Austin Journal of Genetics and Genomic Research

Open Access

Austin Publishing Group

Research Article

Towards a Full-Length Y-Chromosome DNA Sequence of Napoléon the First: Beyond the E-M34 SNP Sub-Haplogroup

Gérard Lucotte[1]*, Peter Hrechdakian[2] and Denis Savard[3]

[1]Institute of Molecular Anthropology, Paris, France
[2]Unifert Group S.A, Belgium
[3]469 Chemin Thériault / Bertrand, N.-B. E1W1K5, Canada

***Corresponding author:** Gérard Lucotte, Institute of Molecular Anthropology, 44 Monge Street, 75 005 Paris, France

Received: December 22, 2015; **Accepted:** December 28, 2015; **Published:** December 31, 2015

Abstract

We report results obtained for the Y-chromosome DNA sequence of Mike Clovis (M.C.), a living fifth generation descendant of Lucien who was one of Napoléon the First (1769-1821)'s brothers. We have previously shown that the Y-haplogroup of Napoléon is E1b1b2a1, SNP (Single Nucleotide Polymorphism) M34 being the terminal genetic marker of this differentiation (Y sub-haplogroup E-M3, in abbreviate). Results obtained here show that there are six successive SNPs (Z841, L791, Y4971, Y4976, Y4970 and PH3893) under M34 in the Y-DNA sequence of M.C. Dating of the corresponding sub-haplogroups shows that E-PH3893 had a TMRCA (Time of the Most Common Recent Ancestor) of about 2700 ybp (years before present). This study permits a better understanding of the paternal ancestry of Napoléon in historical times.

Keywords: Napoléon the First paternal ancestry; Y-chromosome DNA sequence; Single nucleotide polymorphisms; Y-chromosome sub-haplogroups; Time of the most common recent ancestor

Introduction

Mike Clovis (M.C.), is a living fifth generation descendant (Figure 1) of Lucien, one of Napoléon Bonaparte's brothers. As part of Napoléon I Genome (NIG) project we have sequenced the DNA of the Y-chromosome of M.C., in order to obtain pertinent information about Napoléon's Y-chromosome.

Our work on Napoléon Y-DNA markers had begun [1] with the obtaining of genomic DNA from authentic Napoléon's hairs conserved in the Vivant Denon reliquary [2]. Ten successive SNPs (Single Nucleotide Polymorphisms) of the Y-chromosome (M125, M174, M33, M35, M78, M81, M123, M34, M84 and M290) were used to determine Napoléon's Y-haplogroup, that is E1b1b1c1 according to the human Y-chromosomal haplogroup tree [3]; further studies showed that the SNP M34 is the terminal genetic marker of this differentiation, so [4] the corresponding Napoléon Y-chromosome sub-haplogroup was at that time E1b1b1b2a1 (E-M34 in abbreviate).

In this initial study [1] on Napoléon Y-chromosome, only three Y-STRs (Short Tandem Repeats) were used: DYS19 and YCAII.a and .b (palindromics), the respective allelic values obtained being = 13, 19 and 22. Computed with the haplogroup predictor program [5], these three values considered together estimate that Napoléon I corresponds (with a probable value of 76.5%) to a subject belonging to the previously described Y-haplogroup cluster E1b1b [6]. In that study [1] we determine also allelic values for 37 Y-STRs from buccal smear genomic DNA extracted from Charles Napoléon, the living 4th generation descendant of Jérôme Bonaparte (Figure 1), Napoléon I's youngest brother. Results showed the allelic values for DYS19 and for YCAII.a and .b are the same for Charles Napoléon and for Napoléon I. Computed with [5], Charles Napoléon's Y-STR profile is very highly indicative (99.9%) of the E1b1b1c1 Y-haplogroup ; as Charles Napoléon is M34+ , his Y-sub-haplogroup is also E-M34.

In our second study on Y-polymorphisms in Napoléon I's family [7] we have compared, on a total set of one-hundred and thirty three different Y-STRs, the Y-STR profile of Charles Napoléon to that of Alexandre Colonna Walewsky (who is also E-M34), the living 5th generation direct descendant of Napoléon I (Figure 1). Four non-palindromic and two palindromic STRs have different allelic values between Alexandre Colonna Walewsky for 106 Y-STRs. In our third study [8] comparing Mike Clovis (also E-M34), Charles Napoléon and Alexandre Colonna Walewsky for 106 Y-STRs, we found variable allelic values between these three subjects for 7 STRs only: for DYS442, DYS447, DYS454, DYS481, DYS635, DYS712 and CDY. a

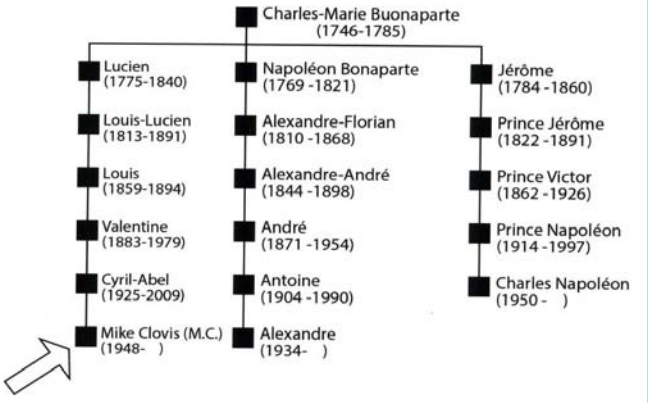

Figure 1: Chain of transmission (seven successive generations of paternal ancestry) from the ancestor Charles-Marie Buonaparte (Napoléon's father) to the *propositus* (arrow) Mike Clovis (M.C.). Are also indicated in the tree the successive generations of paternal ancestry for Alexandre Colonna-Walewski and for Charles Napoléon.

Citation: Lucotte G, Hrechdakian P and Savard D. Towards a Full-Length Y-Chromosome DNA Sequence of Napoléon the First: Beyond the E-M34 SNP Sub-Haplogroup. Austin J Genet Genomic Res. 2015; 2(2): 1015.

Table 1: Main Characteristics of the M.C.'s Y-chromosome sequence obtained.

BAM file size	01Gb
Read numbers	7601387
Mapped reads	All
Length coverage	14173443 bp
Minimal and maximal depth coverage	1x-7999x
Mean depth coverage	57.76x
Median depth coverage	39x

(palindromic). All the observed identical STRs allelic values between the three living subjects, were used to predict the corresponding deduced allelic values for Napoléon I's STRs. A direct determination of allelic values for 16 STRs of Napoléon [9], determinations that were done on Napoléon I genomic DNA extracted from dandruffs adherent to a Napoléon's lock of hair dating from the year 1811, allow us to obtain the real allelic values of Napoléon's STRs for 15 supplementary STRs (allelic value at DYS19 confirmed) others than the three found initially [1].

We have constructed an isofrequency map of the Y-SNP M34 [10], using a large set of genomic DNAs of unrelated males originating from various countries in Europe, Northern Africa and the Near East. The genetic landscape obtained is that of a complex pattern of clines and local decreases and increases in M34 frequencies in some areas. In Armenia, the M34 percentage is 4.2%. The maximal peak in frequency (31%) observed in this map, located in the Dead Sea region of Jordany, indicates the location of origin of M34 in the corresponding region of the Levant. We have reconstructed past migratory routes of M34 chromosomes from this center of origin to West Europe via: Turkey (6.9% in Southeastern Turkey; 7.8% in The Istanbul region), Greece (3.8% for Athens; 1.8% for Northern Greece), the Balkans (1.6% for Croatia) and the continental part of Italy (1.5%).

A secondary peak (9 on 111 subjects = 8.1%) of M34 frequencies, in Ajaccio (Napoléon's native town), can be observed in the partial isofrequency map we have constructed for Corsica and the surrounding regions; the Calabrian region, located in the south of continental Italy, had the most important M34 frequency (5.5%) after that of Ajaccio.

We know 18 generations of Napoléon's paternal ancestor's [11]. His remote ancestor, Gianfardo, was born and lived in Sarzana (a small Italian town on the south of La Magra, a river separating Liguria and Toscany) between the end of XIIth and the beginning of the XIIIth Century. Giovanni (11th generation) was the first paternal ancestor of Napoléon to leave Sarzana for Corsica. None of the nine Ajaccian subjects (and not the only Bastian subject) we have studied -bearing the M34 marker- has the Bonaparte's patronym.

Material and Methods

Buccal swab samples for M.C. were collected, with informed consent. His Y-chromosome sequence was studied in the framework of the "Big Y" project [12], conducted by the Family Tree DNA (FTDNA) Company.

Sequencing was performed on the Illumina HiSeq platform, and downstream analysis was conducted with Arpeggi genome analysis technology. After the sequencing procedure the reads were mapped to the Hg19 version of human genome reference [13], followed by post-processing and variant calling ; all of which were performed using a software (Arpeggi engine).

Table 2: Successive M.C. sub-haplogroups below E-M34, and Y-SNPs found on it.

Sub-haplogroups	SNPs found
E-M34	M34 + Y2931, Y2929, Y4143…+ 57 other SNPs
E-Z841	Z841 + Y2961, CTS2339
E-L791	L791 + L792, Z838, Z852… + 86 other SNPs
E-Y4971	Y4971 + Y4973, Y4968, Y4975
E-Y4976	Y4976 + Y4972, K257
E-Y4970	Y4970
E-PH3893	PH3893+PH4283, Y12966, Y12962… + 9 other SNPs

Aligned data concerning M.C. are given in the form of a BAM file (https://www.dropbox.com/s/nir3pqf6073gdnr/4464.bam.zip?dl=0). This BAM file (after YFull treatment) was downloaded from the site of one of us (D.S.), under the reference E-PH3893.

The M.C. Y-chromosome samples were sequenced at an average depth of 57.76x (Table 1). Centered on the male specific region of the Y-chromosome (MSY), the sequenced part corresponds approximately to 55.25% of the total of the Y-chromosome sequence.

Results

A mean of about 30 000 SNPs are studied in this sort of Big Y analysis [14]. At first time, we made sure that M.C. is really M34+ (Table 2).

A total number of 491 STRs were studied in the present study. The 93 STR allelic values of M.C. found were identical to those published [8]. Particularly, the variable M.C. allelic values DYS442=11, DYS454=11, DYS481=27, DYS635=22, DYS712=23, and CDY.a = 34 are the same. This establishes that it is well the Y-chromosome sequence of M.C. that is under study.

Table 2 summarizes the lists of Y-SNPs of the M.C. Y-chromosome sequence found in the same drafts corresponding to (depicted on Figure 2) the six successive Y-sub-haplogroups (E-Z841, E-L791, E-Y4971, E-Y4976, E-Y4970 and E-PH3893) located below E-M34. In these drafts the sub-haplogroup E-Y4970 is characterized by the corresponding SNP only, while there is a set of two supplementary SNPs for E-Z841 (Y2961 and CTS2339) and for E-Y4976 (Y4972 and K257), a set of three (Y4973, Y4968 and Y4975) for E-Y4971, and up to 89 (L792, Z838, Z852 and 86 others) for E-L791.

The last found in this serie (the most recent, see Table 3) of M.C. subhaplogroups is E-PH3893, with PH4283, Y12966, Y-12962 and nine other SNPs.

Figure 2 is a schematization, according to [15], of the filiation –relationships between the six successive M.C. SNPs located under M34. M.C., belonging to sub-haplogroup E-PH3893, is characterized by S7897 and seventeen of other SNPs.

Up to now there are six subjects, belonging to the latest E-PH893 and E-Y4970 sub-haplotypes, that are involved in the Big Y procedure

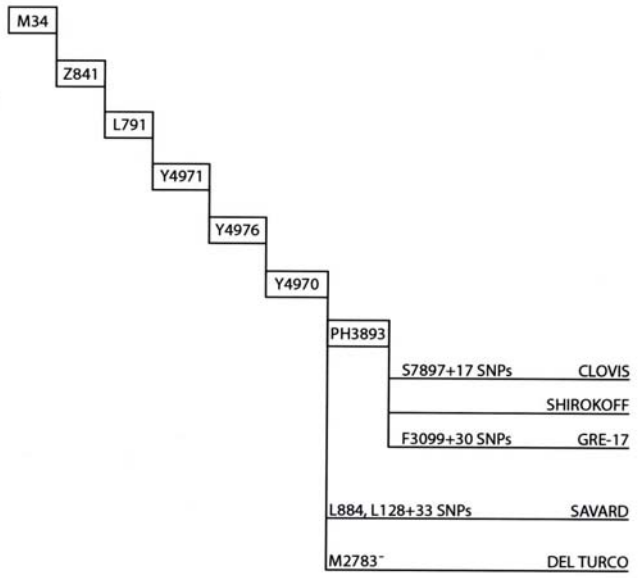

Figure 2: Schematization of the successive Y-SNPs under M34 found in the Y-DNA sequence of Mike Clovis. M. Clovis, Shirokoff and Gre-17 belong to the E-PH3893 Y sub-haplogroup; Savard and del Turco belong to E-Y4970.

[16]. M. Clovis, and Shirokoff (of Russian origin), are two of the three subjects constituting the E-PH3893 group (Figure 2). Shirokoff is characterized by F3099, PH460, PH1181, CTS913, FGC4198 and by twenty-six other YFS-SNPs.

Previous estimates concerning E-M34 datation in Armenia [17] are up to circa 4.8 ±0.8 to 5.1±1.2 kya. Table 3 gives estimates, among [18], concerning dating in ybp (years before present) and TMRCA (Times of the Most Common Recent Ancestor) for the seven M.C. sub-haplogroups below E-M34. Based on the number of SNPs common to M. Clovis and Shirokoff: PH3893, PH4289, Y12961 to Y12969, Y13512 to Y13514, PF498rc (reccurent), M4081rc and Y1829rc, the TMRCA is about 2 700 ybp.

The third individual belonging to sub-haplogroup E-PH3893 studied in the Big Y project is gre-17 (originating from Greece), that is characterized by F3099 and thirty other SNPs (Figure 2).

One of us (D.S.), whose sub haplogroup derives from E-Y4970 (and consequently is the nearest neighbour of the three previous subjects belonging to E-PH3893 sub-haplogroup found in the Big Y project), is of distant Armenian origin. He is characterized by L884, L128, FGC1551, PF3499, Z54459, M7761 and twenty- nine other FS SNPs. Also deriving from E-Y4970 is del Turco (M2783-), whose name suggests an ancient Turkish origin.

Conclusion

One of our studies [1], based from genomic DNA extracted from his remains, permits us to obtain the Y-chromosome haplogroup of Napoléon the First. This Y-haplogroup is E1b1b2a1, according to present–day rules of Y-haplogroup nomenclature. As the terminal SNP marker of the differentiation is M34, we designated this corresponding Y sub-haplogroup (in abbreviate form) as E-M34. Based on a large number of unrelated DNA of subjects originating from Europe, Northern Africa and the Near East [10], we constructed

Table 3: Dating estimates of the six sub-haplogroups below E-M34.

Sub-haplogroups below E-M34	Datations
E-Z841	Formed 14 900 ybp ; TMRCA = 14 000 ybp
E-L791	Formed 14 000 ybp ; TMRCA = 5 600 ybp
E-Y4971	Formed 5 600 ybp ; TMRCA = 5 300 ybp
E-Y4976	Formed 5 300 ybp ; TMRCA = 4 800 ybp
E-Y4970	Formed 4 800 ybp ; TMRCA = 4 700 ybp
E-PH3893	Formed 4 700 ybp ; TMRCA = 2 700 ybp

a global isofrequency map of the M34 genetic marker. The geographic origin of M34 is some region in the Levant. The past migratory routes of M34 chromosomes from this center of origin to West Europe are Turkey, Greece, the Balkans and the continental part of Italy; that links with the surrounding regions of Corsica, in Ajaccio, where Napoléon was born. Early estimates of E-M34 datation [17] concerned a period approximately in the middle of the Neolithic.

Mike Clovis (M.C.) is a living fifth generation descendant of Lucien, one of Napoléon's brother. He is E-M34, and his Y-STR pro le is quasi-similar to those of the direct descendant of Napoléon or of a Napoléon's brother [1,7,8], as well as that of Napoléon's deduced profile. In the present study the full-length Y-chromosome DNA sequence of M.C. was determined by the big Y procedure. Six successive SNPs under M34 were detected in the M.C. Y-chromosome DNA sequence. The corresponding E sub-haplogroups were dated, according to the Time of the Most Common Recent Ancestor (TMRCA) methodology ; this dating evolve from 14000 years before present (ybp) for E-Z841 to 4700 ybp for the antepenultian E-Y4970 sub-haplogroup.

The estimated TMRCA of the most recently found M.C. sub-haplogroup E-PH3893 is about 2700 ybp. That time interval concerns the historical time when Napoléon male ancestors evolved. It seems remarkable that several subjects included in this big Y study belong to the various countries located among the migratory road starting from the M34 focus in the Levant [10] that finally goes to Corsica.

Acknowledgment

We thank Mike Clovis, who furnished to us his personal BAM file. The sequencing work was realized by the Family Tree DNA Company, in the frame of the Big Y project. We thank the Count Alexandre Colonna Walewski for his indefectible financial support.

References

1. Lucotte G, Thomasset T, Hrechdakian P. Haplogroup of the Y chromosome of Napoléon the First. J Mol Biol Res. 2011; 1: 12-19.
2. Lucotte G. A rare variant of the mtDNA HVS1 sequence in the hairs of Napoléon's family. Investig Genet. 2010; 1: 7.
3. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res. 2008; 18: 830-838.
4. http://www.isogg.org/tree/ISOGhapgrE.html
5. Athey W. Haplogroup prediction from Y-STR values using a Bayesian allele frequency approach. J Genet Geneal. 2006; 2: 34-39.
6. http://fr.wikipedia.org/wiki/HaplogroupE?E1b1b1.28E-M35.29
7. Lucotte G, Macé J, Hrechdakian P. Reconstruction of the lineage Y chromosome haplotype of Napoléon the First. Int J Sciences. 2013; 2: 127-139.

8. Lucotte G, Hrechdakian P. New advances reconstructing the Y chromosome haplotype of Napoléon the First based on three of his living descendants. J Mol Biol Res. 2015; 5.

9. Lucotte G, Bouin Wilkinson A. An autosomal STR profile of Napoléon the First. Open J Genet. 2014; 4: 292-299.

10. Lucotte G, Diéterlen F. Frequencies of M34, the ultimate genetic marker of the terminal differenciation of Napoléon the First's Y-chromosome haplogroup E1b1b1c1 in Europe, Northern Africa and the Near East. Int J Anthrop. 2014; 1-2: 27-41.

11. Galantini F. Napoléon et Sarzane, les origines Italiennes des Bonapartes. Editions Michel de Maule (Paris). 2004.

12. Petrone J. Family Tree DNA expands beyond array-based offerings new Y chromosome sequencing service. Genome Web BioArray News. 2013.

13. YFull Experimental Y Tree v3.6.

14. Introduction to Big; Big Y Family Tree DNA (9 pages).

15. www.yfull.com/tree/E-L791/

16. Denis Savard. v16.15.11.2, based on Big Y novel variants, Morley's tree and YFull experimental YTree v2.31, YFull E-L791 group. 2015.

17. Herrera KJ, Lowery RK, Hadden L, Calderon S, Chiou C, Yepiskoposyan L, et al. Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. Eur J Hum Genet. 2012; 20: 313-320.

18. Adamov. Age and estimation algorithm will be presented in a new article from this author.

Austin J Genet Genomic Res - Volume 2 Issue 2 - 2015
ISSN : 2471-030X | www.austinpublishinggroup.com

Citation: Lucotte G, Hrechdakian P and Savard D. Towards a Full-Length Y-Chromosome DNA Sequence of Napoléon the First: Beyond the E-M34 SNP Sub-Haplogroup. Austin J Genet Genomic Res. 2015; 2(2): 1015.