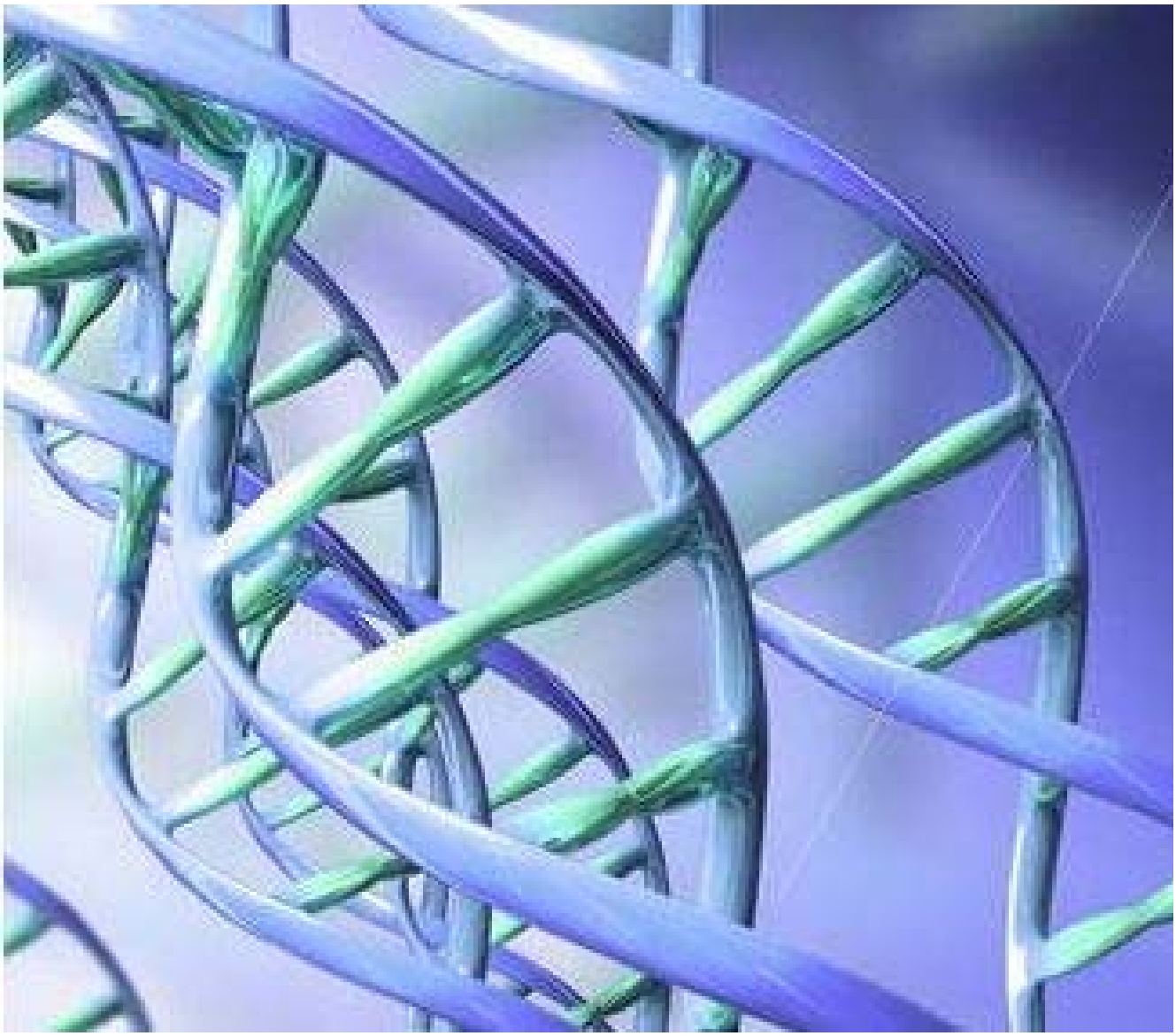


The HAWGOOD FAMILY DNA STUDY

The Hawgood family DNA study aims to connect lines of Hawgoods into the main family tree. The following document explains how this works and the results of the study.

www.hawgoodfamily.co.uk

_____

Where did we come from?

Early Origins

All humans living today have their earliest ancestors around 100,000 years ago in eastern Africa, who then approximately 50,000 years ago migrated to Asia and beyond. Over time, a number of specific genetic groupings emerged defined as Haplotypes. Within each general Haplotype are subgroups and further subgroups which can narrow down specific relationships between surnames.

Hawgood DNA comes from Haplogroup I1-M170, which emerged in Europe about 28,000 years ago. Around 10,000 years later, the ice age had a major impact on our ancestors in Haplogroup I when most of northern and central Europe was uninhabitable. This forced them to retreat to refuge areas in Iberia and the Balkans where living conditions were better.

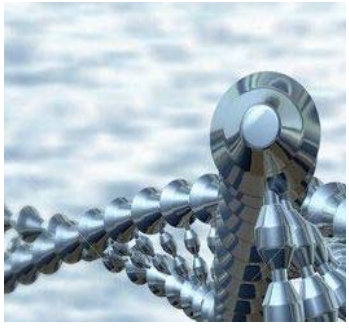
Post ice age

As the Ice Age receded, our ancestors of Haplogroup I1-M170 moved out into the surrounding areas. As Europe was being repopulated, some 8,000 years ago near Denmark, a sub group of I-M170 emerged, I1-M253, of which Hawgoods are a member of.

These ancestors migrated west into the area of the Doggerland land bridge, a piece of land which linked England to Northern Europe, including Denmark and north Germany, and is now covered by the North Sea. It was therefore easy to move into England, where written records of Hawgoods can be traced from around 1500.

Today, the subgroup of I1-M253 is relatively common in Scandinavia at around 35-40% of the population, as shown in the map above, increasing to up to 50% in certain Finnish provinces.

When analysing the specific incidence of the Hawgood values for markers within the overall ‘I’ population, one of the interesting outcomes is that whilst still being in I1-M253, Hawgood DNA has several marker values that are relatively uncommon. Hence finding similar persons that match the other markers, and also with this combination would be quite significant.



What is DNA testing?

An overview

DNA testing can take lots of forms, but genealogical testing only looks at tiny 'junk' fillers in the DNA of the Y-chromosome. These areas are very useful to examine as they rarely change (or mutate) between generations. If these points, or what are known as 'markers', are the same in two males, they are related. The test is known as a Y-STR test, where STR stands for 'short tandem repeat' which is explained below. The test is conducted by sending a cheek swab back to the testing company.

How mutations occur

A mutation is simply a change in the DNA sequence which occurs when a cell is dividing and a certain enzyme fails to copy the DNA correctly. Cells divide through a process called mitosis, where the DNA makes a copy of itself and passes it to the new cell, using an enzyme called DNA Polymerase. This enzyme reads the original code along the chain, and builds the new strand of DNA. The code is made up of a very long pattern of four different nucleotides, abbreviated to A,T,C, and G.

At certain points, the DNA code repeats itself (this is a short tandem repeat, or STR) and sometimes there is a slippage so that for example instead of 7 repeats, we can get 8. This error does not happen very often which makes this kind of analysis very useful to determine relationships and ancestry. An error in these areas does not affect the operation of the cell as it is contained in a junk region of the DNA.

Test result format

The test results themselves are meaningless, except for comparison to another person. The naming convention for each marker is usually a code prefixed by 'DYS', and the test result for each marker is assigned a number equal to the number of times that the DNA sequence is repeated at that location. For example where DYS455=7, the marker point is named DYS455, and the value of 7 means that the DNA code is repeated 7 times. If the value of two people tested at location DYS455 is the same, there is a match. The number of matches, can be used to calculate the TMRCA - the length of time to most recent common ancestor.

The Hawgood family study uses genebase.com examining either 44 or 67 markers.

What do the results tell us?

If two people are closely related, all or nearly all of the markers will be the same. The further apart they are, the more differences, or mutations, will exist. As a rough rule of thumb, if 67 markers are tested, there would be one mutation every 100 or so years.

If we do not know the generation gap between two people, we can calculate the expected gap using the number of mutations and the average mutation rate.

Are you concerned about DNA testing?

Not forensic

Genealogical Y-STR DNA testing is NOT the same as forensic DNA testing undertaken in police investigations. Forensic tests look at genetic profiles that are not held on the Y chromosome and thus are not used in genealogical Y-STR DNA testing.

Not medical

Genealogical Y-STR DNA testing is NOT the same as medical DNA testing which aims at diagnosing genetic disorders looking at active gene portions of our chromosomes.

Not paternity

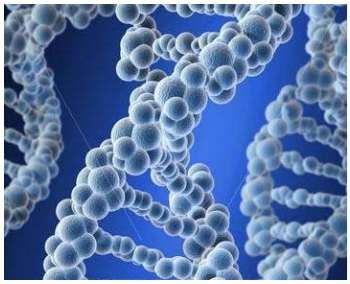
Genealogical Y-STR DNA testing is NOT the same as paternity DNA testing which spreads the test over several chromosomes and is not confined to the Y chromosome as with genealogical testing.

However, genealogical testing can reveal some unexpected male family relationship facts eg two brothers take the same genealogical DNA test, it could be shown that they don't have the same father. Nothing is revealed regarding any family relationships involving females.

Junk fillers

Genealogical Y-STR DNA testing examines tiny sections from the filler (junk) DNA of the Y chromosome which does not yield any direct information about the active genes of the Y chromosome.

Genealogical testing therefore is not used to show genetic disorders caused by abnormalities in genes on other chromosomes or on the Y chromosome.



Mutation rates

Speed is important

The speed of mutation, or what is know as the mutation rate, over which there is much confusion and divided opinion, is the key element in calculating how far apart two people are related.

Early studies (Walsh 2001) indicated an average mutation rate of 0.2% (meaning that in a 44 marker test, a marker would mutate once every 350 years) but later studies a rate of 0.4% (once every 170 years). Some more recent data indicates rates in excess of 0.5% (once every 140 years).

Accuracy is important

The problem with using a rate of 0.2% instead of 0.4% is that the common ancestor calculation will be twice as far apart. Some testing companies however still use the original Walsh value of 0.2% to predict generation gaps. If the input is flawed, then so will be the output. We need therefore to work out an accurate input to be confident of the calculated results.

Some of the difficulties arising are:

- Each marker has a different individual mutation rate, some fast, some slow
- Different testing companies include different markers in their tests, which means that the average rate of mutation from each company will differ. For example there are a larger number of fast mutating markers in the FTDNA 37 marker test, than the Genebase 44 marker test.
- Some studies have small sample sizes and this does not always provide reliable data

Overcoming the inconsistencies

There are tens of thousands of test results and more being added all the time. The good news is that studies are revealing mutation rates which are broadly consistent. A slow mutating marker may have a variance between different studies, but it will still be, broadly speaking a slow mutating marker. By the same token a fast mutating marker in one study does not suddenly become a slow mutating marker in another, it still remains a fast marker across different tests.

Combine multiple databases

We have taken results for individual markers from as many sources as possible, from wide private studies to published testing company data, and then for each individual marker taken the average of all sources. For each marker we have a specific mutation rate which is based on a massive amount of data (some 75-100,000 results). With such a large database, the calculation of the most common recent ancestor becomes much more reliable.

Current average mutation rates

The database is constantly being refined, but currently, the Genebase 44 marker test has an average mutation rate of approx 0.0027, or 0.27%, and the 67 marker test, approx 0.0041, or 0.41%. If you use a different company than Genebase, then you can still use the data on our website for each marker and using your company's specific markers, overlay this data and calculate an average mutation rate for your particular test.

Calculating how many mutations to expect

A simpler explanation

The concept of marker mutations confuses many at first glance. But it is relatively simply and can be explained using the technique of 'expected values', which is the long-run value taken over many independent repetitions.

A lottery

Consider a lottery draw that has numbered balls from 1 to 300. You have bet on number 50. When the draw is held you would expect your number would be unlikely to come up. Your odds are 1/300, or 0.33%

Say that in the lottery, not one, but 44 balls are drawn. The odds of your number coming up are now 44/300 or roughly 1/7. If you then do the same 44 draw every day for a week, you would expect that your number would come up statistically 7 times x 1/7 odds = 1, or in other words you would expect to win once in that week.

Apply to STR markers

Now consider that the 44 balls are in fact the STR markers and that the 7 days are 7 generations. Using the same calculation, you would expect during the 7 generations that 1 mutation would occur on your own 44 markers.

When comparing two people, each one could have 1 mutation, so the total expected mutations must be double that for one person, which is 2 mutations. So comparing two people over 7 generations with 44 markers tested, the expected number of mutations is 2.

This can be distilled to a formula:

No of Markers tested x (No of generations x number of persons tested) x average mutation rate.

Sometimes the phrase 'No of generations x number of persons tested' is rewritten as 'No of transmission events'. The transmission event value is simply the number of generations x the number of persons being compared. So for two 1st cousins, the most recent ancestor would be their grandfather. This would create 2 transmission events each, being grandfather to father to son for each cousin. The total transmission events would be 4.

This formula can be used to create a table of expected mutations. Using the table, for example, in a 63 marker test, using a mutation rate of 0.41%, the number of expected mutations where 11 generations exist between two parties tested, would be 6 (rounded as 5.7 mutations is not possible). The expected result would therefore be 57/63 matching.

Rate	0.27	0.41
Markers	43	63
Generations	Expected mutations	
1	0.2	0.5
2	0.5	1.0
3	0.7	1.6
4	0.9	2.1
5	1.2	2.6
6	1.4	3.1
7	1.6	3.6
8	1.9	4.2
9	2.1	4.7
10	2.3	5.2
11	2.6	5.7
12	2.8	6.2
13	3.0	6.8
14	3.3	7.3
15	3.5	7.8
16	3.8	8.3
17	4.0	8.9

More on the maths

Likelihood of expected outcome

The expected result may not necessarily be the actual result. Throwing six dice should produce one six, but it may not. By the same token, in our 44 marker test over 7 generations with an average mutation rate of 0.33%, we may not see the expected result of 2 mutations.

How likely is the expected result? We can calculated this by using the following formula :

$$(1 - P)^{(T-t)} \times P^{(t)} \times \frac{T!}{(T-t)! \times t!}$$

Where :

P = the mutation rate (quoted as 0.0033 and not a percentage at 0.33)

T= the total number of markers tested multiplied by the number of generations x the number of people in the comparison

t = the number of markers that have mutated

! = factorial (6! is the same as 6x5x4x3x2x1)

To work out the probability of 2 mutations in 44 markers over 7 generations, between 2 people using a mutation rate of 0.33%:

P = 0.0033
T = 44 x 7 x 2 = 616
t = 2

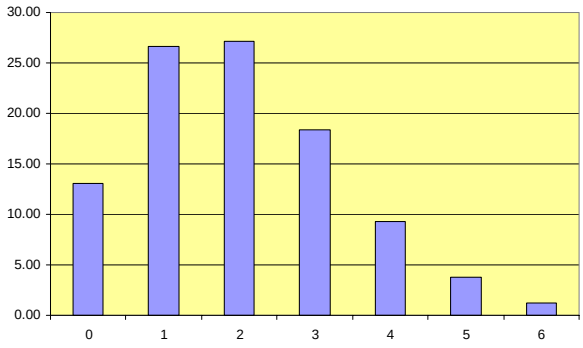
Plugging the numbers in to the equation

= (1-0.0033)^(616-2) x 0.0033^2
x (616!/614!)/2!
= 0.0131394 x 0.00001089
x (616x615)/(2x1)
= 0.0271 or 27.1%

Note that a calculator will not compute 616!, as the output is too big to handle. 616!/614! is the same as 616x615. Using another example, 64!/60! is the same as 64 x 63 x 62 x 61.

The table below shows the percentage chance of 0,1,2,3,4,5, and 6 mutations in a test of 44 markers where the two parties are 7 generations apart. This shows that whilst 2 mutations is the most likely outcome, 1 mutation would not be comparatively unlikely.

Percentage chance of number of mutations



Care with one set of results

Using a small dataset is not always reliable and in order for results to be meaningful, one must consider any one set of results in the context of other results.

The more participants, the more likely that the expected results statistically will pan out in reality. Expected maths always works in the long run but needs a significant dataset.

To avoid manual calculations see the website which provides a calculator for the above.

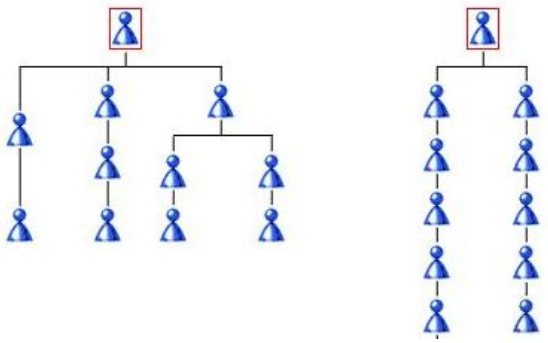
Caution on transmission events

Counting transmission events

There is not need to consider this when two people are being examined. The number of transmission events is simply the number of generations x 2.

Where there are more than two people being examined, the number of transmission events is not necessarily the number of generations multiplied by the number of people tested.

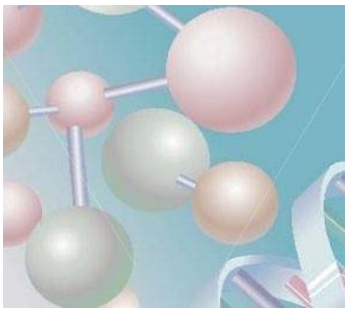
The diagram shows two examples where there are 10 transmission events.



The example on the right has the same common ancestor and the transmission events are simply 5 x 2. However, if those being tested do not have the same earliest common ancestor, as in the left example, then the calculation is not as straight forward. In this example there are four people tested where the transmission events are not 12 (4 x 3) but are still 10.

This is as the third and fourth people share a common ancestor one generation later than the first and second.

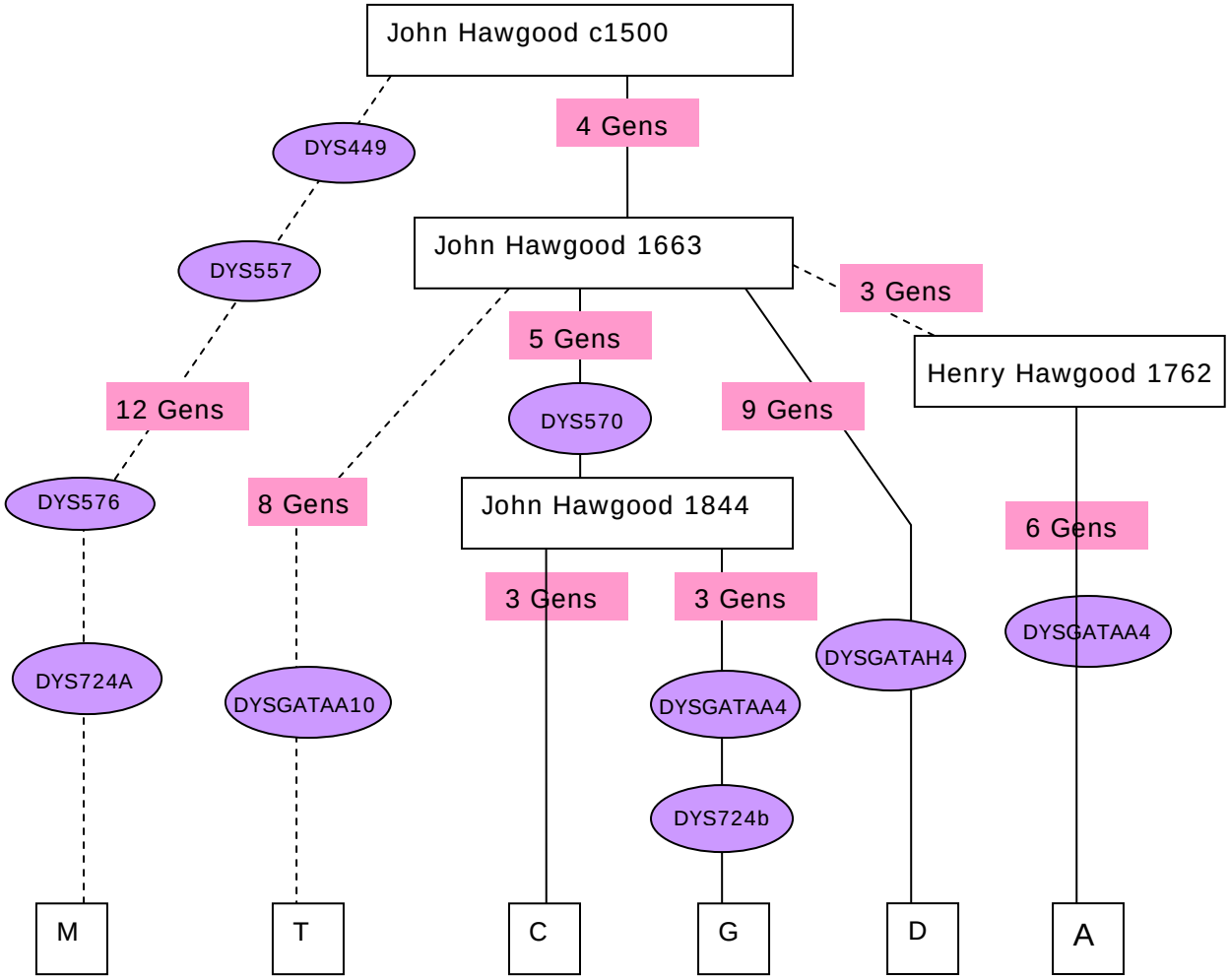
Therefore to calculate the number of transmission events for more than two people, where they do not all share the same common ancestor, the events must be counted manually to ensure that there is no double counting.



The Results

Below are the summary results for six tested in the Hawgood family DNA study.
This diagram shows where and which mutations occurred (purple ovals).

The pink boxes show the number of generations between various ancestors. The dotted lines show links derived from the DNA results, with the unbroken lines showing known links. The test results have enabled us to connect two broken branches into the main family tree. One thing to note is that whilst there are 4 mutations from John 1500 to M, any one of these could have occurred instead between John 1500 and John 1663



- The dotted lines show common ancestors derived from the results
- The unbroken lines show known ancestors
- Pink text boxes show generation gaps (Gens)
- Purple ovals show which markers mutated and in which period
- Note that any of the four mutations between John Hawgood c 1500 and M could have also occurred between John Hawgood c 1500 and John Hawgood 1663

Detailed results

	Marker test	Deduced Core DNA	M	C	T	D	G	A	Haywood	
DYS19	44	15	15		15	15	15	15	15	Same
DYS385a	44	13	13	13	13	13	13	13	13	Same
DYS385b	44	14	14	14	14	14	14	14	14	Same
DYS388	44	14	14	14	14	14	14	14	14	Same
DYS389i	44	13	13	13	13	13	13	13	13	Same
DYS389II	44	30	30	30	30	30	30		30	Same
DYS390	44	22	22	22	22	22	22	22	22	Same
DYS391	44	10	10	10	10	10	10	10	10	Same
DYS392	44	11	11	11	11	11			11	Same
DYS393	44	13	13	13	13	13	13	13	13	Same
DYS426	44	11	11		11	11	11	11	11	Same
DYS437	44	16	16	16	16	16	16	16	16	Same
DYS438	44	10	10	10	10	10	10	10	10	Same
DYS442	44	12	12	12	12	12	12	12	12	Same
DYS445	44	11	11	11	11	11	11	11		Same
DYS446	44	12	12	12	12	12	12	12		Same
DYS447	44	23	23	23	23	23	23	23	23	Same
DYS448	44	20	20	20	20	20	20	20	20	Same
DYS449	44	29	30	29	29	29	29		29	Diff
DYS453	44	11	11	11	11	11	11	11	11	Same
DYS454	44	11	11	11	11	11	11	11	11	Same
DYS455	44	8	8	8	8	8	8	8	8	Same
DYS456	44	14	14	14	14	14	14	14	14	Same
DYS458	44	15	15	15	15	15	15	15	15	Same
DYS459a	44	8	8	8	8	8	8	8	8	Same
DYS459b	44	9	9	9	9	9	9	9	9	Same
DYS460	44	11	11	11	11		11	11	11	Same
DYS461	44	12	12	12	12	12	12	12		Same
DYS462	44	12	12	12	12	12	12	12		Same
DYS468	44	27	27	27	27	27	27	27		Same
DYS484	44	13	13	13	13	13	13	13		Same
DYS522	44	11	11	11	11	11	11	11		Same
DYS527a	44	20	20	20	20	20	20	20		Same
DYS527b	44	21	21	21	21	21	21	21		Same
DYS531	44	11	11	11	11	11	11	11		Same
DYS557	44	16	17	16	16	16	16	16		Diff
DYS588	44	19	19	19	19	19	19	19		Same
GATAA10	44	13	13	13	12	13	13	13		Diff
GATAA4	44	12	12	12	12	12	11	11	11	Diff
GATAC4	44	22	22	22	22	22	22	22		Same
GATAH4	44	11	11	11	11	12		11	11*	Diff
YCAIIa	44	19	19	19	19	19	19	19	19	Same
YCAIIb	44	21	21	21	21	21	21	21	21	Same
DYS464a	44	12	14	12	12	14	12	12	12	
DYS413a	-	23		23		23	23			
DYS413b	-	25					25			
DYS436	67	12	12	12	12		12			Same
DYS444	67	13	13	13	13		13			Same
DYS452	67	31	31	31	31		31			Same
DYS463	67	21	21	21	21		21			Same
DYS464b	67	14	15	14	14		14	14	14	
DYS464c	67	15	16	15	15		15	15	15	
DYS464d	67	16		16	16		16	16	-	
DYS472	67	8	8	8	8		8			Same
DYS481	67	25	25	25	25		25			Same
DYS511	67	9	9	9	9		9			Same
DYS518	67	21	21	21	21		21			Same
DYS520	67	21	21	21	21		21			Same
DYS537	67	11	11	11	11		11			Same
DYS570	67	19	19	18	19		18		19	Diff
DYS576	67	18	17	18	18		18		18	Diff
DYS590	67	8	8	8	8		8			Same
DYS607	67	13	13	13	13		13		13	Same
DYS612	67	35	35	35						Same
DYS614	67	29	29	29	29		29			Same
DYS644	67	16	16	16	16		16			Same
DYS710	67	14	14	14	14		14			Same
DYS711	67	33	33	33	33		33			Same
DYS724a	67	32	31	32	32		32		32 **	Diff
DYS724b	67	36	36	36	36		35		36 **	Diff

Generation gaps - actual versus predicted

We have compared the results of sixparticipants. The actual number of marker differences are shown in the first set of tables, which are used to predict the expected generation gap in the second tables. Finally, the actual generation gap where known is shown in eight instances. Our model correctly predicts the generation gap in 6 out of 8, or 75% of the time. For the other 2, the margin of error is only 2 generations.

44 Test - Number of mutations

	M	C	T	G	D	A
M	-	2	3	3	3	2
C	2	-	1	1	1	1
T	3	1	-	2	2	2
G	3	1	2	-	2	0
D	3	1	2	2	-	2
A	2	1	2	0	2	-

44 Test - Expected generation gap

	M	C	T	G	D	A
M	-	7-10	11-15	11-15	11-15	7-11
C	7-10	-	3-6	3-6	3-6	3-6
T	11-15	3-6	-	7-10	7-10	7-11
G	11-15	3-6	7-10	-	7-10	1-2
D	11-15	3-6	7-10	7-10	-	7-11
A	7-11	3-6	7-11	1-2	7-11	-

Known generation gap

	M	C	T	G	D	A
M	-	?	?	?	?	?
C	?	-	?	3	8	?
T	?	?	-	?	?	?
G	?	3	?	?	8	?
D	?	8	?	8	-	?
A	?	?	?	?	?	-

67 Test - Number of mutations

	M	C	T	G
M	-	5	5	7
C	5	-	2	2
T	5	2	-	4
G	7	2	4	-

67 Test - Expected generation gap

	M	C	T	G
M	-	9-10	9-10	13-14
C	9-10	-	3-4	3-4
T	9-10	3-4	-	7-8
G	13-14	3-4	7-8	-

Known generation gap

	M	C	T	G
M	-	?	?	?
C	?	-	?	3
T	?	?	-	?
G	?	3	?	-

Denotes where the expected generation gap matches the known generation gap

Observations and conclusions

- M is related to all others tested at a distance of most likely 11-12 generations. This would mean a common ancestor of between 1500-1600, being John Hawgood of Church Brampton
- C has 2 mutations with G in the 67 marker test, which implies a gap of 3-4 generations. The known gap is 3 generations, so the result is as expected
- The generation gap of T to others tested is not known. When comparing T to C, G, and D, the results are not entirely consistent. Using the 67 test results, the generation gap between T and the others can be estimated at between 7-10 generations. Using this, and existing knowledge, the common ancestry of T with C&G can be estimated to be around 1663
- The generation gap of D to G and D to C is known to be 8 generations. The result between G & D in the 44 marker test is consistent with the actual gap, but there is one less mutation than expected between D & C
- The same mutation exists on DYS 570, a known fast mutating marker in both C and G, but not the other three tested. This evidences that the mutation occurred prior to the common ancestor of C&G, being John Hawgood b 1844, and after the most common ancestor of both C&G to D, being John Hawgood b 1663. It should be noted that there is a very remote possibility that the same mutation occurred in both C&G after 1844, but statistically highly unlikely

The HAWGOOD mutation rate

As there are more than 2 people being compared - in this case 6, we must count the number of transmission events to calculate the average mutation rate.

For the 6 in the test, there are a total of 2961 transmission events, with 10 mutations. This means that the average Hawgood mutation rate is 10/2691 = 0.34%. If we include DYS464, which is normally excluded from calculations (two people have mutations here), the average rate is closer to 0.40%.

We can also use our calculator on the website to see how likely 10 mutations actually is, over 2961 transmission events.

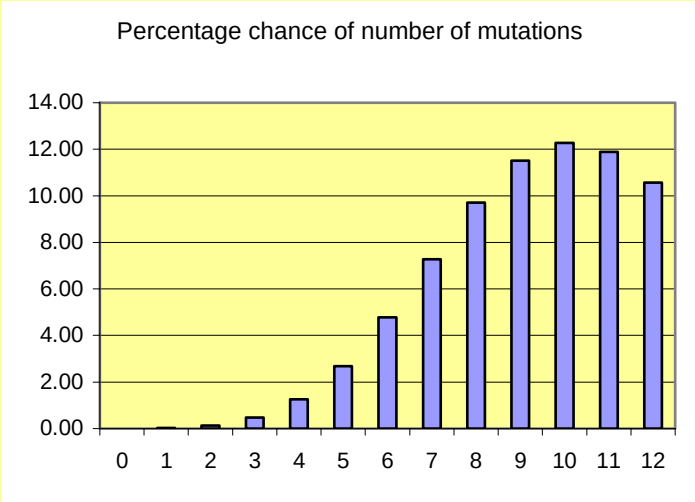
Four people took the 67 marker test, and 2 the 44 marker test. Based on the mutation rate database, the statistical chance of a single mutation is ((4*0.41)+(2*0.27))/6 = 0.36%.

Using 0.36%, we can check to see how likely the actual result of 10 mutations is. The graph below shows that the most likely outcome is infact 10 mutations at 12.27%, although it should be noted that 9 or 11 is almost as likely.

When using the calculation for more than 2 people, the calculator needs to be forced by entering 'markers' as transmission value, and 'generations' & 'numbers tested' as '1' . This then forces the value for 'Total opportunities for mutations' to be the transmission event value. This creates in effect the throwing of the dice 2961 times and calculates how many times a number is likely to come up.

Markers	2961	Chance of 1 mutation *	0.0036
Generations	1	Chance of no mutation	0.9964
Numbers tested **	1		
Transmission events	1	* Chance of 1 mutation is the average mutation rate for your test	
Total opportunities for mutation	2961	The average for a Genebase 67 marker test is 0.0041 or 0.41%	

No Mutations	% Chance
0	0.00
1	0.02
2	0.13
3	0.47
4	1.25
5	2.68
6	4.77
7	7.28
8	9.71
9	11.51
10	12.27
11	11.90
12	10.57



Appendix A - Forecasting table

The table below can be used to estimate the generation gap based on the number of mutations, shown in pink highlight. The rates shown are in percentages, so 0.27% is the same as 0.0027.

Statistically likely generation gaps based on actual mutations

Rate	0.27	0.41
Markers	43	63
Generations	Expected mutations	
1	0.2	0.5
2	0.5	1.0
3	0.7	1.5
4	0.9	2.1
5	1.2	2.6
6	1.4	3.1
7	1.6	3.6
8	1.9	4.1
9	2.1	4.6
10	2.3	5.2
11	2.6	5.7
12	2.8	6.2
13	3.0	6.7
14	3.3	7.2
15	3.5	7.7
16	3.7	8.3
17	3.9	8.8

Actual mutations

	1	2	3	4
1	0	0	43 Markers	
2	0	0	0	0
3	1	1	1	1
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	2	2	2	2
8	2	2	2	2
9	2	2	2	2
10	2	2	2	2
11	3	3	3	3
12	3	3	3	3
13	3	3	3	3
14	3	3	3	3
15	3	3	3	3
16	4	4	4	4
17	4	4	4	4

	1	2	3	4	5	6	7	8	
1	1	1	1	1	63 Markers				1
2	1	1	1	1	1	1	1	1	2
3	2	2	2	2	2	2	2	2	3
4	2	2	2	2	2	2	2	2	4
5	3	3	3	3	3	3	3	3	5
6	3	3	3	3	3	3	3	3	6
7	4	4	4	4	4	4	4	4	7
8	4	4	4	4	4	4	4	4	8
9	5	5	5	5	5	5	5	5	9
10	5	5	5	5	5	5	5	5	10
11	6	6	6	6	6	6	6	6	11
12	6	6	6	6	6	6	6	6	12
13	7	7	7	7	7	7	7	7	13
14	7	7	7	7	7	7	7	7	14
15	8	8	8	8	8	8	8	8	15
16	8	8	8	8	8	8	8	8	16
17	9	9	9	9	9	9	9	9	17

Generation gap

For example, in a 43 Markers test with 3 mutations, at a 0.27% mutation rate, the generation gap is likely to be between 11 and 15 generations.

Appendix B - Individual marker mutation rates

Summary average data | **Individual data surveys**

Genebase Marker test		Average	Median	1 Janzen (Jun 2008)	2 Chandler	3 www.cstl.nist.gov (2007)	4 McDonald	5 SMGF.org 7976 dataset	6 YHRD.org	7 Sorenson	8 Burgarella 2010	9 ASHG 2004	10 Bioinformetrics	11 Ballantyne etc 17 markers 2009	12 Ballantyne etc 186 markers Sep 2010	13 Worldfamilies (McD rates)
DYS19	44	**0.0023**	0.0023		0.0015	0.0025	0.0016	0.0015	0.0023	0.0015	0.0028		0.0030	0.0027	0.0044	0.0016
DYS385a	44	**0.0032**	0.0028	0.0023		0.0021	0.0033	0.0056	0.0021	0.0057		0.0034		0.0021	0.0021	0.0033
DYS385b	44	**0.0034**	0.0033	0.0023		0.0021	0.0033	0.0056	0.0021	0.0057		0.0034		0.0021	0.0041	0.0033
DYS388	44	**0.0005**	0.0005		0.0006		0.0005	0.0004		0.0002	0.0005		0.0008			0.0005
DYS389i	44	**0.0025**	0.0022		0.0019	0.0024	0.0021	0.0022	0.0025	0.0022	0.0022		0.0020	0.0028	0.0055	0.0021
DYS389ii	44	**0.0028**	0.0028		0.0024	0.0035	0.0028	0.0026	0.0036	0.0027	0.0025	0.0011	0.0032	0.0031	0.0038	0.0028
DYS390	44	**0.0034**	0.0032		0.0031	0.0025	0.0045	0.0044	0.0021	0.0053	0.0047	0.0034	0.0028	0.0022	0.0015	0.0045
DYS391	44	**0.0027**	0.0029		0.0027	0.0028	0.0036	0.0032	0.0026	0.0041	0.0020	0.0011	0.0010	0.0030	0.0032	0.0036
DYS392	44	**0.0014**	0.0012		0.0005	0.0008	0.0016	0.0015	0.0041	0.0016	0.0005	0.0023	0.0008	0.0006	0.0009	0.0016
DYS393	44	**0.0015**	0.0014		0.0008		0.0014	0.0011	0.0010	0.0014	0.0026	0.0023	0.0018	0.0010	0.0021	0.0012
DYS426	44	**0.0003**	0.0003		0.0001		0.0005	0.0003			0.0005		0.0001		0.0000	0.0005
DYS437	44	**0.0017**	0.0017		0.0010		0.0020	0.0018	0.0012	0.0021	0.0023		0.0014	0.0013	0.0015	0.0020
DYS438	44	**0.0008**	0.0008		0.0006	0.0007	0.0012	0.0010	0.0003		0.0007		0.0012	0.0006	0.0010	0.0012
DYS442	44	**0.0050**	0.0041		0.0032						0.0019	0.0049			0.0098	
DYS445	44	**0.0026**	0.0025	0.0030				0.0030			0.0025	0.0024			0.0022	
DYS446	44	**0.0035**	0.0031	0.0037			0.0032	0.0031			0.0007	0.0075			0.0027	
DYS447	44	**0.0030**	0.0029		0.0026		0.0045	0.0031		0.0040	0.0007	0.0023			0.0021	0.0045
DYS448	44	**0.0017**	0.0016		0.0014	0.0011	0.0028	0.0024	0.0016		0.0016	0.0011		0.0006		0.0028
DYS449	44	**0.0088**	0.0081		0.0084		0.0056	0.0065		0.0078	0.0096	0.0124			0.0122	0.0075
DYS453	44	**0.0022**	0.0022								0.0022					
DYS454	44	**0.0007**	0.0005	0.0002			0.0005	0.0002		0.0005	0.0021					0.0005
DYS455	44	**0.0007**	0.0005	0.0002			0.0005	0.0003		0.0005	0.0021					0.0005
DYS456	44	**0.0054**	0.0051		0.0074	0.0053		0.0053	0.0042		0.0033	0.0083		0.0044	0.0049	
DYS458	44	**0.0073**	0.0066	0.0081		0.0106	0.0066	0.0058	0.0064	0.0063	0.0048	0.0090			0.0084	0.0066
DYS459a	44	**0.0021**	0.0024		0.0013		0.0014	0.0024		0.0026					0.0027	
DYS459b	44	**0.0021**	0.0024		0.0013		0.0014	0.0024		0.0026					0.0027	
DYS460	44	**0.0034**	0.0028		0.0040		0.0028	0.0029			0.0025	0.0023			0.0062	0.0028
DYS461	44	**0.0023**	0.0026				0.0028	0.0023			0.0030				0.0010	
DYS462	44	**0.0015**	0.0011				0.0005	0.0005			0.0028	0.0011			0.0027	
DYS468	44	**0.0017**	0.0017												0.0017	
DYS484	44	**0.0027**	0.0027								0.0028				0.0026	
DYS522	44	**0.0026**	0.0023				0.0045				0.0023				0.0010	
DYS527a	44	**0.0065**	0.0065												0.0065	
DYS527b	44															
DYS531	44	**0.0012**	0.0010	0.0004							0.0023				0.0010	
DYS557	44	**0.0034**	0.0034	0.0032			0.0036				0.0031				0.0038	
DYS588	44	**0.0004**	0.0004								0.0004					
GATAA10	44	**0.0034**	0.0038	0.0038			0.0045	0.0038			0.0030	0.0011			0.0033	0.0045
GATAA4	44	**0.0048**	0.0048	0.0048	0.0048	0.0061	0.0045	0.0042	0.0052	0.0049	0.0029	0.0068	0.0037	0.0056	0.0038	0.0045
GATAC4	44	**0.0034**	0.0035	0.0024		0.0046	0.0028	0.0023	0.0035		0.0028	0.0045	0.0039		0.0039	
GATAH4	44	**0.0032**	0.0031		0.0029	0.0043	0.0036	0.0030	0.0024		0.0022	0.0034	0.0031	0.0031	0.0032	0.0036
YCAIIa	44	**0.0016**	0.0014		0.0012		0.0014	0.0025								0.0014
YCAIIb	44	**0.0016**	0.0014		0.0012		0.0014	0.0025								0.0014
DYS464a	44	**0.0050**	0.0046		0.0057		0.0035								0.0073	0.0035
Average - 44		**0.0028**	**0.0027**	**0.0028**	**0.0025**	**0.0034**	**0.0027**	**0.0027**	**0.0028**	**0.0033**	**0.0024**	**0.0040**			**0.0036**	**0.0028**
DYS413a	67	**0.0020**	0.0020	0.00202												
DYS413b	67															
DYS436	67	**0.0003**	0.0004	0.00018							0.00044				0.0004	
DYS444	67	**0.0029**	0.0026	0.00321			0.00224	0.003			0.002	0.0012			0.00545	
DYS452	67	**0.0033**	0.0040					0.00174			0.004				0.00402	
DYS463	67	**0.0017**	0.0016				0.00204	0.0016			0.0007	0.0025			0.00151	
DYS464b	67	**0.0050**	0.0046		0.00566		0.0035								0.00727	0.0035
DYS464c	67	**0.0050**	0.0046		0.00566		0.0035								0.00727	0.0035
DYS464d	67	**0.0055**	0.0057		0.00566		0.0035								0.00727	
DYS472	67	**0.0002**	0.0002	0.00001							0.0004					
DYS481	67	**0.0058**	0.0054	0.00544							0.0069				0.00497	
DYS511	67	**0.0017**	0.0015	0.00128							0.0024				0.00152	
DYS518	67	**0.0184**	0.0184												0.0184	
DYS520	67	**0.0024**	0.0024	0.00245			0.00216				0.0023				0.0027	
DYS537	67	**0.0019**	0.0024	0.00057							0.0028				0.0024	
DYS570	67	**0.0075**	0.0079	0.00838	0.0079		0.004778				0.0042				0.0124	
DYS576	67	**0.0086**	0.0079		0.01022		0.00552				0.0042				0.0143	
DYS590	67	**0.0005**	0.0005	0.00054							0.0004					
DYS607	67	**0.0039**	0.0039		0.00411						0.0037					
DYS612	67	**0.0145**	0.0145												0.0145	
DYS614	67	**0.0043**	0.0043												0.00432	
DYS644	67	**0.0032**	0.0032												0.00322	
DYS710	67	**0.0175**	0.0175	0.0175												
DYS711	67															
DYS724a	67	**0.0353**	0.0353		0.03531											
DYS724b	67															
Average - 67		**0.0041**	**0.0040**													

- For DYS724 and 710/11 mutation rates may be over both markers in each set - hence for prudence, only once value is used
- DYS 464 ignored for average rate calculation
- DYS413a FTDNA at https://docs.google.com/viewer?a=v&pid=explorer&chrome=true&srcid=0By9Y3jb2fORNY2ZjZWM4OGItZjI2Yy00NDQwLWIyYzMtZmUwY2ZlYjFiZı

Mutation rate sources

No	Sample Size	Source
1	8430	Chandler values taken from http://www.timjanzen.com/variance_calculator.xls where not in jogg.info/22/chandler
2	15295	http://www.jogg.info/22/Chandler.pdf http://www.jogg.info/22/Chandler.htm
3	5000	http://www.cstl.nist.gov/biotech/strbase/ystr_fact.htm Sample size estimate
4	500	McDonald (2004-6) Sample size estimate
5	7,976	http://www.smgf.org/resources/papers/ASHG2004-4.pdf
6	6364-25306	http://www.yhrd.org/Research/Loci
7	5,000	http://www.smgf.org/resources/papers/ASHG2004-4.pdf (Ysearch.org dataset) Sample size estimate
8	80-13948	https://sites.google.com/site/navascuesresearch/publications-conferences/journalpublications/burgarellanavascues2010
9	864	http://www.smgf.org/resources/papers/ASHG2004-3.pdf
10	3780	http://bioinformatics.oxfordjournals.org/content/26/18/i440.full.pdf+html
11	3384-11900	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2766043/table/Tab2/
12	1700	http://www.cell.com/AJHG/supplemental/S0002-9297(10)00419-2 and at http://www.sciencedirect.com/science/article/B8JDD-50XJR
13		http://www.worldfamilies.net/marker
Total circa	74045	