Asymmetric Male and Female Genetic Histories among Native Americans from Eastern North America

Deborah A. Bolnick, Daniel I. Bolnick,† and David Glenn Smith‡§*

*Department of Anthropology, University of Texas at Austin; †Section of Integrative Biology, University of Texas at Austin; ‡National Primate Research Center, University of California, Davis; and §Department of Anthropology, University of California, Davis

Previous studies have investigated the human population history of eastern North America by examining mitochondrial DNA (mtDNA) variation among Native Americans, but these studies could only reconstruct maternal population history. To evaluate similarities and differences in the maternal and paternal population histories of this region, we obtained DNA samples from 605 individuals, representing 16 indigenous populations. After amplifying the amelogenin locus to identify males, we genotyped 8 binary polymorphisms and 10 microsatellites in the male-specific region of the Y chromosome. This analysis identified 6 haplogroups and 175 haplotypes. We found that sociocultural factors have played a more important role than language or geography in shaping the patterns of Y chromosome variation in eastern North America. Comparisons with previous mtDNA studies of the same samples demonstrate that male and female demographic histories differ substantially in this region. Postmarital residence patterns have strongly influenced genetic structure, with patrilocal and matrilocal populations showing different patterns of male and female gene flow. European contact also had a significant but sex-specific impact due to a high level of male-mediated European admixture. Finally, this study addresses long-standing questions about the history of Iroquoian populations by suggesting that the ancestral Iroquoian population lived in southeastern North America.

Introduction

To date, most phylogeographic studies have used mitochondrial DNA (mtDNA) to reconstruct the demographic and evolutionary histories of populations (Avise 2000; Hare 2001; Ballard and Whitlock 2004). Although this approach has yielded many important insights, inferences from mtDNA may not always be congruent with those based on other genetic loci. Stochastic processes and natural selection can affect individual loci differently than the genome (or population) as a whole, so analyses of only a single locus may result in false inferences about population history or the relationships between populations (Moore 1995; Hare 2001; Williams et al. 2002; Ballard and Whitlock 2004). In addition, because mtDNA is maternally inherited, it reflects only maternal relationships and female patterns of gene flow and migration (Avise et al. 1987). If male and female demographic histories have differed in any way, mtDNA will not tell the complete story.

Several recent studies have investigated whether male and female demographic histories differ in humans. Although one multilocus study detected no male–female differences on a global scale (Wilder et al. 2004), most comparisons of mtDNA and Y chromosome variation have identified sex-specific patterns of genetic diversity. Many of these comparative studies found less mtDNA differentiation than Y chromosome differentiation, suggesting that females have experienced higher rates of migration and gene flow compared with males and/or lower rates of genetic drift due to sex differences in effective population sizes (Seielstad et al. 1998; Perez-Lezaun et al. 1999; Fagundes et al. 2002; Oota et al. 2002; Dupanloup et al. 2003; Malyarchuk et al. 2004; Nasidze et al. 2004). However, some South American populations exhibit no male–female differences (Bortolini et al. 2002; Fuselli et al. 2003), and some North and Central American populations exhibit the opposite pattern, indicating higher male migration and/or larger male effective population sizes (Bortolini et al. 2002). A variety of factors may contribute to these sex-specific patterns, including sex-biased admixture among populations (Mesa et al. 2000; Wilson et al. 2001; Wen et al. 2004; Wood et al. 2005) and population-specific aspects of social structure, such as postmarital residence patterns (Seielstad et al. 1998; Oota et al. 2001; Malyarchuk et al. 2004; Nasidze et al. 2004; Hamilton et al. 2005).

To further elucidate the causes of sex-specific diversity patterns in humans, we compare mtDNA and Y chromosome variation among Native Americans from eastern North America. Eastern North America is an ideal setting for this investigation because the indigenous populations exhibit diverse cultures, languages, and histories. Both matrilineal and patrilineal social systems are common, and 4 principal language families (Algonquian, Siouan, Muskogean, and Iroquoian) are represented (Campbell 1997).

Comparative linguistic and archeological records suggest that Algonquian-speaking people moved into the Great Lakes region from western North America before expanding across eastern North America (Siebert 1967; Denny 1991; Fiedel 1991; Goddard 1994). Siouan populations likely lived in the Southeast or Midwest at first, but the Algonquian expansion pushed them into the Plains (Chafe 1976; Campbell 1997). Muskogean populations are found only in southeastern North America (Haas 1941), whereas Iroquoian-speaking populations are more geographically dispersed. The Iroquoian language family contains a northern branch of several closely related languages that are spoken in the lower Great Lakes region and a distantly related southern branch (Cherokee) in southern Appalachia. Given this geographic distribution, the history of the Iroquoian-speaking populations has been the subject of particularly intense debate. Some scholars cite linguistic and archeological evidence for ancestral Iroquoian lands in the Southeast and a subsequent migration to the Northeast (Parker 1916; Snow 1980, 1995, 1996), whereas others favor a model of in situ development for the northern Iroquoians and a later

Key words: Native Americans, mitochondrial DNA, Y chromosome, European admixture, social structure, Iroquoian history.

E-mail: deborah.bolnick@mail.utexas.edu.

Mol. Biol. Evol. 23(11):2161–2174. 2006
doi:10.1093/molbev/msl088
Advance Access publication August 17, 2006

FIG. 1.—Geographic locations of the sampled populations in early historic times. The dashed line separates the northeastern and southeastern culture areas.

southward migration of the Cherokee (MacNeish 1952; Lounsbury 1978; Crawford and Smith 1996).

Although mtDNA studies have not resolved questions about the history of Iroquoian populations, they have helped reconstruct other aspect of eastern North American prehistory. Malhi et al. (2001) found evidence of extensive gene flow and admixture among Algonquian and Siouan populations in the Northeast, and ancient DNA studies of prehistoric burial populations confirmed that this pattern of regional continuity predates European contact (Stone and Stoneking 1998; Mills 2003; Bolnick 2005; Shook 2005). In contrast, Native American populations from southeastern North America exhibit greater genetic differentiation (Bolnick and Smith 2003). Because the Southeast also exhibits lower overall levels of mtDNA diversity, this differentiation likely reflects genetic drift due to the population decline following European contact (Bolnick and Smith 2003). However, because these inferences are all based only on mtDNA, it is important to assess any differences in male and female demographic histories before drawing firm conclusions about the population history of eastern North America.

In this study, we examine variation on the Y chromosome to explore such differences and to improve our understanding of eastern North American prehistory. The male-specific region of the Y chromosome is strictly paternally inherited, so it can be used to assess paternal relationships and male patterns of gene flow and migration (Hurles and Jobling 2001). Previous studies of Native American Y chromosome variation included few eastern North American populations, and virtually all such studies focused on issues related to the initial peopling of the continent (Underhill et al. 1996; Bianchi et al. 1998; Karafet et al. 1999; Ruiz-Linares et al. 1999; Santos et al. 1999; Lell et al. 2002; Zegura et al. 2004) rather than on the subsequent, region-specific history of the past 10,000 years. This study therefore characterizes Y chromosome variation among Native Americans from eastern North America to investigate 1) the population history of the last 10,000 years, 2) the relative influences of geography, language and culture on genetic structure in this region, and 3) male–female differences in gene flow and migration.

Materials and Methods

Population Samples

We obtained samples from 605 individuals, representing 16 eastern North American populations (fig. 1 and table 1). All individuals reported having at least 75% ancestry from one of these populations. The samples were previously analyzed for mtDNA markers (Lorenz and Smith 1996, 1997; Smith et al. 1999; Malhi et al. 2001; Bolnick and Smith 2003; Shook 2005), and detailed information about their sources can be found in those publications. The Office of Human Research Protection at the University of California, Davis approved all sampling protocols.

The populations included in this study can be divided into 2 culture areas, the Northeast and the Southeast. Populations from the northeastern culture area generally exhibited patrilineal kinship systems and patrilocal postmarital residence (male philopatry), whereas populations from the southeastern culture area exhibited matrilineal kinship systems and matrilocal postmarital residence (female philopatry). Historically, these populations also spoke languages belonging to the 4 principal language families in eastern North America (Algonquian, Siouan, Iroquoian, and Muskogean).

Table 1
Y chromosome Haplogroup Frequencies in Eastern North American Populations

Population	Region	Language Family	N	Q-M3*	Q-M242*	C-M130	R-M173	P-M45*	DE-YAP	Other
Turtle Mountain Chippewa	Northeast	Algonquian	51	0.059	0.000	0.000	0.549	0.000	0.020	0.373
Wisconsin Chippewa	Northeast	Algonquian	37	0.081	0.135	0.000	0.541	0.000	0.027	0.216
Minnesota Chippewa	Northeast	Algonquian	9	0.222	0.222	0.444	0.111	0.000	0.000	0.000
Cheyenne/Arapaho[a]	Northeast	Algonquian	53	0.151	0.472	0.151	0.151	0.019	0.038	0.019
Shawnee	Northeast	Algonquian	1	1.000	0.000	0.000	0.000	0.000	0.000	0.000
Micmac	Northeast	Algonquian	1	0.000	0.000	0.000	1.000	0.000	0.000	0.000
Kickapoo	Northeast	Algonquian	2	1.000	0.000	0.000	0.000	0.000	0.000	0.000
Fox	Northeast	Algonquian	1	1.000	0.000	0.000	0.000	0.000	0.000	0.000
Sisseton/Wahpeton Sioux	Northeast	Siouan	27	0.259	0.259	0.296	0.148	0.000	0.000	0.037
Omaha	Northeast	Siouan	1	1.000	0.000	0.000	0.000	0.000	0.000	0.000
Oklahoma Red Cross Cherokee	Southeast	Iroquoian	27	0.333	0.259	0.037	0.259	0.037	0.037	0.037
Stillwell Cherokee	Southeast	Iroquoian	35	0.314	0.114	0.000	0.400	0.029	0.029	0.114
Chickasaw	Southeast	Muskogean	6	0.667	0.167	0.000	0.167	0.000	0.000	0.000
Choctaw	Southeast	Muskogean	12	0.583	0.333	0.000	0.083	0.000	0.000	0.000
Creek	Southeast	Muskogean	15	0.333	0.267	0.067	0.133	0.000	0.000	0.200
Seminole	Southeast	Muskogean	3	0.667	0.000	0.000	0.000	0.000	0.000	0.333
Northeast			183	0.153	0.213	0.109	0.339	0.005	0.022	0.158
Southeast			98	0.388	0.204	0.020	0.255	0.020	0.020	0.092
Eastern North America (Total)			281	0.235	0.210	0.078	0.310	0.011	0.021	0.135

[a] This sample includes 44 Cheyenne individuals studied by Zegura et al. (2004).

Genetic Analyses

We extracted DNA from 200 μl of serum using the QIAamp DNA Blood Kit (Qiagen, Valencia, CA). Because the sex of each individual was not recorded when the samples were collected anonymously, we identified males by screening part of the amelogenin gene that differs in size between the X and Y chromosomes (Sullivan et al. 1993). We then amplified 8 binary polymorphisms (M19, M3, M242, M173, M45, M130, YAP, and Tat) to define Y chromosome haplogroups. Polymerase chain reaction (PCR) amplifications were performed in a 25 μl volume with 1–5 μl of DNA template, 67 mM Tris-Cl (pH 8.8), 16 mM $(NH_4)_2SO_4$, 0.01% Tween-20, 0.05 mM of each dNTP, 1.7 mM $MgSO_4$, 0.2 mM of each primer, and 0.025 units/μl of Platinum Taq (Invitrogen, Carlsbad, CA). All PCRs were performed with an initial denaturation at 95 °C for 3 min and a final extension at 72 °C for 5 min. Interim cycling conditions and primer sequences are provided in the Supplementary Material online. PCR products containing the M3, M242, M45, M130, and Tat polymorphisms were digested with the appropriate restriction enzyme (Supplementary Material online), and digestion products were visualized with ethidium bromide on 6% polyacrylamide gels. PCR products containing the M19 and M173 polymorphisms were purified with *Exo*I (0.25 units/μl) and Montage PCRμ96 plates (Millipore, Billerica, MA) before being sequenced at the CA&ES Genomics Facility at the University of California, Davis.

These binary polymorphisms define 7 haplogroups, which are named using the nomenclature recommended by the Y Chromosome Consortium (2002). The phylogenetic relationships among these haplogroups are shown in figure 2. Q-M3*, Q-M242*, and C-M130 are thought to be founding Native American haplogroups (Zegura et al. 2004), whereas Q-M19 represents a post-colonization mutational event that occurred in the Americas (Bortolini et al. 2003). The other haplogroups likely represent recent non–Native American admixture (Zegura et al. 2004 and discussed below).

To define haplotypes within each haplogroup, we amplified 1 trinucleotide-repeat microsatellite (DYS392) and 9 tetranucleotide-repeat microsatellites (DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, and DYS393, DYS439) for 214 individuals using the Y-PLEX™ 12 amplification kit (Reliagene, New Orleans, LA). PCR products were separated on an ABI 310 Genetic Analyzer and analyzed using the GENESCAN software package (version 3.1.2, Applied Biosystems, Foster City, CA). We obtained comparative data from Bortolini et al. (2003), Zegura et al. (2004), and the Y-STR Haplotype Reference Database (YHRD; www.ystr.org), which

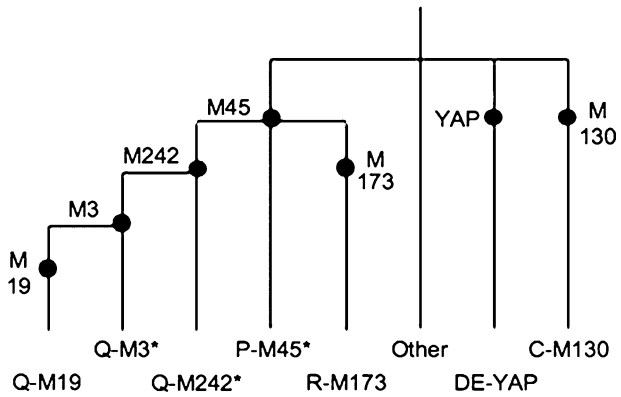

FIG. 2.—Phylogenetic relationships among the Y chromosome haplogroups.

contained a worldwide sample of 32,196 haplotypes from 271 populations (Release 16; Roewer et al. 2001).

Statistical Analyses

Unless otherwise noted, we excluded individuals with suspected non–Native American admixture from statistical analyses and performed analyses only on population samples containing more than 2 individuals. Some populations were therefore represented by small samples, but because most statistical analyses take sample size into account, any significant results should be robust to low sample sizes.

Population haplogroup frequencies were compared in pairs using exact tests of population differentiation to test the null hypothesis of identical haplogroup frequencies in the 2 populations being compared. Exact tests were performed using the Arlequin 2.000 software package (Schneider et al. 2000), and P values based on 10,000 Markov steps are reported. P values for the pairwise exact tests were corrected for multiple comparisons with a sequential Bonferroni correction (Rice 1989). We also used Arlequin 2.000 to estimate haplogroup and haplotype diversities (h), mean numbers of pairwise differences among haplotypes (p), and pairwise R_{ST} values. Both h and p were estimated as measures of microsatellite haplotype diversity to ensure that both haplotype frequencies and the molecular differences among haplotypes were taken into consideration. Correlations among the 3 estimates of genetic diversity were measured using Pearson correlations with Bonferroni-corrected probabilities, and the values of h and p for the 2 culture areas were compared using 2-sample t-tests (Zar 1999).

Following Qamar et al. (2002), we calculated weighted means of within-haplogroup R_{ST} (WMWH-R_{ST}) values as a measure of population differentiation that is less influenced by differences that accumulated between haplogroups before colonization of the Americas. Multidimensional scaling (MDS) analysis was performed using the SYSTAT 9 software package (SPSS), with WMWH-R_{ST} values as variables, to investigate the genetic relationships among populations.

We constructed haplotype median-joining networks for each haplogroup using the NETWORK 3.0 program, which assumes a stepwise mutation model for microsatellite evolution (Bandelt et al. 1999). Because DYS389II includes DYS389I, we subtracted the DYS389I repeat size from DYS389II to derive DYS389b. A 6-fold weighting scheme was used when constructing the networks, and the weights assigned to microsatellites were specific for each haplogroup. The following weights were used: variance 0–0.19, weight 6; variance 0.2–0.29, weight 5; variance 0.3–0.39, weight 4; variance 0.4–0.59, weight 3; variance 0.6–0.89, weight 2; and variance ≥ 0.9, weight 1.

To assess the relative influences of language and culture on genetic structure, we performed an analysis of molecular variance (AMOVA) using Arlequin 2.000. We also performed Mantel tests in Arlequin to assess the impact of geography and language on genetic structure. Great-circle geographic distances between populations were calculated from latitude–longitude data, and linguistic distances were calculated following the method of Zegura et al. (2004) but with the language classification of Campbell (1997). Geographic and linguistic distances are given in the Supplementary Material online. It should be noted that significant P values for the partial Mantel tests should be treated with caution because there is some debate over their accuracy (Raufaste and Rousset 2001; Castellano and Balletto 2002; Rousset 2002).

Differences in the genetic structure of males and females were assessed using the parameter $N\nu$, which is calculated as $(1/\Phi_{ST}) - 1$ according to the island migration model for haploid systems (Cavalli-Sforza and Bodmer 1971). This parameter incorporates migration rate, mutation rate, and effective population size, but because the contribution of mutation rate to the $N\nu$ parameter may be considered negligible for the genetic systems studied here (Destro-Bisol et al. 2004), different $N\nu$ values reflect different migration rates and/or effective population sizes between populations.

In addition, maximum likelihood estimates of male and female migration rates between populations were calculated using the program MIGRATE (Beerli and Felsenstein 2001). We estimated male migration rates using the Y chromosome microsatellite data and a Brownian motion approximation of the stepwise mutation model for microsatellites. MIGRATE performed 10 replicate runs using the adaptive heating scheme with 5 heated chains. Each run contained 10 primary short chains (4,000 genealogies per chain) and 3 primary long chains (40,000 genealogies per chain), and migration rate estimates were averaged across runs. We estimated female migration rates using the previously published mtDNA hypervariable region I sequences from these populations (Lorenz and Smith 1997; Malhi et al. 2001; Bolnick and Smith 2003). Similar MIGRATE conditions were used, but with 15 short chains of 100,000 genealogies each and 3 long chains of 1,000,000 genealogies each. For each pair of populations, we calculated the effective number of migrants per generation (N_em) by summing the unidirectional N_em estimates provided by MIGRATE.

Results

Y Chromosome Haplogroups

Of the 605 samples examined, 261 males and 344 females were identified. Twenty-four male samples contained insufficient quantities of DNA for further analysis, so we analyzed 237 Y chromosomes. We identified 6 of the 7 possible haplogroups, with 37 individuals falling into the "Other" category. The remaining 200 individuals represent 59 members of haplogroup Q-M3*, 39 of Q-M242*, 15 of C-M130, 80 of R-M173, 3 of P-M45*, and 4 of DE-YAP. Haplogroup frequencies for all eastern North American populations are given in table 1, and haplogroup diversities (h) are given in table 2. The Northeast exhibits significantly higher levels of haplogroup diversity than the Southeast ($P < 0.001$).

Four haplogroups occur at frequencies greater than 5% (Q-M3* = 23.5%, Q-M242* = 21.0%, C-M130 = 7.8%, and R-M173 = 31.0%), and together they comprise 83.3% of the eastern North American sample. Haplogroup R-M173 likely represents recent (post-1492) European

Table 2
Haplogroup and Haplotype Diversity Estimates for Eastern North America[a]

Population	Haplogroup *N*	Haplogroup Diversity (*h*)	Haplotype *N*	Number of Haplotypes	Haplotype Diversity (*h*)	Mean Number of Pairwise Differences (*p*)
Turtle Mountain Chippewa	3	0.000	3	1	0.000	0.000
Wisconsin Chippewa	8	0.536	7	5	0.905	0.340
Minnesota Chippewa	8	0.714	8	6	0.893	0.529
Cheyenne/Arapaho	41	0.566	3	3	1.000	0.533
Sisseton/Wahpeton Sioux	22	0.697	20	16	0.974	0.594
Stillwell Cherokee	15	0.419	13	12	0.987	0.505
Oklahoma Red Cross Cherokee	17	0.581	16	15	0.992	0.551
Chickasaw	5	0.400	4	4	1.000	0.533
Choctaw	11	0.509	11	10	0.982	0.503
Creek	10	0.644	9	9	1.000	0.617
Northeast	87	0.650	45	33	0.984	0.585
Southeast	60	0.495	55	47	0.993	0.543
Eastern North America	147	0.619	100	80	0.995	0.587

[a] Based only on haplogroups Q-M3*, Q-M242*, and C-M130. Populations with $N < 3$ are not listed separately but are included in the diversity estimates for the Northeast, Southeast, and Eastern North America as a whole.

admixture, as may P-M45* (Tarazona-Santos and Santos 2002; Bosch et al. 2003; Zegura et al. 2004). Y chromosomes belonging to haplogroup DE-YAP probably result from recent admixture with individuals of African or European ancestry (Karafet et al. 1999; Lell et al. 2002). Following Bosch et al. (2003), these findings suggest that at least 34.2 ± 3% of eastern North American Y chromosomes result from recent admixture. This figure may actually be as high as 47.7 ± 3%, if the 37 "Other" Y chromosomes represent additional European or African lineages rather than previously unidentified founding Native American lineages. Admixture estimates for the Northeast (36.6–52.5% depending on whether the "Other" lineages represent admixture) are higher than those for the Southeast (29.6–38.8%) due to extremely high estimates for 2 of the Chippewa populations (56.9–94.1% for the Turtle Mountain Chippewa and 56.8–78.4% for the Wisconsin Chippewa).

Frequencies of the founding Native American haplogroups (Q-M3*, Q-M242*, and C-M130) also differ between the 2 areas. Populations in the southeastern culture area exhibit similar haplogroup frequencies, with both Muskogean and Iroquoian populations showing high frequencies of haplogroup Q-M3*, moderate frequencies of haplogroup Q-M242*, and low frequencies of haplogroup C-M130 (if present at all). In the northeastern culture area, populations generally exhibit lower frequencies of Q-M3* and higher frequencies of C-M130.

Exact tests of population differentiation confirm these patterns. The haplogroup frequencies of the northeastern and southeastern culture areas are significantly different ($P = 0.001$). When individuals are grouped by language family, all comparisons yield statistically significant differences ($P \leq 0.010$) except the Muskogean–Iroquoian ($P = 0.891$) and Algonquian–Siouan ($P = 0.155$) comparisons. The Wisconsin Chippewa also differ significantly from the Seminole ($P < 0.001$), and the Cheyenne/Arapaho differ significantly from the Stillwell Cherokee ($P < 0.001$). Other comparisons of individual populations are not statistically significant, perhaps because the small sample sizes for some populations provide relatively low statistical power. Alternatively, the Bonferroni correction for multiple comparisons may be too conservative: it assumes that all tests are statistically independent, but these exact tests involve the repeated use of each population.

Y Chromosome Microsatellite Haplotypes

To obtain a more detailed picture of Y chromosome variation in eastern North America, we also genotyped 214 Y chromosomes for 10 Y-specific microsatellite loci. These microsatellites define 173 different haplotypes, of which 147 (84%) are unique. Two haplotypes are shared by Y chromosomes belonging to different haplogroups, resulting in 175 combination haplotypes (see Supplementary Material online). Haplotype diversity (*h*) is generally high (table 2), with the Southeast exhibiting a statistically higher estimate than the Northeast ($P < 0.001$). In contrast, the mean number of pairwise differences among haplotypes (*p*) is higher in the Northeast, but not significantly so ($P = 0.494$). Overall, number of pairwise differences is significantly correlated with haplogroup diversity ($r = 0.828$, $P = 0.010$), but haplotype diversity is not ($r = 0.570$, $P = 0.404$).

Haplotype median-joining networks show that some population-specific substructure exists, but haplotypes generally cluster by culture area, with different populations from the same culture area sharing closely related haplotypes. For example, the haplotype network for haplogroup Q-M3* (fig. 3) exhibits one population-specific cluster, containing 2 Sioux haplotypes and several area-specific clusters of haplotypes. In the haplotype network for haplogroup Q-M242* (fig. 4), all 19 southeastern individuals cluster together, as do all but 3 northeastern individuals. The haplogroup C-M130 network (fig. 5) exhibits similar groupings: the haplotype shared by 2 southeastern populations is 10 mutational steps away from a cluster of 7 closely related northeastern haplotypes. This northeastern cluster contains 2 smaller population-specific clusters, one representing 8 Sioux and the other representing 4 Minnesota Chippewa.

In contrast to the Q-M3*, Q-M242*, and C-M130 networks, which all exhibit area-specific clusters of haplotypes,

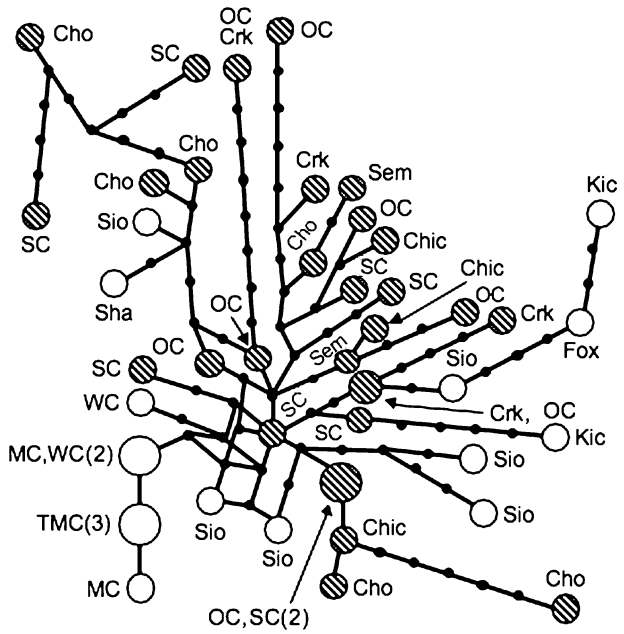

FIG. 3.—Haplotype network for haplogroup Q-M3*. Circle size reflects the frequency of that haplotype in this sample. White circles represent northeastern haplotypes, striped circles represent southeastern haplotypes, and small black circles represent intermediate haplotypes not identified in these samples. MC, Minnesota Chippewa; WC, Wisconsin Chippewa; Sio, Sisseton/Wahpeton Sioux; OC, Oklahoma Red Cross Cherokee; SC, Stillwell Cherokee; Chic, Chickasaw; Cho, Choctaw; Crk, Creek.

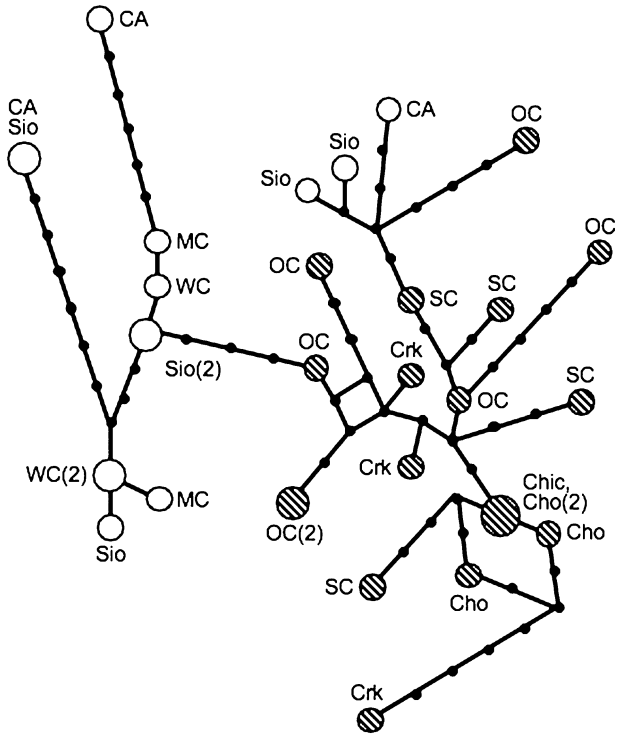

FIG. 4.—Haplotype network for haplogroup Q-M242*. Circle size reflects the frequency of that haplotype in this sample. White circles represent northeastern haplotypes, striped circles represent southeastern haplotypes, and small black circles represent intermediate haplotypes not identified in these samples. MC, Minnesota Chippewa; WC, Wisconsin Chippewa; Sio, Sisseton/Wahpeton Sioux; CA, Cheyenne/Arapaho; OC, Oklahoma Red Cross Cherokee; SC, Stillwell Cherokee; Chic, Chickasaw; Cho, Choctaw; Crk, Creek.

the R-M173 network (fig. 6) shows no clear-cut patterns. Haplotypes in this network do not cluster by population, culture area, language family, or geography (fig. 6*a*). This lack of structure is consistent with the hypothesis that haplogroup R-M173 represents recent (post-1492) European admixture in eastern North America rather than a founding Native American lineage (Tarazona-Santos and Santos 2002; Bosch et al. 2003; Zegura et al. 2004). Comparative data from the YHRD, Zegura et al. (2004), and Bortolini et al. (2003) further support this conclusion: the most common R-M173 haplotype in eastern North America is also the most common R-M173 haplotype in Europe, but this haplotype is rare in Asia (and therefore unlikely to be a founding lineage). In addition, the 62 other R-M173 haplotypes found in eastern North America appear to be rare or absent in Asia, whereas 43 of them are common in Europe and 14 others are one mutational step away from known European haplotypes (fig. 6*b*). Five haplotypes are 4–8 mutational steps away from European haplotypes, but they probably also stem from recent European admixture: the median-joining network suggests that they derive from separate European haplotypes (fig. 6*b*), whereas one would expect them to cluster together in a distinct clade if they represented a founding lineage from Asia. Thus, few, if any, R-M173 haplotypes in eastern North America are likely derived from founding lineages of the Americas.

Comparative data from the YHRD also suggest that most individuals in the "Other" category have European Y chromosome haplotypes. Seventeen of the 27 "Other" haplotypes are shared with Europeans, and 8 additional haplotypes are one mutational step away from known European haplotypes. These findings indicate that the "Other" lineages most likely represent recent non–Native American admixture rather than additional founding Native American lineages. Similarly, the haplotypes in haplogroup DE-YAP match or are closely related to African and European haplotypes, suggesting that they also represent recent non–Native American admixture.

MDS Analysis

Figure 7 shows the results of the MDS analysis. A low stress value was obtained (0.085), indicating a good

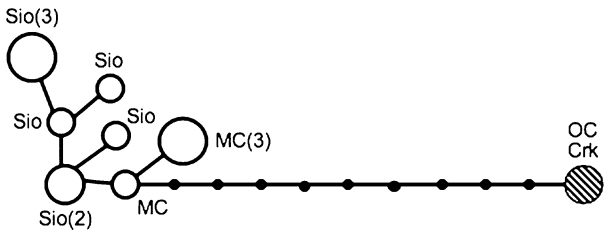

FIG. 5.—Haplotype network for haplogroup C-M130. Circle size reflects the frequency of that haplotype in this sample. White circles represent northeastern haplotypes, striped circles represent southeastern haplotypes, and small black circles represent intermediate haplotypes not identified in these samples. MC, Minnesota Chippewa; Sio, Sisseton/Wahpeton Sioux; OC, Oklahoma Red Cross Cherokee; Crk, Creek; Sem, Seminole.

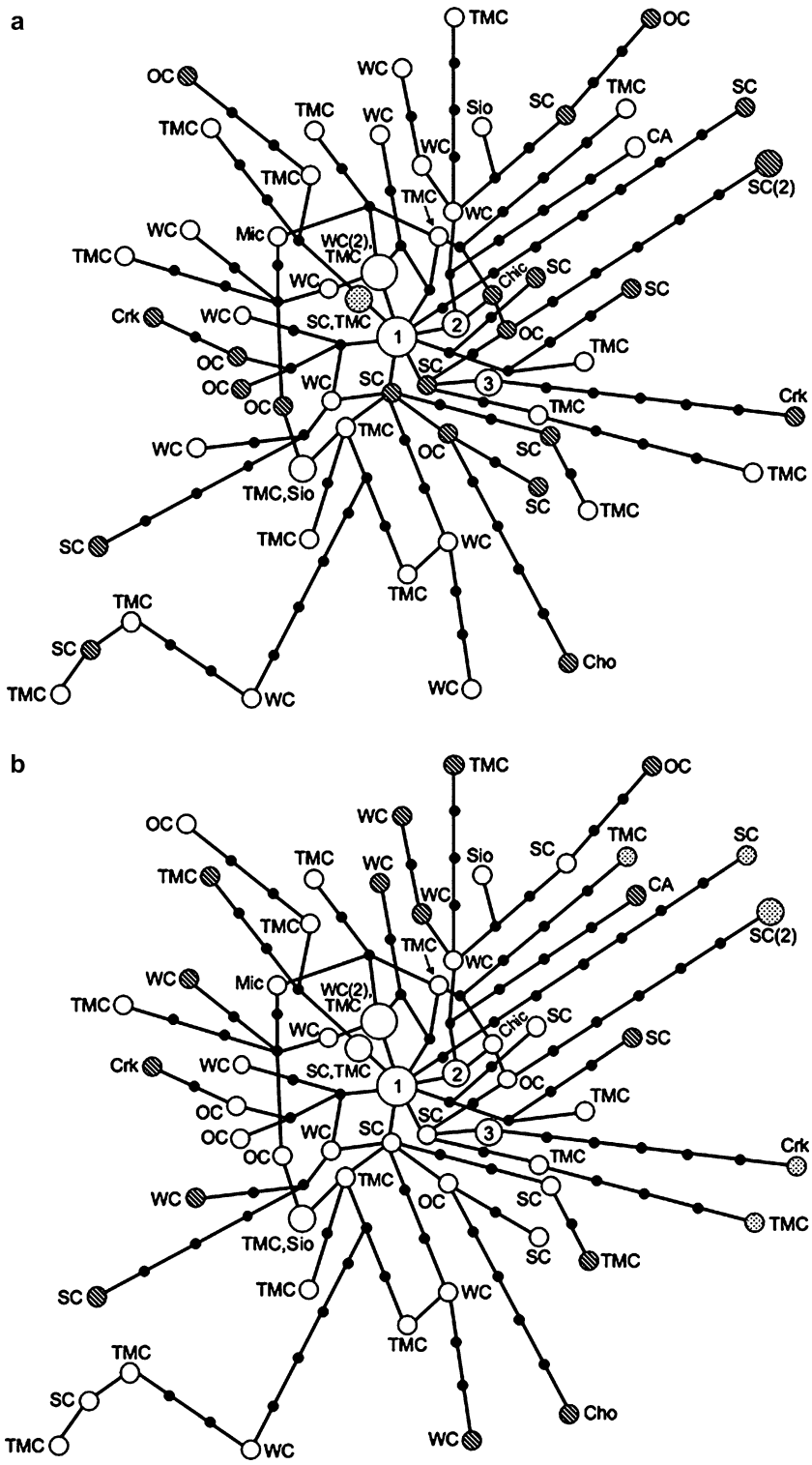

FIG. 6.—Haplotype network for haplogroup R-M173. Haplotype 1 was found in 1 Minnesota Chippewa, 2 Wisconsin Chippewa, and 3 Turtle Mountain Chippewa. Haplotype 2 was found in 2 Turtle Mountain Chippewa. Haplotype 3 was found in 2 Turtle Mountain Chippewa. (*a*) White circles represent northeastern haplotypes, striped circles represent southeastern haplotypes, and dotted circles represent haplotypes shared by both northeastern and southeastern individuals. (*b*) White circles represent haplotypes shared with Europeans, striped circles represent haplotypes that are one mutational step away from European R-M173 haplotypes, and dotted circles represent haplotypes not shared with Europeans. TMC, Turtle Mountain Chippewa; WC, Wisconsin Chippewa; Sio, Sisseton/Wahpeton Sioux; CA, Cheyenne/Arapaho; Mic, Micmac; OC, Oklahoma Red Cross Cherokee; SC, Stillwell Cherokee; Chic, Chickasaw; Crk, Creek; Sem, Seminole.

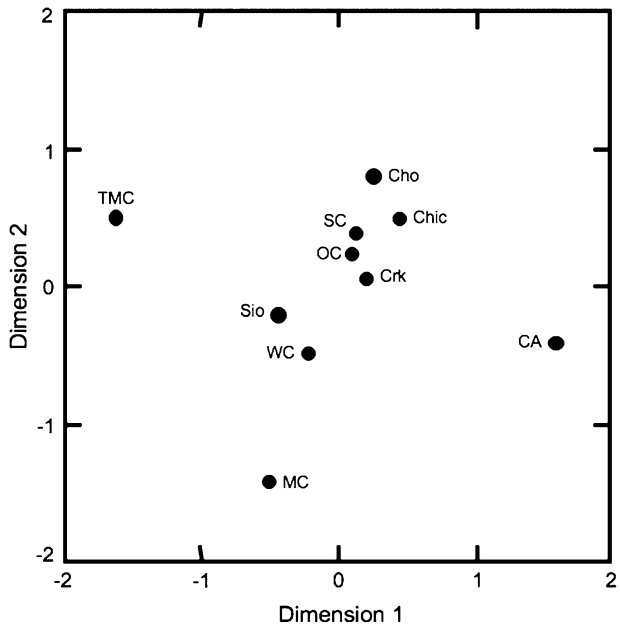

FIG. 7.—MDS plot based on WMWH-R_{ST} values. TMC, Turtle Mountain Chippewa; MC, Minnesota Chippewa; WC, Wisconsin Chippewa; Sio, Sisseton/Wahpeton Sioux; CA, Cheyenne/Arapaho; OC, Oklahoma Red Cross Cherokee; SC, Stillwell Cherokee; Chic, Chickasaw; Cho, Choctaw; Crk, Creek.

fit between the 2-dimensional plot and the source data (WMWH-R_{ST} values). All 5 populations from the southeastern culture area (Choctaw, Chickasaw, Creek, Stillwell Cherokee, and ORC Cherokee) cluster together in the MDS plot, whereas only 2 populations from the northeastern culture area do (the Sioux and Wisconsin Chippewa; the other 3 northeastern populations appear as separate outliers). The MDS analysis therefore suggests that populations in the southeastern culture area share closer paternal relationships than populations in the northeastern culture area.

Analysis of Molecular Variance and Mantel Tests

We performed an AMOVA to investigate the relative influences of culture and language on the genetic structure of eastern North American populations. AMOVAs based on haplogroup frequencies are shown in table 3. When the populations are grouped by culture area (2 groups), most genetic variation is found within populations, but a significant portion of the total genetic variance is due to differences between the 2 culture areas. Variation among populations within the same culture area is not statistically significant. When the populations are grouped by language family (4 groups), within-population variation is again significant, but the language groupings themselves do not account for a statistically significant portion of the genetic variance in eastern North America. AMOVAs based on microsatellite haplotypes generally yield similar results (table 3), but they differ in suggesting that a significant amount of variation exists among populations within the same culture area or language family.

Because culture areas also comprise geographic regions in this study, the effects of culture and geography may be conflated in these AMOVAs. Consequently, we performed Mantel tests based on WMWH-R_{ST} values to calculate the partial correlation between genetic and geographic distances while controlling for the confounding effects of both language and culture (table 4). Genetics and geography are significantly correlated across eastern North America when language is taken into account. However, when culture is also taken into consideration by performing this analysis separately for each culture area, no significant correlations are observed. Consequently, no isolation by distance is detectable within either culture area that cannot be explained by linguistic or cultural differences. Mantel tests also suggest that genetics and language are not correlated in eastern North America (table 4).

Sex-Specific Estimates of Migration and Effective Population Size

In eastern North America, paternally and maternally inherited loci show opposite patterns of genetic differentiation (Φ_{ST}) among populations. Y chromosomes indicate high differentiation in the Northeast ($\Phi_{ST} = 0.124$) and low differentiation in the Southeast ($\Phi_{ST} = 0.044$), whereas mtDNA indicates low differentiation in the Northeast ($\Phi_{ST} = 0.062$) and high differentiation in the Southeast ($\Phi_{ST} = 0.192$). The Φ_{ST} values for the Northeast produce a ratio of mtDNA to Y chromosome $N\nu$ of 2.142, indicating a female migration rate and/or effective population size that is more than twice that of males. In contrast, the Φ_{ST} values for the Southeast produce a $N\nu$ ratio of 0.194, which suggests that males in the Southeast have had a migration rate and/or effective population size that is approximately 5 times greater than that of females.

Table 3
AMOVAs

Data Set	Grouping	Percentage of Variation (P value)		
		Among Groups	Among Populations Within Groups	Within Populations
Y chromosome haplogroups[a]	Culture areas[c]	10.89 ($P = 0.030$)	3.96 ($P = 0.088$)	85.14 ($P < 0.001$)
	Language families[d]	5.16 ($P = 0.235$)	6.33 ($P = 0.064$)	88.52 ($P < 0.001$)
Y chromosome haplotypes[b]	Culture areas[c]	6.48 ($P = 0.019$)	6.72 ($P = 0.030$)	86.79 ($P < 0.001$)
	Language families[d]	0.22 ($P = 0.333$)	10.51 ($P = 0.033$)	89.27 ($P < 0.001$)

[a] Only haplogroups Q-M3*, Q-M242*, and C-M130 were included in this analysis.
[b] Only haplotypes in haplogroups Q-M3*, Q-M242*, and C-M130 were included.
[c] Culture areas: Northeast and Southeast.
[d] Language families: Algonquian, Siouan, Iroquoian, and Muskogean.

Table 4
Mantel Tests

Data Set	Comparison[a]	r[b]	P
Eastern North America	Genetics versus geography	0.207	0.048
	Genetics versus language	−0.023	0.586
	Genetics versus geography (language)	0.268	0.047
	Genetics versus language (geography)	−0.175	0.835
Northeastern culture area	Genetics versus geography	0.147	0.428
	Genetics versus language	−0.088	0.559
	Genetics versus geography (language)	0.122	0.401
	Genetics versus language (geography)	−0.029	0.541
Southeastern culture area	Genetics versus geography	0.502	0.013
	Genetics versus language	0.255	0.177
	Genetics versus geography (language)	0.548	0.067
	Genetics versus language (geography)	−0.353	0.721

[a] For each data set, the third and fourth comparisons refer to partial Mantel tests, in which the factor in parentheses was held constant.
[b] Correlation coefficient.

We also calculated maximum likelihood estimates of migration rates (table 5). Although the migration rates between specific pairs of populations are quite variable, male migration is generally higher in the Southeast than in the Northeast (average Southeast N_em = 1.429 vs. average Northeast N_em = 0.768), whereas female migration is generally higher in the Northeast than in the Southeast (average Northeast N_em = 11.071 vs. average Southeast N_em = 3.860).

Table 5
Sex-Specific Migration Rates Estimated Using MIGRATE[a]

Pair of Populations	Region	Male N_em	Female N_em
Chippewa–Cheyenne/Arapaho	NE	0.273	7.324
Sioux–Cheyenne/Arapaho	NE	0.587	0.000
Sioux–Chippewa	NE	1.445	25.888
Cheyenne/Arapaho–Chickasaw	NE–SE	0.175	0.000
Cheyenne/Arapaho–Choctaw	NE–SE	0.395	2.305
Cheyenne/Arapaho–Creek	NE–SE	1.064	0.000
Cheyenne/Arapaho–Cherokee	NE–SE	0.597	0.653
Chippewa–Chickasaw	NE–SE	1.616	1.129
Chippewa–Choctaw	NE–SE	0.280	0.000
Chippewa–Creek	NE–SE	1.506	0.000
Chippewa–Cherokee	NE–SE	2.307	0.565
Sioux–Chickasaw	NE–SE	1.068	0.000
Sioux–Choctaw	NE–SE	2.110	0.000
Sioux–Creek	NE–SE	2.419	0.000
Sioux–Cherokee	NE–SE	1.129	1.628
Chickasaw–Choctaw	SE	0.890	19.346
Chickasaw–Creek	SE	0.730	3.816
Chickasaw–Cherokee	SE	1.937	0.000
Choctaw–Creek	SE	1.869	0.000
Choctaw–Cherokee	SE	1.071	0.000
Creek–Cherokee	SE	2.079	0.000

[a] N_em, effective number of migrants per generation; NE, northeast; SE, southeast.

Discussion

This study presents the first detailed survey of Y chromosome variation among Native Americans from eastern North America. The populations analyzed here represent both the northeastern and southeastern culture areas, as well as the 4 major language families found in eastern North America. When combined with mtDNA data previously collected from these samples (Lorenz and Smith 1996, 1997; Smith et al. 1999; Malhi et al. 2001; Bolnick and Smith 2003; Shook 2005), these Y chromosome results shed new light on 1) the effects of language, culture, and geography on genetic variation in this region, 2) the differences between male and female patterns of gene flow and migration, 3) the history of Iroquoian populations, and 4) the impact of European contact in eastern North America.

Effects of Language, Culture, and Geography on Genetic Variation

Because Native American populations from eastern North America can be divided into 4 language families and 2 culture areas, we examined the relative importance of language and culture in shaping the observed patterns of genetic variation. Although languages clearly evolve differently than genes (Bateman et al. 1990; Renfrew 2000), similar languages may indicate common ancestry between populations, and language differences can act as barriers to gene flow (Barbujani 1991). However, based on the AMOVAs and Mantel tests, language does not seem to have had a significant impact on the structure of genetic variation in eastern North America. Four of the 6 exact tests did indicate statistically significant differences between language family groupings, but those 4 comparisons all involved populations from different culture areas. Language family groupings from the same culture area were not statistically different, suggesting that the effects of language and culture may have been confounded in these tests. Consequently, these analyses provide no clear evidence that language has influenced genetic variation in eastern North America.

Culture, on the other hand, has played an important role in shaping the observed patterns of variation. Populations from the same culture area exhibit more similar Y chromosomes than populations from different areas, and differences between the 2 culture areas account for a significant portion of the total genetic variance in eastern North America. Although culture and geography are partially confounded in this study, the Mantel tests suggest that geography has had no clear effect on the genetic structure of eastern North American populations that is independent of culture and language.

Within the southeastern culture area, populations exhibit similar Y chromosome haplogroup frequencies, share identical or closely related haplotypes, and cluster together in the MDS plot, indicating close paternal relationships among them. A close relationship among the Muskogean populations is not unexpected: archeological, linguistic, and ethnographic evidence suggest that they formed from people who had been living in the Southeast for thousands of years, and they share a variety of cultural traits, closely related languages, and a high level of intermarriage

(Swanton 1922, 1946; Haas 1941; Hudson 1976, 1997; Knight 1994; Galloway 1995).

Within the northeastern culture area, populations exhibit fewer similarities: some Y chromosome haplotypes are closely related, but the populations do not all cluster together in the MDS plot. These results suggest that Native American populations from the Northeast do not share as close paternal relationships.

Male–Female Differences in Gene Flow and Migration

With both Y chromosome and mtDNA data from these populations, it is possible to directly evaluate sex-specific patterns of gene flow and migration. These data show that male and female demographic histories have clearly differed in eastern North America but in opposite ways in the 2 culture areas. In the southeastern culture area, males have experienced higher rates of gene flow/migration than females, whereas females have experienced greater gene flow/migration than males in the northeastern culture area.

Significantly, these patterns of genetic variation correlate precisely with past postmarital residence patterns in eastern North America. Populations from the southeastern culture area exhibited matrilocality (Hudson 1976), in which males moved from their own household or community to that of their wife after marriage. Matrilocal systems therefore facilitated male movement and gene flow each generation while fostering mtDNA differentiation. On the other hand, the populations from the northeastern culture area that we sampled all exhibited patrilocality (Mason 1981), which facilitated female movement and gene flow and augmented Y chromosome differentiation. Thus, postmarital residence clearly had a significant impact on the genetic structure of eastern North American populations. Because the postmarital residence practices of many populations changed during the 19th and 20th centuries (Eggan 1937; Perdue 1989), these results suggest that the effects of past postmarital residence systems may be detectable for at least 5–6 generations after patterns have shifted (assuming a generation length of 30 years). It is important to note, though, that the relationship between postmarital residence and genetic variation will vary across human populations: a clear correlation may be observed in some regions, such as eastern North America, but the pattern of genetic variation in other regions may instead reflect other factors, such as local cultural practices like endogamy (Kumar et al. 2006).

Iroquoian Population History

Although Cherokee mtDNAs differ from those of other southeastern populations (Bolnick and Smith 2003), the Y chromosome data presented here show broad similarities between the Cherokee and the Muskogean-speaking populations of the Southeast. Both mtDNA and Y chromosome data show differences between the Cherokee and northeastern populations. Because Cherokee Y chromosomes are closely related to those of populations that have lived in southeastern North America for several thousand years, the Cherokee must have lived in the Southeast long enough to develop such extensive similarities through gene flow. These findings suggest that Iroquoian populations may have first lived in the Southeast, with a later Northern Iroquoian migration to the Northeast. mtDNA differences between the Cherokee and other southeastern populations are compatible with this hypothesis because the Cherokee's matrilocal social structure would have restricted female gene flow with other populations and increased mtDNA differentiation. However, to confirm this hypothesis, Northern Iroquoian Y chromosomes would have to be studied to detect parallel similarities with the southeastern Y chromosomes as well.

Impact of European Contact

European contact affected genetic variation in eastern North America in at least 2 different ways: by causing a population decline that resulted in genetic drift and by introducing non–Native American lineages through admixture. Earlier studies suggested that the population decline following European contact did not significantly impact genetic variation in North America (Stone and Stoneking 1998; O'Rourke et al. 2000; Kaestle and Smith 2001), but 2 recent studies concluded otherwise. Bolnick and Smith (2003) found evidence that a genetic bottleneck associated with this decline altered patterns of mtDNA variation in southeastern North America, and Wang et al. (2004) identified further evidence of a relatively recent bottleneck in the southeastern Choctaw population by analyzing 175 genome-wide markers. This bottleneck does not seem to have drastically affected Y chromosome variation in southeastern North America, but the Y chromosome data presented here are nevertheless compatible with such an event for 2 reasons. First, bottlenecks are not expected to affect all loci equally (Hoelzel 1999). Second, because the southeastern populations are matrilocal, effective population sizes for their Y chromosomes are much larger than those for their mtDNA, making genetic drift less likely to affect Y chromosome variation than mtDNA variation in this region.

European contact also influenced genetic variation in indigenous eastern North American populations by introducing non–Native American lineages. In most cases, there is widespread agreement about whether a particular haplogroup represents an ancient Native American lineage or post-1492 admixture, but the status of haplogroup R-M173 has recently been subject to some debate. Some authors have argued that this haplogroup represents a founding Native American lineage (Lell et al. 2002; Bortolini et al. 2003), whereas others suggest that it instead reflects recent European admixture (Tarazona-Santos and Santos 2002; Bosch et al. 2003; Zegura et al. 2004). In eastern North America, the pattern of haplotype variation within this haplogroup supports the latter hypothesis: R-M173 haplotypes do not cluster by population or culture area, as haplotypes in the other founding haplogroups do, and most match or are closely related to R-M173 haplotypes that are common in Europe but rare in Asia. This pattern is opposite than expected if the Native American R-M173 haplotypes were descended from Asian haplotypes and suggests that recent European admixture is responsible for the presence of haplogroup R-M173 in eastern North America. This conclusion is consistent with

evidence that European contact also introduced the DE-YAP and "Other" lineages into eastern North America.

If this interpretation is correct, eastern North America exhibits a high level of non–Native American admixture (47.7%). This estimate reflects the extremely high level of admixture in the Chippewa populations, which is consistent with ethnographic evidence indicating frequent interactions between the Chippewa and European traders (Rhodes 1982; Camp 1990). Because of the high representation of Chippewa samples in our Y chromosome data set, this estimate of non–Native American admixture in eastern North America may be somewhat exaggerated. Our estimate is substantially greater than the estimate of 17 ± 2% paternal non–Native American admixture of Zegura et al. (2004) across North, Central, and South America and is particularly surprising because almost all of the sampled individuals reported being full blooded for a single Native American population. Of the 134 individuals with European, African, or "Other" Y chromosomes, only 3 (all R-M173) reported having less than 100% Native American ancestry.

Interestingly, the Y chromosome estimates of non–Native American admixture differ sharply from those based on mtDNA. mtDNA studies of the same samples revealed that few exhibit non–Native American mtDNA (Lorenz and Smith 1996; Smith et al. 1999; Malhi et al. 2001; Bolnick and Smith 2003; Shook 2005). Because the vast majority of non–Native American Y chromosomes appear to be European in origin, this difference suggests a high level of sex-biased admixture in eastern North America, involving European males and Native American females. These results are consistent with ethnographic evidence and with previous comparisons of Y chromosome and mtDNA variation in South America, which identified a similarly asymmetric pattern of mating (Carvajal-Carmona et al. 2000; Mesa et al. 2000; Carvalho-Silva et al. 2001).

Because many Native Americans in eastern North America did not acknowledge their paternal European ancestry, such admixture may predate current genealogical records. Male-mediated European gene flow may have therefore had a greater impact on eastern North American populations than was previously thought. Alternatively, some of the Native Americans who provided samples for this study may have been aware of such ancestry but interpreted phrases like "full bloodedness" and "percentage ancestry" as reflecting their level of social identification with a particular tribe rather than their genetic history (K TallBear, personal communication). Consequently, differences between the Native American and scientific concepts of kinship and ancestry should be explored further before reaching any firm conclusions about the discrepancy between reported and actual genetic ancestry in this study.

Finally, present-day genetic variation in eastern North America may also reflect other events from the centuries following European contact, such as forced migrations, the formation of new Native American communities, and the establishment of the reservation system. Some populations became more isolated as a result of these events: the rigidity of the reservation system, for example, restricted gene flow among the various Chippewa reservations (Shook 2005). In other cases, historical events facilitated population interactions: new coalescent societies formed in both the northern Plains and the Southeast following periods of population decline (Taylor 1977; Hudson 1997; Bolnick and Smith 2003), and the geographic proximity of many reservations in Oklahoma may have increased gene flow among the southeastern populations. Historical events may have therefore contributed to the observed patterns of Y chromosome variation (e.g., closer paternal relationships among populations from the Southeast than among those from the Northeast). However, such events cannot explain the opposite pattern observed in the mtDNA data (Bolnick and Smith 2003), so they must have had less effect on eastern North American genetic variation than past patterns of postmarital residence.

Conclusion

This study demonstrates the importance of using multiple loci when reconstructing population history. In eastern North America, male and female demographic histories differ significantly as a result of postmarital residence patterns and European influences. Geography may have also played a role in shaping population interactions. Although the correlation between geography and culture area in this study makes it difficult to completely separate the two, this research suggests that sociocultural factors have had the greatest impact on the genetic structure of indigenous eastern North American populations.

Supplementary Material

1) The primers, PCR conditions, and restriction enzymes used to analyze the Y chromosome binary markers, 2) the geographic and linguistic distances used in the Mantel tests, and 3) the binary marker/microsatellite haplotypes observed in each population are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are indebted to the Native Americans who donated samples for genetic analysis. We thank Beth Shook, Brad Shaffer, and Kim TallBear for valuable discussions and comments on this manuscript and 3 anonymous reviewers for helpful suggestions. This research was supported by an NSF Dissertation Improvement Grant, an National Science Foundation Graduate Research Fellowship, and a Dissertation-Year Fellowship from the University of California, Davis (all to D.A.B.).

Literature Cited

Avise JC. 2000. Phylogeography: the history and formation of species. Cambridge: Harvard University Press.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel NE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu Rev Ecol Syst 18:489–522.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. Mol Ecol 13:729–44.

Bandelt H-J, Forster P, Roehl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37–48.

Barbujani G. 1991. What do languages tell us about human microevolution? Trends Evol Ecol 6:151–5.

Bateman R, Goddard I, O'Grady R, Funk VA, Mooi R, Kress WJ, Cannell P. 1990. Speaking of forked tongues: the feasibility of reconciling human phylogeny and the history of language. Curr Anthropol 31:1–24.

Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. Proc Natl Acad Sci USA 98:4563–8.

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, Lopez-Camelo JS. 1998. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. Am J Hum Genet 63:1862–71.

Bolnick DA. 2005. The genetic prehistory of eastern North America: evidence from ancient and modern DNA [dissertation]. Davis: University of California.

Bolnick DA, Smith DG. 2003. Unexpected patterns of mitochondrial DNA variation among Native Americans from the Southeastern United States. Am J Phys Anthropol 122:336–54.

Bortolini MC, Salzano FM, Bau CHD, et al. (11 co-authors). 2002. Y-chromosome biallelic polymorphisms and Native American population structure. Ann Hum Genet 66:255–9.

Bortolini MC, Salzano FM, Thomas MG, et al. (21 co-authors). 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. Am J Hum Genet 73: 524–39.

Bosch E, Calafell F, Rosser ZH, Norby S, Lynnerup N, Hurles ME, Jobling MA. 2003. High level of male-biased Scandinavian admixture in Greenland Inuit shown by Y-chromosomal analysis. Hum Genet 112:353–63.

Camp GS. 1990. Working out their own salvation: the allotment of land in severalty and the Turtle Mountain Chippewa Band, 1870-1920. Am Indian Cult Res J 14:19–38.

Campbell L. 1997. American Indian languages: the historical linguistics of Native America. New York: Oxford University Press.

Carvajal-Carmona LG, Soto ID, Pineda N, et al. (11 co-authors). 2000. Strong Amerind/White sex bias and a possible Sephardic contribution among founders of a population in Northwest Colombia. Am J Hum Genet 67:1287–95.

Carvalho-Silva DR, Santos FR, Rocha J, Pena SDJ. 2001. The phylogeography of Brazilian Y-chromosome lineages. Am J Hum Genet 68:281–6.

Castellano S, Balletto E. 2002. Is the partial Mantel test inadequate? Evolution 56:1871–3.

Cavalli-Sforza LL, Bodmer W. 1971. The genetics of human populations. San Francisco: Freeman.

Chafe WL. 1976. The Caddoan, Iroquoian, and Siouan languages. The Hague: Mouton.

Crawford GW, Smith DG. 1996. Migration in prehistory: Princess Point and the Northern Iroquoian case. Am Antiq 61:782–90.

Denny JP. 1991. The Algonquian migration from Plateau to Midwest: linguistics and archaeology. In: Cowan W, editor. Papers of the 22nd Algonquian conference. Ottawa: Carleton University. p 103–24.

Destro-Bisol G, Donati F, Coia V, Boschi I, Verginelli F, Caglià A, Tofanelli S, Spedini G, Capelli C. 2004. Variation of female and male lineages in Sub-Saharan populations: the importance of sociocultural factors. Mol Biol Evol 21:1673–82.

Dupanloup I, Pereira L, Bertorelle G, Calafell F, Prata MJ, Amorim A, Barbujani G. 2003. A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. J Mol Evol 57:85–97.

Eggan F. 1937. Historical changes in the Choctaw kinship system. Am Anthropol 39:34–52.

Fagundes NJR, Bonatto SL, Callegari-Jacques SM, Salzano FM. 2002. Genetic, geographic, and linguistic variation among South American Indians: possible sex influence. Am J Phys Anthropol 117:68–78.

Fiedel SJ. 1991. Correlating archaeology and linguistics: the Algonquian case. Man in the Northeast 41:9–32.

Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, Pettener D. 2003. Mitochondrial DNA diversity in South America and the genetic history of Andean Highlanders. Mol Biol Evol 20:1682–91.

Galloway PK. 1995. Choctaw genesis, 1500-1700. Lincoln: University of Nebraska Press.

Goddard I. 1994. The west-to-east cline in Algonkian dialectology. In: Cowan W, editor. Actes du 25 congres des Algonquinistes. Ottawa: Carleton University. p 187–211.

Haas MR. 1941. The classification of the Muskogean languages. In: Spier L, editor. Essays in memory of Edward Sapir. Menasha: Banta Publishing. p 41–56.

Hamilton G, Stoneking M, Excoffier L. 2005. Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. Proc Natl Acad Sci USA 102:7476–80.

Hare MP. 2001. Prospects for nuclear gene phylogeography. Trends Ecol Evol 16:700–6.

Hoelzel AR. 1999. Impact of population bottlenecks on genetic variation and the importance of life-history: a case study of the northern elephant seal. Biol J Linn Soc 68:23–39.

Hudson CM. 1976. The Southeastern Indians. Knoxville: University of Tennessee Press.

Hudson CM. 1997. Knights of Spain, warriors of the sun: Hernando De Soto and the South's ancient chiefdoms. Athens: University of Georgia Press.

Hurles ME, Jobling MA. 2001. Haploid chromosomes in molecularecology: lessons from the human Y. Mol Ecol 10: 1599–613.

Kaestle FA, Smith DG. 2001. Ancient mitochondrial DNA evidence for prehistoric population movement: the Numic expansion. Am J Phys Anthropol 115:1–12.

Karafet TM, Zegura SL, Posukh O, et al. (14 co-authors). 1999. Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. Am J Hum Genet 64:817–31.

Knight VJ. 1994. The formation of the creeks. In: Hudson C, Tesser CC, editors. The forgotten centuries: Indians and Europeans in the American South, 1521-1704. Athens: University of Georgia Press. p 373–92.

Kumar V, Langstieh BT, Madhavi KV, Naidu VM, Singh HP, Biswas S, Thangaraj K, Singh L, Reddy BM. 2006. Global patterns in human mitochondrial DNA and Y-chromosome variation caused by spatial instability of the local cultural processes. PloS Genet 2:420–4.

Lell JT, Sukernik RI, Starikovskaya YB, Su B, Jin L, Schurr TG, Underhill PA, Wallace DC. 2002. The dual origin and Siberian affinities of Native American Y chromosomes. Am J Hum Genet 70:192–206.

Lorenz JG, Smith DG. 1996. Distribution of four founding mtDNA haplogroups among Native North Americans. Am J Phys Anthropol 101:307–23.

Lorenz JG, Smith DG. 1997. Distribution of sequence variation in the mtDNA control region of Native North Americans. Hum Biol 69:749–76.

Lounsbury FG. 1978. Iroquoian languages. Archaeol E Nor Am 15: 334–43.

MacNeish RS. 1952. Iroquois pottery types: a technique for the study of Iroquois prehistory. Ottawa: National Museum of Canada. National Museum of Canada Bulletin 124, Anthropological Series No. 31.
Malhi RS, Schultz BA, Smith DG. 2001. Distribution of mitochondrial DNA lineages among Native American tribes of Northeastern North America. Hum Biol 73:17–55.
Malyarchuk B, Derenko M, Grzybowski T, Lunkina A, Czarny J, Rychkov S, Morozova I, Denisova G, Miscicka-Sliwka D. 2004. Differentiation of mitochondrial DNA and Y chromosomes in Russian populations. Hum Biol 76:877–900.
Mason RJ. 1981. Great Lakes archaeology. New York: Academic Press.
Mesa NR, Mondragon MC, Soto ID, et al. (13 co-authors). 2000. Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: re-and post-Columbian patterns of gene flow in South America. Am J Hum Genet 67:1277–86.
Mills LA. 2003. Mitochondrial DNA analysis of the Ohio Hopewell of the Hopewell Mound Group [dissertation]. Columbus: The Ohio State University.
Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718–26.
Nasidze I, Ling EYS, Quinque D, et al. (17 co-authors). 2004. Mitochondrial DNA and Y-chromosome variation in the Caucasus. Ann Hum Genet 68:205–21.
Oota H, Kitano T, Jin F, Yuasa I, Wang L, Ueda S, Saitou N, Stoneking M. 2002. Extreme mtDNA homogeneity in continental Asian populations. Am J Phys Anthropol 118:146–53.
Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29:20–1.
O'Rourke DH, Hayes MG, Carlyle SW. 2000. Spatial and temporal stability of mtDNA haplogroup frequencies in native North America. Hum Biol 72:15–34.
Parker AC. 1916. The origin of the Iroquois as suggested by their archaeology. Am Anthropol 18:479–507.
Perdue T. 1989. Cherokee women and the trail of tears. J Womens Hist 1:14–30.
Perez-Lezaun A, Calafell F, Comas D, et al. (12 co-authors). 1999. Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. Am J Hum Genet 65:208–19.
Qamar R, Ayub Q, Mohyuddin A, Helgason A, Mazhar K, Mansoor A, Zerjal T, Tyler-Smith C, Qasim Mehdi S. 2002. Y-chromosomal DNA variation in Pakistan. Am J Hum Genet 70:1107–24.
Raufaste N, Rousset F. 2001. Are partial Mantel tests adequate? Evolution 55:1703–5.
Renfrew C. 2000. Genetics and languages in the Americas and beyond: variations at the micro and the macro level. In: Renfrew C, editor. America past, America, present: genes and languages in the Americas and beyond. Cambridge: McDonald Institute for Archaeological Research. p 3–15.
Rhodes R. 1982. Algonquian trade languages. In: Cowan W, editor. Proceedings of the 13th Algonquian conference. Ottawa: Carleton University. p 1–10.
Rice WR. 1989. Analyzing tables of statistical tests. Evolution 43:223–5.
Roewer L, Krawczak M, Willuweit S, et al. (38 co-authors). 2001. Online reference database of Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118:103–11.
Rousset F. 2002. Partial Mantel tests: reply to Castellano and Balletto. Evolution 56:1874–5.
Ruiz-Linares A, Ortiz-Barrientos D, Figueroa M, et al. (12 co-authors). 1999. Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. Proc Natl Acad Sci USA 96:6312–7.
Santos FR, Pandya A, Tyler-Smith C, Pena SDJ, Schanfield M, Leonard WR, Osipova L, Crawford MH, Mitchell RJ. 1999. The Central Siberian origin for Native American Y chromosomes. Am J Hum Genet 64:619–28.
Schneider S, Roessli D, Excoffier L. 2000. Arlequin ver. 2.000: a software for population genetics data analysis. Geneva: Genetics and Biometry Laboratory, University of Geneva.
Seielstad MT, Minch E, Cavalli-Sforza LL. 1998. Genetic evidence for a higher female migration rate in humans. Nat Genet 20:278–80.
Shook BAS. 2005. Ancient DNA and the biological history and prehistory of northeastern North America [dissertation]. Davis: University of California.
Siebert FT. 1967. The original homeland of the Proto-Algonquian people. Contributions to anthropology: linguistics 1 (Algonquian). Anthropological Series 78, National Museum of Canada Bulletin. Ottawa, Canada: National Museum of Canada 214: 13–47.
Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA. 1999. Distribution of mtDNA haplogroup X among Native North Americans. Am J Phys Anthropol 110:271–84.
Snow DR. 1980. The archaeology of New England. New York: Academic Press.
Snow DR. 1995. Migration in prehistory: the Northern Iroquoian case. Am Antiq 60:59–79.
Snow DR. 1996. More on migration in prehistory: accommodating new evidence in the Northern Iroquoian case. Am Antiq 61:791–6.
Stone AC, Stoneking M. 1998. mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. Am J Hum Genet 62:1153–70.
Sullivan KM, Mannucci A, Kimpton CP, Gill P. 1993. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 15:636–41.
Swanton J. 1922. Early history of the Creek Indians and their neighbors. Washington, DC: Smithsonian Institution. Bureau of American Ethnology Bulletin 73.
Swanton J. 1946. The Indians of the southeastern United States. Washington, DC: Smithsonian Institution. Bureau of American Ethnology Bulletin 137.
Tarazona-Santos E, Santos FR. 2002. The peopling of the Americas: a second major migration? Am J Hum Genet 70:1377–80.
Taylor JF. 1977. Sociocultural effects of epidemics on the Northern Plains: 1734-1850. West Can J Anthropol 7:55–81.
Underhill PA, Lin AA, Zemans R, Oefner PJ, Cavalli-Sforza LL. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93:196–200.
Wang N, Zhou X, Tan FK, Foster MW, Arnett FC, Chakraborty R. 2004. Genetic signatures of pre-expansion bottleneck in the Choctaw population of Oklahoma. Am J Phys Anthropol 124:373–9.
Wen B, Xie X, Gao S, et al. (13 co-authors). 2004. Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in Southern Tibeto-Burmans. Am J Hum Genet 74:856–65.
Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF. 2004. Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. Nat Genet 36:1122–5.
Williams SR, Chagnon NA, Spielman RS. 2002. Nuclear and mitochondrial genetic variation in the Yanomamö: a test case for ancient DNA studies of prehistoric populations. Am J Phys Anthropol 117:246–59.

Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB. 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. Proc Natl Acad Sci USA 98:5078–83.

Wood ET, Stover DA, Ehret C, et al. (11 co-authors). 2005. Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. Eur J Hum Genet 13:867–76.

Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–48.

Zar JH. 1999. Biostatistical analysis. 6th ed. Upper Saddle River: Prentice Hall.

Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. 2004. High resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Mol Biol Evol 21:164–75.

Anne Stone, Associate Editor

Accepted August 14, 2006