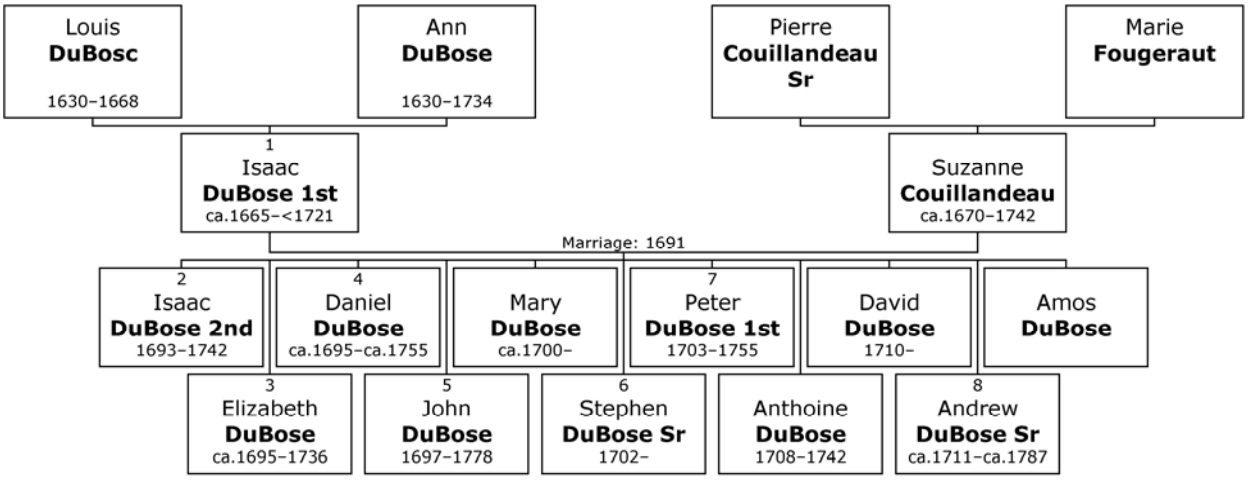
Genetic Genealogy
Analysis of Y-DNA Simple Tandem Repeats (STRs)
For The DuBose-DuBois DNA Project
by Steven James Coker
SJCoker1@gmail.com
28 April 2017

A. Introduction.

Genetic genealogy provides information that helps guide and validate research conducted by genealogists. This paper will examine clade grouping for The DuBose-DuBois DNA Project based on Y-DNA Simple Tandem Repeat (STR) data. STRs are also referred to as Short Tandem Repeats.

Isaac DuBose was born about 1665 in St Jacques Parish, Dieppe, Normandy, France. Isaac died in French Santee, Charleston County, South Carolina, before 1721. Isaac DuBose married Suzanne Couillandeau about 1691 in South Carolina. The couple were French Huguenots and they are paternal 6th Great Grandparents of the writer. Suzanne Couillandeau was born in La Tremblade Provenec, Saintoge, Or, France, about 1670. Suzanne reached 72 years of age and died in Santee, Charleston, South Carolina, on June 12, 1742. She was the daughter of Pierre Couillandeau Sr and Marie Fougeraut.

In 1685, King Louis XIV revoked the Edict of Nantes which had been published in 1598 to halt the French Wars of Religion (1562-98) between Catholics and Protestants. The Edict granted religious liberty in France to the Huguenots and other French Protestants. The revocation in 1685 caused a flood of emigrants leaving France to avoid religious persecution. Isaac and Suzanne were among the many French Huguenots who came to South Carolina at that time.



Family of Suzanne and Isaac DuBose

I. Isaac DuBose 2nd was born in St James Parish, South Carolina, about 1693.

II. Elizabeth DuBose was born in Jamestown, Berkeley, South Carolina, about 1695. She died in Christ Parish, Charleston, South Carolina, in 1736 at the age of 41. Elizabeth was buried in Awendaw, Charleston County, South Carolina.

III. Daniel DuBose was born in Jamestown, Berkeley, South Carolina, about 1695. He died in St James, Clarendon, South Carolina, about June 1755 at the age of 60.

IV. John DuBose was born in Jamestown, Berkeley, South Carolina, in 1697. He died at Lynches Creek, South Carolina, in 1778 at the age of 81.

V. Mary DuBose was born in South Carolina, about 1700.

VI. Stephen DuBose Sr was born in Jamestown, Berkeley, South Carolina, on May 13, 1702. He died in South Carolina. He is the writer's 5th Great Grandfather.

VII. Peter DuBose 1st was born in Jamestown, Berkeley County, South Carolina, in 1703. He died in Christ Church Parish, Berkeley County, South Carolina, on June 9, 1755, at the age of 52.

VIII. Anthoine DuBose was born in Jamestown, Berkeley, South Carolina, about 1708.

IX. David DuBose was born in Jamestown, Berkeley County, South Carolina, in 1710. He died at Lynches Creek, Darlington County, South Carolina.

X. Andrew DuBose Sr was born in Jamestown, Berkeley, South Carolina, about 1711. He died at Lynches Creek, Darlington County, South Carolina, about 1787 at the age of 76.

XI. Amos DuBose.

The descendants of Suzanne and Isaac DuBose increased and became early and prominent settlers of South Carolina and throughout much of the United States. This paper will not attempt to name their many descendants or discuss their history. Rather it will attempt to describe the DNA of a few of the male descendants.

1731 Map of Carolina

B. Statement of Hypothesis.

The hypothesis which will be examined is that the STR data compiled in the DuBose-DuBois DNA Project can be taken as evidence of the Y-DNA clade representing male descendants of Isaac DuBose.

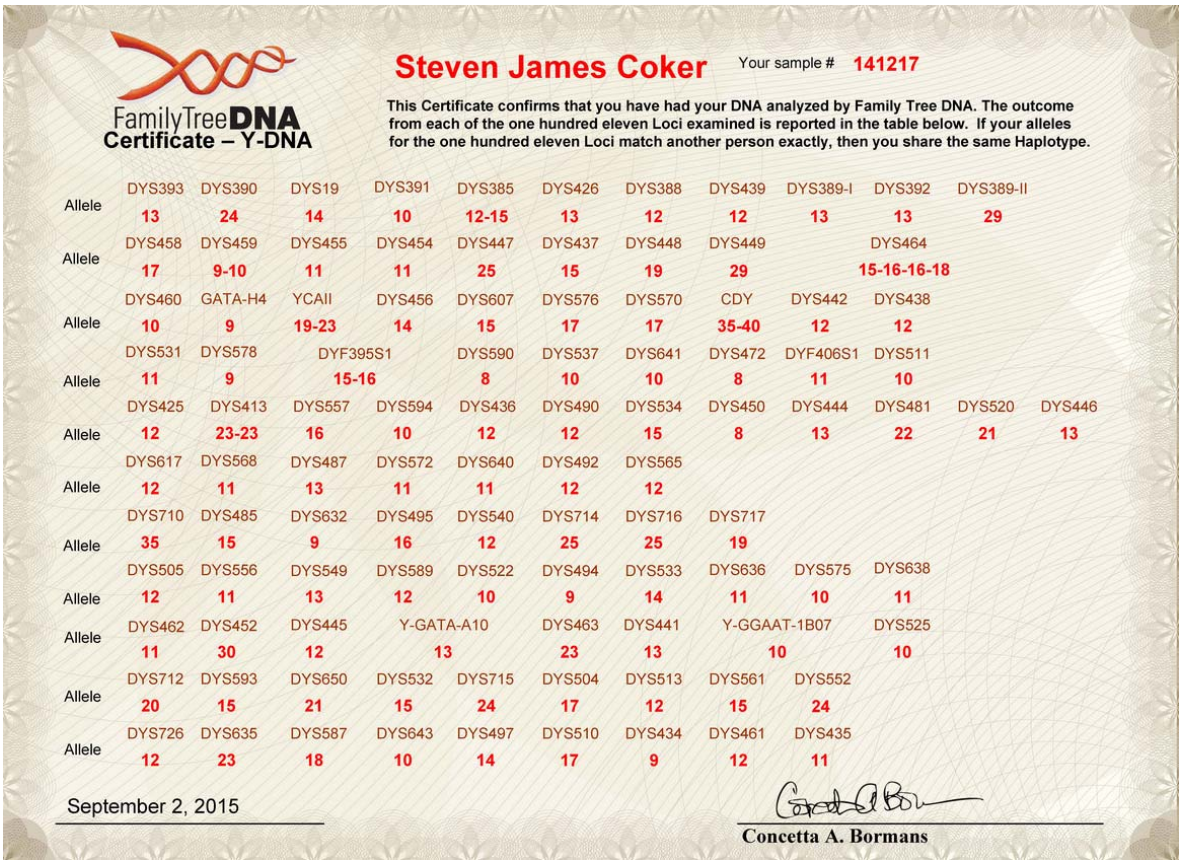
C. Methods.

1. DNA collection. The DuBose-DuBois DNA project was started in 2004 and has grown slowly as more interested descendants have joined. The project currently has 82 members who have submitted Y-DNA samples. Of those, 29 have been identified as descendants of Isaac DuBose. The remainder include descendants from other lineages. As well as men who descended from a DuBose female and thus did not inherit the DuBose Y-DNA. The DNA project includes data on STRs and Single Nucleotide Polymorphisms (SNPs) for Y-DNA. The project also includes mtDNA and autosomal DNA data. This analysis will be limited to the Y-DNA data.

2. Data Testing. The DNA samples for this project have been collected and tested by the Family Tree DNA (FTDNA) company in Houston, Texas. All samples and data are managed, stored, and reported by FTDNA. The data is available for public viewing at the following web site.

https://www.familytreedna.com/groups/du-bose-du-bois/dna-results

However, some data may be restricted from view by the DNA owner. FTDNA does not take ownership of the data from individual samples. The person who submits the sample controls viewing access for data from their sample. All data is accessible to the project administrator. The writer is administrator for the project.



FamilyTreeDNA
Certificate – Y-DNA

Steven James Coker Your sample # **141217**

This Certificate confirms that you have had your DNA analyzed by Family Tree DNA. The outcome from each of the one hundred eleven Loci examined is reported in the table below. If your alleles for the one hundred eleven Loci match another person exactly, then you share the same Haplotype.

Allele	DYS393	DYS390	DYS19	DYS391	DYS385	DYS426	DYS388	DYS439	DYS389-I	DYS392	DYS389-II	
	13	24	14	10	12-15	13	12	12	13	13	29	
Allele	DYS458	DYS459	DYS455	DYS454	DYS447	DYS437	DYS448	DYS449		DYS464		
	17	9-10	11	11	25	15	19	29		15-16-16-18		
Allele	DYS460	GATA-H4	YCAII	DYS456	DYS607	DYS576	DYS570	CDY	DYS442	DYS438		
	10	9	19-23	14	15	17	17	35-40	12	12		
Allele	DYS531	DYS578	DYF395S1		DYS590	DYS537	DYS641	DYS472	DYF406S1	DYS511		
	11	9	15-16		8	10	10	8	11	10		
Allele	DYS425	DYS413	DYS557	DYS594	DYS436	DYS490	DYS534	DYS450	DYS444	DYS481	DYS520	DYS446
	12	23-23	16	10	12	12	15	8	13	22	21	13
Allele	DYS617	DYS568	DYS487	DYS572	DYS640	DYS492	DYS565					
	12	11	13	11	11	12	12					
Allele	DYS710	DYS485	DYS632	DYS495	DYS540	DYS714	DYS716	DYS717				
	35	15	9	16	12	25	25	19				
Allele	DYS505	DYS556	DYS549	DYS589	DYS522	DYS494	DYS533	DYS636	DYS575	DYS638		
	12	11	13	12	10	9	14	11	10	11		
Allele	DYS462	DYS452	DYS445	Y-GATA-A10		DYS463	DYS441	Y-GGAAT-1B07		DYS525		
	11	30	12	13		23	13	10		10		
Allele	DYS712	DYS593	DYS650	DYS532	DYS715	DYS504	DYS513	DYS561	DYS552			
	20	15	21	15	24	17	12	15	24			
Allele	DYS726	DYS635	DYS587	DYS643	DYS497	DYS510	DYS434	DYS461	DYS435			
	12	23	18	10	14	17	9	12	11			

September 2, 2015

Concetta A. Bormans

Example of Individual Report of STR Data.

3. Data. For this paper, I used Y-DNA STR data. I further restricted the data set to only samples that had tested 111 markers. Lastly, I used only samples in the R Haplogroup. The result of these filters left 43 subjects with data samples for study.

5. Phylogenetic Methodology. The Y-DNA STR data was compiled into an Excel spreadsheet from which a FASTA data file was created. The FASTA data was imported into MEGA. The data was aligned in MEGA using MUSCLE with Neighbor Joining clustering method. Phylogenetic analysis was done in MEGA of the aligned data using both the Neighbor-Joining and the Maximum Likelihood methods. Similar results were obtained from both methods.

6. For estimating the Time to Most Recent Common Ancestor (TMRCA) of the descendants of Isaac DuBose, I used the program by McGee at http://www.mymcgee.com/tools/yutility111.html

D. Discussion and Procedures.

1. Converting STR Counts to ATCG.

The STR results are reported by FTDNA as repetition counts. For this analysis, I wanted to convert the counts into ATCG nucleotide sequences for which the counts are a user-friendly shorthand. This turned out to be much harder than I anticipated. I spent a considerable portion of the hours devoted to this report in attempting to discover reliable information which could be used to translate the counts into ATCG sequences. I achieved partial success in this effort. Enough to complete the analysis.

FTDNA does not publish or provide information for converting the counts into ATCG format. After considerable searching, the best conversions I found were given in the NIST table shown below. The repeat motifs in the table were sufficient for many of the STRs. But, 36 of the 111 STRs were not included in the data table. This most likely is because the table was published in 2008 and has not been updated. The repeat motifs for six of the STRs in the table were not usable because I could not find explanation of the code used in the table.

There were 112 STR data marker values available in the data set. Including multi-value markers. I found usable ATCG conversions for 70 markers. Thus, the analysis was limited to 70 STR markers.

Summary List of Y Chromosome STR Loci and Available Fact Sheets

See also Butler, J.M., Kline, M.C., Decker, A.E. (2008) Addressing Y-chromosome short tandem repeat (Y-STR) allele nomenclature. *Journal of Genetic Genealogy* 4(2): 125-148

Marker Name	Allele Range* (repeat numbers)	Repeat Motif	GenBank Accession	Reference Allele
DYS19	10-19	TAGA	AC017019 (r&c)	15
DYS385 a/b	7-28	GAAA	AC022486 (r&c)	11
DYS389 I DYS389 II	DYS389I: 9-17 DYS389II:24-34	(TCTG) (TCTA) (TCTG) (TCTA)	AC004617 (r&c)	12, 29
DYS390	17-28	(TCTA) (TCTG)	AC011289	24
DYS391	6-14	TCTA	AC011302	11
DYS392	6-17	TAT	AC011745 (r&c)	13
DYS393	9-17	AGAT	AC006152	12
YCAII a/b	11-25	CA	AC015978	23
DYS388	10-18	ATT	AC004810	12
DYS425	10-14	TGT	AC095380	10
DYS426	10-12	GTT	AC007034	12
DYS434	9-12	TAAT (CTAT)	AC002992	10
DYS435	9-13	TGGA	AC002992	9
DYS436	9-15	GTT	AC005820	12
DYS437	13-17	TCTA	AC002992	16
DYS438	6-14	TTTTC	AC002531	10
DYS439	9-14	AGAT	AC002992	13
DYS441	12-18	TTCC	AC004474	14
DYS442	10-14	$(TATC)_2(TGTC)_3(TATC)_{12}$	AC004810	17
DYS443	12-17	TTCC	AC007274	13
DYS444	11-15	TAGA	AC007043	14
DYS445	10-13	TTTA	AC009233	12
DYS446	10-18	TCTCT	AC006152	14
DYS447	22-29	TAAWA	AC005820	23
DYS448	20-26	AGAGAT	AC025227	22
DYS449	26-36	TTTC	AC051663	29
DYS450	8-11	TTTTA	AC051663	9
DYS452	27-33	YATAC	AC010137	31
DYS453	9-13	AAAT	AC006157	11
DYS454	10-12	AAAT	AC025731	11
DYS455	8-12	AAAT	AC012068	11
DYS456	13-18	AGAT	AC010106	15
DYS458	13-20	GAAA	AC010902	16
DYS459 a/b	7-10	TAAA	AC010682	9

http://www.cstl.nist.gov/biotech/strbase/ystr_fact.htm

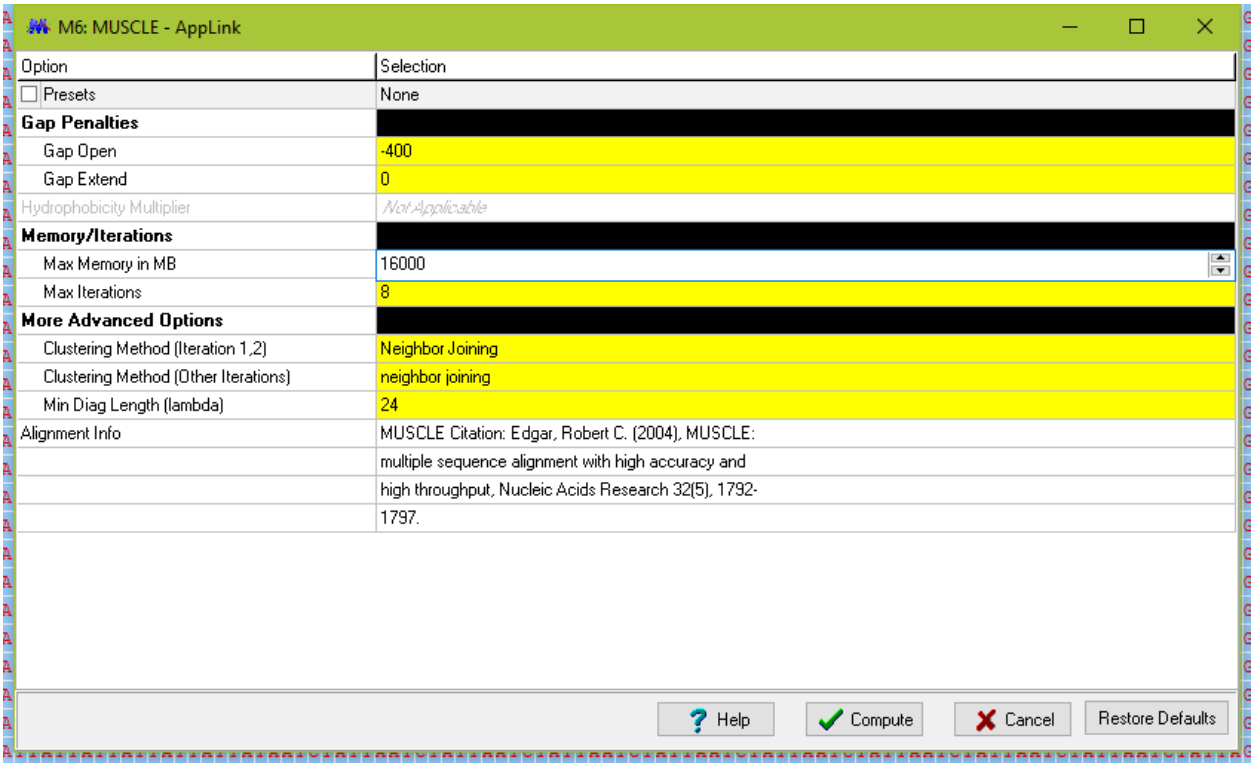
2. Pruning Data.

There were a total of 82 Y-DNA STR data samples in the project. Of these, 14 had only tested to the 37 marker level, 22 had tested to 67 markers, and 45 had tested fully to 111 markers. After repeated tries, I found that including 111 markers with 37 or 67 markers in MEGA gave some illogical results in the phylogenetic analysis. Therefore, I eliminated the 37 and 67 marker samples and worked only with the 45 remaining 111 marker samples.

Two of the 111 marker samples were not in the Y-DNA R Haplogroup. I attempted running the MEGA phylogenetic analysis with both R and the two non-R haplogroup types included. But, I found that inclusion of the non-R data invariably resulted in trees that had logical flaws. One possible factor is that these two non-R samples had null data for one marker. It is apparent from the SNP data that all the known descendants of Isaac DuBose are in the R Haplogroup. Thus, I eliminated the data samples that were of non-R haplogroup.

3. Aligning the Data in MEGA.

I ran many simulations in MEGA. I found that the most logical phylogenetic result was obtained when the data was aligned using MUSCLE with Neighbor Joining as the clustering method.



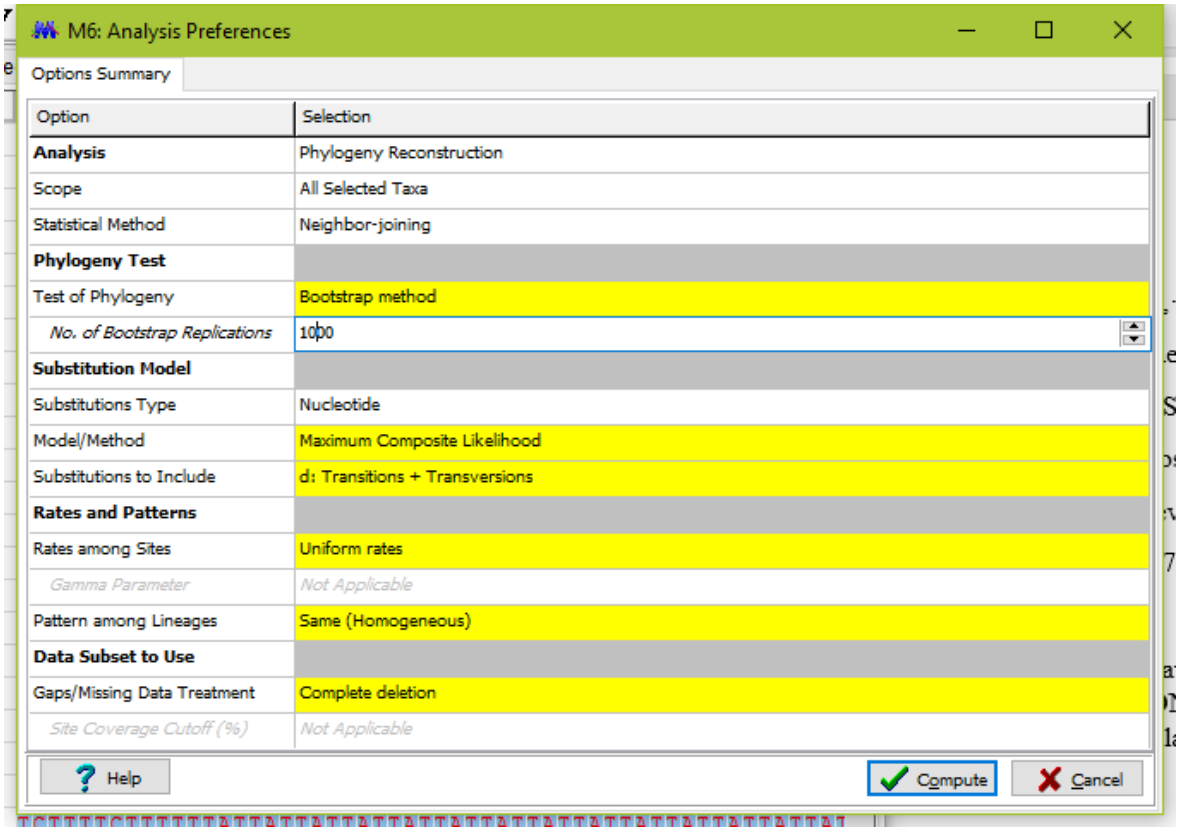
Alignment Procedure

E. Results

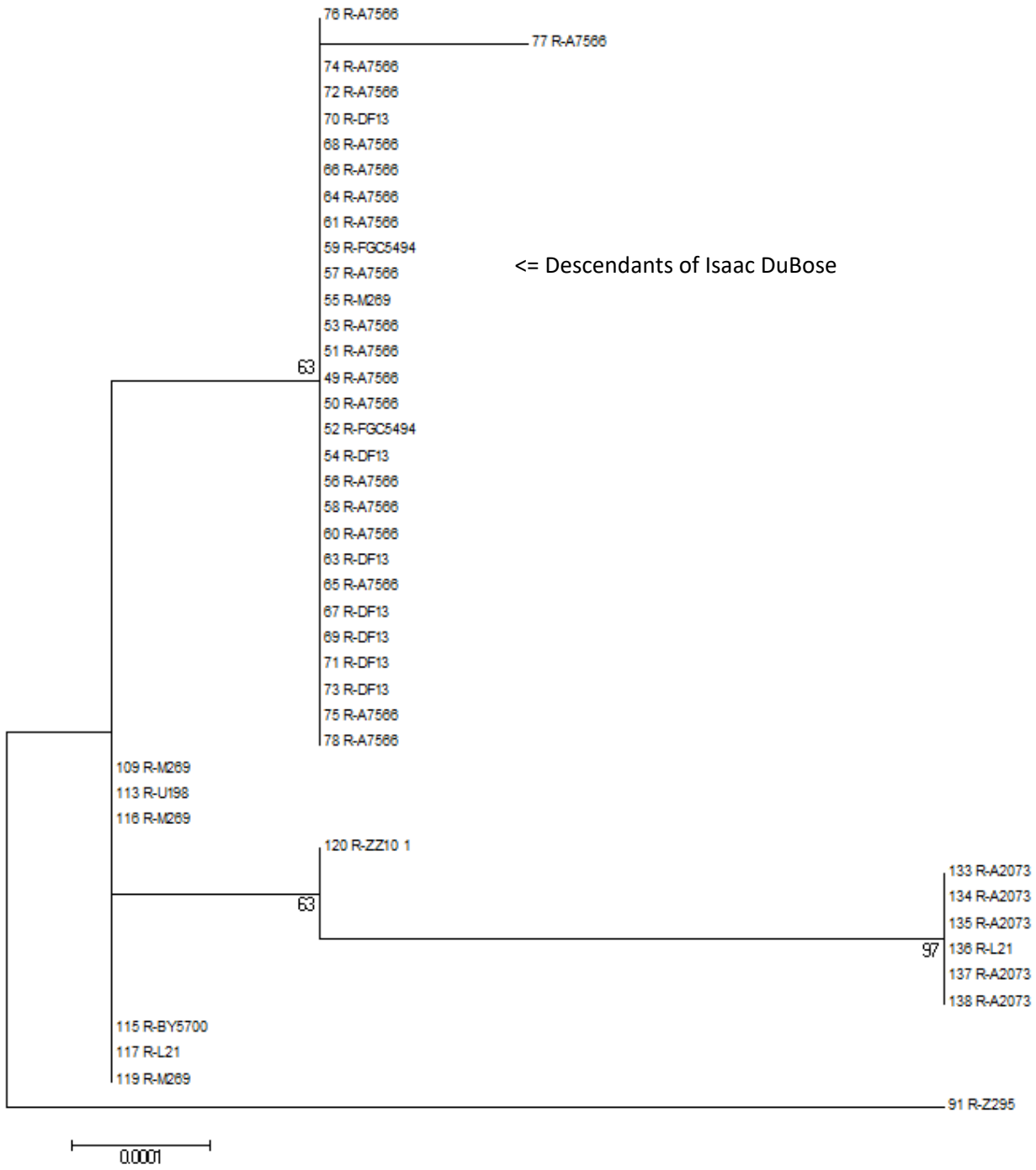
1. Phylogenetic Analysis in MEGA.

I tried various phylogenetic methods in MEGA. The results obtained using either Neighbor Joining or Maximum Likelihood methods appeared equivalent. In the trees shown below, the data samples are numbered with the Haplogroup and Terminal SNP shown after the sample number.

The Y-DNA clade representing descendants of Isaac DuBose are samples 49-78. Sample 62 was omitted from the analysis because it was only tested to 37 marker level.

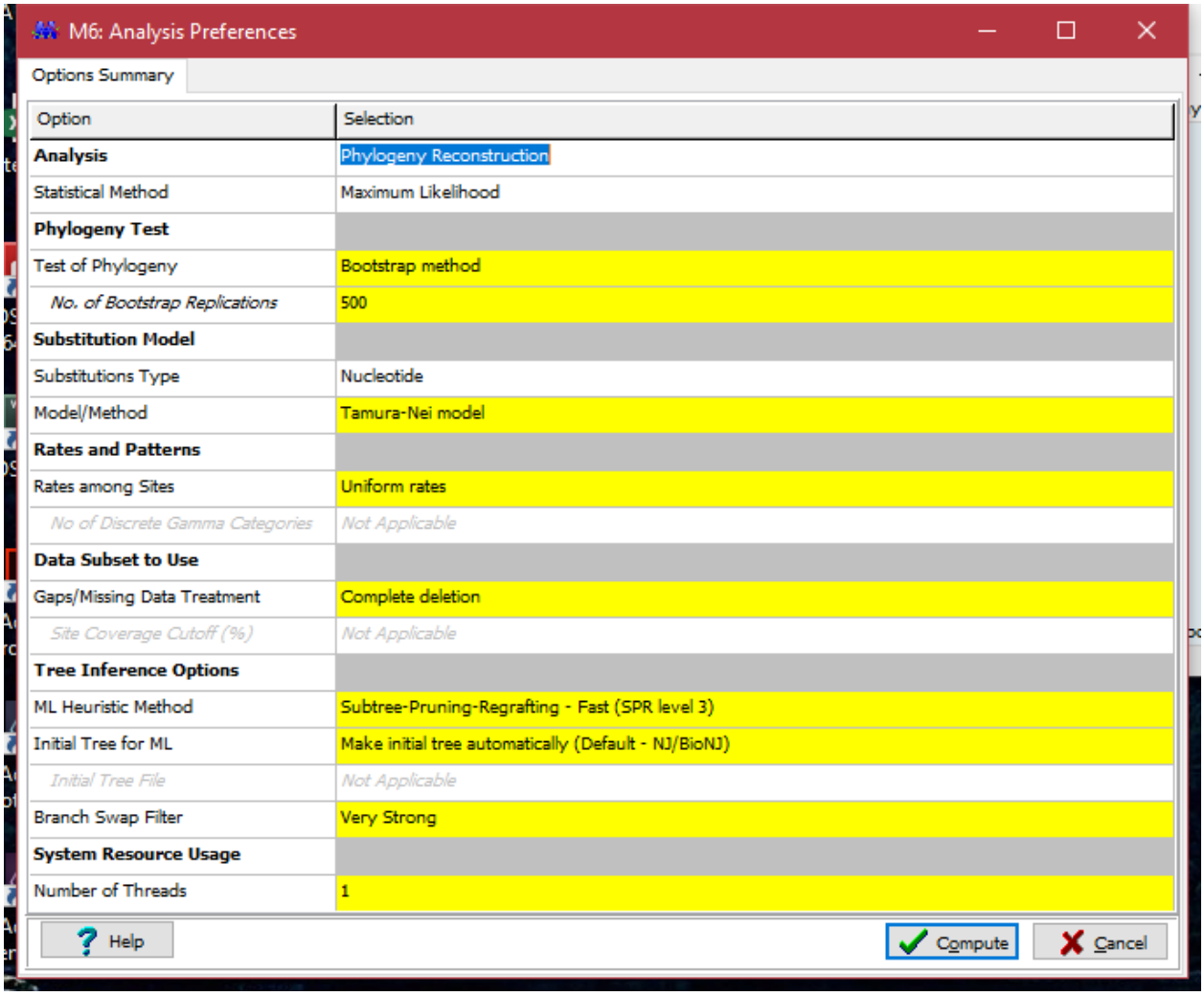


Phylogenetic Analysis by Neighbor-Joining

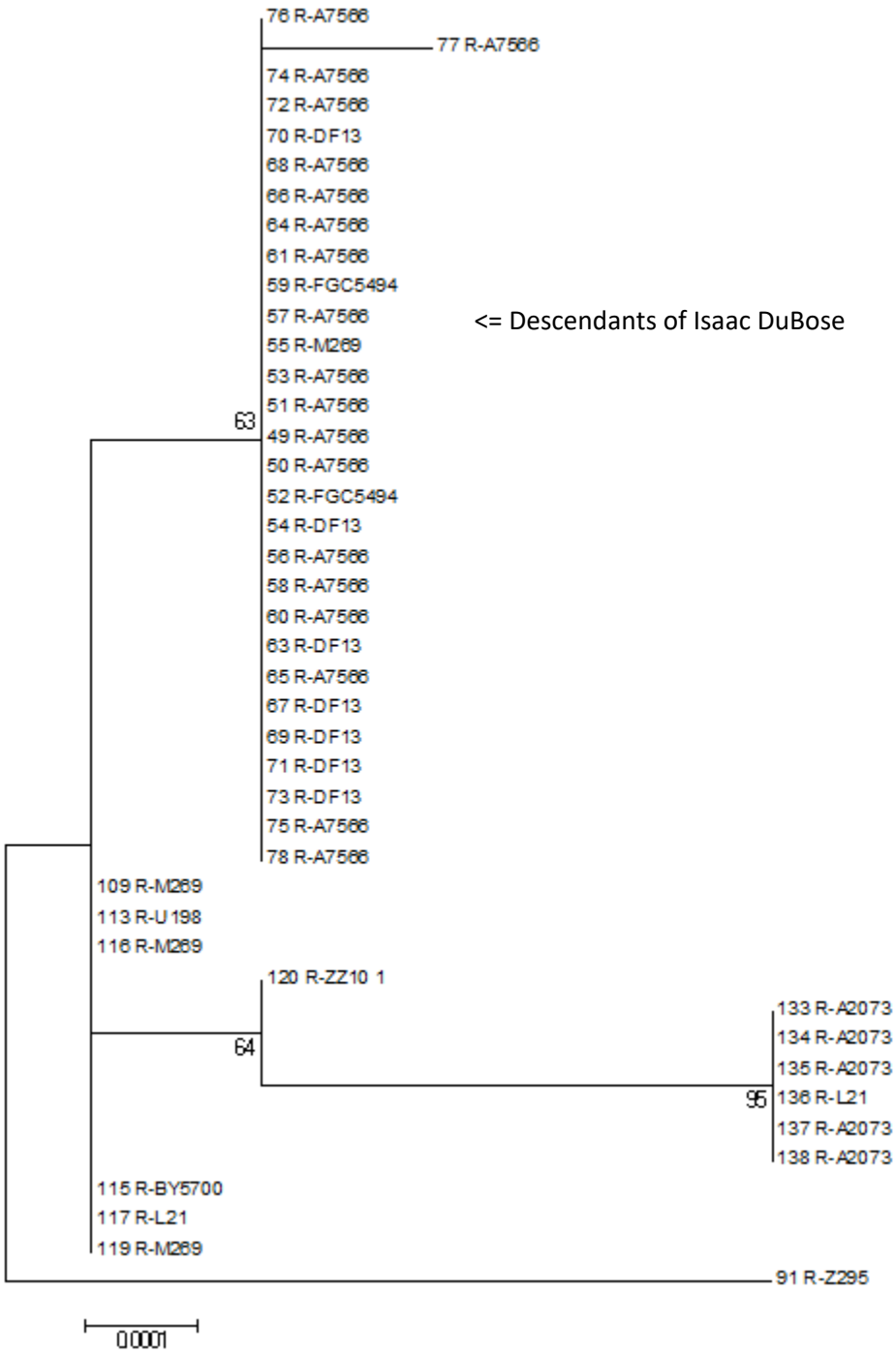


<= Descendants of Isaac DuBose

Tree Using Neighbor Joining



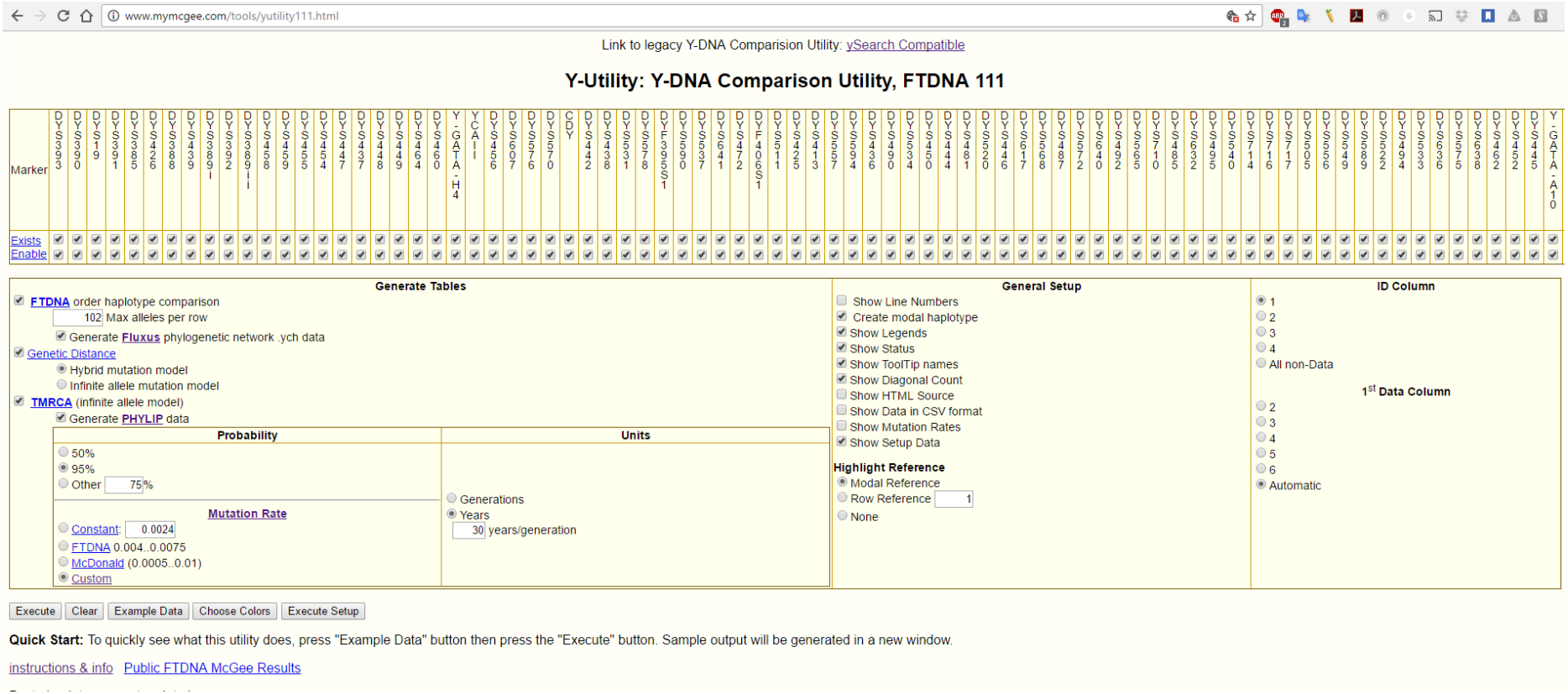
Phylogenetic Analysis by Maximum Likelihood



Tree Using Maximum Likelihood

2. Estimating Time to Most Recent Common Ancestor (TMRCA)

For estimating the TMRCA of the descendants of Isaac DuBose, I used the program provided by McGee at http://www.mymcgee.com/tools/yutility111.html



Genetic Distance

ID	modal	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78
modal	111	3	5	5	5	5	2	2	2	3	2	2	1	2	0	4	3	2	0	0	2	1	2	5	4	5	7	5	4	3	3
49	3	111	6	8	8	8	5	5	5	5	5	5	4	5	2	7	6	5	3	3	5	4	5	8	7	7	9	7	7	6	4
50	5	6	111	10	10	10	7	7	5	6	7	7	6	7	3	9	8	7	5	5	7	6	7	10	9	8	8	6	7	8	6
51	5	8	10	111	10	10	7	5	7	8	5	5	4	5	3	9	8	7	5	5	7	6	7	10	9	10	12	10	6	6	6
52	5	8	10	10	111	0	7	7	7	8	7	7	6	5	1	9	8	7	5	5	7	6	7	10	7	10	12	10	9	6	8
53	5	8	10	10	0	111	7	7	7	8	7	7	6	5	1	9	8	7	5	5	7	6	7	10	7	10	12	10	9	6	8
54	2	5	7	7	7	7	111	4	4	5	4	4	3	4	1	6	5	4	2	2	4	1	4	7	6	5	7	5	6	5	5
55	2	5	7	5	7	7	4	111	4	5	4	4	3	4	1	6	5	4	2	2	4	3	4	7	6	7	9	7	5	5	5
56	2	5	5	7	7	7	4	4	111	3	4	4	3	4	1	6	5	4	2	2	4	3	4	7	6	7	9	7	6	5	5
57	3	5	6	8	8	8	5	5	3	111	5	5	4	5	2	7	6	5	3	3	5	4	5	8	7	7	9	7	7	6	6
58	2	5	7	5	7	7	4	4	4	5	111	2	1	2	2	6	5	4	2	2	4	3	4	7	6	5	9	7	4	3	3
59	2	5	7	5	7	7	4	4	4	5	2	111	1	2	1	6	5	2	2	2	4	3	4	7	6	7	9	7	4	3	3
60	1	4	6	4	6	6	3	3	3	4	1	1	111	1	1	5	4	3	1	1	3	2	3	6	5	6	8	6	3	2	2
61	2	5	7	5	5	5	4	4	4	5	2	2	1	111	1	6	5	4	2	2	4	3	4	7	6	7	9	7	4	3	3
62	0	2	3	3	1	1	1	1	1	2	2	1	1	1	37	0	0	0	0	0	0	0	1	1	1	4	3	3	4	3	3
63	4	7	9	9	9	9	6	6	6	7	6	6	5	6	0	111	1	6	4	4	6	5	6	3	4	9	7	9	8	7	7
64	3	6	8	8	8	8	5	5	5	6	5	5	4	5	0	1	111	5	3	3	5	4	5	2	3	8	8	8	7	6	6
65	2	5	7	7	7	7	4	4	4	5	4	2	3	4	0	6	5	111	2	2	4	3	4	7	6	7	9	7	6	5	5
66	0	3	5	5	5	5	2	2	2	3	2	2	1	2	0	4	3	2	111	0	2	1	2	5	4	5	7	5	4	3	3
67	0	3	5	5	5	5	2	2	2	3	2	2	1	2	0	4	3	2	0	111	2	1	2	5	4	5	7	5	4	3	3
68	2	5	7	7	7	7	4	4	4	5	4	4	3	4	0	6	5	4	2	2	111	3	4	7	6	7	9	7	6	5	5
69	1	4	6	6	6	6	1	3	3	4	3	3	2	3	0	5	4	3	1	1	3	111	3	6	5	4	6	4	5	4	4
70	2	5	7	7	7	7	4	4	4	5	4	4	3	4	1	6	5	4	2	2	4	3	111	7	6	7	9	7	6	5	5
71	5	8	10	10	10	10	7	7	7	8	7	7	6	7	1	3	2	7	5	5	7	6	7	111	5	10	10	10	9	8	8
72	4	7	9	9	7	7	6	6	6	7	6	6	5	6	1	4	3	6	4	4	6	5	6	5	111	9	9	9	6	7	7
73	5	7	8	10	10	10	5	7	7	7	5	7	6	7	4	9	8	7	5	5	7	4	7	10	9	111	4	2	7	8	8
74	7	9	8	12	12	12	7	9	9	9	9	9	8	9	3	7	8	9	7	7	9	6	9	10	9	4	111	2	9	10	10
75	5	7	6	10	10	10	5	7	7	7	7	7	6	7	3	9	8	7	5	5	7	4	7	10	9	2	2	111	7	8	8
76	4	7	7	6	9	9	6	5	6	7	4	4	3	4	4	8	7	6	4	4	6	5	6	9	6	7	9	7	111	5	5
77	3	6	8	6	6	6	5	5	5	6	3	3	2	3	3	7	6	5	3	3	5	4	5	8	7	8	10	8	5	111	4
78	3	4	6	6	8	8	5	5	5	6	3	3	2	3	3	7	6	5	3	3	5	4	5	8	7	8	10	8	5	4	111

Related | Probably Related | Possibly Related

FTDNA's Interpreting Genetic Distance for 12 Markers
FTDNA's Interpreting Genetic Distance for 25 Markers
FTDNA's Interpreting Genetic Distance for 37 Markers
FTDNA's Interpreting Genetic Distance for 67 Markers
FTDNA's Interpreting Genetic Distance for 111 Markers

- Hybrid mutation model is used
- Values on the diagonal indicate number of markers tested

The diagonal elements of the table indicate the number of allele data existing for that haplotype. The calculations can be of two types:

Hybrid Mutation Model
The target of this model is to match that used by Ysearch and FTDNA. It uses the stepwise mutation model for all alleles except DYS464 and YCA which use the infinite allele model.

The stepwise model says that each mutation is allowed to change the allele value by exactly one, so a difference of two means that two mutations occurred and a difference of three means that three mutations occurred.

Infinite Allele Mutation Model
The infinite allele model says that the entire difference between allele values, no matter how large, is the result of one mutation.

Based on Genetic Distance	12 Markers	25 Markers	37 Markers	67 Markers	111 Markers
Very Tightly Related	N/A	N/A	0	0	0
Tightly Related	N/A	N/A	1	1-2	1-2
Related	0	0-1	2-3	3-4	3-5
Probably Related (within 15 generations)	1	2	4	5-6	6-7
Possibly Related (maybe over 15 generations)	2	3	5	7	8-10
Not Related	>2	>3	>5	>7	>10

Source: https://www.familytreedna.com/learn/y-dna-testing/y-str/expected-relationship-match/

Time to Most Recent Common Ancestor (Generations)																															
ID	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	MIN
49	111	22	26	26	26	19	19	19	19	19	19	16	19	28	22	19	19	14	14	16	16	19	24	24	24	29	24	24	22	16	14
50	22	111	31	26	26	24	24	19	22	24	24	22	24	35	29	26	24	19	19	22	22	24	31	26	26	26	22	22	24	22	19
51	26	31	111	31	31	24	19	24	26	19	19	16	19	35	29	26	24	19	19	22	22	24	29	29	31	36	31	22	22	22	16
52	26	26	31	111	5	22	24	24	26	24	24	22	19	21	26	24	24	19	19	22	19	24	29	22	26	26	24	26	22	26	19
53	26	26	31	5	111	22	24	24	26	24	24	22	19	21	26	24	24	19	19	22	19	24	29	22	26	26	24	26	22	26	19
54	19	24	24	22	22	111	16	16	19	16	16	14	16	21	22	19	16	11	11	14	10	16	24	22	19	24	19	22	19	19	10
55	19	24	19	24	24	16	111	16	19	16	16	14	16	21	22	19	16	11	11	14	14	16	22	22	24	29	24	19	19	19	11
56	19	19	24	24	24	16	16	111	14	16	16	14	16	21	22	19	16	11	11	14	14	16	24	19	24	29	24	19	19	19	11
57	19	22	26	26	26	19	19	14	111	19	19	16	19	28	24	22	19	14	14	16	16	19	26	22	24	29	24	22	22	22	14
58	19	24	19	24	24	16	16	16	19	111	11	10	11	28	22	19	16	11	11	14	14	16	22	22	19	29	24	16	14	14	10
59	19	24	19	24	24	16	16	16	19	11	111	10	11	21	22	19	11	11	11	14	14	16	22	22	24	29	24	16	14	14	10
60	16	22	16	22	22	14	14	14	16	10	10	111	10	21	19	16	14	10	10	11	11	14	19	19	22	26	22	14	11	11	10
61	19	24	19	19	19	16	16	16	19	11	11	10	111	21	19	16	16	11	11	14	14	16	19	19	24	26	24	16	14	14	10
62	28	35	35	21	21	21	21	21	28	28	21	21	21	37	13	13	13	13	13	13	13	21	21	21	42	35	35	42	35	35	13
63	22	29	29	26	26	22	22	22	24	22	22	19	19	13	111	10	22	16	16	19	19	22	14	16	29	24	29	26	24	24	10
64	19	26	26	24	24	19	19	19	22	19	19	16	16	13	10	111	19	14	14	16	16	19	11	14	26	26	26	24	22	22	10
65	19	24	24	24	24	16	16	16	19	16	11	14	16	13	22	19	111	11	11	14	14	16	24	22	24	29	24	22	19	19	11
66	14	19	19	19	19	11	11	11	14	11	11	10	11	13	16	14	11	111	5	10	10	11	19	16	19	24	19	16	14	14	10
67	14	19	19	19	19	11	11	11	14	11	11	10	11	13	16	14	11	5	111	10	10	11	19	16	19	24	19	16	14	14	10
68	16	22	22	22	22	14	14	14	16	14	14	11	14	13	19	16	14	10	10	111	11	14	22	19	22	26	22	19	16	16	10
69	16	22	22	19	19	10	14	14	16	14	14	11	14	13	19	16	14	10	10	11	111	14	22	19	16	22	16	19	16	16	10
70	19	24	24	24	24	16	16	16	19	16	16	14	16	21	22	19	16	11	11	14	14	111	24	22	24	29	24	22	19	19	11
71	24	31	29	29	29	24	22	24	26	22	22	19	19	21	14	11	24	19	19	22	22	24	111	19	31	31	31	26	24	24	11
72	24	26	29	22	22	22	22	19	22	22	22	19	19	21	16	14	22	16	16	19	19	22	19	111	29	29	29	22	24	24	14
73	24	26	31	26	26	19	24	24	24	19	24	22	24	42	29	26	24	19	19	22	16	24	31	29	111	16	11	24	24	26	11
74	29	26	36	26	26	24	29	29	29	29	29	26	26	35	24	26	29	24	24	26	22	29	31	29	16	111	11	29	29	31	11
75	24	22	31	24	24	19	24	24	24	24	24	22	24	35	29	26	24	19	19	22	16	24	31	29	11	11	111	24	24	26	11
76	24	22	22	26	26	22	19	19	22	16	16	14	16	42	26	24	22	16	16	19	19	22	26	22	24	29	24	111	16	19	14
77	22	24	22	22	22	19	19	19	22	14	14	11	14	35	24	22	19	14	14	16	16	19	24	24	24	29	24	16	111	16	11
78	16	22	22	26	26	19	19	19	22	14	14	11	14	35	24	22	19	14	14	16	16	19	24	24	26	31	26	19	16	111	11
MIN	14	19	16	19	19	10	11	11	14	10	10	10	10	13	10	10	11	10	10	10	10	11	11	14	11	11	11	14	11	11	

0-9 Generations	10-19 Generations	20-29 Generations	30-39 Generations	Close ancestor excluded

12 Gen = Mean of Mins
10 Gen = Mode of Mins

The above table gives the 95% probability that the MRCA was no longer than the specified generations. The algorithm was taken from Bruce Walsh paper, Estimating the Time to the Most Recent Common Ancestor for the Y chromosome or Mitochondrial DNA for a Pair of Individuals.

Values on the diagonal indicate number of markers tested. Generations assume 30 years on average per generation. Calculations use the average mutation rate for all the markers common between the pair of haplotypes being compared. Assumed mutation rates are shown in the following table.

The TMRCA analysis shows most samples share a common male ancestor within 10 to 11 generations. The adjusted average minimum generations to the common male ancestor is 12. The most common (mode) of the minimums is 10 generations. This result agrees with the approximately 350 years that have passed since Isaac was born and 325 years since Isaac and Suzanne began their family.

It is not unexpected that some samples have a larger TMRCA value and some smaller. This may be caused by variability of mutations and differences in length of average generations between lineages. Also, the assumed mutation rates are based on population studies and actual mutation rates for this group may be somewhat different.

Mutation Rates - Custom values															
DYS393	DYS390	DYS19	DYS391	DYS385	DYS426	DYS388	DYS439	DYS389i	DYS392	DYS389ii					
0.0009	0.0033	0.0015	0.0030	0.0023	0.0003	0.0004	0.0042	0.0020	0.0011	0.0024					
DYS458	DYS459	DYS455	DYS454	DYS447	DYS437	DYS448	DYS449	DYS464							
0.0053	0.0011	0.0003	0.0001	0.0033	0.0014	0.0019	0.0060	0.0032							
DYS460	Y-GATA-H4	YCAII	DYS456	DYS607	DYS576	DYS570	CDY	DYS442	DYS438						
0.0030	0.0030	0.0012	0.0053	0.0056	0.0078	0.0061	0.0102	0.0030	0.0009						
DYS531	DYS578	DYF395S1	DYS590	DYS537	DYS641	DYS472	DYF406S1	DYS511	DYS425	DYS413	DYS557	DYS594			
0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0020	0.0024	0.0024	0.0024			
DYS436	DYS490	DYS534	DYS450	DYS444	DYS481	DYS520	DYS446	DYS617	DYS568	DYS487	DYS572	DYS640	DYS492	DYS565	
0.0024	0.0024	0.0024	0.0024	0.0005	0.0024	0.0024	0.0008	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	
DYS710	DYS485	DYS632	DYS495	DYS540	DYS714	DYS716	DYS717	DYS505	DYS556	DYS549	DYS589	DYS522	DYS494	DYS533	
0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	
DYS636	DYS575	DYS638	DYS462	DYS452	DYS445	Y-GATA-A10	DYS463	DYS441	Y-GGAAT-1B07	DYS525	DYS712	DYS593	DYS650	DYS532	
0.0024	0.0024	0.0024	0.0025	0.0032	0.0014	0.0008	0.0020	0.0020	0.0015	0.0024	0.0024	0.0024	0.0024	0.0024	
DYS715	DYS504	DYS513	DYS561	DYS552	DYS726	DYS635	DYS587	DYS643	DYS497	DYS510	DYS434	DYS461	DYS435		
0.0024	0.0024	0.0024	0.0024	0.0024	0.0024	0.0027	0.0024	0.0024	0.0024	0.0024	0.0024	0.0011	0.0024		
Mutation rates not yet identified have a light green background and use the value of 0.0024 from the constant rate box															

Custom Mutation Rates Used

The default mutation rate is 0.0024 mutations/allele/generation which represents the 60 total mutations during 24870 total allele meioses as given in Y-Chromosomal Microsatellite Mutation Rates: Differences in Mutation Rate Between and Within Loci by B.Myhre Dupuy, M.Stenerson, T.Egeland, and B.Olaisen; Human Mutation 23:117-124 (2004).

The second mutation rate selection uses the FTDNA derived mutation rates. This includes a rate of 0.00399 for the first 12 markers, 0.00481 for markers 13 through 25, and 0.00748 for the markers 26 through 37.

The third mutation rate selection uses the values determined by Doug McDonald as derived from the Sorenson database. Markers not included in the Sorenson database are derived by Doug through other means. Values not included in the McDonald rates use the value in the default of 0.0024.

The Custom Values, use the McDonald Rates.

F. Conclusions.

Based on the analysis discussed above, I conclude that The DuBose-DuBois DNA Project has correctly identified the clade of men who are descended from Isaac DuBose.

1. Genealogical research showing the ancestry of certain individuals in the project demonstrates they descend from Isaac DuBose.

2. The Y-DNA SNP haplogroup signature of men identified as descendants of Isaac DuBose is

R1b > L21 > DF13 > FGC5494 > Z16502 > A7566
See https://www.familytreedna.com/groups/du-bose-du-bois/dna-results

3. As presented, the phylogenetic analysis of STR data demonstrates common ancestry for the identified clade of descendants of Isaac DuBose.

4. The TMRCA analysis demonstrates that members of the identified clade share a common ancestor within the approximate expected number of generations.