

Mitochondrial and Y-chromosomal Analysis of an Indigenous Mexican Population, the Filomeno Mata, to Investigate Admixture with the Spanish

K.H. Ritchie¹, S. Harmon¹, J.E. Gomez¹, N.M. Myres¹, U.A. Perego¹, J.E. Ekins¹, L.A.D. Hutchison^{1,3}, M.L. Lunt¹, S.S. Masek¹, A.A. Nelson¹, M.E. Nelson¹, K.L. Pennington¹, J.L. Peterson¹, T.A. Tolley¹, S.R. Woodward¹ 1) Sorenson Molecular Genealogy Foundation (SMGF), Salt Lake City, UT, USA; 2) Dept. of Microbiology, Brigham Young University, Provo, UT, USA; 3) Computational Biology Research Group, CSAIL, Massachusetts Institute of Technology, MA, USA

Abstract

According to oral tradition, the Filomeno Mata population of eastern Mexico has remained in relative isolation since its initial encounters with Spanish conquerors during the Conquistador era. Genetic analysis can support the legitimacy of such historical accounts as they are reflected in the extant genetic composition of a population. We report the estimated degree of European admixture present in the Filomeno Mata population based on concurrent analysis of Y-chromosome (Y-ch) and mitochondrial DNA (mtDNA). The differential histories of the maternal and paternal genetic records are also examined. Samples were collected from 117 females and 52 males from the Filomeno Mata village. All samples were d-loop sequenced, and typed for the characteristic Native American mtDNA haplogroups, A-D and X. Males were also tested for three Y-ch specific SNPs DYS199 (M3), M242, and RPS4Y711 and 36 STR loci. Y-ch haplotypes were compared with published Spanish and Portuguese datasets to assess the degree of admixture. The combination of mtDNA and Y-ch data demonstrate a complex genetic history of the Filomeno Mata.



Background

The Amerindian village of Filomeno Mata (FMM) is inhabited by indigenous peoples whose language is primarily Totonac. The Totonacs emerged as a farming culture around 800 C.E. in the regions of present day Veracruz and the northern mountains of Puebla, Mexico. In 1480, the Aztec empire subdued and consequently ruled the Totonacs until the Sixteenth Century when Hernán Cortés arrived. The Totonacs were some of the first indigenous groups to accept him and the Spaniards. Now the Totonacs numbering ~100,000 inhabit Puebla and Veracruz (Figure 1), and are industrious farmers. The Filomeno Mata village has been mostly secluded since the 1830s when the borders of Veracruz were re-divided and the Totonac isolated themselves.



Mitochondrial Data

Native American haplogroup frequencies and percentages are shown in Table 1. All of the mtDNA lineages are Native American. Two previous investigations studied populations spatially close to the FMM population. Torroni et al. (1994) also reported 100% Native American inclusion. Green et al. (2000) reports an average of 91% for two North-Central Mexican populations. Table 1 demonstrates that about 60% of the mtDNA lineages are members of haplogroup A. Both Green et al. (2000) and Torroni et al. (1994) also found haplogroup A to be the highest proportion (38% and 65.5% respectively). Mahli et al. (2002) demonstrated the highest concentration of haplogroup A is in California and spreads east, north and south. Thus it is not surprising to find it in such a high frequency in Mexico. The Network clade (Figure 2) shows extensive haplotypic diversity within haplogroup A, and to a lesser degree haplogroup B. It is of interest to note haplogroup C is in higher frequency than haplogroup B, but contains less diversity. Mahli et al. (2002), investigated North American natives, also saw the most diversity within haplogroup A types followed by haplogroup B. Also discovered were samples who had the characteristic control-region mutations of both haplogroup A and B but HVSI and II sequence characteristic of haplogroup A. Both Green et al. (2000) and Torroni et al. (1994) also discovered individuals who belong to haplogroup A that also have the 8 bp deletion of haplogroup B. It is hypothesized that this occurs in Mesoamerica but further research is required to confirm.

Figure 2: Network of Filomeno Mata mtDNA sequences. The number of individuals in each cluster is shown.

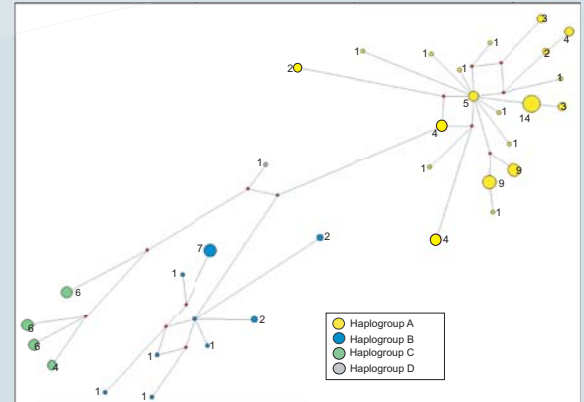


Table 1: Haplogroup frequencies and percentages of mtDNA in Filomeno Mata

Population	Haplogroup (%)					
	A	B	C	D	X	other
Filomeno Mata	118 (56%)	37 (18%)	9 (4%)	42 (20%)	3 (1%)	0 (0%)
FMM reduced*	65 (81%)	16 (14%)	3 (3%)	22 (20%)	1 (1%)	0 (0%)

*Filomeno Mata dataset with only one member of a known maternal relationship
** Samples are members of haplogroup A but also have the 8 bp deletion

Methods

163 samples were sequenced for both hypervariable regions including base pairs 16063-16569 and 1-306. Native American haplogroups were tested according to previously published methods (Torroni 1994, Smith 1999). The phylogenetic analysis of mtDNA sequences was performed by means of the MJ-network method (Bandelt et al. 1999) as implemented in Network (see the Network website). Only one individual from known maternal relationships was included in the Network clade. 54 males were typed for 36 Y-ch specific STRs as follows: DYS385, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS394, DYS426, DYS437, DYS438, DYS439, DYS441, DYS442, DYS444, DYS445, DYS446, DYS447, DYS448, DYS449, DYS452, DYS454, DYS455, DYS456, DYS458, DYS459, DYS460, DYS461, DYS462, DYS463, GGAAT1B07, YCAII, YGATAA10, YGATAC4, and YGATAH4. The 50 most common 8 locus haplotypes from Amerind and Eskimo Aleut populations from the Y Chromosome Haplotype reference Database (YHRD) (found at the YHRD website) were used as Native American reference populations. Six Spanish datasets and one Portuguese dataset (Gamerio 2002, Lopez 2004, Martin 2004, Zarrabeitia 2001, Fernandes 2003) comprising of 707 samples were used as Spanish/Portuguese reference populations. Y-ch SNPs were tested according to published methods (Bortolini 2003, Seielstad 2003, Underhill 2001). The genetic structure of populations was investigated by the analysis-of-molecular-variance approach (AMOVA), using Arlequin software (Schneider et al. 2000). Arlequin was also used to generate a Neighbor-joining tree of the Filomeno Mata population using 36 STRs. LEADMIX (Wang 2003) software (see the LEADMIX website) was used to estimate admixture levels of Spanish Y-chromosomes in the FMM population. Only 8 loci were used in Arlequin and LEADMIX to maximize the amount of overlapping loci and samples.

Y-chromosome Data

Whether the male founders of FMM share haplotypes with individuals living in Spain or Portugal was investigated. Y-ch STR data was compared to 6 Spanish datasets and one Portuguese dataset from the literature (Gamerio 2002, Lopez 2004, Martin 2004, Zarrabeitia 2001, Fernandes 2003) to identify identical haplotypes at 6 loci. Only two FMM family groups shared a 6 locus haplotype with any members from the Spain/Portugal datasets. Both of these were resolved when additional markers were compared (8 and 10 markers). AMOVA was used to calculate the amount of variance within and between populations (Table 2). More variance was seen between the Spanish/Portuguese and FMM populations than FMM and Amerind/Eskimo Aleut populations. LEADMIX was also used to evaluate this question. The Spanish/Portuguese and Amerind/Eskimo Aleut datasets were used as parent populations to the FMM and the proportion of admixture was determined. Table 3 shows no Spanish influence. It was expected that the FMM population would be more closely related to the Amerind/Eskimo Aleut populations but no influence was not expected. These findings were confirmed with Y-ch SNPs. Three SNPs were used to categorize samples into either Native American or Asian haplogroups. No samples fell into the Asian haplogroup and 64 of 70 samples (91%) were determined to be a part of the Native American haplogroup Q according to the nomenclature recommended by the YCC2003 Tree (Jobling 2003). Using DYS199 (M3) 60 of these samples (94%) were further characterized into haplogroup Q3. An explanation of the 9% that are not paternally Native American is they are from a different European Source population than Spain or Portugal. The STR data from these individuals supports an European origin. The diversity of the complete 36 locus haplotypes can be seen in Figure 3. Excluding known paternal relationships 11 haplotypes are shared and 30 are unique.

Table 2: AMOVA results of the Filomeno Mata population compared with Spanish/Portuguese and Amerind/Eskimo Aleut datasets.

Populations	Number of Groups	Number of Populations	Variance Components		P (a)
			Among Populations	Within Populations	
FMM-(Amerind/EskimoAleut)	2	3	-1.32	87.75	<0.000001
FMM-(Spain/Portugal)	2	8	19.06	76.58	<0.000001

(a) calculated by 10⁵ permutation tests

Figure 3: Neighbor-Joining tree of Filomeno Mata 36 locus Y-ch STR haplotypes. Blue box indicates inclusion into the Native American haplogroup Q.

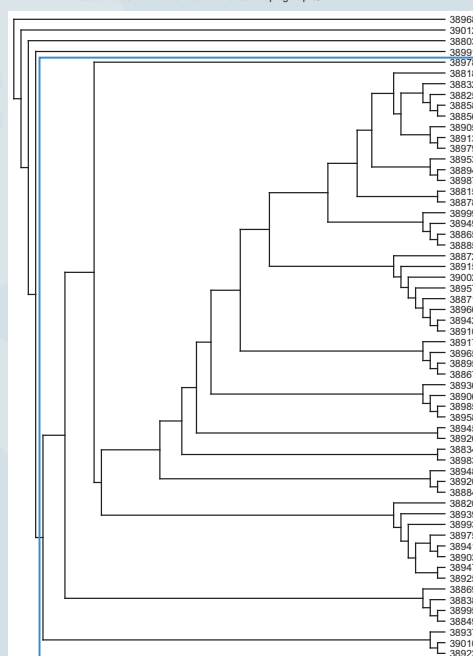


Table 3: Admixture of Spanish/Portuguese and Amerind/Eskimo Aleut in Filomeno Mata.

Populations	Admixture Proportion	Standard Deviation	Confidence Intervals	
			-95%	+95%
FMM-(Amerind/EskimoAleut)	1.0000	0.2809	0.8692	1.12016
FMM-(Spain/Portugal)	<0.1686	0.2809	-0.7016	0.3988
FMM-(Amerind/EskimoAleut)	1.4300	1.2261	-1.000	3.8202
FMM-(Spain/Portugal)	<0.1105	0.2328	-0.5652	0.3698

SD and 95%CI obtained from 1000 bootstrapping

References

Bandelt H-J, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 16:37-48.
Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nassan SPK, Bau CHD, et al. 2003. Y-Chromosome Evidence for Differing Ancient Demographic Histories in the Americas. *Am J Hum Genet.* 73:524-539.
Carvajal-Carmona LG, Soto ID, Pineda N, Ortiz-Barricatos D, Duque C, Ospina-Duque J, et al. 2000. Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *Am J Hum Genet.* 67:1287-1295.
Fernandes A, Behr M. 2003. Y-chromosome STR haplotypes in the Azores Archipelago (Portugal). *Forensic Sci Int.* 135:239-242.
Gamerio J-L, Romero J-L, Gonzalez J-L, Carvalho M, Anjos M-J, Real FC, Vide M-C. 2002. Y-chromosome STR haplotypes in a southern Spain population sample. *Forensic Sci Int.* 125:86-89.
Green LD, Derr JN, Knight A. 2000. mtDNA affinities of the peoples of North-Central Mexico. *Am J Hum Genet.* 66:989-998.
Jobling M, Tyler-Smith C. 2003. The Human Y Chromosome: an Evolutionary Marker Comes of Age. *Nature.* 4:598-612.
Lopez AM, Alvarez S, Gusmao L, Alves C, Mesa MS, Albornos A, et al. 2004. Population data for 16 Y-chromosome STRs in four populations from Pynenes (Spain). *Forensic Sci Int.* 140:125-129.
Mahli R, Eshleman JA, Greenberg JA, Weiss DA, Shook BAS, Kacstle FA, et al. 2002. The structure of Diversity within New World mitochondrial DNA haplogroups: Implications for the prehistory of North America.
Martin P, Garcia-Hirschfeld J, Garcia O, Gusmao L, Garcia P, Albarran C, et al. 2004. A Spanish population study of 17 Y-chromosome STR loci. *Forensic Sci Int.* 139:231-235.
Mesa NR, Mondragon MC, Soto ID, Parra MV, Duque C, Ortiz-Barricatos D, et al. 2000. Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: pre- and post-Columbian patterns of gene flow in South America. *Am J Hum Genet.* 67:1277-1286.
Santos FR, Pandya A, Tyler-Smith C, Pena SD, Schanfield M, Leonard WR, et al. 1999. The central Siberian origin of native American Y chromosomes. *Am J Hum Genet.* 64:619-628.
Schneider S, Kueffer J-M, Roessli D, Excoffier L. 2000. Arlequin version 2.000: a software for population genetic analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva.
Seielstad M, Yuldashova N, Singh N, Underhill P, Oefner P, Shen P, Wells RS. 2003. A Novel Y Chromosome Variant Puts an Upper Limit on the Timing of First Entry into the Americas. *Am J Hum Genet.* 73:700-705.
Smith DG, Mahli RS, Eshleman J, Lorenz JG, Kacstle FA. 1999. Distribution of mtDNA haplogroup X Among Native North Americans. *Am J Phys Anthropol.* 110:271-284.
Torroni A, Chen Y-S, Semino O, Santachiara-Beneccheri AS, Scott CR, Lott MT et al. 1994. mtDNA and Y-Chromosome Polymorphisms in Four Native American Populations from Southern Mexico. *Am J Hum Genet.* 54:303-318.
Underhill P, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA et al. 2001. The Phylogeography of Y Chromosome Binary Haplotypes and the Origins of Modern Human Populations. *Ann Hum Genet.* 65:43-62.
Wang J. 2003. Maximum-likelihood estimation of admixture proportions from genetic data. *Genetics.* 164:747-765.
Zarrabeitia MT, Rancho JA, Sanchez-Diz P, Sanchez-Velasco P. 2001. 7-locus Y chromosome haplotype profiling in a northern Spain population. *Forensic Sci Int.* 123:78-80.
Electronic Database Information:
YHRD <http://www.yhrd.org/index.html>
Arlequin <http://anthro.unige.ch/arlequin/>
LEADMIX <http://www.zoo.cim.ac.uk/izoz/software.htm>

Conclusions

Our results do not show much of a difference in the population origin of the female and male founders of Filomeno Mata. The maternal lineages are exclusively Native American. Compared to a 9% quantity of Old World influence in the male lineages. Such low proportions of outside influence are not uncommon in the literature. Out of 5 populations studied in Mesa et al. (2000) one population found 3% European admixture. Torroni et al. (1994) also reports all of the maternal ancestors of 2 Mexican populations are native. A common phenomenon is sex-biased admixture (Carvajal-Carmona 2000, Green 2000, Mesa 2000, Santos 1999). This bias is explained as being a result of "directional mating" (Wen 2004). Admixture occurred with European men and native women due to the sex ratio of European colonizers (Carvajal-Carmona 2000, Wen 2004). Spanish records show a relatively small number (10% between 1509-1559) of women migrated to the Americas (Carvajal-Carmona 2000). While there is a difference between the maternal and paternal origins, it is not as significant as other reports. This suggests this particular population was quite isolated and did not merge with the Spanish as much as other Totonacs.

