Mitochondrial DNA Analysis of Mongolian Populations and Implications for the Origin of New World Founders

Connie J. Kolman,* Nyamkhishig Sambuughin*,†,[1] and Eldredge Bermingham*,[1]

**Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Panamá and †National Centre for Anthropology, Mongolian Medical University, Ulaanbaatar, Mongolia*

Manuscript received June 23, 1995
Accepted for publication December 22, 1995

ABSTRACT

High levels of mitochondrial DNA (mtDNA) diversity were determined for Mongolian populations, represented by the Mongol-speaking Khalkha and Dariganga. Although 103 samples were collected across Mongolia, low levels of genetic substructuring were detected, reflecting the nomadic lifestyle and relatively recent ethnic differentiation of Mongolian populations. mtDNA control region I sequence and seven additional mtDNA polymorphisms were assayed to allow extensive comparison with previous human population studies. Based on a comparative analysis, we propose that indigenous populations in east Central Asia represent the closest genetic link between Old and New World populations. Utilizing restriction/deletion polymorphisms, Mongolian populations were found to carry all four New World founding haplogroups as defined by WALLACE and coworkers. The ubiquitous presence of the four New World haplogroups in the Americas but narrow distribution across Asia weakens support for GREENBERG and coworkers' theory of New World colonization via three independent migrations. The statistical and geographic scarcity of New World haplogroups in Asia makes it improbable that the same four haplotypes would be drawn from one geographic region three independent times. Instead, it is likely that founder effects manifest throughout Asia and the Americas are responsible for differences in mtDNA haplotype frequencies observed in these regions.

ALTHOUGH Mongolians have long been identified as possible founders of the New World, few data are available to investigate this theory. TURNER (1984) proposed that migration(s) to the New World originated in the area of northern China and southeastern Siberia based on the spread of Sinodontic dental characters throughout northeast Asia and the New World and a notched unifacial spear point (similar to Clovis spear points excavated in the southwestern United States) discovered in Novopetrovka (DEREVIANKO 1969). Phylogenetic analysis of protein polymorphisms in populations distributed throughout the world (29 loci from 26 populations) led NEI and ROYCHOUDHURY (1993) to suggest that migrations out of east Central Asia might have colonized the New World, Japan and the Polynesian and Micronesian Islands. Mitochondrial DNA (mtDNA) studies have identified a COII/tRNA^{Lys} 9-bp intergenic deletion that is widespread in New World populations and present in Mongol, Altai and Tibetan populations, but absent in Siberian groups (SAMBUUGHIN *et al.* 1991; SHIELDS *et al.* 1992, 1993; TORRONI *et al.* 1993b, 1994b). Based on the 9-bp deletion and three restriction fragment length polymorphisms (RFLPs), recent molecular genetic studies have identified four mtDNA haplotypes in the New World that are proposed to represent founding lineages (TORRONI *et al.* 1992) (Table 1). Although postcolonization events such as genetic drift and founder effects have undoubtedly played a role in determining current haplotype distributions in Asia and the New World, a comparative analysis of key mtDNA polymorphisms such as those that define the four founding haplogroups can be used as a molecular genetic footprint to search for possible New World ancestors in Asian populations.

Notwithstanding their geographic proximity to the New World, Siberian populations appear to be eliminated as founders due to their lack of the 9-bp deletion, thus moving the focus to Mongolia as the next nearest neighbor to the Americas. Archaeological and anthropological data reveal that modern Mongolians are the descendants of ancient nomads whose remains have been found throughout the territories of Mongolia dating to at least the end of the Neolithic period (TUMEN 1992). The continuity of Mongols suggests there may be a direct genetic link between the contemporary populations available for study and the Mongolian populations that existed at the time of the New World migration(s). Despite the fact that the Mongolian populace inhabits a vast territory [~2 million people (1989 population census) dispersed over 1.5 million km^2], very low levels of genetic differentiation have been detected between populations from different geographical regions (BATSUURI 1986; SAMBUUGHIN *et al.* 1992). This most

Corresponding author: Connie J. Kolman, Smithsonian Institution, Conservation Analytical Laboratory/Museum Support Center, 4210 Silver Hill Rd., Suitland, MD 20746. E-mail: cjk@cal.si.edu

[1] *Mailing address:* Smithsonian Tropical Research Institute, Unit 0948, APO, AA 34002-0948 USA.

likely reflects the continual genetic mixing produced by the nomadic lifestyle of the Mongols and the relatively recent ethnic differentiation of some contemporary Mongolian tribes [between 300 and 1000 years before present (YBP) (GONGOR 1970)].

To investigate mtDNA diversity within Mongolia and identify possible New World founding populations, we studied two Mongolian indigenous populations: the Khalkha and the Dariganga, representing 80 and 1.5% of the Mongolian populace, respectively. The Khalkha are the largest cultural group in Mongolia and number ~1.6 million individuals (1989 population census). They are distributed throughout the territory of Mongolia and, like all Mongol-speakers, speak a language assigned to the Mongolian branch of the Altaic linguistic family (BADAMKHATAN 1987). The Dariganga are the most recently diverged of the Mongol-speaking groups, formed more than 300 years ago (GONGOR 1970; BADAMKHATAN 1987). They number ~30,000 (1989 population census) and currently inhabit the southeast region of Mongolia.

We present the mtDNA control region I sequence and analysis of seven additional polymorphic sites for 103 Mongolians. By utilizing both DNA sequence and restriction/deletion site analysis, we place our data in a comparative framework constructed from the large number of mtDNA control region sequence studies (HORAI and HAYASAKA 1990; WARD *et al.* 1991; GINTHER *et al.* 1993; HORAI *et al.* 1993; SHIELDS *et al.* 1993; WARD *et al.* 1993; BATISTA *et al.* 1995; KOLMAN *et al.* 1995; MERRIWETHER *et al.* 1995) and restriction/deletion studies (BALLINGER *et al.* 1992; TORRONI *et al.* 1993a,b, 1994a,b) of east Asian and New World populations. Based on these comparisons and on data presented herein, we propose that Mongolians and surrounding populations in east Central Asia represent the closest genetic link between Old and New World populations.

MATERIALS AND METHODS

Population samples: Blood samples were collected from individuals representing 15 geographic regions distributed throughout the country of Mongolia (Figure 1). One hundred and three samples were collected from the Khalkha, Durvet, Dariganga, Buriat and Darkhad Mongol-speaking groups comprising 80, 2.6, 1.5, 1.2 and 0.1% of the Mongolian population, respectively (BADAMKHATAN 1987). Collection of blood samples was designed to investigate the Mongol-speaking populace as a whole and also to focus on the Dariganga and Khalkha cultural groups. Concentrated sampling of regions 8 and 17 resulted in 36 Dariganga and 41 Khalkha samples, respectively. One to four individuals were sampled from each of 13 additional regions to represent the Khalkha (all regions except 1 and 6, $n = 23$), Durvet (region 1, $n = 1$), Buriat (region 6, $n = 1$) and Darkhad (region 3, $n = 1$) at their actual frequencies.

Blood samples were collected from students while they attended university in the Mongolia capital, Ulaanbaatar. Extensive biographical information was taken for each individual including designation of cultural group and names and birthplaces of parents and grandparents. All individuals were chosen for study based on a family history suggesting no obvious, *i.e.*, recent, admixture with other cultural groups or with individuals from other geographic regions.

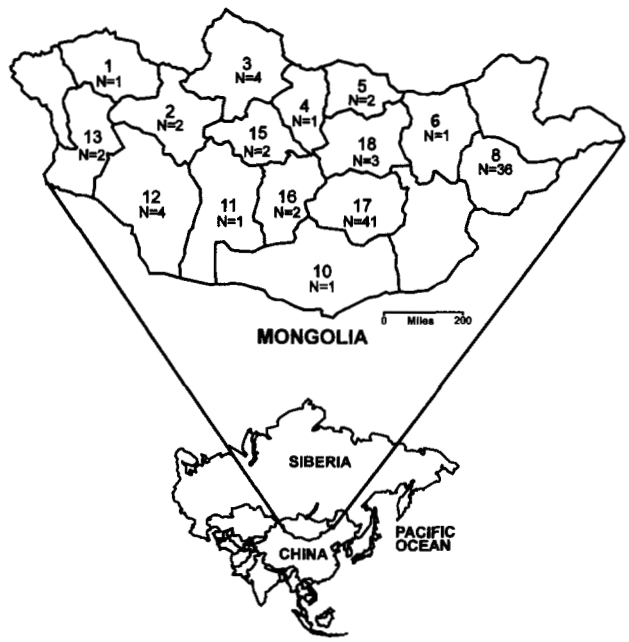


FIGURE 1.—Location of collection sites in Mongolia. The 18 political regions (aimak) of Mongolia are outlined in black, and the 15 regions that were sampled are indicated in addition to the sample size from each collected region. Khalkha populations are distributed throughout Mongolia, and the Dariganga are concentrated in region 8. Region numbers were incorporated into sample labels to identify the region from which each sample originated with exceptions as noted in Figure 2.

Isolation of DNA: Ten milliliters of blood were collected into Vacutainer tubes containing EDTA to prevent coagulation (Becton Dickinson). Leucocytes and erythrocytes were separated by allowing the collection tubes to stand, undisturbed, for several hours. DNA was isolated from leucocytes using an overnight proteinase K digestion followed by phenol-chloroform extraction and alcohol precipitation of the DNA as previously described (SAMBUUGHIN *et al.* 1992).

DNA sequence analysis: Primers L15997:H16401 (WARD *et al.* 1991) were used to amplify region I of the mitochondrial control region (nomenclature from VIGILANT *et al.* 1989). [The numbers in the primer designations identify the 3′ ends of the primers according to the reference sequence (ANDERSON *et al.* 1991), while "L" and "H" designate the light and heavy strands of the mtDNA molecule, respectively.] Nucleotides 16020–16400 were sequenced for each individual. A single amplification reaction was carried out for 30 cycles using the following reaction conditions: 67 mM Tris-HCl pH 8.8, 2 mM $MgCl_2$, 250 μM each of dATP, dCTP, dTTP and dGTP, 250 μg/ml bovine serum albumin (BSA), 0.4 μM of each primer and 0.625 units *Thermus aquaticus* DNA polymerase (USB). The thermal cycler conditions were 94° for 30 sec, 55° for 45 sec and 72° for 30 sec. Lambda exonuclease digestion of the PCR product produced single-stranded DNA using the following reaction conditions: 77.5 mM glycine, 2.78 mM KOH, 0.58 mM $MgCl_2$, 0.58 mg/ml BSA and 0.75 units lambda exonuclease diluted 1:10 in 1 mM Tris pH 7.5, and 0.01 mM EDTA. The lambda exonuclease reaction was performed at 37° for 30 min. After incubation at 96° for 5 min to denature the enzyme, the samples were purified through Centricon-30 microconcentrators (Amicon Inc.).

Seven microliters of the Centricon-30 retentate were sequenced using Sequenase Version 2.0 (US Biochemical). Primers L15997 (WARD *et al.* 1991), L16106 (BATISTA *et al.* 1995) and L16191 (BATISTA *et al.* 1995) were used to sequence the light strand, and primers H16401 (WARD *et al.* 1991) and H16226 (BATISTA *et al.* 1995) were used to sequence the heavy strand of each sample. Reaction products were separated by electrophoresis through 6% Long Ranger polyacrylamide gels containing 7 M urea.

A T-to-C mutation at position 16189 created a homopolymer of 10 cytosines that caused the polymerase to stutter, resulting in poor sequence beyond the homopolymer. Thus, using the PCR primers, we could only confidently read sequence in each direction until position 16189 in the 24 individuals carrying the 10-cytosine homopolymer. An internal primer (L16191) downstream of the homopolymer allowed us to read sequence of the light strand from position 16230 to 16400. In the remaining 79 individuals, the sequences analyzed were completely verified through full overlap of light and heavy strands. Length differences in the adenosine homopolymer at positions 16180–16183 and in the cytosine homopolymer at positions 16184–16188 were also reported in association with the 16189:T-to-C transition. However, because of their nonindependent association with the transition (HORAI *et al.* 1993), the length difference mutations were excluded from all analyses and are not reported here (complete DNA sequences, including the length difference mutations, were deposited in the GenBank database).

The sequences reported in this paper have been deposited in the GenBank data base (accession numbers U33336–U33418).

RFLP analysis: Limited RFLP analysis was performed on the sequenced individuals to permit comparison of our data to the four mtDNA haplogroups defined by TORRONI *et al.* (1992, 1993a). Six sets of primers were used in balanced PCR reactions (run for 29 cycles) to screen for seven polymorphic sites located outside of mtDNA control region I.

1. *Hae*III: bp 663, primers L577:H743 (KOLMAN *et al.* 1995), annealing temperature = 62°.
2. *Alu*I: bp 5176, primers L5099:H5333 (BATISTA *et al.* 1995), annealing temperature = 62°.
3. COII/$tRNA^{Lys}$ intergenic deletion [9-bp deletion relative to the reference sequence (ANDERSON *et al.* 1981)]:bps 8272–8289, primers L8215:H8297 (WARD *et al.* 1991), annealing temperature = 55°.
4. *Dde*I: bp 10394 and *Alu*I:bp 10397, primers L10284 (5′-CCATGAGCCCTACAAACAACTAACC-3′):H10484 (5′-GTAA-ATGAGGGGCATTTGGTAAATAT-3′), annealing temperature = 61°.
5. *Alu*I: bp 13262, primers L13232:H13393 (WARD *et al.* 1991), annealing temperature = 49°.
6. *Hae*III: bp 16517, primers L16453 (5′-CCGGGCCCATAA-CACTTGGG-3′):H408 (5′-CTGTTAAAAGTGCATACCGCCA-3′), annealing temperature = 57°.

Thermal cycler and PCR conditions were as described in KOLMAN *et al.* (1995) except that L577:H743 and L5099:H5333 used the following temperature profile: 94° for 30 sec, 62° for 15 sec and 72° for 15 sec. PCR products were electrophoresed through 15% polyacrylamide gels or 4% MetaPhor (FMC BioProducts) agarose gels.

Diversity, phylogenetic and statistical analyses: Haplotype diversity, h (NEI and ROYCHOUDHURY 1974) and two measures of nucleotide diversity, π (NEI and TAJIMA 1981) and $E(\nu) = K/[0.577 + \log_e (n - 1)]$ (K = number of polymorphic sites and n = sample size) (WATTERSON 1975), were calculated using DNA sequence data.

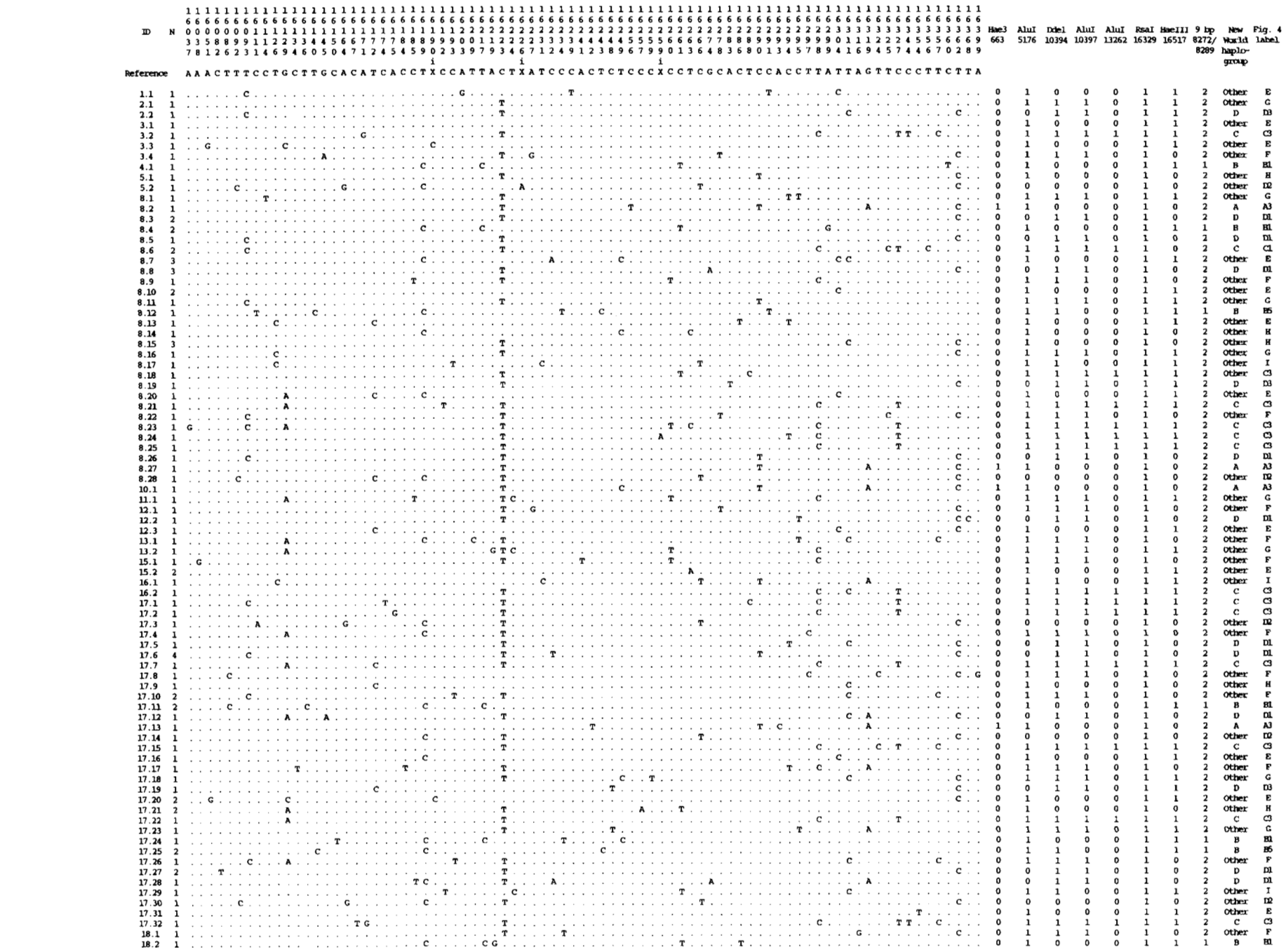
Phylogenetic trees of Mongolian DNA sequence haplotypes were constructed using the Phylogenetic Analysis Using Parsimony package (PAUP) (SWOFFORD 1989). One African Pygmy (sequence #2) (VIGILANT *et al.* 1991) and one Papua New Guinea (sequence #81) (VIGILANT *et al.* 1991) mtDNA sequence were used as outgroups based on MADDISON *et al.*'s (1992) identification of these sequences as two of the most divergent in the human dataset compiled by VIGILANT *et al.* (1991). To demonstrate correspondence of the Mongolian haplotypes defined as "New World" with actual New World haplotypes, four haplotypes from each of the four New World haplogroups were also tested in parsimony analyses with the Mongolian sequences. The New World haplotypes tested are listed in the APPENDIX.

Heuristic searches utilizing random addition of the input sequences, the tree-bisection-reconnection branch-swapping algorithm (subtree pruning-regrafting was tested in one search) and saving 50 minimal trees per replication were performed. Approximately 20–30 replications per search were performed. For weighted searches, a transition:transversion ratio of 18:1 was utilized based on an empirical calculation of simple transitions and transversions in the Mongolian dataset. Five unweighted and four weighted searches were performed with the Mongolian sequence dataset and five unweighted and three weighted searches were performed with Mongolian + New World sequence dataset. A bootstrap search of 100 replications was also performed on the Mongolian data set. Clustering of group A, B, C and D haplotypes was performed by manually switching terminal branches using MacClade 3.0 (MADDISON and MADDISON 1992).

Minimum spanning trees (MSTs) were constructed by hand following the recommendations of EXCOFFIER *et al.* (1992) and used either DNA sequence or restriction/deletion data. MST-based analyses of molecular variance (AMOVA) were performed using a program provided by L. EXCOFFIER. Mongolian populations were analyzed on the basis of cultural identification (*e.g.*, Dariganga *vs.* Khalkha), geographic location (*e.g.*, east *vs.* central) and ecological classification (*e.g.*, steppe *vs.* high mountain *vs.* semi-desert *vs.* forest/mountain). All AMOVAs utilized DNA sequence data. The significance of variance terms was tested by permuting the original data 1000 times.

RESULTS

Mitochondrial diversity in Mongolian populations: One hundred and three Mongolian samples were analyzed by determining the DNA sequence of mitochondrial control region I and by assaying seven additional polymorphic mitochondrial markers (six RFLPs and one 9-bp deletion) for each individual (Figure 2). Eighty-three mtDNA sequence haplotypes were observed among the 103 Mongolians. Of the 381 nucleotide positions sequenced, 81 were polymorphic. Seventy of the polymorphic positions were transitions with a strong bias toward pyrimidine transitions [30 C-toT, 25 T-to-C, 12 A-to-G and three G-to-A mutations, relative to the human reference sequence (ANDERSON *et al.* 1981)] and four were transversions (1 A-to-C, 1 C-to-A and 2 A-to-T mutations). Four instances of a transition and transversion at the same site were also observed: positions 16232 and 16111 showed both C-to-T and C-to-A mutations, position 16129 showed both G-to-A and G-to-C mutations and position 16263 showed both T-to-C and T-to-A mutations. Three single-base insertions were detected.


ID N Reference
Hae3 663
AluI 5176
DdeI 10394
AluI 10397
AluI 13262
RsaI 16329
HaeIII 16517
9 bp 8272/8289
New World haplogroup
Fig. 4 label

Nucleotide diversity, π, was determined to be 0.018 for the complete collection of 103 samples and 0.016 and 0.018 for the Dariganga and Khalkha samples, respectively. Long-term nucleotide diversity, $E(\nu)$, was 15.6 for the complete set of 103 samples and 10.4 and 14.3 for the Dariganga and Khalkha samples, respectively. Haplotype diversity, h, was 0.99 for all 103 samples and 0.98 and 0.99 for the Dariganga and Khalkha samples, respectively.

AMOVA was performed to test for population subdivision within the Mongolian populace. No significant interpopulation variance was detected between any of the three Mongolian collections representing Mongolia as a whole and the Khalkha and Dariganga cultural groups. However, when the molecular information was excluded, significant interpopulation variance was detected between the Khalkha and Dariganga (1.1% interpopulation variance, $P = 0.002$). Eight combinations of haplotypes from the 15 sampled regions were also analyzed to contrast cultural, geographic and ecological classifications. All analyses (calculated with molecular information) yielded minimal levels of 1–2% interpopulation variance. However, population differentiation was significant only in the comparison of regions 1–3, 8, 11–13 (western Mongolia + Dariganga) to regions 4, 5, 10, 15–18 (central Mongolia) (1.8% interpopulation variance, $P = 0.038$).

New World haplogroups in Mongolian populations: Restriction/deletion data and control region I sequence data were used independently to identify New World haplogroups in the Mongolian data set. Restriction/deletion sites were described by TORRONI *et al.* (1993a) and polymorphic sites within the control region were defined by HORAI *et al.* (1993) (Table 1). Although evidence suggests that several of the restriction/deletion polymorphisms have evolved more than a single time in human populations, additional cooccurring polymorphisms that define New World populations have also been detected in the Mongolian populations. For example, the combinations of 9-bp deletion/loss of *Dde*I/10394 and presence of *Alu*I/13262/C:16325 provide further definition of haplogroups B and C, respectively, in the New World and have also been detected in Mongolian populations, suggesting that the restriction/deletion polymorphisms assayed in this study can be used to describe haplotypes outside of the New World. Based on restriction/deletion data (seven restriction sites and a 9-bp deletion), 13 haplotypes were determined for the 103 Mongolian samples (Figure 2). All four New World haplogroups were detected in the collection of 103 Mongolian samples and also in each of the Dariganga and Khalkha collections. In addition, six haplotypes not defined by a single restriction site were identified. Of these six haplotypes, three had been previously identified in New World populations but most likely represented admixture in the New World groups with nonindigenous populations (TORRONI *et al.* 1993a).

Restriction/deletion analysis identified one group A haplotype (3.9% of the total sample), two group B haplotypes (9.7%), two group C haplotypes (15%) and two group D haplotypes (19%). Using control region sequence data, five group A haplotypes (4.8% of the total

TABLE 1

Mitochondrial DNA analysis of founding New World haplogroups in Mongolians

Site[a]	Haplogroup A	B	C	D
Restriction/deletion sites				
*Hae*III:663	1	0	0	0
9-bp	2	1	2	2
*Alu*I:13262	0	0	1	0
*Alu*I:5176	1	1	1	0
*Dde*I:10394	0	0	1	1
*Alu*I:10397	0	0	1	1
*Rsa*I:16329	1[b]/0	1	1	1
*Hae*III:16517	1/0	1	1/0	1/0
Total[c]	4	10	15	20
Control region I polymorphisms				
16189	T/c	C	T	T/c
16217	T	C	T	T
16223	T	C	T	T
16290	T	C	C	C
16298	T	T	C	T
16319	A	G	G	G
16325	T	T	C/t	C
16327	C	C	T	C
16362	C	T	T	C
Total[c]	5	10	14	ND

[a] Restriction/deletion defining sites are provided by TORRONI *et al.* (1993a), and control region I polymorphisms are provided by HORAI *et al.* (1993). Presence and absence of a restriction site are indicated by 1 and 0, respectively, and 1/0 signifies that the site is polymorphic in that haplogroup. One and two copies of the 9-bp repeat are indicated by 1 and 2, respectively. For control region I DNA sequence data, the defining sites are underlined. X/y indicates that X is the predominant nucleotide at that position although Y does occur at a low frequency.

[b] *Rsa*I/16329 is polymorphic only in Na-Dene populations.

[c] Number of haplotypes in each haplogroup in the Mongolian dataset based on restriction/deletion or control region I sequence data.

FIGURE 2.—Polymorphic nucleotides and restriction sites, numbered according to ANDERSON *et al.* (1981), and frequency of the 83 haplotypes reported in this study. New World haplogroup identification is based on DNA sequence and restriction/deletion data. Each haplotype is identified by the region in which it occurred with the following exceptions: haplotype 8.4 occurred in regions 8 and 17, haplotype 8.15 occurred two times in region 8 and one time in region 17, haplotype 15.2 occurred in regions 15 and 8, haplotype 17.6 occurred three times in region 17 and one time in region 6 and haplotype 17.27 occurred in regions 17 and 12. *Rsa*I/16329 is presented as a separate site although this position was scored by the DNA sequence analysis.

sample, including two individuals who lacked T:16223), seven group B haplotypes (9.7%, including three individuals who lacked C:16217) and 13 group C haplotypes (14%) were identified. None of the Mongolian haplotypes carried all three control region polymorphisms for haplogroup D although 23 haplotypes (not counting haplotypes previously identified as haplogroup A, B or C) carried two out of three of the mutations: a C-to-T transition at bp 16223 and a T-to-C transition at bp 16362.

One sample, 16.1, was identified as a group A haplotype (lacking T:16223) by DNA sequence analysis but was classified as a non-New World haplotype by restriction/deletion analysis. Sample 8.18 was identified as a group C haplotype based on restriction/deletion analysis but was placed in a non-New World haplogroup by DNA sequence analysis. By minimizing the number of changes required to reconcile the two data sets, both samples were determined to be non-New World haplotypes, *i.e.*, "other" haplotypes. With the exception of samples 16.1 and 8.18, there was exact congruence in the identification of haplogroups A, B and C in the Mongolian samples using DNA sequence or restriction/deletion data.

Phylogenetic analysis of Mongolian populations: Phylogenetic analyses were performed with a dataset consisting of 83 Mongolian haplotypes and 16 New World haplotypes (Figure 3) (New World sequences are listed in the APPENDIX). All unweighted heuristic searches resulted in trees of length 182 (2100 trees) while all weighted searches revealed trees of length 284 (1500 trees). All searches demonstrated that Mongolian haplotypes designated as "New World types" clustered with New World haplotypes from the same haplogroup, with the exception of group D haplotypes that did not cluster regardless of origin. Consensus trees from weighted and unweighted searches had the same basic topology: all group A haplotypes formed a single clade with the exception of Mongolian sample 17.13, group B haplotypes clustered although Mongolian samples 8.12 and 17.25 were excluded 50% of the time, group C haplotypes were consistently monophyletic and group D haplotypes were never monophyletic.

Phylogenetic analyses were also performed on the Mongolian data set without New World sequences. Regardless of branch swapping algorithm used (tree bisection-reconnection or subtree pruning-regrafting, with or without steepest descent), all unweighted searches resulted in trees of length 167 (3350 trees) and all weighted searches revealed trees of length 269 (2050 trees). Consensus trees from all searches shared the same topology described above with New World sequences: group A haplotypes clustered with the exception of 17.13, group B haplotypes clustered with the exception of 8.12 and 17.13, and, in two instances, 17.24, group C haplotypes were consistently monophyletic and group D haplotypes were never monophyletic. No consensus tree maintained all four New World haplogroups as monophyletic clusters, although one analysis resulted in a consensus tree that clustered haplogroups A, B and C with the addition of one step (length = 168). When the analysis was constrained by the requirement that all four haplogroups be monophyletic, 11 steps were added to the tree. Bootstrap analysis supported the clustering of group B haplotypes 4.1, 8.4, 18.2; group B haplotypes 8.12, 17.25; group C haplotypes 3.2, 17.15, 17.32; and group C haplotypes 8.21, 17.7, 17.22 at the 50% confidence level. All other groupings were supported at <50% confidence levels. Phylogenetic analysis was also used to test for population subdivision among the Mongolian populations. None of the analyses provided any distinction between populations based on geographic or cultural affiliation.

An MST of restriction/deletion haplotypes from Mongolian, Siberian, southeast Asian, Tibetan and several New World populations was also constructed (Figure 4). As with the phylogenetic analyses, the MST failed to resolve the Khalkha and Dariganga populations into distinct clusters (data not shown) and, furthermore, only resolved group A haplotypes into a monophyletic cluster while group B, C and D haplotypes were polyphyletic. Although multiple minimum-length trees from both the parsimony and MST analyses were found, the conclusions remained unchanged; neither analysis provided any distinction between geographic or cultural groups or defined all four New World haplogroups as monophyletic clusters.

DISCUSSION

Mongolia lies in the heart of east Central Asia, a region first inhabited by modern humans 20,000–25,000 years ago (FIEDEL 1992) and proposed to represent the origin of colonizing migrations to the north and south (NEI and ROYCHOUDHURY 1993). Evidence, both molecular and nonmolecular, reveals that Mongolians may also represent the first colonists of the New World (TURNER 1984; NEEL *et al.* 1994). Archaeological data document the stable presence of Mongols in the area for at least 10,000 years (TUMEN 1992), suggesting that Mongolians represent a group central to our understanding of the evolutionary and population genetic relationships of east Asian and New World indigenous populations.

Although mtDNA diversity was studied previously in Mongolian populations (SAMBUUGHIN 1992), it was not possible to correlate those conclusions with other mtDNA results because the markers that have come to dominate population studies of Asian and New World groups were not assayed. We set out to provide an mtDNA database on Mongolians that would permit us to place our results in a comparative framework based on both DNA sequence and restriction/deletion site data from Old and New World populations. This strat-

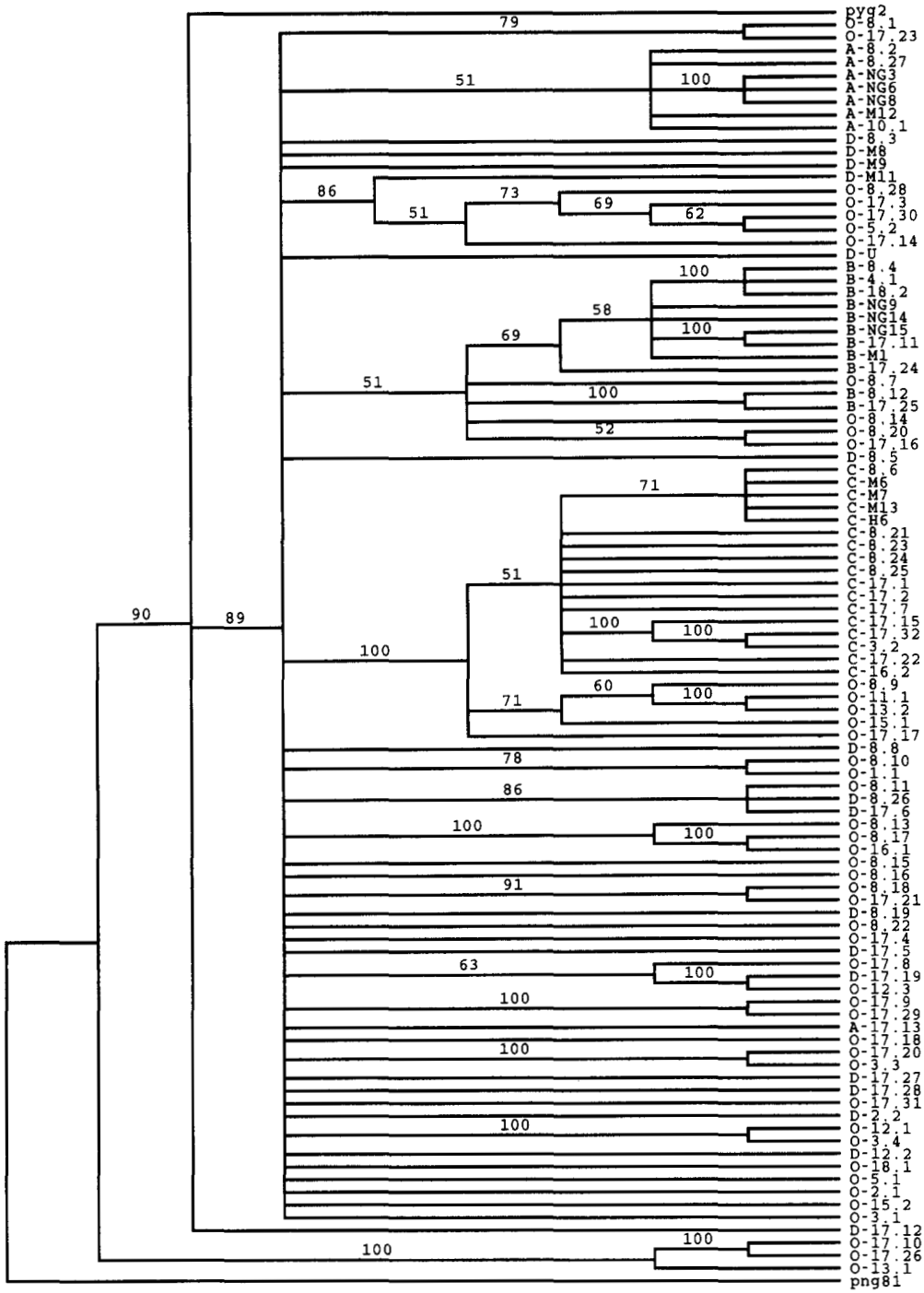


FIGURE 3.—Majority-rule consensus (50%) of 500 trees with length 182 generated by PAUP package. The data matrix of mtDNA sequences consisted of 101 taxa (83 Mongolian, 16 New World and two outgroup sequences) and 87 characters. Rohlf's consistency index for the consensus tree was 0.166. The single letter preceding each sample name indicates the haplogroup to which that sample belongs: A, B, C, D and O (for other). Mongolian haplotypes are labeled as in Figure 2 and New World haplotypes are labeled as in Figure 2 from KOLMAN *et al.* (1995) (New World sequences are listed in the APPENDIX). Numbers indicate the percentage of trees (out of 500) that support each branch point.

egy allowed us to compare our results with data from 18 eastern Asian populations representing 708 individuals and 15 New World populations representing 1170 individuals (Figure 5).

Mitochondrial diversity in Mongolian populations: mtDNA data were used to test the theory, based on archaeological evidence, of a long-term presence by Mongolians in east Asia. If accurate, high levels of genetic diversity in Mongolian populations relative to other east Asian groups would be expected. Furthermore, if ancient Mongols colonized the New World, elevated levels of mtDNA diversity of Mongolians rela-

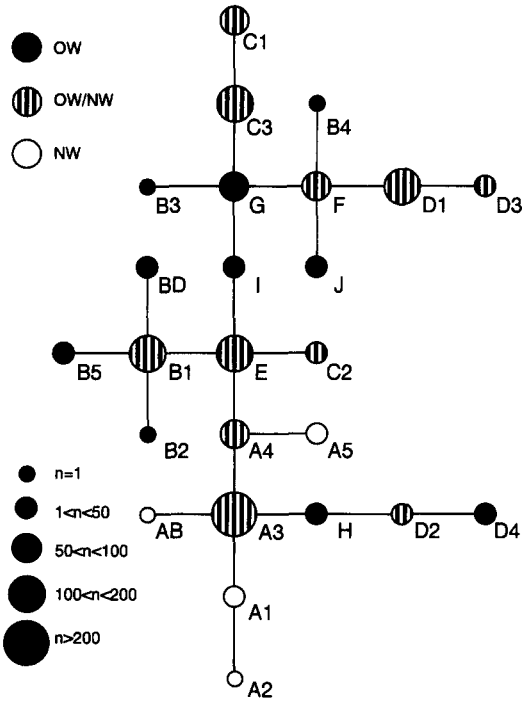


FIGURE 4.—Minimum spanning tree of RFLP haplotypes from Old and New World populations (weighted by haplotype frequency). The size of each circle reflects the frequency of that haplotype and different circle fill patterns indicate haplotypes from Old and/or New World populations (designated OW and NW). Each circle is separated from adjacent circles by a single mutation. Data on populations is as described in Figure 5 except that New World Eskimos, Aleuts and Mapuche are excluded because of insufficient restriction site information, Nuu-Chah-Nulth data are added (WARD *et al.* 1991), and Haida data are from TORRONI *et al.* (1993a). Group A, B, C and D haplotypes are identified by their respective letter and are numbered consecutively. Haplotype AB has the defining polymorphisms for haplogroups A and B and haplotype BD has the defining polymorphisms for haplogroups B and D. Non-New World haplotypes are designated E–J [haplotypes E and F were detected in some New World populations (TORRONI *et al.* 1993a) but were assumed to reflect nonindigenous admixture].

tive to their New World descendants would also be anticipated. When comparing levels of nucleotide diversity, as defined by π, no clear distinctions between Old and New World populations could be made. Mongolians, whether tested on a geographic or population basis, exhibited high levels of nucleotide diversity; π = 0.018 for all Mongolian samples, π = 0.016 for the Dariganga and π = 0.018 for the Khalkha. Although they represented a broader geographic sampling than the Mongolians, the Asian populations in HORAI and HAYASAKA's (1990) study carried a similar level of nucleotide diversity (π = 0.014 for a sample composed of 86% Japanese and 14% southeast Asians). To provide additional populations for comparison, we calculated a nucleotide diversity value of 0.014 for Siberians (representing five cultural groups) and 0.017 for the linguistically divergent Altai (two pooled populations) of southwest Siberia based on mtDNA control region I sequence data (SHIELDS *et al.* 1993; TORRONI *et al.* 1993b). Levels of mtDNA diversity in New World indigenous groups showed a wide range of values and were surprisingly high given the presumed history of small founder populations in the New World. For example, we calculated a π of 0.016 for the Nuu-Chah-Nulth of the Pacific Northwest (WARD *et al.* 1991) and 0.009 for the Kuna of eastern Panamá (BATISTA *et al.* 1995).

However, EXCOFFIER and LANGANEY (1989) have suggested that to speculate on the evolutionary history of a group one should use an estimator of nucleotide diversity, such as $E(\nu)$, which is independent of haplotype frequencies and measures a long-term average rather than a current generation value. Calculations of $E(\nu)$ revealed a dichotomy in diversity levels between Mongolians and non-Mongolians that was not apparent using a frequency-based measure of nucleotide diversity. For Mongolians, $E(\nu)$ = 15.6 for all Mongolian samples, 14.4 for the Khalkha and 10.4 for the Dariganga, while $E(\nu)$ = 7.4 for Siberians and 8.0 for the Altai. In the New World, the reduction in mtDNA diversity was even more apparent with $E(\nu)$ = 5.5 for the Emberá and 2.1 for the Kuna. These calculations suggest that much of the non-Mongolian mtDNA diversity may be due to the current distribution of haplotypes whereas Mongolians may have historically carried greater levels of mtDNA diversity.

Geographic differentiation of Mongolian genetic diversity (based on blood group, allozyme and nuclear RFLP markers) was studied previously by BATSUURI (1987) and SAMBUUGHIN *et al.* (1992) who, despite extensive sampling of Mongolia's 18 political regions (>30 samples from each region), found no significant differentiation between regions. The present study investigated mtDNA diversity at the level of cultural differentiation as represented by the Mongol-speaking Khalkha and Dariganga. No genetic differentiation was detected between the two groups when compared to each other or to Mongolia as a whole. Slight, but significant, levels of interpopulation variance between the Khalkha and Dariganga were detected when the molecular information was excluded from the analysis (1.1%, P = 0.002). The fact that only frequency data, and not DNA sequence data, revealed significant levels of genetic differentiation suggests that the subdivision may reflect a more recent, demographic effect rather than a long-term evolutionary split between these groups.

Slight, but significant levels of interpopulation variation were also detected when comparing regions 1–3, 8, 11–13 with regions 4, 5, 10, 15–18 (1.8%, P = 0.038). The first group of regions represents the Dariganga (region 8) and their geographic area of ancestry in western Mongolia (regions 1–3 and 11–13). Historical records show that more than 300 years ago several groups migrated from northwest Mongolia to the previously uninhabited region of southeast Mongolia (BA-

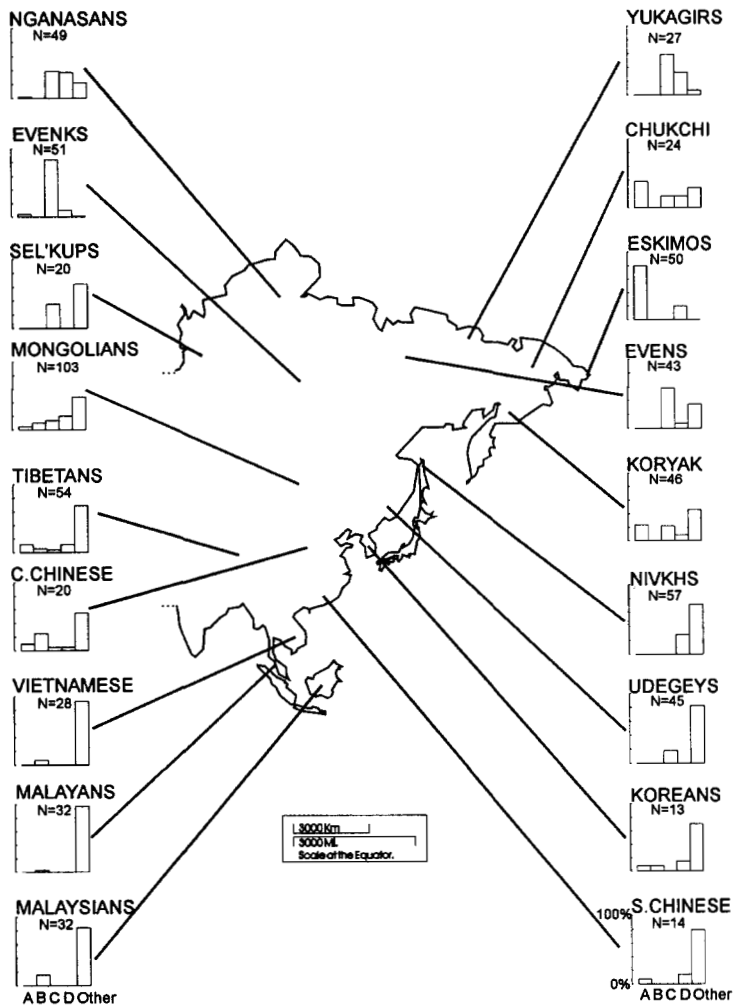


FIGURE 5.—Histogram representation of the frequency of New World haplogroups detected in 18 Asian and 16 New World groups. Frequency distributions are based on RFLP data using the definitive restriction sites described by TORRONI *et al.* (1993a). All haplotypes not defined by TORRONI *et al.* (1993a) are pooled as other haplotypes. For each histogram, sample sizes are indicated and the range of the Y axis is 0 to 100%. The location of each indigenous group is shown on the maps. Esk-Aleut and Na-Dene groups are indicated while all other groups on the New World map are Amerind. Data for the Nganasans, Sel'kups, Evenks, Yukagirs, Chukchi, Old World Eskimos, Evens, Koryak, Nivkhs and Udegeys are by TORRONI *et al.* (1993b), for the central Chinese, Vietnamese, Malayan and Malaysian aborigines, Koreans and south Chinese are by BALLINGER *et al.* (1992), for the Tibetans are by TORRONI *et al.* (1994), for the New World Eskimos and Aleut are by MERRIWETHER *et al.* (1995), for the Haida are by WARD *et al.* (1993) and for the Mapuche are by GINTHER *et al.* (1993). Data on Amerind and Na-Dene groups (excepting the Haida and Mapuche) are by TORRONI *et al.* (1993a), KOLMAN *et al.* (1995) and BATISTA *et al.* (1995) and are pooled according to the following linguistic classifications: N. Athapaskan, Dogrib; S. Athapaskan, Navajo and Apache; N. Amerind, Ojibwa and Bella Coola; Macro-Panoan, Marubo and Mataco; Chibhcan, Ngöbé, Boruca, Kuna, Guaymi and Bribri-Cabecar; Equatorial, Piaroa and Wapishana; Macro-Carib, Makiritare and Macushi.

DAMKHATAN 1987). It is thought that the Dariganga represent the modern descendants of these peoples. The observed subdivision of the Dariganga with western populations *vs.* the intervening central Mongolian populations would tend to support this view and suggests that a genetic record of the migration may exist within contemporary Mongolia populations.

No population structure was detected in any other regional grouping of Mongolian haplotypes although multiple combinations were tested. Comparisons tested were based on geographic (west *vs.* central), ecological (steppe *vs.* high mountain *vs.* semi-desert *vs.* forest/mountain) and cultural designations. The low degree of interpopulation genetic variation most likely reflects the nomadic lifestyle and recent divergence of Mongolian ethnia. Linguistic evidence also supports a lack of population subdivision as reflected in the fact that languages spoken by Mongol-speaking groups in Mongolia are generally mutually intelligible. Notwithstanding the two statistically significant but low levels of interpopulation variance presented above, these results suggest that from a population genetic perspective Mongolia can be viewed as a single, genetically homogeneous population.

Comparative analyses utilizing the four New World haplogroups: We now turn our attention to the possible role that Mongolians have played in the peopling

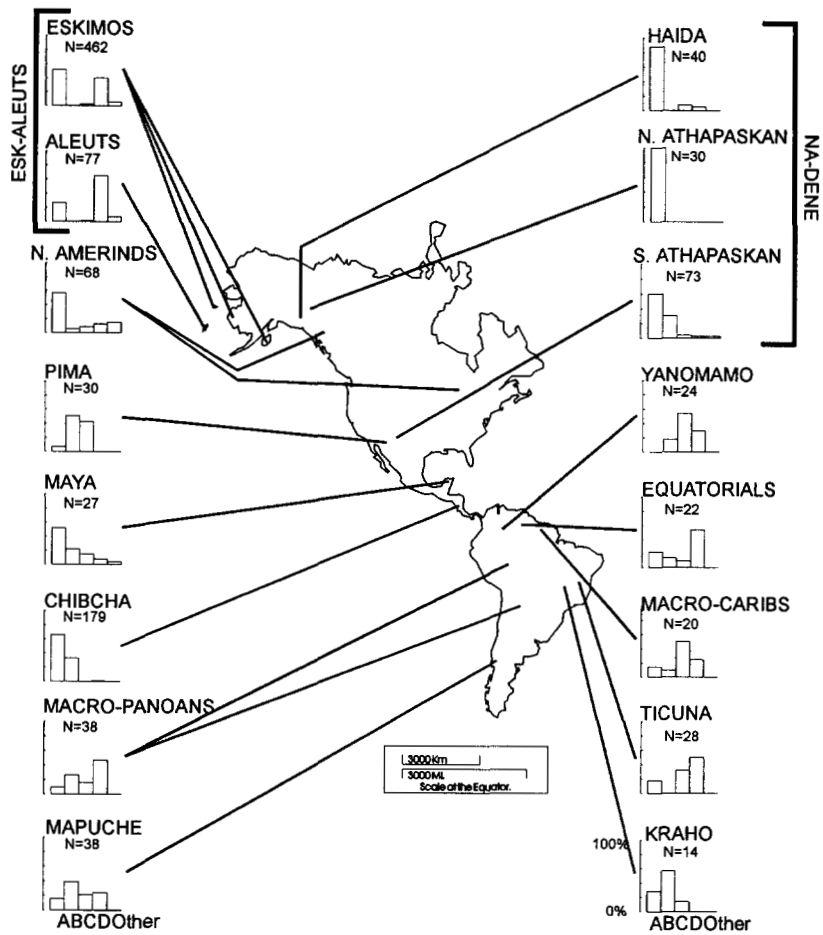


FIGURE 5.—(*Continued*)

of both Asia and the New World. The four New World haplogroups and their defining sites have come to dominate studies of both Old and New World genetic diversity (HERTZBERG *et al.* 1989; BAILLIET *et al.* 1993; HORAI *et al.* 1993; TORRONI *et al.* 1993a,b, 1994a,b) (Table 1). DOUG WALLACE's group first proposed that restriction/deletion data could be used to define the four haplogroups (designated A–D) that have been detected in populations distributed throughout the New World. Populations belonging to the three New World linguistic classes described in GREENBERG *et al.*'s (1986) three-wave hypothesis of New World colonization show characteristic frequencies of the four haplogroups. Eskimo-Aleuts exhibit high frequencies of haplogroups A and D and carry virtually no B and C haplotypes. Na-Dene populations display high frequencies of haplogroup A and low frequencies of haplogroups B, C and D. On average, Amerinds carry significant frequencies of all four haplogroups, although many individual Amerind groups lack one of the four New World haplogroups.

There is, however, some evidence that the four haplogroup scenario may not completely describe New World mtDNA diversity. For example, BAILLIET *et al.* (1994) proposed that the non-New World haplotypes identified in previous studies as nonindigenous admixture in New World groups may be evidence of a fifth founding haplotype. The picture becomes less clear when Asian haplotypes are considered. The premise of four founding haplotypes assumes that all haplotypes within a haplogroup derive from a single ancestral haplotype originating in Asia. This implies that haplotypes from any one haplogroup will cluster when Old and New World haplotypes are analyzed. However, a MST based on restriction/deletion haplotypes from Asian, Na-Dene and Amerind groups does not show the expected clustering (Figure 4). The five group A haplotypes resolve into a monophyletic cluster, but group B, C and D haplotypes are polyphletic occurring in three, two and two locations of the tree, respectively. Furthermore, although parsimony analysis revealed many trees of equal length, none were found that preserved all New World haplogroups as monophyletic clusters (Figure 3). When the analysis was constrained by the requirement that the four haplogroups be monophyletic, 11 steps were added to the minimal tree.

Notwithstanding some of the problems outlined above regarding the four haplogroup scenario, we have utilized the haplogroup defining sites to compare Old and New World populations. The parsimony analysis shown in Figure 3 reveals that Mongolian haplotypes

designated as New World types cluster with New World haplotypes from the same haplogroup, with the exception of group D haplotypes that do not cluster regardless of origin. This result suggests that the haplogroup defining sites may be useful in describing populations from both the Old and New World. Figure 5 shows the frequency of New World haplogroups throughout Asia and the New World based on restriction/deletion polymorphisms; +*Hae*III/663, + 9-bp deletion, +*Alu*I/13262 and −*Alu*I/5176. Haplotypes that did not carry any of the defining sites were grouped together under "other". Assuming that only four haplogroups were present in the populations that colonized the New World, the presence of other haplotypes in New World populations was presumed to reflect admixture with nonindigenous individuals, while other haplotypes in Asian populations were simply a measure of the non-New World haplotypes in that population.

Mongolian populations exhibit the highest percentage (48%) of New World haplogroups of any Asian population that carries all four haplogroups. The other two Asian populations that carry all four haplogroups are the central Chinese (designated Taiwanese Han in BALLINGER *et al.* 1992) with 45% New World haplotypes and the Tibetans with 31% New World haplotypes (TORRONI *et al.* 1994b). However, the four haplogroups are not well represented in most east Asian populations. For example, populations in southeast Asia carry haplogroup B almost exclusively of the four New World haplogroups. Siberians, on the other hand, carry all New World haplogroups except haplogroup B. The extensive geographic and numeric sampling of Siberian populations (10 cultural groups, $n = 411$) (TORRONI *et al.* 1993b) makes it unlikely that the inability to detect B haplotypes in Siberia was due to inadequate sampling.

The four haplogroups are also nonrandomly distributed throughout the New World. To the extent that there is little admixture in Amerind groups, even between neighboring populations (KOLMAN *et al.* 1995), each population can be thought of as representing a small sample from a small pot of founding haplotypes. It is, therefore, not surprising that there are sampling effects and that they appear to be picked up again and again as one moves away from Central Asia. One clear example of this is exhibited by the Chibcha of lower Central America. In contrast to the majority of Amerind groups, six Chibchan populations, with a sample size of 179, have been found to carry significant frequencies of only A and B haplotypes (TORRONI *et al.* 1993b; SANTOS *et al.* 1994; BATISTA *et al.* 1995; KOLMAN *et al.* 1995) (Figure 5). KOLMAN *et al.* (1995) proposed that this reduction in mtDNA diversity reflected the founder effect caused by divergence of the proto-Chibcha population from the Amerind migration. General frequency clines of the four haplogroups are also evident throughout the New World. The most obvious of these is the north-to-south decrease in group A haplotypes. In Amerind groups, there is a less striking increase in haplogroups B, C and D as one moves north to south.

New World colonization theories: The distribution of New World haplogroups across Asia and the Americas can be used to address various scenarios regarding colonization of the New World. One of the most vigorously defended colonization hypotheses was advanced by GREENBERG *et al.* (1986) who proposed three temporally and genetically distinct migrations to the New World based on linguistic, dental morphology and protein polymorphism evidence. Recent support for the three-wave hypothesis has come from mtDNA studies that suggest that mtDNA diversity differences between Amerind and Na-Dene indicate two distinct migrations for these groups (TORRONI *et al.* 1993a).

However, a review of the published literature reveals that all four New World mtDNA haplogroups have been detected in Esk-Aleut, Na-Dene and Amerind populations with no haplogroup unique to any of the three linguistic classes (Table 2). In other words, groups assigned to the three proposed migrations simply exhibit different frequencies of the same set of mtDNA haplogroups. The characteristic haplogroup frequencies of the three classes of populations may reflect the effects of genetic drift and founder events rather than traces of independent waves of migration. For example, the virtual absence of C and D haplotypes in the Chibcha has been proposed to indicate a population bottleneck at Chibchan ethnogenesis since an independent Chibchan migration to the New World represents an unlikely scenario (BATISTA *et al.* 1995; KOLMAN *et al.* 1995). The large variance in haplogroup frequencies observed in New World populations provides further evidence of the sampling effects manifest throughout the Americas, *e.g.*, haplogroup A ranges from 0–66% in Amerind groups alone. When the focus is broadened to include Asia, one realizes that founder effects as illustrated by the four haplogroups are not limited to the New World but extend into the Old World, *e.g.*, haplogroup B reaches fixation in the Polynesian Islands (HERTZBERG *et al.* 1989) while it is absent in Siberia (Table 2).

Mitochondrial DNA support for three distinct waves of migration is further weakened when viewed from the perspective of Asian populations. Examination of Figure 5 reveals that the four New World haplogroups are not common or widespread in Asia. Mongolia, Tibet and central China circumscribe the only region in Central Asia with populations that carry all four haplogroups. The statistical and geographic scarcity of the New World haplogroups in Asia makes it extremely unlikely that the same four haplogroups would be drawn out of one geographic region three independent times to form New World Esk-Aleut, Na-Dene and Amerind groups. Viewed as a simple sampling exercise, the mtDNA data appear to support a single colonizing migration to the New World with subsequent founder effects and genetic drift largely responsible for the dif-

TABLE 2

Frequencies (×100) of New World haplogroups in Old and New World populations

Population(s)[a]	Haplogroup A	B	C	D	Other
Siberians					
Nganasans	2.0%	0	39%	37%	22%
Evenks	3.9	0	84	10	2.0
Sel'kups	0	0	35	0	65
Yukagirs	0	0	59	33	7.4
Chukchi	38	0	17	17	29
Eskimos	80	0	0	20	0
Evens	0	0	58	7.0	35
Koryak	24	0	22	8.7	46
Nivkhs	0	0	0	28	72
Udegeys	0	0	18	0	82
Central Asia					
Mongolians	4.8	9.7	14	20	52
Tibetans	11	5.6	3.7	11	68
C. Chinese	10	25	5.0	5.0	55
SE Asia					
Vietnamese	0	7.1	0	0	93
Malayans	0	3.1	0	0	97
Malaysians	0	16	0	0	84
Koreans	7.7	7.7	0	15	69
S. Chinese	7.1	0	0	14	78
Esk-Aleuts					
Eskimos	51	0.6	2.4	40	5.4
Aleuts	27	0	1.3	65	6.5
Na-Dene					
Haida	88	0	7.5	5.0	0
N. Athapaskan	100	0	0	0	0
S. Athapaskan	60	30	4.1	2.7	2.7
Amerinds					
N. Amerinds	57	5.9	8.8	13	15
Pima	6.7	50	43	0	0
Maya	52	22	15	7.4	3.7
Chibcha	66	33	0	0.6	0
Macro-Panoans	10	26	16	47	0
Mapuche	16	39	21	24	0
Yanomamo	0	17	54	29	0
Equatorials	23	14	9.1	54	0
Macro-Caribs	15	10	50	25	0
Ticuna	18	0	32	50	0
Kraho	28	57	14	0	0

[a] References and linguistic classifications for populations are described in Figure 5.

ferences in haplotype frequencies observed in contemporary indigenous groups.

Founders of the New World: Examination of the distribution of the four New World haplogroups across Asia and the New World is equally revealing when applied to the search for New World ancestors. Detection of all four haplogroups in an Asian population would seem to be a first criterion in the identification of a possible New World founder. Mongolian, Tibetan and central Chinese populations are the only groups in Central Asia that carry all four haplogroups and, therefore, appear to be the strongest candidates for New World founder populations. Tibetans are known to have migrated from northern China several thousand years ago (CAVALLI-SFORZA *et al.* 1994) and, therefore, would have been northern Chinese at the time of New World colonization. This fact would seem to restrict possible New World founders to the narrow strip of east Central Asia that extends from Mongolia eastward to the coast.

Lack of group B haplotypes would appear to eliminate Siberians as possible New World founders despite their geographic proximity to the Americas. In fact, evidence exists that suggests that Siberia may have been colonized subsequent to the New World based on more recent divergence dates for Siberian C and D haplotypes relative to New World C and D haplotypes (13,500–27,000 YBP *vs.* 18,750–37,500 YBP) (TORRONI *et al.* 1993b). An additional challenge to Siberian ancestry of New World groups comes from several virology studies that detected human T-cell lymphotrophic retrovirus (HTLV-II) in 11 of 38 Amerind groups tested (*e.g.*, LAIRMORE *et al.* 1990; MALONEY *et al.* 1992) and also in Mongolians (HALL *et al.* 1994) but not in 10 Siberian groups assayed (NEEL *et al.* 1994).

The fact remains, however, that A, C and D haplotype frequencies are elevated in many Siberian populations relative to other Asian groups (Figure 5) suggesting an historical relation between Siberians and New World indigenous groups. SZATHMARY (1994) confirmed the close genetic relationship between circumarctic populations in a UPGMA analysis of 15 blood group and allozyme markers that revealed that Siberian, New World Eskimo and Athapaskan (subset of Na-Dene) populations clustered with each other and not with other Asian or Amerind groups. Thus, Siberians and other circumarctic populations may share a similar genetic origin or their close genetic relationship may reflect long-term genetic exchange.

Conclusions: The results presented here establish a broad geographical and numerical database of Mongolian mtDNA diversity. Using mtDNA control region I DNA sequence and seven additional restriction/deletion sites, Mongolian populations were found to exhibit high levels of interpopulations variance, whereas little to no intrapopulation variance was detected. The four New World founding haplogroups, although rare in Asia, were detected in the two Mongolian populations. A comparative analysis of the four haplogroups across Asia and the New World provided evidence of founder effects manifest throughout the entire area that may be responsible for differences in mtDNA haplogroup frequencies observed in contemporary populations.

Based on the current distribution of mtDNA haplogroups, we propose that populations in east Central Asia represent the closest genetic link between the Old World and the New World. All four New World haplogroups have been detected in Mongolian, central Chinese and Tibetan populations that delineate the only region in Asia where all four haplogroups exist and no population

lacking any one of the haplogroups occurs. Thus, the narrow strip of east Central Asia that extends from Mongolia to the Pacific coast may have served as the starting point for the human migration that led to colonization of the New World. Furthermore, presence of the four New World haplogroups throughout the Americas, but a restricted distribution in Asia, suggests a single sampling of these haplotypes. The emerging mtDNA picture of genetic diversity in the Americas appears to support a single migration, perhaps sustained over a period of time, of modern humans that gave rise to all contemporary New World populations. This scenario still allows for the possibility of other ancient migrations whose populations did not survive or at least left no maternal, *i.e.*, mtDNA, record of their occupation.

We gratefully acknowledge the participation of the Mongolian people in our study. M. WALKER and O. BATISTA assisted in the DNA sequence and RFLP analysis. We owe a special debt of thanks to S. MCCAFFERTY for laboratory and analytical assistance. P. SMOUSE, C. KOBAK and C. CHEVILLON collaborated in the MST analysis of DNA sequence and RFLP haplotypes. D. SWOFFORD provided invaluable assistance in the parsimony analysis of DNA sequence haplotypes. C. J. K. was supported by the Smithsonian Tropical Biology and Molecular Evolution Fellowship programs. N.S. was supported by an American Association of University Women International Fellowship and a United Nations Educational, Scientific and Cultural Organization Short-term Fellowship. We acknowledge the Smithsonian Molecular Evolution program and a James Smithson Award to E. B. and N. TUROSS for research support.

LITERATURE CITED

ANDERSON, S., A. T. BANKEIR, B. G. BARRELL, M. H. L. DEBRUIJN, A. R. COULSON *et al.*, 1981 Sequence and organization of the human mitochondrial genome. Nature **290:** 457–465.

BADAMKHATAN, S., 1987 *Ethnography of Mongolian People's Republic.* Mongolian State Publ., Vol. 1, Ulaanbaatar, Mongolia.

BAILLIET, G., F. ROTHHAMMER, F. R. CARNESE, C. J. BRAVI and N. O. BIANCHI, 1994 Founder mitochondrial haplotypes in Amerindian populations. Am. J. Hum. Genet. **54:** 27–33.

BALLINGER, S. W., T. G. SCHURR, A. TORRONI, Y. Y. GAN, J. A. HODGE *et al.*, 1992 Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongoloid migrations. Genetics **130:** 139–152.

BATISTA, O., C. J. KOLMAN and E. BERMINGHAM, 1995 Mitochondrial DNA diversity in the Kuna Amerinds of Panamá. Hum. Mol. Genet. **4:** 921–929.

BATSUURI, J., 1986 Genetic polymorphism and gene geography of Mongolian population. Ph.D. thesis, N. I. Vavilov Institute of General Genetics, Moscow, Russia.

CAVALLI-SFORZA, L. L., P. MENOZZI and A. PIAZZA, 1994 *The History and Geography of Human Genes.* Princeton University Press, Princeton, NJ.

DEREVIANKO, A. P., 1969 The Novopetrovka blade culture on the middle Amur. Arctic Anthropol. **6:** 119–127.

EXCOFFIER, L., and A. LANGANEY, 1989 Origin and differentiation of human mitochondrial DNA. Am. J. Hum. Genet. **44:** 73–85.

EXCOFFIER, L., P. E. SMOUSE and J. M. QUATTRO, 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application of human mitochondrial DNA restriction data. Genetics **131:** 479–491.

FIEDEL, S. J., 1992 *Prehistory of the Americas.* Cambridge University Press, Cambridge.

GINTHER, C., D. CORACH, G. A. PENACINO, J. A. REY, F. R. CARNESE *et al.*, 1993 Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes, pp. 211–219 in *DNA Fingerprinting: State of the Science,* edited by S. D. J. PENA, R. CHAKRABORTY, J. T. EPPLEN and A. J. JEFFREYS. Birkhäuser Verlag, Basel, Switzerland.

GONGOR, D., 1970 *Khalkha Mongolians and Imperia of Khalkha (YIII-XYII cent.).* Acad. Sci. Mongolia, Ulaanbaatar, Mongolia.

GREENBERG, J. H., C. G. TURNER II and S. L. ZEGURA, 1986 The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. Curr. Anthropol. **27:** 477–497.

HALL, W. W., S. W. ZHU, P. HORAL, Y. FURUTA, G. ZAGAANY *et al.*, 1994 HTLV-II infection in Mongolia. AIDS Res. Hum. Retroviruses **10:** 443 (abstr.).

HERTZBERG, M., K. N. P. MICKLESON, S. W. SERJEANTSON, J. F. PRIOR and R. J. TRENT, 1989 An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. Am. J. Hum. Genet. **44:** 504–510.

HORAI, S., and K. HAYASAKA, 1990 Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. Am. J. Hum. Genet. **46:** 828–842.

HORAI, S., R. KONDO, Y. NAKAGAWA-HATTORI, S. HAYASHI, S. SONODA *et al.*, 1993 Peopling of the Americas, founded by four main lineages of mitochondrial DNA. Mol. Biol. Evol. **10:** 23–47.

KOLMAN, C. J., E. BERMINGHAM, R. COOKE, R. H. WARD, T. D. ARIAS *et al.*, 1995 Reduced mtDNA diversity in the Ngöbé Amerinds of Panamá. Genetics **140:** 275–283.

LAIRMORE, M. D., S. JACOBSON, F. GARCIA, B. K. DE, L. CASTILLO *et al.*, 1990 Isolation of human T-cell lymphotropic virus type 2 from Guaymi Indians in Panamá. Proc. Natl. Acad. Sci. USA **87:** 8840–8844.

MADDISON, D. R., M. RUVOLO and D. L. SWOFFORD, 1992 Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. Syst. Biol. **41:** 111–124.

MALONEY, E. M., R. J. BIGGAR, J. V. NEEL, M. E. TAYLOR, B. H. HAHN *et al.*, 1992 Endemic human T cell lymphotropic virus type II infection among isolated Brazilian Amerindians. J. Infect. Dis. **166:** 100–107.

MERRIWETHER, D. A., F. ROTHHAMMER and R. E. FERRELL, 1995 Distribution of the four founding lineage haplotypes in native Americans suggests a single wave of migration for the New World. Am. J. Phys. Anthropol. **98:** 411–430.

NEEL, J. V., R. J. BIGGAR and R. I. SUKERNIK, 1994 Virologic and genetic studies relate Amerind origins to the indigenous people of the Mongolia/Manchuria/southeastern Siberia region. Proc. Natl. Acad. Sci. USA **91:** 10737–10741.

NEI, M., and A. K. ROYCHOUDHURY, 1974 Sampling variance of heterozygosity and genetic distance. Genetics **76:** 379–390.

NEI, M., and A. K. ROYCHOUDHURY, 1993 Evolutionary relationships of human populations on a global scale. Mol. Biol. Evol. **10:** 927–943.

NEI, M., and F. TAJIMA, 1981 DNA polymorphism detectable by restriction endonucleases. Genetics **97:** 145–163.

SAMBUUGHIN, N. 1992 Molecular genetics of Mongolian populations by RFLP analysis of nuclear and mitochondrial genomes. Ph.D. thesis, N. I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia.

SAMBUUGHIN, N., V. N. PETRISHCHEV and Y. G. RYCHKOV, 1991 DNA polymorphism in a Mongolian population: analysis of restriction endonuclease polymorphism of nuclear DNA. Genetika **27:** 2137–2142.

SAMBUUGHIN, N., Y. G. RYCHKOV and V. N. PETRISHCHEV, 1992 Genetic differentiation of Mongolian population. The geographical distribution of mtDNA RFLPs, mitotypes and population estimation of mutation rate for mitochondrial genome. Genetika **28:** 136–153.

SANTOS, M., R. H. WARD and R. BARRANTES, 1994 mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. Hum. Biol. **66:** 963–977.

SHIELDS, G. F., K. HECKER, M. I. VOEVODA and J. K. REED, 1992 Absence of the Asian-specific region V mitochondrial marker in native Beringians. Am. J. Hum. Genet. **50:** 758-765.

SHIELDS, G. F., A. M. SCHMIECHEN, B. L. FRAZIER, A. REDD, M. I. VOEVODA *et al.*, 1993 mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am. J. Hum. Genet. **53:** 549–562.

SWOFFORD, D. L., 1989 *PAUP, Phylogenetic Analysis Using Parsimony, Version X.* Illinois Natural History Survey, Champaign, IL.

SZATHMARY, E. J. E., 1994 Modelling ancient population relationships from modern populations genetics, pp. 117–130 in *Method*

and Theory for Investigating the Peopling of the Americas, edited by R. BONNICHSEN and D. G. STEELE. Center for the Study of the First Americans, Oregon State Univ., Corvallis, OR.

TORRONI, A., T. G. SCHURR, C.-C. YANG, E. J. E. SZATHMARY, R. C. WILLIAMS *et al.,* 1992 Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics **130:** 153–162.

TORRONI, A., T. G. SCHURR, M. F. CABELL, M. D. BROWN, J. V. NEEL *et al.,* 1993a Asian affinities and continental radiation of the four founding native American mtDNAs. Am. J. Hum. Genet. **53:** 563–590.

TORRONI, A., R. I. SUKERNIK, T. G. SCHURR, Y. B. STARIKOVSHAYA, M. F. CABELL *et al.,* 1993b mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am. J. Hum. Genet. **53:** 591–608.

TORRONI, A., Y.-S. CHEN, O. SEMINO, A. S. SANTACHLARA-BENECERETTI, C. R. SCOTT *et al.,* 1994a mtDNA and Y-chromosome polymorphisms in four native American populations from southern Mexico. Am. J. Hum. Genet. **54:** 303–318.

TORRONI, A., J. A. MILLER, L. G. MOORE, S. ZAMUDIO, J. ZHUANG *et al.,* 1994b Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am. J. Phys. Anthropol. **93:** 189–199.

TUMEN, D., 1992 Anthropology of the contemporary population of Mongolian People's Republic, Ph.D. thesis, M. L. Lomonosov State University of Moscow, Moscow, Russia.

TURNER, C. G., 1984 Advances in the dental search for native American origins. Acta Anthropogenetica **8:** 23–78.

VIGILANT, L., R. PENNINGTON, H. HARPENDING, T. D. KOCHER and A. C. WILSON, 1989 Mitochondrial DNA sequences in single hairs from a southern African population. Proc. Natl. Acad. Sci. USA **86:** 9350–9354.

VIGILANT, L., M. STONEKING, H. HARPENDING, K. HAWKES and A. C. WILSON, 1991 African populations and the evolution of human mitochondrial DNA. Science **253:** 1503–1507.

WARD, R. H., B. L. FRAZIER, K. DEW-JAGER and S. PÄÄBO, 1991 Extensive mitochondrial diversity within a single Amerindian tribe. Proc. Natl. Acad. Sci. USA **88:** 8720–8724.

WARD, R. H., A. REDD, D. VALENCIA, B. FRAZIER and S. PÄÄBO, 1993 Genetic and linguistic differentiation in the Americas. Proc. Natl. Acad. Sci. USA **90:** 10663–10667.

WATTERSON, G. A., 1975 On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7:** 256–276.

Communicating editor: W.-H. LI

APPENDIX

New World haplotypes tested with Mongolian haplotypes in PAUP analysis

ID	Figure 2 label	Reference	GenBank accession number[a]
Group A haplotypes			
NG3	NG3	KOLMAN *et al.* 1995	L39329
NG6	NG6	KOLMAN *et al.* 1995	L39332
NG8	NG8	KOLMAN *et al.* 1995	L39334
Lineage 18	M12	GINTHER *et al.* 1993	T:16223, T:16290, A:16319, C:16362
Group B haplotypes			
NG9	NG9	KOLMAN *et al.* 1995	L39335
NG14	NG14	KOLMAN *et al.* 1995	L39340
NG15	NG15	KOLMAN *et al.* 1995	L39341
Lineage 1	M1	GINTHER *et al.* 1993	C:16189, G:16207, C:16217, T:16291
Group C haplotypes			
Lineage 9	M6	GINTHER *et al.* 1993	C:16189, T:16223, C:16298, C:16325, C:16327
Lineage 10	M7	GINTHER *et al.* 1993	G:16051, T:16223, C:16298, C:16325, C:16327
Lineage 11	M13	GINTHER *et al.* 1993	T:16223, C:16298, C:16325, C:16327
Lineage 34	H6	WARD *et al.* 1993	L20148
Group D haplotypes			
Lineage 12	M8	GINTHER *et al.* 1993	T:16223, T:16242, C:16311, C:16325, C:16362
Lineage 13	M9	GINTHER *et al.* 1993	T:16187, T:16223, C:16325, C:16362
Lineage 17	M11	GINTHER *et al.* 1993	C:16092, T:16187, C:16189, T:16223, C:16362
Lineage 21	U	WARD *et al.* 1991	M76011

[a] If a GenBank accession number is unavailable, the mutations relative to the ANDERSON *et al.* (1981) reference sequence are listed.