

Group H Mutation Rate - Butler Surname DNA Project

Jerry L. Butler, March 26, 2007

The DNA results of Group H of the Butler Surname DNA Project appears to have a high mutation rate among certain subgroups. It is also difficult to see how the different subgroups relate to each other. A literature review was undertaken to determine if there was a methodology that could be applied to the analysis of the Group H results to provide more insight from the DNA markers.

The tutorial section of www.familytreedna.com provided a link to Charles Kerchner's great reference material at www.kerchner.com/dna-information.htm. Charles is conducting a research project to collect the mutation rates of the various surname projects. Most of the analysis presented here is based on the methodology of Charles Kerchner.

The DNA markers for Group H were copied into an Excel spreadsheet. From this spreadsheet, all the results were excluded that only had 12 markers leaving 25 members of Group H. All had at least 37 markers except for one that had only 25. Columns were added for the known ancestors of each project member starting with the earliest Proven Butler Ancestor down to the project member. In several cases, I extended the earliest Proven Butler Ancestor with likely ancestors by comparing the information in the different family histories provided by the project members – the histories were most helpful in doing this. In a couple of cases, project members provided additional information to confirm that the descendant lines were correct.

The spreadsheet was then sorted in descendant order which grouped people with common lines together. This analysis yielded eight subgroups of Group H grouped by the patriarch of that line.

Subgroup	Patriarch	Birth	No.
H-1	Benjamin		1
H-2	Daniel Mangrum	1832 SC	1
H-3	Edmund	1694-1704	4
H-4	Gabriel	1779 SC	1
H-5	James Anthony	1740 SC	1
H-6	James	1730 Ireland	2
H-7	Joseph James	1820 SC	3
H-8	Zachariah	1736 VA	10

The Zachariah subgroup was further subdivided into three subgroups.

Subgroup	Patriarch	Birth	No.
H-8a	Joel	1766 VA	1
H-8b	James	1758 VA	5
H-8c	George Snellings	1808 GA	4

These groups have been renumbered from the earlier incomplete analysis. The additional analysis presented here warrants a restart of the numbering for the various subgroups of Group H.

The triangulation method developed by Charles Kerchner was used to deduce the Ancestral Haplotype of the patriarch of each group. This method looks at each marker of people with a common ancestor and selects the marker that is most likely for the common ancestor. If the markers match, then the resulting markers are of high confidence. Otherwise, it is just the most likely haplotype for the Most Recent Common Ancestor. Please consult Appendix A for the results of the triangulation method for each subgroup and the resulting MRCA haplotypes.

For each subgroup, the markers for each member were compared with those of the patriarch to determine the number of Unique Mutation Events. Some of the mutations were reflected in the markers of multiple project members so each mutation was only counted one time to compile an accurate counting of Unique Mutation Events. The number of Unique Transmission Events was also counted to determine the number of opportunities for mutation. The lines of the project members overlapped in many cases so the Unique Transmission Events were also only counted a single time.

With this data, the mutation rates were calculated using the formula:

$$\text{Mutation Rate} = UME \div UTE \div HS$$

Where **UME** is the number of Unique Mutation Events and **UTE** is the number of Unique Transmission Events and **HS** is the Haplotype Size or the number of markers tested which is 37 in this case.

The mutation rates were calculated for the various subgroups as well as the overall rate for group H and are presented in the following table:

Subgroup	Patriarch	Birth	No.	Mutation Rate
H-1	Benjamin		1	
H-2	Daniel Mangrum	1832 SC	1	
H-3	Edmund	1694-1704	4	0.45%
H-4	Gabriel	1779 SC	1	
H-5	James Anthony	1740 SC	1	
H-6	James	1730 Ireland	2	0.25%
H-7	Joseph James	1820 SC	3	0.30%
H-8	Zachariah	1736 VA	10	0.31%
H-8a	Joel	1766 VA	1	0.77%
H-8b	James	1758 VA	5	0.00%
H-8c	George Snellings	1808 GA	4	0.54%
Overall			19	0.33%

Historically, the mutation rate has been estimated to be 0.2%. Family Tree DNA has reported that they have observed a rate closer to 0.4% in their studies. The most common reported value in Charles Kerchner's study is between 0.35% and 0.45%. The overall rate for Group H of 0.33% is within the expected range based on these other studies. Please note that the subgroups vary widely in their results. Some of this is due to the small sample sizes involved and some is due to the normal variation across different lines and populations. Charles Kerchner reports observed values between 0.05% and 0.95%.

The Deduced Ancestral Haplotypes of Appendix A allow speculation as to how the various subgroups of Group H are related to each other. Appendix B provides one possible way that the subgroups could be related to each other. Please note that subgroups H-3, H-4, and H-8 have exactly the same deduced Haplotype. H-7 has a single marker mutation from the Haplotype of the three subgroups. These four subgroups are likely the most closely related of the eight subgroups. The chart in Appendix B is drawn to reflect this possibility.

If the triangulation method is applied to the total H Group, the following Deduced Ancestral Haplotype is indicated:

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

This is the most likely Haplotype for Group H based on the data available but it is not yet known if this is the actual Haplotype. It is presented here for reference and to stimulate further discussion on the best way to move forward with Group H research.

Conclusions

The mutation rates for Group H of the Butler Surname DNA project vary from line to line but overall are within the norms of the expected mutation rates.

More project members and more research is needed to establish the relationship between the different subgroups of Group H.

What's Next?

We need to continue our research to connect the various subgroups that make up Group H. The most useful thing at this point would be to add members to the subgroups that only have a single member. New project members that can connect through a different line to the earliest known ancestor will enable the deduction of the Ancestral Haplotype for that line and allow for the refinement of this study.

Appendix A

Deduced Ancestral Haplotype using the Triangulation Method

Subgroup H-1 – Benjamin

Member – M-0111

Triangulation - There is only one member of this group so triangulation is not possible. The Haplotype is assumed to be the same as the single member.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	32	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	15
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-2 – Daniel Mangrum, b. 1832 SC

Member – M-0036

Triangulation - There is only one member of this group so triangulation is not possible. The Haplotype is assumed to be the same as the single member.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	38
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	15
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-3 – Edmund, b. 1694-1704

Members – M-0053, M-0077, M-0101, M-0120

Triangulation - Three members of this subgroup have Edmund as MRCA with the fourth having a later MRCA. Triangulation was applied to the three members with Edmund as MRCA. The statistics would be stronger here with more members of this subgroup.

DYS# 449 has Allele value 31 for two project members and 32 for one so Allele value 31 is picked as the most likely value for Edmund. DYS# 576 has Allele value 18 for two project members and 19 for one so Allele value 18 is picked. DYS# 442 has Allele value 14 for two project members and 15 for one so Allele value 14 is picked.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-4 – Gabriel, b. 1779 SC

Member – M-0089

Triangulation - There is only one member of this group so triangulation is not possible. The Haplotype is assumed to be the same as the single member.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-5 – James Anthony, b. 1740 SC

Member – M-0020

Triangulation - There is only one member of this group so triangulation is not possible. The Haplotype is assumed to be the same as the single member.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	17	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	30	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-6 – James, b. 1730 Ireland

Members – M-0016, M-0026

Triangulation - There are only two members of this subgroup so triangulation is not possible. It is not possible to deduce the Haplotype for this subgroup.

There are two markers different between the two members so these were counted as two mutations but we need more members of this subgroup to be confident in this.

Deduced Haplotype

Subgroup H-7 – Joseph James, b. SC

Members – M-0024, M-0051, M-0058

Triangulation - Two of the members of this subgroup have DYS# 442 with an Allele value of 15 and one with 14. 15 is picked because it is a match between two people with Joseph James as the MRCA. The Allele value 14 is the mutation. Thanks to William A. Hughes for his analysis of this subgroup.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	15
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-8 – Zachariah, b. 1736 VA

Members – M-0003, M-0009, M-0010, M-0019, M-0049
M-0056, M-0074, M-0075, M-0078, M-0129

Triangulation - DYS# YCA11b has Allele value 23 except for one member so Allele value 23 is picked. For DYS# 576, the Joel line has 18, the James line has 18 and all of the George Snellings line has 19 except one that has 20, so Allele value 18 is picked. For DYS# CDYb, the Joel line has 38, the James line has 39 and all of the George Snellings line has 39 except one that has 38, so Allele value 39 is picked. For DYS# 442, the Joel line has 13, the James line has 14 and the George Snellings line has 14 so Allele value 14 is picked.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-8a – Zachariah>Joel

Members – M-0074

Triangulation - There is only one member of this group so triangulation is not possible. The Haplotype is assumed to be the same as the single member.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	38
385b	15	459a	9	464a	14	YCA II b	23	442	13
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-8b – Zachariah>James

Members – M-0009, M-0010, M-0019, M-0049, M-0128

Triangulation - All Allele values of this subgroup match so no triangulation is needed.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	18
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Subgroup H-8c – Zachariah>Patrick>Peter Patrick>George Snellings

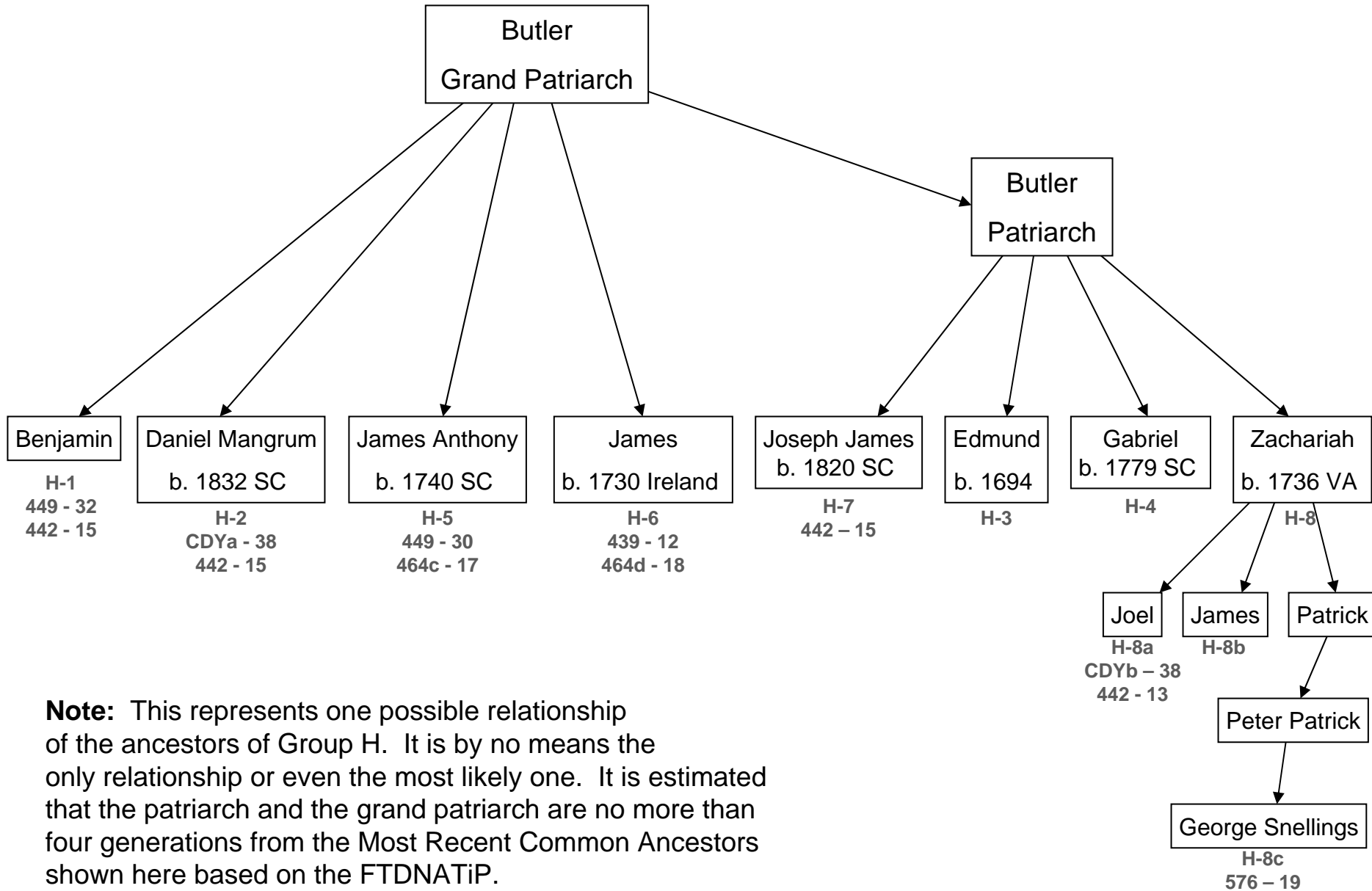
Members – M-0003, M-0056, M-0075, M-0078

Triangulation - For this subgroup, the four members that have George Snellings as Most Recent Common Ancestor were used for the triangulation. DYS# YCA11b has Allele value 23 for three of the members and 24 for one so 23 is picked. DYS# 576 has Allele value 19 for three of the members and 20 for one so 19 is picked. DYS# CDYb has Allele value 39 for three of the members and 38 for one so 39 is picked.

Deduced Haplotype

DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles	DYS#	Alleles
393	13	439	13	454	11	464c	16	607	14
390	24	389-1	13	447	25	464d	17	576	19
19	14	392	13	437	15	460	11	570	16
391	11	389-2	29	448	19	GATA H4	11	CDY a	37
385a	11	458	18	449	31	YCA II a	19	CDY b	39
385b	15	459a	9	464a	14	YCA II b	23	442	14
426	12	459b	10	464b	15	456	17	438	12
388	12	455	11						

Appendix B



Note: This represents one possible relationship of the ancestors of Group H. It is by no means the only relationship or even the most likely one. It is estimated that the patriarch and the grand patriarch are no more than four generations from the Most Recent Common Ancestors shown here based on the FTDNATiP.