The Värmland-Hedmark Cluster

No. 5 (July 2021)

A newsletter with updates about the genetic genealogy project "The Värmland-Hedmark Cluster." The project is run by a group of hobbyists who try to learn about the early history of their patrilineal ancestors in southern Värmland (Sweden) and in Hedmark (Norway). What can be said about the ancestors' whereabouts before they start to appear in the written records, i.e., in the 1500s and earlier? Where did the ancestors live before they arrived in Värmland and Hedmark? In technical terms, the project concerns the study of the Y chromosome haplogroup R1b → M269 → U106 → Z18 → S11601 → Y112538 → Y130179 and surrounding clades.

Contact information: The newsletter is written by Johan N.M. Lagerlöf (email: johan.lagerlof@econ.ku.dk; website: www.johanlagerlof.com). *This issue and previous ones are available at www.johanlagerlof.com/gengen. A blog (called "The VHC Blog"), where the content of the newletters can be discussed, is available here: https://www.johanlagerlof.com/the-vhc-blog.*

En sammanfattning på svenska finns i slutet av nyhetsbrevet.

WELCOME to the fifth issue of the newsletter. I hope all of you are well. The pandemic appears to be slowing down (or, at least, it should be soon), as an increasing number of people get vaccinated. Parts of Sweden have been quite slow with vaccinating their residents, but I will get my second jab on August 3.

The first half of 2021 has been unusually hectic for me, and I have not had as much time as I would have wanted to interact with you, among other things. My apologies for any late replies to your emails! However, I have recently spent a fair amount of time on something that is closely related to the VHC project, namely to try to compute my own age estimates for the ancestors A1-A8 in our tree. Unfortunately I have not yet been able to solve some remaining problems, and I am therefore not ready to present any results now. Hopefully, though, I can do that in the next issue of the newsletter. One reason why I want to do my own age estimations is that I think one can do it in a statistically and genealogically more appealing way, compared to how YFull is doing it. I also think one should carry out some formal hypothesis testing, and for this I cannot rely on YFull's analysis.

Still, there are plenty of news to report in the present issue of the newsletter. We have two new testers in the tree (Tester 14 and Tester 15). A third person, whose lineage is from Nor in Värmland, has tested recently with Dante Labs, but unfortunately there are problems with the quality of the results. Yet I have not given up, and one way or another we will get a decent test result for this tester (or some other representative of his lineage) and get him into the tree. One of the two new testers in the tree (namely, Tester 14) has together with an existing tester (Tester 12) formed a new branch, identifying an ancestor called A8 in the tree. The new branch connects a man living in the 1700s in Södra Rådom in Nyed parish in Värmland with a man born in 1885 in Grue in Hedmark, Norway. The common ancestor A8 might have lived in the Middle Ages (although this is, as always, very hard to estimate accurately). Among other news is that Tester 9 (whose oldest known patrilineal ancestor is from Ireland) has submitted his BAM file from Family Tree DNA to YFull, which is great. This should help us get the highest possible accuracy for the age estimate of ancestor A4 (our Norman/Norwegian branch). Moreover, Tester 5 has upgraded his Big Y-500 test to a Big Y-700. This is very useful, as it can potentially help us identify a new branch (the 700 test can find more SNPs). Tester 5 has not formed a new branch with anyone, but hopefully he will in the future, as new testers arrive.

The Värmland-Hedmark Cluster and a Neighboring Branch, July 2021

Confidence intervals for age estimates

The age estimates in the tree are from YFull (YTree v9.03.00; note that Testers 9 and 15 are not yet included). YFull also provides the following confidence intervals at the 95% level:

- **A1:** AD 921 — AD 1621
- **A2:** AD 621 — AD 1621
- **A3:** 379 BC — AD 621
- **A4:** AD 221 — AD 1371
- **A5:** AD 921 — AD 1746
- **A6:** AD 921 — AD 1771
- **A7:** AD 621 — AD 1621
- **A8:** AD 621 — AD 1621

SNPs that define the branches

The Värmland-Hedmark cluster is a twig on the R1b branch of the human Y chromosome haplotree: R1b → M269 → U106 → Z18 → S11601 → Y112538 → Y130179. For more information, see www.johanlagerlof.com/gengen. The branches shown in the tree are defined by the SNPs listed below (an asterisk indicates that the SNP is located outside of the so-called combBED region of the Y chromosome, and it is therefore not used by YFull for age estimation; an "(H)" means that the SNP is located in a homologous region—i.e., one that is similar to other regions on the Y chromosome or on other chromosomes—and therefore is less reliable).

- **A1:** Y86344, Y107658*
- **A2:** Y100873*, Y128033, Y101814*, Y103246, Y130179, Y83455, Y125815*
- **A3:** Y112538, Y129064
- **A4:** BY71612, Y159325, BY116275, BY146620, BY78202, BY99722, BY100142, BY102070
- **A5:** FT146431
- **A6:** BY12164
- **A7:** A25847*
- **A8:** A25843, A25844, A25845, Y126692(H)*

A3 (~ AD 184)

A4 (~ AD 901)

A2 (~ AD 1309)

A7 (~ AD 1304)

A1 (~ AD 1305)

A8 (~ AD 1232)

A6 (~ AD 1463)

A5 (~ AD 1450)

Kilkenny, Ireland
William Walsh (c. 1814-1896)

Somerset, England
Henry Peadon (?-1866)

George Peadon (1822-1893)

Levanger, Norway
Andreas Johanesen (c. 1820 - ?)

Sør-Odal, Norway
Ole Peterson (1823-?)

Grue, Norway
Botolf Magnussen (1885-1967)

Södra Rådom, Nyed
Bengt Olsson (c. 1725-1800)

Ölmbäck, Väse
Christoffer Andersson (1600-1678)

Anders Christoffersson (1657-1730)

Anders Andersson (1683-1772)

Per Andersson (1696-1757)

Östra Glänne, Frykerud
Torsten Pålsson (?-1737)

Pål Torstensson (1711-1777)

Nils Pålsson (1756-1832)

Nedre Ullerud
Anders Andersson (1790-1873)

Gersheden, Ransäter
Per Jonsson (1750-1809)

Skived, Grava
Per Månsson (late 1500s to 1600s)

Magnus Lagerlöf (1778-1844)

Rosensjö, Karlskoga
Nils Nilsson (1742-?)

Jordkullen, Kroppa
Jon Pädhersson (1676-1733)

Anders Jonsson (1708-1775)

Per Andersson (1743-1810)

Stockholm, Sweden
unknown

Private (1915-1989)

Tester 9

Tester 8

Tester 7 (BY-500)

Tester 6 (BY-700)

Tester 12 (Y37, 30X WGS)

Tester 14 (Y111, 30X WGS)

Tester 3 (BY-700)

Tester 4 (BY-500, SNP test)

Tester 5 (BY-700)

Tester 2 (BY-700)

Tester 13 (Y67, 30X WGS)

Tester 1a (BY-500, 30X WGS)

Tester 1b (test of two SNPs)

Tester 11 (BY-700)

Tester 10 (BY-700)

Tester 15 (Y37, 30X WGS)

Discussion of the Latest Version of the Tree

THE new version of the tree is shown on page 2. Let me talk us through this tree, to remind us about what it shows and what we think we can, and cannot, infer from the tree structure and our other results. The yellow boxes at the bottom of the tree represent our testers. These are people who live today and who, in their roles as hobby genealogists, have taken a genealogical DNA test.[1] All of them (except for Tester 1b, who we do not need it for) have purchased at least a basic Y chromosome test from the American company Family Tree DNA—a so-called STR test, typically with 37 (but sometimes 67 or 111) markers. In addition, they have all (again with the exception of Tester 1b) taken a more ambitious test that studies which SNPs on the Y chromosome the tester is positive and negative for, respectively. (If you need a reminder about the difference between STR markers and SNPs, study the box below.) The more ambitious tests—i.e., the ones that study SNPs—have been purchased at either Family Tree DNA (a so-called Big Y test) or at the Italian company Dante Labs (a whole genome sequencing, or WGS, test).

About STR markers and SNPs

Our DNA consists of a long series of "letters" A, C, G, and T (to call them letters is a simplification but is good enough now for our genealogy purposes). A segment of our DNA (say, on the Y chromosome) may thus look like this:

AGCCTACAGGCAGGCAGGCAGGCTGGACAGTACTGA

A SNP (single-nucleotide polymorphism), pronounced "snip", is simply a change of one of these letters to another letter. For example, the first letter above, an A, may change to a T.

What is an STR (short tandem repeats) marker? Study the series of letters above. If you look carefully, you will see that the combination of letters AGGC is repeated four times in the middle of the series (below the same sequence is shown again, but with the four appearances of AGGC being highlighted):

AGCCTACAGGCAGGCAGGCAGGCTGGACAGTACTGA

An STR marker is simply the number of such repeats (so, in the example, the value of the STR marker is 4). The reason why a certain combination of letters is repeated several times is that there was a DNA copying error. It would have sufficed if Mother Nature had included the combination once, but by mistake it was added a few additional times (without affecting anything biologically with the body, as far as I understand).

The value of an STR marker tends to change quite often. Sometimes the number of repeats goes up and sometimes it goes down. A SNP, in contrast, is much more likely to stay put, once it has occurred. For example, at the conception of some patrilineal ancestor of the individuals in the tree on page 2, the SNP Z18 occurred (this is a change from a G to a A at position #12,879,820 on the Y chromosome, if we use the Hg38 reference genome). This happened about 3000 or 4000 years ago, and the SNP has not reverted in any of the generations after that ancestor. As the SNPs are more stable and reliable, they are used for building the tree. The STR markers have historically been easier and less expensive to test, which is one reason why they have been used a lot over the years in forensic science, genealogy, and other fields. They are also useful for us in the project, partly because they help us identify people who are likely to belong to the Värmland-Hedmark cluster. Yet, to confirm this and to be able to put these individuals in the tree, a fairly advanced SNP test is needed (e.g., a Big Y-700 or a 30X WGS test).

The greenish boxes in the tree represent a selection of the testers' known patrilineal ancestors, the

[1] The Wikipedia entry for "genealogical DNA test" is useful; see `https://en.wikipedia.org/wiki/Genealogical_DNA_test`.

information coming from church records or other written sources. The genealogical research has been carried out by the testers themselves and, in some cases, by me. As is evident from the tree, the oldest known patrilineal ancestors of almost all testers on the right side of the tree (i.e., all testers except for #7, #8, and #9) lived in the southern part of Värmland, Sweden, or in Hedmark, Norway, which in this context is a very small geographic region. The only exception is the oldest known patrilineal ancestor of our new Tester 15.[2] However, this tester's paternal grandfather, born in Stockholm, Sweden, had an unknown father. We know that the mother of the tester's paternal grandfather (she lived between 1880 and 1975) resided in Stockholm at the time when her son was conceived. I will do my best, together with Tester 15, to identify the unknown father. I will be extremely surprised if it turns that the unknown father's lineage does not originate from Värmland or Hedmark. The testers on the left side of the tree (Testers 7, 8, and 9) have oldest known patrilineal ancestors from Ireland, the south-west of England, and Norway (in particular from Levanger, which is located about 400 km north of Hedmark). Still, we know that this group of testers share patrilineal ancestors with the testers who have a Värmland-Hedmark origin. Hence, the presence of the branch made up of by Testers 7-9 gives us a hint about how the migratory paths might have looked like. Another important piece of information, which is not indicated in the tree, is that two of the testers on the left, namely Testers 8 and 9, carry surnames with a Norman origin (see the discussion on pp. 8-9 in issue No. 4 of the newsletter).

The almond-colored boxes at the top of the tree represent patrilineal ancestors that are unknown. A few of them have been given the names A1, A2, A3, and so on, up to A8. Note that although these ancestors cannot be found in any written records, they were real people that must have existed, and they must have had the patrilineal descendants that are indicated by the tree branches. We can also make inferences, or at least intelligent guesses, about the geographic region in which they lived, as well as when they lived. The most recently introduced branch in our tree is the one involving ancestor A8, which according to the tree is a patrilineal ancestor to both Tester 12 and Tester 14, but to no one else among our testers—nor, for that matter, to anyone else among all known Y chromosome testers in the world. We know this because both Testers 12 and 14 are positive for the SNPs called A25843, A25844, A25845, and Y126692,[3] which all other known testers are negative for. It certainly looks as if the ancestor A8 lived somewhere in (what is at least today referred to as) Värmland or Hedmark, because A8 is in turn a patrilineal descendant of A2; moreover, all of A2's known descendant lineages (except, as far as we know, the one of Tester 15—see the discussion above) lead to either Värmland or Hedmark. The company YFull carries out age estimations, based on the number of novel SNPs (i.e., SNPs that are not shared by anyone else) that each tester carries. For example, Tester 12 has six novel SNPs and Tester 14 has four novel SNPs (in both cases, this is the number of novel SNPs in the so-called combBED region of the Y chromosome, which YFull uses for age estimation). This has lead YFull to come up with the estimated birth year 1232 for A8, with the following 95% confidence interval: 621–1621 (so a range of a thousand years).

So, what have we, so far, learned from the tree structure and the associated data? In the previous issue of the newsletter, I offered my attempt at answering that question, and I refer the reader to that text for a more elaborate discussion. In brief, I argued that there were at least three things that we have learned with a reasonably large degree of confidence: the extended family defined by the patrilineal descendants of A2 in the tree must have been quite large already in the Middle Ages; these family members must have lived in Värmland/Hedmark for quite a long time before the 1600s (say, at least one and probably several centuries); and, within the family, there were multiple migrations between Värmland and Hedmark.

In the previous issue of the newsletter, I also discussed how the likely migration paths might have looked like—for example, where from A2's lineage came when it entered the Värmland/Hedmark region. I chose to be agnostic about these questions. I listed a few hypotheses (mutually incompatible),

[2]Note that this is not the "Tester 15" that I mentioned in issue No. 4 of the newsletter but another person. The Tester 15 that I mentioned in that issue was the person with patrilineal ancestors from Nor, whose test at Dante Labs so far has failed. Given that that person is still in the pipeline, I have assigned the number 15 to this other tester.

[3]The names of three first SNPs have been assigned by the German company YSEQ, which discovered these SNPs (after I had sent them the files with raw data for remapping to a newer reference genome). The prefix A stands for, I believe, Astrid, which is the name of one of the two people running this family business (Astrid Krahn and Thomas Krahn).

and I argued that we do not have enough data to rule any of them out. In particular, I kept open both the possibility that A2's lineage entered from the British Isles (perhaps in connection with timber trade, missionary activities, or a Brit being taken as slave and brought to Scandinavia) and the possibility that the lineage entered from Norway. After I wrote that text, I heard objections that it seems more likely that the lineage entered from Norway (and that it therefore is inappropriate to be agnostic in the choice between the two competing hypotheses).[4] The main argument that I have seen for why entrance from Norway would be more likely is that Z18 and other SNPs upstream of the ancestor A3 (so in a part of the haplotree that is not included in the tree on page 2) are believed to have originated in Scandinavia. My response to that is that the British Isles hypothesis is fully consistent with Z18 (and other upstream SNPs) having originated in Scandinavia. The reason is that upstream of A3 there are *both* branches that lead to testers with oldest know patrilineal ancestors in Scandinavia *and* branches with testers having oldest know patrilineal ancestors on the British Isles (I have not had time create a figure that illustrates this, but I refer the reader to the public haplotrees at FTDNA and YFull, in order to verify my claim). That is, to me it seems that we have clear and convincing support in the data for the idea that, after the formation of Z18 (some three or four thousand years ago), there were several migrations, back and forth, between Scandinavia and the British Isles. Thus, the ancestor A3 might have lived on the British Isles, while still having patrilineal ancestors himself that lived in Scandinavia. We do not have any support in the data for ruling that scenario out.

However, since I wrote the text in the previous issue of the newsletter, I have discovered one piece of information that makes me willing to be a bit more sympathetic toward the view of those who find the Norway hypothesis more convincing than the British Isles hypothesis. In an essay published in the book "Naboliv,"[5] Peter Olausson (2005, p. 33) points out that Erik Fernow (1773-79, Vol. 1, p. 90) claimed, when writing in the 1700s, that several noble families in Värmland originate from Norway; they emigrated, at least according to Olausson, due to either political or economic reasons. Olausson mentions two examples. First, the family Roos af Hjelmsäter is said to have branched off in the 1300s from the family of the Norwegian king Harald Hårfager (in English often written Harald Hairfair) and some landowners in Värmland.[6] Second, the noble family Bratt is said to have migrated from Norway to Värmland in the 1400s. I find this information interesting and suggestive. It is true that the Värmland-Hedmark cluster has no ties to nobility (to our knowledge anyway). Moreover, the fact that family A came from Norway does not mean that so did family B. Still, if it is true that there was a reasonably large number of families migrating to Värmland from Norway in the Middle Ages, our case would fit well into this pattern. The timing (say, the 1300s or 1400s) certainly makes sense, as it is consistent with our age estimates and what we know from genealogical data. Furthermore, even if the members of the Värmland-Hedmark cluster were not nobility, they could have been reasonably well off—thereby qualifying into a nearby social category; in addition, it is conceivable that although Fernow talks exclusively about noble families, non-noble families might have exhibited a similar migratory pattern but been thought of as less interesting to write about.[7]

All in all, I find the information about this alleged migratory pattern intriguing, and perhaps we can find more information along these lines that will shed some light also on the origin of the Värmland-Hedmark cluster. Still, we should bear in mind that, as evidence, this information is circumstantial. We do not have anything concrete that links the Värmland-Hedmark cluster to a Middle Age migration from Norway to Värmland.

It is also worth noting a couple of other weaknesses and limitations with the Norway hypothesis. First, the reason why we have started to think about the possibility that the ancestor A3 might have lived in Norway is the presence in our tree of Tester 7, whose oldest known patrilineal ancestor is from Levanger in Norway. That is a single data point! This observation shows how fragile any conclusions

[4] See the post "Can We Reject the Hypothesis That A3 Lived on the British Isles?" (February 1, 2021) on the VHC Blog.

[5] This title is in Norwegian, meaning "Neighbor Life"—the texts in the book are either in Norwegian or Swedish.

[6] This is supposed to be documented, and Olausson cites a paper by himself, Olausson (1997), which I have not yet read.

[7] Elsewhere in his essay (p. 36), Olausson (2005) argues that "[h]uvudströmmen för den ekonomiskt betingade migrationen har åtminstone sedan 1600-talet gått från Sverige till Norge." That is, at least from the 1600s, the migration in the geographic region that the essay concerns (Värmland and Hedmark) was primarily in the direction of Norway. This supports the idea that the migratory movements that we observe in our tree also were from Värmland to Hedmark, and not in the opposite direction.

drawn from our current data are. If Tester 7 never had submitted his test, we might not have had this discussion at all. Reversely, it is quite possible that there are people out there who have not yet submitted their saliva samples and if they did, we would start to formulate quite different hypotheses. Second, the Norway hypothesis (as I formulated it in the previous issue) is complex, in the sense that it involves migration not directly from Norway to the British Isles but via Normandy (in order to rationalize the presence of Norman surnames in our tree). It is of course quite possible that this is how the migratory path looked like—life is sometimes complex and takes strange roundabouts (and, in this case, there is historical evidence that people migrated along that route at that time). Yet, this complexity might make fans of Occam's razor feel uneasy, and they might prefer to wait until some new tester with roots in Normandy shows up before they are willing to accept the non-direct migratory path (as direct-to-consumer DNA tests are banned in France, this could be a futile hope). Both these points underscore the fact that more data would help us a lot. It sometimes feels as if the arrival of new testers is way too slow. Still, I am an optimist. DNA testing (in particular advanced tests of the Y chromosome) is a relatively new thing, and our project has been going on only for a few years. It seems likely that in, say, five or ten years from now we will be in a much better position and know significantly more than we do today.

A New but Unknown Tester in FTDNA's Tree

THE previous section ended on a somewhat pessimistic note, lamenting the fact that we do not have as much data as we would want. Here, however, is some information that I think should give us hope, although at first blush it might create the opposite impression. In January 2021, I discovered that a new tester belonging to our part of the haplotree had appeared in Family Tree DNA's so-called block tree. In particular, this tester is supposed to be positive for one of the two SNPs that define our ancestor A3 (namely, the SNP Y129064) but negative for the other one (Y112538)—see Figure 1.

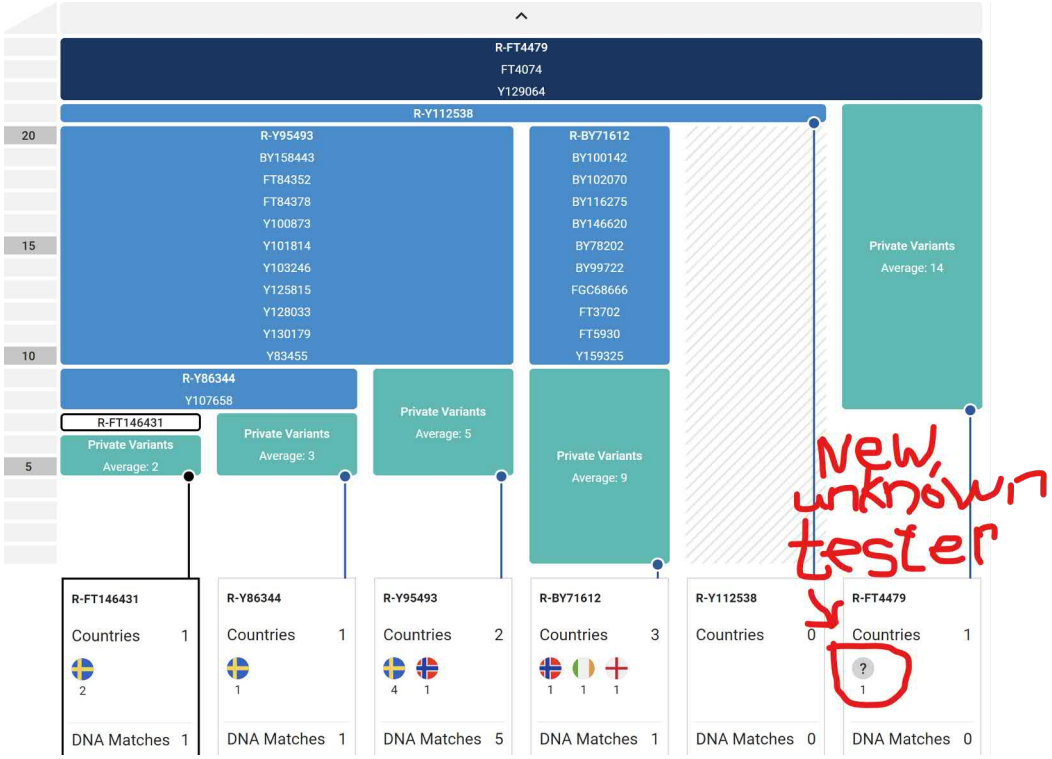


Figure 1: A screenshot of the relevant part of FTDNA's block tree (not showing any names).

Unfortunately, though, this person has not added any information about his ancestor's country of origin or any other information. I do not even have an email address to him, or any other contact

information. I have tried to get in contact with the tester via an administrator in a project that this tester has joined (the administrator was happy to forward my message), but so far I have not had any response from the tester. Now, after half a year, I have almost given up hope. It looks to me as if it will be very difficult for us to learn anything more about this person. But this experience at least serves as a reminder for us that it is possible to find new testers who can help us understand better, for example, where A3 lived. It would be incredibly useful for us to at least know this unknown tester's own country of birth. At the moment, we have no idea. Below I copy in the questions I asked the administrator in the project to forward to the tester.

1. *If you are interested, I would like to let you know about a haplogroup project that is called the Värmland-Hedmark Cluster. You belong to a branch immediately above the group of testers that make up this cluster. The relevant part of the haplotree is shown in the figure on this link:* `https://dl.dropboxusercontent.com/s/nvy82myb1kyzlnh/VHC-tree-only-Dec-2020-v1.pdf?dl=0`. *On the following webpage you can find links to a newsletter with information about the project (four issues have been published so far, the latest one in Dec. 2020):* `https://www.johanlagerlof.com/gengen.html`

2. *What is the name and birth country of your oldest known ancestor along the direct paternal line (your father's father's father's...)? If you don't know, it would be very useful if you could tell me at least where you were born yourself.*

3. *I am curious about whether you really are negative for the SNP Y112538, as Family Tree DNA says you are. Do you know yourself if your Big Y test has confirmed this? If there is any doubt about this, I would be happy to pay for a test of that particular SNP, to have it confirmed or disconfirmed (one can do such at a test at the German company YSEQ).*

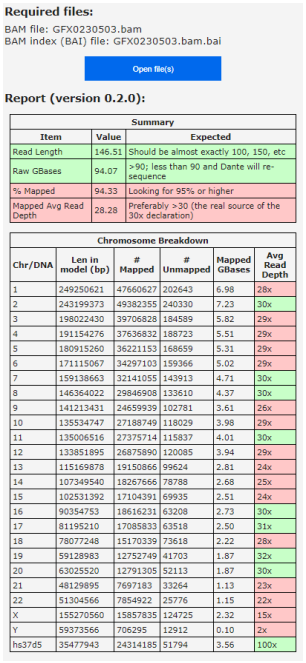
4. *Please consider submitting your BAM file to the company YFull (I presume you have done a Big Y test?). I would be happy to pay YFull's fee (49 dollars, which also includes analysis of an mt-test, if you have one). If you submitted there, it would help us with age estimates, enable me to include you in the tree I linked to above, and be useful to us all for many other reasons. If you get in contact with me, I would be happy to explain how to do the submission practically (what file you need etc.).*

Problems with a Dante Labs WGS Test

As mentioned in the introduction, we have had problems with one of the 30X WGS tests that we recently have submitted to Dante Labs. The test person is the one with a lineage from Nor in Värmland:

- **Lineage L1 (Nor)**. *Oldest known patrilineal ancestor: Anders Bengtsson, born in 1740 in Nor in southern Värmland and dead in 1809.* The test person has previously done a Y37 test with the company Family Tree DNA. The genetic distance to me (Tester 1a) is 4 at the 37-markers level.

The test sample was received by Dante Labs on September 21, 2020, and it was registered as received by the company on October 20. On December 9, the kit status updated to "Sequencing Started," and on January 5, 2021 the results were ready. However, a quality check by me showed that the number of reads on the Y chromosome appeared to be too low—so low that it would be problematic, or impossible, to use it for our kind of research. The left-hand side of Figure 2 shows a report generated by the service `https://qual.iobio.io/`, which indicates that the average read depth on the Y chromosome is only two (i.e., each position on the Y chromosome has, on average, been read two times). The check relies only on a sample of all the information in the raw data files. Still, I also inspected a number of key SNPs in the data, and I found that the number of reads was indeed worryingly low. In my notes I have written: "In the part of the Y chromosome haplotree that is most relevant for our project, there are 12 SNPs that define the branches. By inspecting the BAM file, I have checked the number of reads for



Required files:

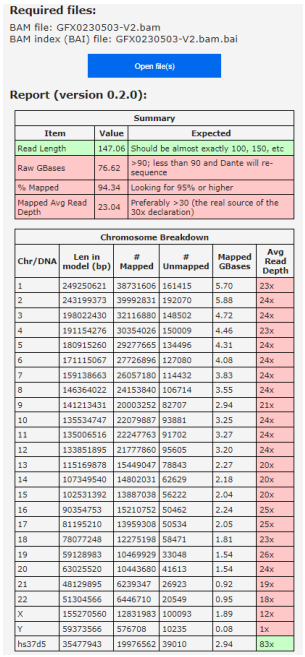
BAM file: GFX0230503.bam
BAM index (BAI) file: GFX0230503.bam.bai

Open file(s)

Report (version 0.2.0):

Summary		
Item	Value	Expected
Read Length	146.51	Should be almost exactly 100, 150, etc
Raw GBases	94.07	>90; less than 90 and Dante will re-sequence
% Mapped	94.33	Looking for 95% or higher
Mapped Avg Read Depth	28.28	Preferably >30 (the real source of the 30x declaration)

Chromosome Breakdown					
Chr/DNA	Len in model (bp)	# Mapped	# Unmapped	Mapped GBases	Avg Read Depth
1	249250621	47660627	202643	6.98	28x
2	243199373	49382355	240330	7.23	30x
3	198022430	39706828	184589	5.82	29x
4	191154276	37636832	188723	5.51	29x
5	180915260	36221153	168659	5.31	29x
6	171115067	34297103	159366	5.02	29x
7	159138663	32141055	143913	4.71	30x
8	146364022	29846908	133610	4.37	30x
9	141213431	24659939	102781	3.61	26x
10	135534747	27188749	118029	3.98	29x
11	135006516	27375714	115837	4.01	30x
12	133851895	26875890	120085	3.94	29x
13	115169878	19150866	99624	2.81	24x
14	107349540	18267666	78788	2.68	25x
15	102531392	17104391	69935	2.51	24x
16	90354753	18616231	63208	2.73	30x
17	81195210	17085833	63518	2.50	31x
18	78077248	15170339	73618	2.22	28x
19	59128983	12752749	41703	1.87	32x
20	63025520	12791305	52113	1.87	30x
21	48129895	7697183	33264	1.13	23x
22	51304566	7854922	25776	1.15	22x
X	155270560	15857835	124725	2.32	15x
Y	59373566	706295	12912	0.10	2x
hs37d5	35477943	24314185	51794	3.56	100x



Required files:

BAM file: GFX0230503-V2.bam
BAM index (BAI) file: GFX0230503-V2.bam.bai

Open file(s)

Report (version 0.2.0):

Summary		
Item	Value	Expected
Read Length	147.06	Should be almost exactly 100, 150, etc
Raw GBases	76.62	>90; less than 90 and Dante will re-sequence
% Mapped	94.34	Looking for 95% or higher
Mapped Avg Read Depth	23.04	Preferably >30 (the real source of the 30x declaration)

Chromosome Breakdown					
Chr/DNA	Len in model (bp)	# Mapped	# Unmapped	Mapped GBases	Avg Read Depth
1	249250621	38731605	161415	5.70	23x
2	243199373	39992831	192070	5.88	24x
3	198022430	32116880	148502	4.72	24x
4	191154276	30354026	150009	4.46	23x
5	180915260	29277665	134496	4.31	24x
6	171115067	27726896	127080	4.08	24x
7	159138663	26057180	114432	3.83	24x
8	146364022	24153840	106714	3.55	24x
9	141213431	20003252	82707	2.94	21x
10	135534747	22079887	93881	3.25	24x
11	135006516	22247763	91702	3.27	24x
12	133851895	21777860	95605	3.20	24x
13	115169878	15449047	78843	2.27	20x
14	107349540	14802031	62629	2.18	20x
15	102531392	13887038	56222	2.04	20x
16	90354753	15210752	50462	2.24	25x
17	81195210	13959308	50534	2.05	25x
18	78077248	12275198	58471	1.81	23x
19	59128983	10469929	33048	1.54	26x
20	63025520	10443680	41613	1.54	24x
21	48129895	6239347	26923	0.92	19x
22	51304566	6446710	20549	0.95	18x
X	155270560	12831983	100093	1.89	12x
Y	59373566	576708	10235	0.08	1x
hs37d5	35477943	19976562	39010	2.94	83x

Figure 2: Left: The original test results (January 5, 2021). Right: The second-round test results (June 18, 2021). The most important piece of information in each table is the number in the right-most column and in the second to last row. This shows that the average number of reads on the Y chromosome was two in the January results and one in the June results.

these SNPs. For six of them, the number of reads is 1; for three SNPs, the number is 2; and the remaining three SNPs have 3, 6, and 7 reads."[8] With these typical numbers of reads, it will also be very hard to identify novel SNPs, not only to check whether the tester is positive or negative for known SNPs.

I complained to Dante Labs and they agreed to re-sequence the sample (although they never explicitly acknowledged that they had made some mistake, nor did they explain to me what they thought had happened). I received the new results on June 18. Again I did a quality check—and the problem remained (and appeared to be even worse)! The report from the service `https://qual.iobio.io/` is shown on the right-hand side of Figure 2.

I tried to communicate with Dante Labs about this issue on their official customer support Facebook page, but their representatives have ignored me. So that is where we are at the moment. I have downloaded all the files and I keep them in a safe place, but I have not yet done anything more with them. One possibility would be to merge the two sets of files (after having checked that they are independent), realign the merged raw data file to the latest reference genome, and then at least *try* to submit it to YFull. However, I am worried that this would not give us an acceptable quality, and that therefore the time, energy, and money required to do these things would be wasted. An alternative, which I am leaning toward, is to simply test this person again, and also this time use Dante Labs (I have two Dante Labs test kits left in my home). Indeed, I have recently asked the test person about this possibility (or, rather, I have asked the close relative of his that I am communicating with and who manages the test person's other tests).

Overall, this experience with Dante Labs has been extremely disappointing. However, my attitude toward these things is pragmatic. The tests that I purchased from Dante Labs were indeed very inexpensive, and one can argue that it is logical that the company therefore must compromise somewhere else (like with their service or their consistency with the quality provided). All the other Dante Labs tests that I have been involved with have eventually been delivered with a good or reasonably good

[8]The reads suggest that this tester is indeed positive for the SNPs defining A2, and thus that he belongs to the Värmland-Hedmark cluster. Moreover, the reads indicate that he is negative for the known branches downstream of A2. Because of the low number of reads, this information is of course not very reliable. Still, given the geographical information that we also have available, I think we can feel very confident that the tester belongs to the Värmland-Hedmark cluster.

quality. Still, the level of clumsiness and lack of professionalism that I have seen from this company has often seemed unnecessary and hard to excuse with their low prices. I have today quite mixed feelings about the company.

If I indeed use one of my two remaining Dante Labs test kits to re-test the representative for lineage L1 (the one discussed above), then there is one test kit left. I think I will try to use this to test someone from lineage L7:

- **Lineage L7 (Östra Tolerud, Grava).** *Oldest known patrilineal ancestor: Anders Jonsson (1745-1819), who died in Östra Tolerud, Grava parish, in the south of Värmland.* This test person has done a Y37 STR test at Family Tree DNA. At the 37-markers level, he matches me (i.e., Tester 1a) with a genetic distance of 3.

I have failed before to get in contact with the Y-37 tester himself, but I have now found several other representatives of the lineage and I will probably contact one of them and ask that person if I may test him. I have made a lot of progress with my research into this lineage, and in the spring I read a useful book about Grava parish, "Grava and Forshaga: Kulturhistoriska anteckningar" (G.M. Sandin, 1930). In that book, the son of Anders Jonsson, Jonas Andersson (1789-1857), is mentioned a number of times.

If we manage to get those two lineages discussed above (i.e., L1 from Nor and L7 from Östra Tolerud in Grava) into our tree, then we will at that point have two lineages left among the ones we know about and suspect belong to the Värmland-Hedmark cluster. These lineages have been discussed in previous issues of the newsletter, but for convenience I list them here again (I feel quite sure that L4 belongs to the Värmland-Hedmark cluster, but I am much less sure about L5):

- **Lineage L4 (Ölme)**. *Oldest known patrilineal ancestor: Erland Eriksson Falk (1863-1937), born in Ölme in south-eastern Värmland.* The person in this lineage who shows up in my match list has done a Y37 test. The genetic distance to me (Tester 1a) is 3 at the 37-markers level.

- **Lineage L5 (Växjö)**. *Oldest known patrilineal ancestor: Peter Engström (1764-1855), born in Telestad, Växjö, Småland.* The person in this lineage who shows up in my match list has done a Y37 test at FTDNA and the genetic distance between him and me is only 1 at the 37-markers level.

Known Errors in No. 4 of the Newsletter

ALERT readers (and in some cases myself) have found some errors in the previous issue of the newsletter. I have already listed these in the VHS Blog (see `https://www.johanlagerlof.com/the-vhc-blog/known-errors-in-no-4-of-the-newsletter`). However, to facilitate for any future readers of the newsletter, I include the list here as well.[9]

- In the tree on page 2, the youngest one of Tester 5's ancestors should be "Nils Pålsson" and not "Nils Persson". (Thanks, Norm, form pointing this out.)

- On page 9, when describing hypothesis H1, I write: "[...] and gave rise to ancestor A2 and eventually to Testers 1-6 and 12-13". Instead of "12-13" it should be "10-13". That is, all testers who have A2 as an ancestor should be included in the list (so all testers in the tree except for Testers 7-9). The same error shows up under H2a-c a bit below on the same page. (Thanks, Bruno, for pointing these errors out.)

- In the tree on page 2 (the box on the top right), Y112538 is incorrectly mentioned twice in the list of SNPs that define A3.

- In the tree on page 10 (showing the Östra Tolerud branch): The two men Jonas Jonsson (1824-?) and Petter Jonsson (1825-1923) are shown as sons to Jonas Andersson (1789-1857). This is wrong.

[9] I have also found some other typos and language issues, but as I do not think they are likely to create misunderstandings, I will not list them here.

Figure 3: Pictures from Kalvbergsåsen in Ransäter parish, not very far from Gersheden where Tester 13's patrilineal ancestors lived. The pictures are taken by me on a bicyle trip on July 24, 2021. Gersheden was actually my target destination, but I failed to find the right roads leading there. I will make a new attempt some other time.

> These two men were born in the previous marriage of Jonas Andersson's wife Kajsa Pehrsdotter (1802-1871). I now have a much more complete version of the tree (where the errors mentioned above of course also are corrected), although it is not yet finished and sufficiently well proofread. I will probably include the new version of the tree in some future issue of the newsletter.

References

Fernow, Erik. 1773-79. *Beskrifning öfwer Wärmeland, 1-3*.

Olausson, Peter. 1997. "Värmlänningar i rörelse: några anteckningar om migrationsrörelser, migrationsforskning och källor till migrationsforskningen avseende Värmland från 1500-talet fram till våra dagar, med tonvikt på emigrationen till Amerika." Forskningsrapport 97:14, Högskolan i Karlstad, Samhällsvetenskap.

Olausson, Peter. 2005. *Naboliv. Värmland, Hedmark. Arkiven berättar.*. Chap. Svenskeradder och baggeruker.: om några källor till gränsflyttarnas historia 31–46: Värmlandsarkiv.

Sandin, J.M. 1930. *Grava och Forshaga: Kulturhistoriska anteckningar*.: Eget förlag, Forshaga.

Svensk sammanfattning

Värmland-Hedmark-klustret är benämningen på en – vad det verkar – väldigt stor släkt som levde i Värmland och Hedmark under medeltiden och kanske även ännu längre tillbaka i tiden. Idag är det många människor, inte minst i Värmland, som härstammar från denna släkt på sitt raka fäderne. En av Värmland-Hedmark-klustrets många grenar leder till exempel till den värmländska släkten Lagerlöf (som författaren till de här raderna råkar tillhöra).

Dessa saker har vi upptäckt med hjälp av DNA-test av nu levande personer. Framförallt har vi varit behjälpta av test av Y-kromosomen, som bara män har och som ärvs från far till son. Arbetet är i full gång med att samla mer kunskap om Värmland-Hedmark-klustret. Det här nyhetsbrevet utkommer med jämna mellanrum (hittills har det blivit i juli och december varje år), och det rapporterar och diskuterar vad som har hänt inom projektet sedan sist.

Exempel på frågor som vi försöker förstå: Hur stor var den här släkten? Var i Värmland och Hedmark fanns den? Hur långt tillbaka i tiden har medlemmar i släkten funnits i Värmland och/eller Hedmark? Varifrån kom släktmedlemmarna när de anlände dit? Kan resultaten hjälpa oss att förstå frågor kring social mobilitet?

- ✔ För fjärde gången i följd har Värmland-Hedmarks-klustret fått en ny gren, jämfört med det föregående numret av nyhetsbrevet. Den här gången kopplar den nya grenen ihop, å ena sidan, två testpersoner som har sin äldsta kända ana (längs det raka fädernet) från Södra Rådom i Nyeds socken med, å andra sidan, en testperson med äldsta kända ana (längs det raka fädernet) från Grue i Hedmark i Norge. Man kan se den senaste versionen av trädet på sidan 2 i nyhetsbrevet. Anfadern som bildar en ny gren har där fått beteckningen A8, och det ser ut som om denne var född under medeltiden någon gång.
- ✔ Vi har ytterligare en ny testperson i trädet, som även han har visat sig tillhöra Värmland-Hedmark-klustret. Testarens farfars far är dock okänd (och spåren slutar i Stockholm), så mer släktforskningsarbete återstår innan vi vet om hans linje går tillbaka till antingen Värmland eller Hedmark.
- ✔ I den granngren till det egentliga Värmland-Hedmarks-klustret som syns till vänster i bilden på sidan 2 har den sista av de tre testarna nu skickat in sin rådata-fil till YFull. Detta är mycket glädjande, då det innebär att det nu kommer att finnas ett bättre underlag för YFulls åldersuppskattningar (eftersom inskickandet gjordes nyligen, har dock YFulls åldersuppskattningar ännu inte tagit denne testare i beaktande).
- ✔ Jag har också själv arbetat med mina egna åldersuppskattningar, men jag är tyvärr inte klar ännu. Skälet till att jag vill göra egna åldersuppskattningar är att jag i alla avseenden inte är riktigt nöjd med det sätt som YFull gör dem på. (Man kan göra den statistiska analysen bättre, menar jag. Det är också önskvärt att inkorporera våra genealogiska data i estimationen. Dessutom skulle jag vilja utföra formella statistiska test av olika hypoteser som vi formulerar, och detta är inte möjligt enbart genom att använda YFulls analys.)
- ✔ Det här numret av nyhetsbrevet innehåller en fortsättning från föregående nummer av en diskussion om Värmland-Hedmark-klustrets ursprung. Bland annat tas möjligheten upp att det kanske går att dra paralleller med andra värmländska medeltida släkter (som Roos af Hjelmsäter och Bratt) som man tror invandrade från Norge. Det är en intressant parallell, men tyvärr finns det fortfarande nog alldeles för lite data för att kunna säga särskilt mycket i frågan.
- ✔ En av personerna som vi har försökt testa hos företaget Dante Labs i Italien är en man med patrilinjära anor från Nors socken i Värmland. Tyvärr har vi haft problem med det testet. De resultat som vi har fått från Dante Labs har alldeles för få läsningar på just Y-kromosomen. Möjligen kommer jag att försöka med ett nytt test av testpersonen (om han är med på det själv, förstås).