Am. J. Hum. Genet. 55:7–11, 1994

INVITED EDITORIAL
mtDNA and Native Americans: a Southern Perspective

Rebecca L. Cann

Department of Genetics and Molecular Biology, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu

Introduction

Bailliet et al.'s (1994) article "Founder Mitochondrial Haplotypes in Amerindian Populations," in this issue of the *Journal,* magnifies the importance of understanding the genetic links between modern Native American populations, their ties to past populations, and their connections to the aboriginal people of Asia. In particular, human geneticists are now poised to break the circular use of linguistic, dental, and archaeological evidence and accurately dissect the last continental expansion of anatomically modern people. Contrary to expectation that the slow accumulation of evidence from enough loci would eventually resolve contradictions, we see instead that the key to unambiguous results is, once again, the careful attention to how and where populations are sampled (Morton 1993).

Although a number of landmark papers in anthropological genetics using new mtDNA markers have appeared in the past 3 years (Ward et al. 1991; Ginther et al. 1993; Horai et al. 1993; Shields et al. 1993; Torroni et al. 1993*a*, 1993*b*), the interpretations of these data sets have not been straightforward. At face value, any date 40,000–15,000 years before the present (ybp) is consistent with the arrival of the first Native Americans. Not only did authors publishing about aboriginal Americans disagree on the number of migrations, they also parted company on when people began moving, and from where.

Readers here may be relieved to know that their frustration in understanding the extensive dialogue on Native American genetic diversity that has taken place between research laboratories is widely shared among anthropologists and linguists as well (Flemming 1993; Meltzer 1993; Szathmary 1993*b*): "one could argue, of course, that there have always been foolish people and there always will be" (Flemming 1993, p. 17). If the questions did not interest so many at a variety of levels, we might simply ignore the lack of consensus.

Received April 12, 1994; accepted for publication April 13, 1994.

Address for correspondence and reprints: Dr. Rebecca L. Cann, Department of Genetics and Molecular Biology, Biomedical Sciences Building, 1960 East-West Road, Honolulu, HI 96822.

This article represents the opinion of the author and has not been peer-reviewed.

0002-9297/94/5501-0002$02.00

What were the genotypes of the first Americans, their time of arrival, or their genetic population structures? Where exactly did they stem from, and how fast did they expand? Most human geneticists have believed that there were at least three different migratory waves from north-central Asia, corresponding to Paleoindians of North and South America (roughly equivalent to Amerind), Na Dene peoples now largely restricted to North America, and the Aleut-Eskimo populations of the far north who now occupy both sides of the Bering Straits (Cavalli-Sforza et al. 1988). Genetic, linguistic, and morphological data on Native Americans were summarized for a general audience during the mid 1980s in a widely cited multidisciplinary paper (Greenberg et al. 1986). Instead of development of diversity in situ, these authors placed the major origins of diversity in three different founding groups that took separate evolutionary trajectories in time and space. According to their model, three groups of migrants were significantly divergent in language, morphology, and genetic markers before they crossed Beringia and extended their range into the Americas. The time scale associated with first colonization was open but was assumed to be related to Clovis-age artifacts <13,000 years old. Owing to the relatively recent appearance of RFLP and DNA sequence data, the authors relied on more readily available protein polymorphism studies.

Although there has been a significant amount of dissension among linguists regarding Greenberg's (1987) startling classification of American language groups into only three major phyla, he is also supported by many specialists. The close correlation between some linguistic groups and their gene pools (Renfew 1991) has meant that geneticists, on the whole, have been less critical of Greenberg's ideas. One break with this tradition among geneticists has been the recent paper of Ward et al. (1993), which questions the equivalence of demes with linguistic units in the Pacific Northwest. I also refer readers to Szathmary's (1993*a*) excellent summary of the history of anthropological issues raised by previous American colonization models—and of how new genetic information was supposed to support or refute those models.

The Beginning of the Dialogue

Early work with maternally inherited mtDNA RFLP patterns in Pima, Maya, and Ticuna Indians was interpreted to support the suggestion of a significant popula-

tion bottleneck in the peopling of North America, with a secondary barrier developing in Central America that partially isolated the populations of South America. This supposedly accounted for the different frequencies of known lineage clusters found there (Wallace et al. 1985; Schurr et al. 1990). If the bottleneck in Beringia or North America was prolonged, genetic diversity might be purged and surviving modern lineages could appear to trace to a single geographic homeland. The task was to clearly identify the homeland.

In the Amerind donors sampled, only four major RFLP haplotype clusters (A–D) were recognized by these authors; and the implications of a dramatic founder event for epidemiological genetics were profound. In passing through the Arctic filter, some Old World genotypes and associated diseases might have been lost, but others could have been accentuated. Diabetes, hypertension, alcoholism, or any other current public health problem that had genetic as well as socioeconomic components might be shown to have simple genetic correlates in indigenous peoples. Genetic screening for individuals potentially at risk could conceivably be cheap as well as comprehensive. In trying to solve the colonization puzzle, it looked as if researchers might also score a major advance for genetic epidemiology.

Torroni et al. (1993*b*) have recently suggested that the homeland of Amerinds is in eastern Siberia, on the basis of comparing 10 aboriginal populations in Siberia and the Russian Far East with Native Americans. Three major mtDNA groups are present in this region. The authors acknowledged that modern Siberians lack a maternal genetic lineage cluster (termed "B"), now present at highest frequency in South Americans, Pacific Islanders, and Indonesians (Hertzberg et al. 1989; Shields et al. 1993; Lum et al. 1994). They presume that this lineage came into the Americas in a second, later wave of Amerind migrants (Torroni et al. 1994).

In fact, the B lineage cluster is present along the Pacific coast of North and South America. It was also reported to be present in an ancient Colorado Paleoindian mummified donor dated 8,000 ybp (A. Stone, personal communication). Now, Bailliet et al. document that this lineage cluster, as well as four others, can be found in unadmixed Argentine Native Americans. Researchers are asking, just how robust can the Siberian-model ancestor for all Amerinds be—and just how robust can the time scale associated with timing Paleoindian entry into North America be—if we cannot yet agree on the number of major lineage groups present in the initial colonization wave?

On a theoretical front, disputes arose around the idea that Native Americans had ever undergone a significant genetic population bottleneck (Chakraborty and Weiss 1991). This challenge gained momentum when mtDNA hypervariable sequences were reported for a single Amerind tribe (Nuu-Chah-Nulth) from the Pacific Northwest. The authors of this independent study failed to support the limited-diversity/restricted-origin model (Ward et al. 1991). Instead of only four haplotypes, that study found 28 discrete maternal lineages in just 63 donors. The confusion between haplotypes as one lineage and haplotype clusters as seen in sequencing leads to a resolution problem. By focusing on a single segment of the mitochondrial hypervariable control region in one tribe, those authors discovered maternal genetic diversity that was equivalent to ~62% of that found in modern Africans and to ~81% of that present in urban Japan. The Amerind maternal lineages, moreover, traced back to a coalescent female who was projected to have lived ~60,000 ybp. Far from supporting the hypothesis of a genetic population bottleneck in the founding of Amerindians, this study argued that comparatively large groups were involved in the colonization of the New World, on the basis of both the large number of Nuu-Chah-Nulth lineages surviving today and the implied effective population size.

Critics countered that this tribe was merely an amalgamated group of epidemic smallpox survivors (aren't we all!) with questionable linguistic affinities, not a local population equivalent to previously tested groups. According to others, multilocus electrophoretic or RFLP patterns from nuclear loci would resolve issues in human population expansion that mtDNA studies could not (Livshits and Nei 1990; Szathmary 1993*a*). Reasoning that mtDNA samples represented only a single locus with large stochastic variation, they asserted that few definitive results would ever be expected from using such a system. This criticism might have intimidated some people, but it should be noted that the most recent synthetic summary of human population genetics from 29 polymorphic nuclear loci in 26 populations with 121 alleles (and no missing comparisons, making it unlike the presentation of Cavalli-Sforza et al. 1988) could not even state, with statistical significance, whether modern Eskimo populations are more closely related to North American or to South American aboriginal groups (Nei and Roychoudhury 1993)!

New hypervariable mtDNA sequencing studies by Ginther et al. (1993), Horai et al. (1993) and Torroni et al. (1993*a*) fully support the view that Amerinds have an ancient maternal coalescent point, with many genetically distinct lineages. They can also be interpreted to show that most of the identified lineages correspond to one of four major clusters in phylogenetic analysis, vindicating previous RFLP surveys. Each cluster has a relatively deep coalescent, predating the Clovis artifacts associated with Paleoindian expansion. These studies, however, intensify the mystery both of migratory waves and of the associated time scales. The statistical power of the phylogenetic analysis in all cases is low, because of the large number of lineages examined and the relatively few number of substitutions. Bailliet et al.'s new study shows that the evolutionary pattern of Amerinds in the south had not yet been fully resolved by sequencing of hypervariable domains. High resolution is, unfortunately, the key to understanding a

time scale, as well as to modeling the process of colonization.

First, Bailliet et al. eliminate, once and for all, the idea that a severe population genetic bottleneck took place in the process of continental colonization in the New World. They chose southern and geographically distinct representatives of the Amerind language phylum, according to the Greenberg classification. In focusing on three Argentine tribal populations (Mapuche, Huilliches, and Atacamenos) having minuscule to no evidence of genetic admixture with Europeans or African ethnic groups, they roughly doubled the number of new maternal genetic lineage clusters. These lineage groups are present in modern Asian populations and should have been present in the Americas, if modern Asians represent the logical source for most of the lineages that made up the gene pool of the first Americans. In fact, some of these lineages had been uncovered by previous workers but had been dismissed as either probably not authentic or enigmatic. Lineages that do not clearly fit the simple pattern (ABCD) are present in these indigenous peoples. Forced to confront the fact of differential lineage survival, we should have predicted as much from the dropout of the B lineage cluster in some North Americans and Siberians. However, why is the B cluster always lost?

A full 10% of the Mapuche would not be considered "authentic" by the criteria that some assert should be used to recognize founder lineages. These individuals do not fit the predicted pattern (ABCD) and yet are present in a tribal group that, by geography alone, could represent some of the earliest arriving Amerind colonists. Owing to both the nature of the mutations uncovered and the rate of mutation for the hypervariable region sequenced, it is unlikely that >30% of the lineages discovered in these three tribes arose in parallel in both the Old World and South America.

Bailliet et al. ask us to consider the effects that epidemic diseases, forced relocation and displacement that accompanied both indigenous and European expansion of dominant cultures, and active warfare would have had on the aboriginal American gene pool. Why shouldn't some groups be characterized by social structures that lead to high lineage extinction or explosive bursts of local expansion? Wouldn't nomadic hunters be expected to experience higher rates of transient lineage extinction than would settled agriculturists or local foragers? The notion that founding lineages will always be widespread and shared between tribes, will always be present in modern Asiatics, and will occupy a basal position in phylogenetic analysis may be considered unrealistic in light of present knowledge. In addition to this, what would be the consequences of a recent explosive burst of more closely related lineages with extensive gene flow?

Limited Resolution: Now Possible

Anatomically modern people might have successfully entered North America through Beringia anytime in the past 100,000 years, during the Wisconsin glacial (Wright 1991). If they came by land, the Alberta corridor, an ice-free route of dispersal between the Rocky Mountains and Hudson Bay, is estimated by Wright to have been open during the period of 55,000–18,000 ybp, to have closed for ~6,000 years during the late Wisconsin glacial, and to slowly reopen as it is now. This corridor to the south also roughly corresponds to a time when the Beringian land bridge existed (65,000–13,500 ybp).

On the basis of estimates of productivity and available resources, it has been suggested that the time when it was best to actually cross Beringia was 40,000–30,000 ybp (Aigner 1984). Winter ice before 65,000 ybp and after 13,500 ybp would also have allowed people to move freely, according to seasonal constraints. Migrants would still have to trek across Siberian and Alaskan Arctic deserts and/or pass through bogs, but it is clear that the real time period of potential dispersal is enormous, as was the home range that founding tribal populations might have occupied.

Archaeological evidence under such conditions should be rare, as populations densities would be expected to be very low. Once in North America, colonizing groups could go farther inland or follow the coast south. The coastal route is considered high risk, because dangerous glaciers were thought to block clear passage to the south. However, with our knowledge of both the antiquity of the use of water craft and the plentiful supply of marine resources, as well as of the emotional familiarity of aboriginal peoples from southern Asia with dangerous marine environments and their mammals, coastal passages should probably be reconsidered as a viable alternative. The coastal route is a key factor in asking whether Amerind populations cut off from this region by the ice corridor were effectively isolated from tribes to the south, as well as from the coast, for 6,000–10,000 years. Pacific populations using this route of entry into the Americas might well show a different suite of genotypes, such as those in the B cluster, if they were in migratory equilibrium with coastal foragers now displaced by modern Siberians.

Undisputed archaeological evidence of human occupation in the Americas is confined to the late Wisconsin glacial 23,000–10,000 ybp. Much of this evidence is <15,000 years old, and some of the oldest sites are found in South America, such as Monte Verde in south-central Chile and Pachamachay cave in highland Peru (Dillehay and Meltzer 1991). For every advocate of the north-Asian/late-Wisconsin stage/three-migration-wave model, it is possible to find a reputable researcher who supports, instead, a single sustained population pulse from a heterogeneous and predominantly north Asian gene pool, now largely incorporated into modern cultural units. Can we dissect a one-wave from the many possible wavelets and infer their arrival? If the level of migration between groups, postexpansion, is low (<10 females/generation), then simulations suggest that we can (Harpending et al. 1993).

Geneticists should now consider a provocative idea that was recently proposed in order to focus debate on the early/late colonization models (Beaton 1991) to help predict patterns associated with human use of an immense, uncharted landscape. In Beaton's view, one should contrast the world of transient explorers with that of estate settlers, who will have different demographic characteristics. The very slow or nearly stationary growth rate of high-mobility transient exploring groups should lead to rapid extinction of genetic lineages. In contrast, the high fecundity low-mobility settlers, still maintaining ties with their home groups and periodically exchanging mates, should show minimal lineage loss. These predictions will be familiar to those following the Rogers and Harpending (1992) pairwise-difference projections for population comparisons, as recently applied to both empirical and theoretical populations (Harpending et al. 1993).

The observation that many Amerinds, especially in the north, have lost the B lineage cluster as well as other unrelated clusters can have a number of explanations. First, as Torroni et al. (1994) imply, we could simply have a second migratory wave of settlers carrying the B cluster, which comes in after a major expansion of most Amerind groups. The time of the Amerind expansion could be early or late. The absence of B lineages in Beringia today says that, if this expansion were late, then the second expansion group was also recently displaced by a modern Siberian population, more closely related by lineage frequencies to the original founders. The fact that B clusters are present at high frequency in the south in a presumed ancestral group contradicts this simple model.

Second, North American Amerinds could be a second wavelet of the huge Amerind wave that began >20,000 ybp. A blockage of the Alberta corridor isolated Amerinds to the south, who continued to expand demographically, with minimal lineage extinction. Many lineages were lost in their northern cousins, who have secondarily expanded, after lineage loss, with the reopening of the corridor. This hypothesis accounts for some early sites in South America, higher genetic complexity there, and relative isolation between some Amerind tribal groups as judged by traditional population genetic parameters. A prediction of this model is that Amerinds should be clustered into linguistic subphyla by their retention of the mtDNA B and other lineage clusters (Alberta south), versus cluster ACD (Alberta north), and that these subphyla would now show tighter linguistic/genetic affinities. Lower diversity in the B cluster contradicts the prediction that B is an early colonizing cluster.

A third possibility includes the idea that the lineage B cluster is contributed either by a continual trickle of colonists using the coastal route or by direct contact across the Pacific ocean. Amerindian groups most likely to be in migration equilibrium with such a source of lineages are those closest to the tropical west coast. The B cluster of lineages have a coalescent of almost 30,000 years, so they might represent a separate source population isolated in the south of Asia. A coastal route in equilibrium along the entire Pacific Rim does not yet account for the geographic gradient that is seen in B lineage frequencies, which are highest always in the south. Pacific voyagers could have contributed this lineage separately to the Americas, without ever going through Beringia. A prediction of this model is that the B lineage cluster should be seen as intrusive archaeologically, confined to a time scale when we know that active voyaging was taking place in Remote Oceania. This time period corresponds to the spread of the Lapita cultural complex and has an antiquity of only 6,000 years (Kirch and Hunt 1988). Thus, the observation of an 8,000-year-old Paleoindian with a B lineage might seem to invalidate this model. However, B is a diverse lineage cluster, and, if it is retained ancestry versus intrusive, unique mutations found today in Remote Oceanic lineages should be missing in this sample when it is subjected to more extensive analysis.

Conclusion

The inadequacies of the archaeological record require us to face the facts that "traditional" evidence (i.e., a recurrent pattern of stratigraphic sequences based on radiocarbon dates in defined cultural settings) supporting the idea that the Americas were colonized >14,000 ybp is not strong. A simple, late expansion of three waves into the New World is dead, however. Active exploration and documentation of archaeological sites in the Americas is continuing, and at any moment the time depth for first occupation may change. Massive disease epidemics that followed cultural displacement were thought to have plagued aboriginal population genetic reconstructions, and they may not have been severe for Amerinds (Stone and Stoneking 1993). South American populations will have central importance in quantifying loss of diversity versus groups in the north, because of the opportunity to check for lineage extinction against skeletal remains. We will never be able to recover archaeological sites lost to coastal flooding, shopping-mall development, modern agricultural practices, or repatriation of stolen human remains. As limited as our opportunities may be, reconstructions of past human population diversity that are based on inferences of DNA sequence variability are the only independent way to scientifically approach the questions of Native American genealogical relationships.

References

Aigner JS (1984) The Asiatic-New World continuum in late Pleistocene times. In: Whyte RO (ed) Evolution of the East Asian environment. Center for Asian Studies, University of Hong Kong, Hong Kong, pp 915–937

Bailliet G, Rothammer F, Carnese FR, Bravi CM, Bianchi NO (1994) Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55:000–000

Beaton JM (1991) Colonizing continents: some problems from Australia and the Americas. In: Dillehay TD, Meltzer DJ (eds) The first Americans: search and research. CRC Press, Boca Raton, pp 209–230

Cavalli-Sforza LL, Piazza A, Menozzi P, Mountain J (1988) Reconstruction of human evolution: bringing together genetic, archaeological and linguistic data. Proc Natl Acad Sci USA 85: 6002–6006

Chakraborty R, Weiss KM (1991) Genetic variation of the mitochondrial DNA genome in American Indians is in mutation-drift equilibrium. Am J Phys Anthropol 86:497–506

Dillehay TD, Meltzer DJ (eds) (1991) The first Americans: search and research. CRC Press, Boca Raton

Flemming HC (1993) Toward a definitive classification of the world languages. Mother Tongue 20:4–30

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, et al (1993) Genetic variation among the Mapuche indians from the Patagonian region of Argentina: mitochondrial DNA sequence, nuclear variation, and allele frequencies of several nuclear genes. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffrey AJ (eds) DNA fingerprinting: state of the science. Birkhauser, Basel, pp 211–219

Greenberg JH (1987) Language in the Americas. Stanford University Press, Palo Alto, CA

Greenberg JH, Turner CG II, Zegura SL (1986) The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. Curr Anthropol 27:477–498

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. Curr Anthropol 34:483–496

Hertzberg M, Mickleson SW, Serjeantson SW, Prior JF, Trent RJ (1989) An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. Hum Genet 44:504–510

Horai S, Kondo R, Nakagawa-Hattorri Y, Hayashi S, Sonoda S, Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10:23–47

Kirch PV, Hunt TL (1988) Problems and issues in Lapita archaeology. In: Kirch PV, Hunt TL (eds) Archaeology and the Lapita cultural complex. Thomas Burke Memorial Washington State Museum research rep 5. Washington State Museum, Seattle, pp 157–165

Livshits G, Nei M (1990) Relationships between intrapopulational and interpopulational genetic diversity in man. Ann Hum Biol 17:501–513

Lum JK, Rickards O, Ching C, Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum Biol 66:567–590

Meltzer DJ (1993) The Pleistocene peopling of the Americas. Evol Anthropol 1:157–169

Morton N (1993) Genetic epidemiology. Annu Rev Genet 27: 523–538

Nei M, Roychoudhury AK (1993) Evolutionary relationships of human populations on a global scale. Mol Biol Evol 10:927–943

Renfew C (1991) Before Babel: speculations on the origins of linguistic diversity. Camb Arch J 1:3–23

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, et al (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613–623

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, Reed JK, Ward RH (1993) mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am J Hum Genet 53:549–562

Stone AC, Stoneking M (1993) Ancient DNA from a pre-Columbian Amerindian population. Am J Phys Anthropol 92:463–471

Szathmary EJE (1993*a*) Genetics of aboriginal North Americans. Evol Anthropol 1:202–220

——— (1993*b*) mtDNA and the peopling of the Americas. Am J Hum Genet 53:793–799

Torroni A, Neel J, Barrantes R, Schurr TG, Wallace DC (1994) Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci USA 91:1158–1162

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993*a*) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, Comuzzie AG, et al (1993*b*) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53:591–608

Wallace DC, Garrison K, Knowler WC (1985) Dramatic founder effects in Amerindian mitochondrial DNAs. Am J Phys Anthropol 68:149–155

Ward RH, Frazier BL, Dew-Jager K, Paabo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 88:8720–8724

Ward RH, Redd A, Valencia D, Frazier B, Paabo S (1993) Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 90:10663–10667

Wright HE (1991) Environmental conditions for Paleoindian immigration. In: Dillehay TD, Meltzer DJ (eds) The first Americans: search and research. CRC Press, Boca Raton, pp 113–13