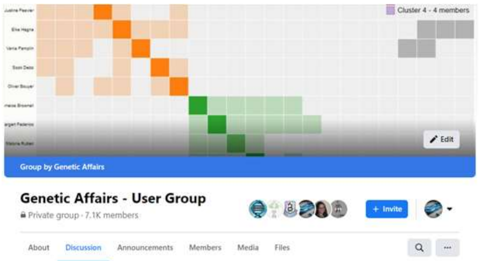
- **AutoTree** reconstructs genealogical trees based on trees shared by (FTDNA) DNA matches
  - **AutoTree** integration for AutoKinship combines best of both worlds: genealogical trees and reconstructed trees based on DNA
- **AutoSegment** provides shared segments for AutoKinship branches
- **AutoKinship** reconstructs trees based on shared DNA between shared matches
  - Automatically for 23andme matches
  - Semi automatically for MyHeritage
  - Will be available on GEDmatch (tier 1)

Additional resources

Genetic genealogy blogs

https://patriciacolemangenealogy.com/

https://dna-explained.com/category/genetic-affairs/



Genetic Affairs - User Group

Private group · 7.1K members

www.facebook.com/groups/GeneticAffairs

www.geneticaffairs.com

info@geneticaffairs.com

https://twitter.com/AffairsGenetic