

**AN IBERIAN SUB-CLUSTER IS REVEALED  
IN A PHYLOGENETIC TREE ANALYSIS  
OF  
THE Y-CHROMOSOME E3b HAPLOGROUP**

**By**

**Robert L. Tarín, Jr.**

**November 23, 2005**

## **Introduction**

Recent analysis of 290 individuals in Y-DNA Haplogroup E3b using the Fluxus Phylogenetic Network Analysis<sup>1</sup> software from [www.fluxus-engineering.com](http://www.fluxus-engineering.com) has revealed a separate cluster of mostly Iberian origin haplotypes which seem to represent a North African entry into Iberia distinct from the E3b in Europe that may have arisen from Neolithic or other migratory events. A modal value for this Iberian sub-cluster has been determined from the data used and it suggests a match to either the E3b1-beta cluster (E-M78) or E3b2 (E-M81). There were no available SNP results for the haplotypes making up the dataset. It is unknown whether this finding reflects relatively recent gene flow from the Islamic rule of Spain or an older influx possibly from the Phoenicians. The Phoenicians founded and settled many cities in North Africa and the Iberian Peninsula. Haplotype matching at [www.yhrd.org](http://www.yhrd.org) reveals significant results for Tunisia, North Africa with this Iberian sub-cluster.

## **Background and Methods**

This study was initiated from an exchange of email between this author and Victor Villarreal, both members of The Genealogy of Mexico DNA Surname Project<sup>2</sup> headed up by Gary Felix. Victor ran some analyses of the data from the E3b Project using the PHYLIP (Phylogeny Inference Package)<sup>3</sup> software and noted the initial appearance of what seemed to be two main groups. I asked him if he would run the same tree using the Iberian origin E3b data from my private database and once again it seemed to show two groups. Wanting to pursue this further, I ran the E3b data from my database using the Fluxus software and saw a definite and separate cluster. Data from [www.ysearch.org](http://www.ysearch.org) was added for E3b haplotypes listing an origin in England and Germany. The resulting Fluxus tree revealed the sub-cluster to be mainly Iberian with those from England and Germany in the main cluster with the overall E3b group modal. More data was added for other countries with the same basic results. At this point all the available E3b data with a country origin was collected and categorized into arbitrary regions. Using Iberia as its own region, the others were set as Western Europe, Central Europe, Northern Europe, Southern Europe, Eastern Europe, North Africa, and the Middle East. The dataset was comprised of haplotypes from these regions with 12, 25 and 37 markers in FTDNA order. Duplicate haplotypes within surnames were deleted from the final dataset. An analysis was run on each set of markers. The full 12 marker set was used in the first analysis without regard to fast mutation rates. In the 25 and 37 marker sets the fast mutators were not used in the analyses. Those not included were DYS458, 449, 464a,b,c,d in the 25 marker set and DYS456, 576, 570, and CDYa,b in the 37 marker set. Each node in the resulting trees was examined as to its record number(s) in the dataset and color-coded according to region. Haplotypes within the Iberian sub-cluster were identified and modal values were determined for each set of 12, 25 and 37 markers.

## **Results**

An analysis was completed on the full 12 marker dataset and the final color-coded tree clearly showed the separate cluster being mostly Iberian. As a precaution, another analysis was run on the 12 marker dataset with the fast mutators DYS385a,b and 439 being removed but no significant difference was noted between the two so it was decided to keep the full 12 marker dataset intact. Subsequent analyses of the 25 and 37 marker datasets produced similar results. The separate sub-cluster contained majority Iberian origin haplotypes and included a few Western European and Central European individuals along with a North African.

The modal values for this sub-cluster were compared to values from other studies and favorably identify them as matching North African sequences. Note: As stated earlier, marker values for fast mutators in the 25 and 37 marker datasets were not used in the calculations of the Fluxus trees but were included in the determination of the modals and are shown here to provide some possible values.

MODAL VALUES FOR E3b IBERIAN SUB-CLUSTER												
<b>12 marker modal</b>												
393	390	19	391	385a	385b	426	388	439	389-1	392	389-2	
13	24	13	9	13	14	11	12	10	14	11	30	
<b>25 marker modal</b>												
458	459a	459b	455	454	447	437	448	449	464a	464b	464c	464d
18	9	9	11	12	23	14	20	32	14	16	16	17
<b>37 marker modal</b>												
460	H4	YCAIIa	YCAIIb	456	607	576	570	CDYa	CDYb	442	438	
11	11	19	22	16	13	18	22	36	41	12	10	
* Fast mutators identified in red												

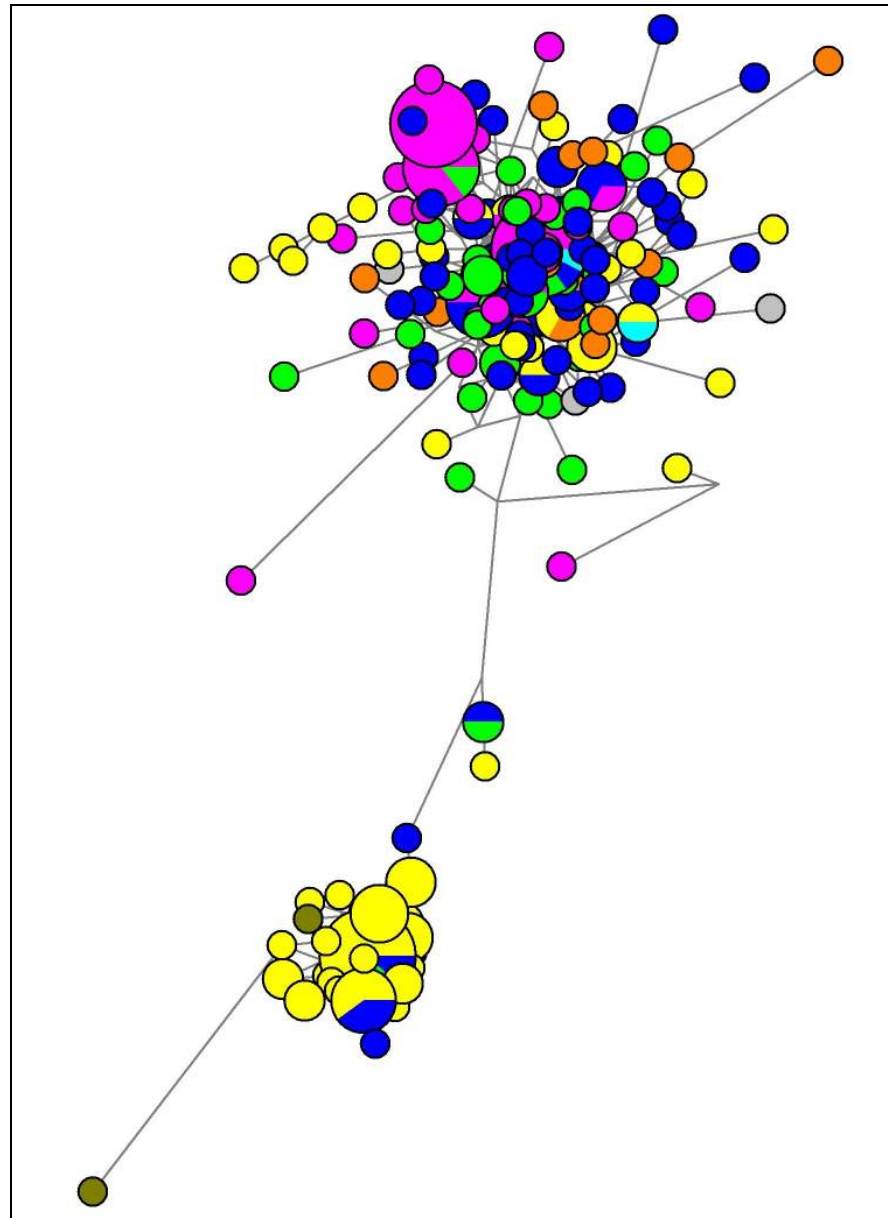
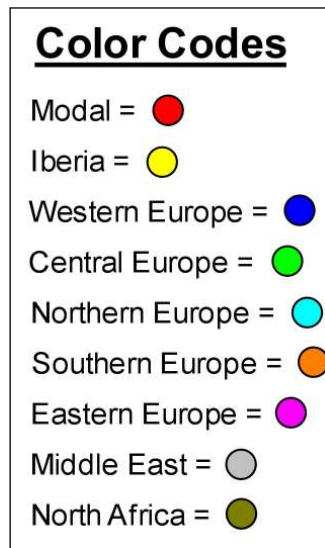
Utilizing the modal values for markers DYS19, 389-1, 389-2, 390, 391, 392, 393, 385a,b, 438, and 439 in an online search of the YHRD – Y Chromosome Haplotype Reference Database (Release 17) at [www.yhrd.org](http://www.yhrd.org) produced significant results from Tunisia, North Africa including Arabs and Berbers. Most of the remaining matches were either in Iberia or countries with an Iberian origin ancestry.

Population	Metapopulation	%	#	Sample
Zriba, Tunisia	Africa	61.29%	19	31
Tunisia [Andalusian Arabs]	Africa	6.87%	9	131
Tunisia [Berber]	Africa	6.67%	2	30
Madrid, Spain	Europe	1.97%	3	152
Tunis, Tunisia	Africa	1.85%	1	54
Central Portugal	Europe	1.84%	9	489
Rio de Janeiro, Brazil [European]	Latin America	1.59%	2	126
Andalucía/Extremadura, Spain	Europe	1.30%	5	386
Somalia	Africa	1.00%	2	201
Northern Portugal	Europe	0.89%	5	564
Lyon, France	Europe	0.80%	1	125
Pyrenees, Spain	Europe	0.75%	1	134
Argentina [European]	Latin America	0.65%	4	614

More evidence for a presence in Tunisia is seen in Capelli et al. 2005<sup>4</sup> where the modal values for DYS388, 393, 392, 19, 390, and 391 match haplotype 44 in Table A1. Out of 22 samples matching this haplotype, 15 were in Tunisia. Regarding an E3b subclade determination, according to Cruciani et al. 2004<sup>5</sup>, a rare 10 repeat allele at DYS439 is one characteristic of the E3b1-beta cluster (E-M78). Garvey<sup>6</sup> states that a value of DYS391 = 9 is seen in both the E3b1-beta cluster and E3b2. Referencing a study by Bosch et al 2001<sup>7</sup>, Garvey also has modal values for E3b2 that match the modal in this Iberian sub-cluster.

## Fluxus Trees

Shown below are the resulting median joining networks for each analysis and a detailed view of the 37 marker analysis sub-cluster. The color codes apply to all of the figures shown.



**12 marker analysis**

### Color Codes

Modal = ●

Iberia = ●

Western Europe = ●

Central Europe = ●

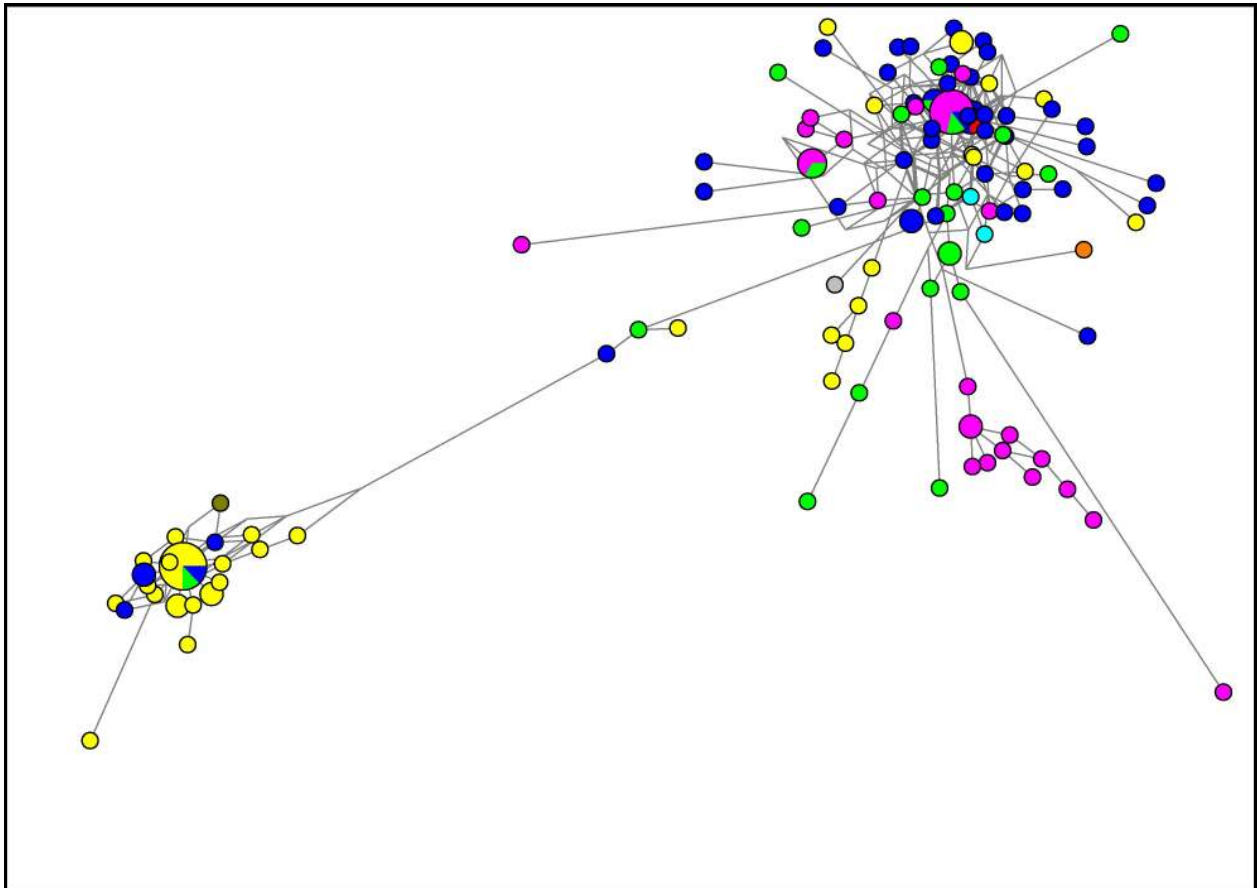
Northern Europe = ●

Southern Europe = ●

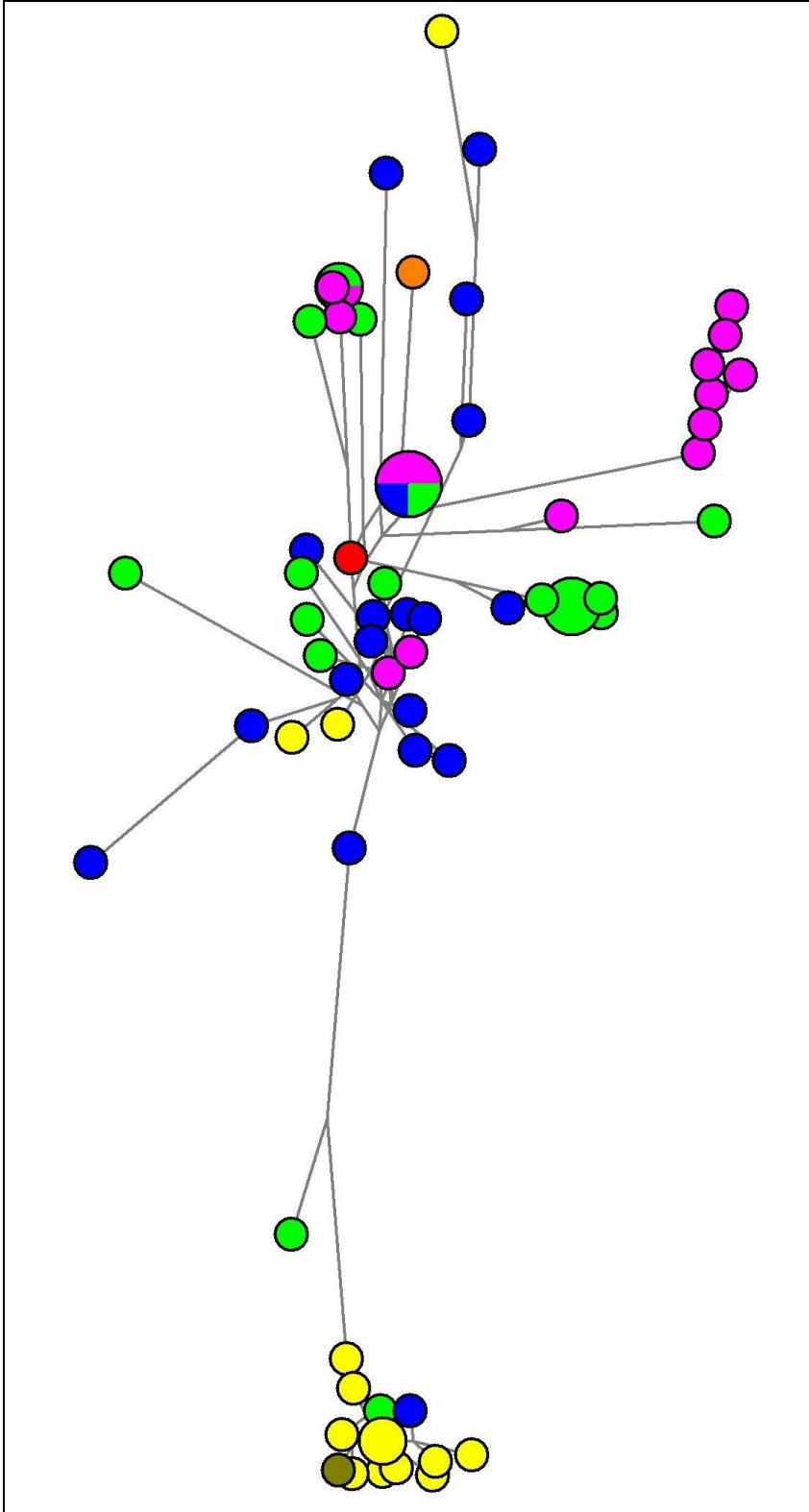
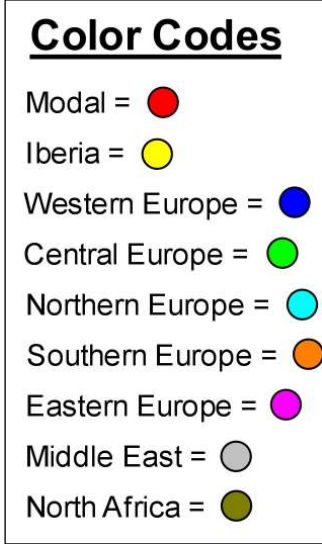
Eastern Europe = ●

Middle East = ●

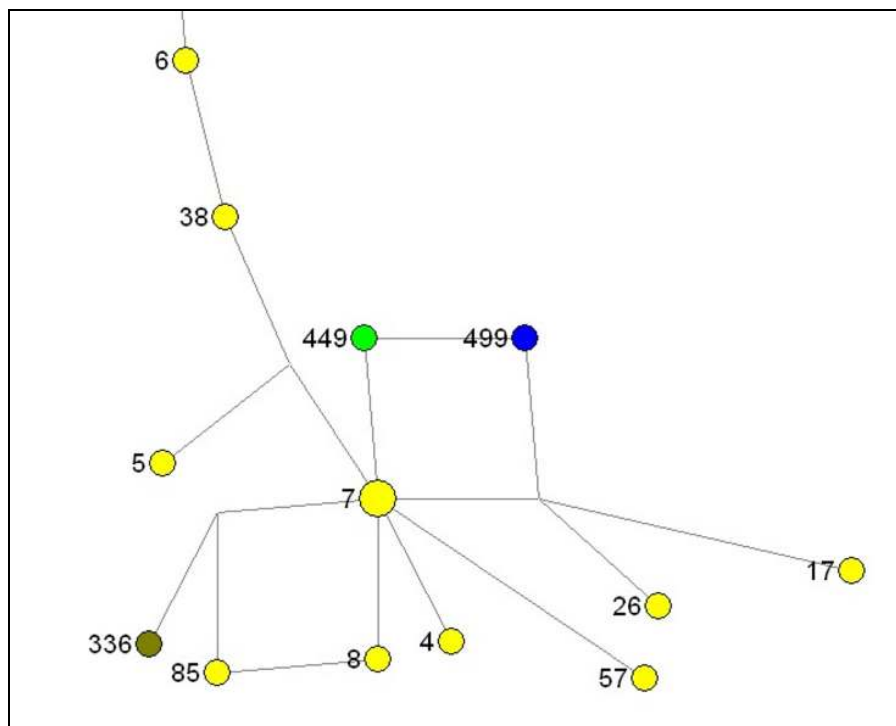
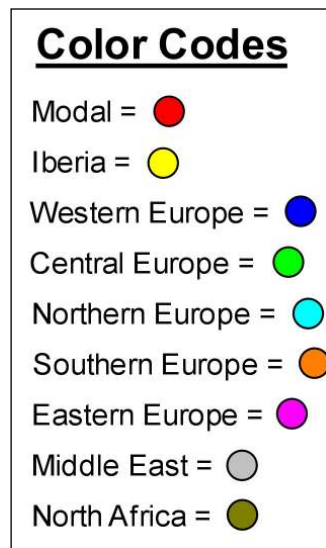
North Africa = ●



25 marker analysis



37 marker analysis



**Sub-cluster detail from 37 marker analysis with record numbers**

## Discussion

While findings of North African genetic links to the Iberian Peninsula are not new, this analysis helps to confirm those links and serves to highlight the separation of this Iberian sub-cluster from that of the majority European E3b. It also helps to advance our knowledge for the extended marker modal values of this sub-cluster which can be referenced for future studies. As testing for various subclades of Haplogroup E3b becomes more common, we may then be able to make a clear determination for this Iberian sub-cluster. The dataset is available upon request. Links for some common database searches using the modal values are provided here:

[www.ysearch.org](http://www.ysearch.org) (12 marker) <http://tinyurl.com/7rl59>

[www.smgf.org](http://www.smgf.org) (12 marker) <http://tinyurl.com/awqv2>

## References

1. Fluxus Phylogenetic Network Analysis, [www.fluxus-engineering.com](http://www.fluxus-engineering.com)
2. The Genealogy of Mexico DNA Surname Project, Gary Felix, administrator, <http://members.tripod.com/~GaryFelix/index63.htm>
3. Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. *Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.*
4. Capelli et al. 2005. Population Structure in the Mediterranean Basin: A Y Chromosome Perspective. *Department of Biology, University College of London, London, UK.*
5. Cruciani et al. 2004. Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. *Am. J. Hum. Genet. 74:1014–1022, 2004.*
6. Garvey, D. Y Haplogroup E3b. [http://freepages.genealogy.rootsweb.com/~dgarvey/DNA/hg/YCC\\_E3b.html#E3b1](http://freepages.genealogy.rootsweb.com/~dgarvey/DNA/hg/YCC_E3b.html#E3b1)
7. Garvey, D. [http://freepages.genealogy.rootsweb.com/~dgarvey/DNA/y\\_bosch.htm](http://freepages.genealogy.rootsweb.com/~dgarvey/DNA/y_bosch.htm)