

The human colonization of the Americas: population genetics

Simon G. Southerton

This chapter documents the genetic evidence for the colonization of the Americas, a topic already introduced archaeologically in chapter 8. On this occasion, the two chronologies (archaeological and genetic) appear to be in relatively good agreement.

Important new insights into the nature of human migration into the Americas have emerged from the study of genetic variation in Native American and Siberian populations. The Asian origin of essentially all Native Americans was confirmed by the middle of the 20th century, using classical genetic markers such as blood group and enzyme polymorphisms (O'Rourke & Raff 2010). During the last two decades, much higher resolution molecular studies have shed considerable light on the Asian source populations of these ancestral groups, the timing of migrations into the Americas, the number of colonizing populations, and the routes these people took as they entered North and South America (Schurr 2004). Discoveries from these studies are contributing to significant revisions in our understanding of Native American prehistory.

As the ancestors of Native Americans separated from neighboring Asian populations they gradually became genetically distinct. The earliest ancestral groups carried a subset of the genetic variation contained in the surrounding Asian populations. Mutations unique to Native Americans have also arisen in their populations since separation. Most research to date has focused on uniparentally inherited mitochondrial DNA or the non-recombining portion of the Y chromosome (NRY). By analyzing molecular variation in these maternal and paternal lineages, molecular geneticists have been able to detect signals of the patterns and timing of past migrations. Rapid advances in genomics technology are also now making it feasible to study the vast stores of genetic variation present in the autosomes.

Migration to Beringia

The earliest immigrants to the Americas brought with them a subset of the maternal and paternal DNA lineages present in their Asian source populations. Mitochondrial and Y-chromosome DNA variation in North, Central and South America indicate unambiguously that the ancestors of Native Americans originated in Asia. Virtually all modern Native Americans possess an mtDNA lineage that belongs to one of five founding haplogroups, which are all present among native populations of Siberia. These maternal lineages have been designated A, B, C, D, and X (Brown et al. 1998; Schurr et al. 1990). Of these haplogroups, only X is present in both central Asian and European populations; however, the X haplogroup is large and diverse, and the X lineage (X2a) found in Native American populations represents a distinct branch on the Eurasian X lineage tree (Reidla et al. 2003). A small proportion of mtDNA lineages found in indigenous peoples (<1%) are derived from recent non-native (European or African) admixture. Ancient mtDNA analysis has revealed the same founding haplogroups described in extant populations, thus confirming the genetic continuity between extinct and contemporary Native Americans (Stone & Stoneking 1998).

Much higher rates of male-mediated admixture into Native American populations (over 16%) since colonial times has made the analysis of Native American-specific Y chromosomes more complicated. The two most common founding NRY lineages within Native American populations are Q-M3 (also called Q1a3a) and C-3b (Karafet et al. 2008). In addition, recent studies of autosomal genetic variation have revealed a decline in genetic variation with distance from Bering Strait (Wang et al. 2007).

The paternal and maternal ancestors of Native Americans likely derived from central Asia, in the vicinity of southern Siberia and northern Mongolia. Asian populations that have the highest proportions of New World maternal and paternal lineages are generally located in the region immediately surrounding Lake Baikal, encompassing the Altai and Sayan Mountains (Karafet et al. 1999; Starikovskaya et al. 2005). Interestingly, these late Pleistocene migratory parties included domesticated dogs, as mtDNA sequences isolated from ancient dog remains from Latin America and Alaska are most closely related to the DNA lineages present in Old World dogs (Leonard et al. 2002).

Each of the maternal and paternal founding lineage groups found among Native Americans contains lineages with mutations that occurred after the original founders became separated from their Asian source population. Analysis of the degree of variation in each lineage has been used to calculate approximately when central Asian and American populations separated from each other. Recent estimates of coalescence times for each of the five founding New World mtDNA lineages are 25 to 20kya, while estimates based on NRY variability suggest that divergence occurred after 22.5kya, possibly as late as 20 to 15kya (Goebel et al. 2008), thus after the LGM, which peaked at around 20kya. The coalescence data have led some geneticists to conclude that all Native Americans descend from a single founding population that colonized northeast Asia, including Beringia, prior to the LGM (Fagundes et al. 2008), although archaeological evidence for the presence of humans in Beringia around the LGM is currently non-existent (see Goebel et al. 2010). The degree of mtDNA lineage diversity in contemporary Native Americans has also

been used to estimate the size of the original founding population. Although imprecise, between 100 and 1,000 females may have been in the founding groups from which all Native Americans descend (Fagundes et al. 2008).

Migration into North and South America

Geographic patterns in the distribution of the five founding mtDNA lineages in the Americas have been used to shed light on the number of founder migrations and the possible routes these migratory groups took. Haplogroups A–D are observed in extant Amerindian populations from North, Central, and South America (Schurr 2004), which strongly suggests that all four of these mtDNA lineages were present in the original migration(s) to the New World. There are different trends within the three Native American linguistic groups (Amerind, Na-Dené, Eskimo-Aleut) proposed by Greenberg and his colleagues (1986). Na-Dené and Eskimo-Aleut speaking populations have predominantly haplogroup A and D mtDNAs, and their original founders may have lacked haplogroups B and C. Consequently, they are thought to have resulted from a later expansion into North America than that giving rise to the original Amerindians. It was recently shown that a subclade of D2 lineages present in Chukchi, Siberian Inuit, and Aleuts (D2b) has a coalescence date of 8 to 6 kya (Derenko et al. 2007). This implies that the ancestors of modern Eskimo-Aleuts spread from Siberia into the Americas in the middle-Holocene (see chapters 44 and 45), which is in harmony with earlier explanations based on dental evidence (Greenberg et al. 1986).

Analyses of mtDNA subclades within haplogroups has been particularly useful for exploring the nature of Native American migrations from Beringia into the remainder of the Americas. Three subclades of mtDNA subhaplogroup C1 are widely distributed among North, Central, and South American Indians but absent in Asian populations (Tamm et al. 2007). This suggests that these subclades evolved after the central Asian–Native American split, as the first Americans were dispersing from Beringia. The coalescence estimates for them is 16.6–11.2 kya, which suggests that the colonization of the Americas south of the continental ice sheets may have occurred after the LGM. A date of 16–17 kya, as suggested by the genetic evidence, is in agreement with recent archaeological discoveries (e.g. from Monte Verde in Chile and Meadowcroft in Pennsylvania) that predate Clovis lithic sites in North America.

Until recently, it was widely accepted that people belonging to the Clovis culture were the first to enter the Americas, around 13 kya and thus during the period of postglacial climatic amelioration. They were believed to have entered North America through Alaska via an interior ice-free corridor, and then rapidly expanded into the remainder of the Americas. However, the distribution of molecular genetic variation observed in contemporary Native American populations is not consistent with such a late entry or rapid dispersal of humans across the whole of the New World, as the Clovis-First model would require (Fagundes et al. 2008; Perego et al. 2009). A recent statistical analysis of nuclear DNA data was also most consistent with a pre-Clovis settlement model (Ray et al. 2010). Meanwhile, human coprolites recovered from Paisley Cave, Oregon, have been found to predate the Clovis complex by more than 1000 ¹⁴C years (Gilbert et al 2008).

Revised migration models have recently been proposed to account for the accumulating molecular data. Schurr and Sherry (2004) proposed that there were two migrations from Siberia to South America around 20–15 kya. The first migration was along the Pacific coast, bringing lineages from haplogroups A, B, C, and D. This was followed by a second migration (containing haplogroup X) into North America after the ice-free corridor appeared. This model is supported by the geographic distribution of two rare mtDNA lineages: the D4h3 lineage which appears at highest frequency along the Pacific coast of North and South America, and the X2a lineage which is restricted to northeastern North America (Perego et al. 2009). O'Rourke and Raff (2010) recently proposed that pre-LGM coastal migration may have occurred via the northern Beringian coastline. Such a migration scenario along the northern and eastern seaboard of North America may have provided source populations for the development of the Clovis culture in the region where the earliest appearance and highest density of Clovis artifacts occurs. However, there is no archaeological support as yet for such an early entry, especially along the Arctic seaboard.

The initial colonization of South America appears to have taken place along the western coastline (Perego et al. 2009; Rothhammer & Dillehay 2009). Once south of the developing ice masses, coastal populations could have moved more rapidly than migratory groups located in the interior of the continents, reaching the southern regions of South America within a relatively short time period. The highest levels of genetic diversity are observed in western populations, and these are considerably lower in Brazilian populations. Wang et al (2007) also found high genetic similarity between Andean and Mesoamerican populations. This was interpreted as being consistent with an early coastal colonization along the west coast of South America, and a subsequent peopling of the eastern coastline.

Comparisons between molecular data from past and present Amerindian populations have revealed regional and temporal patterns of population settlement and movement within the Americas. In general, ancient DNA studies indicate that once Amerindian populations settled a particular region, they tended to become and remain genetically distinct over thousands of years (O'Rourke et al. 2000). Continuity has been observed between Anasazi and modern Puebloan groups, ancient Tainos and modern Puerto Ricans, and ancient and modern populations in southern Chile and Patagonia (reviewed by Schurr 2004). However, evidence of regional migration has also been detected. For example, clear differences in haplogroup frequencies between ancient and historical populations from the Great Basin are thought to be due to the arrival of Numic peoples in that area (Kaestle & Smith 2001). Ancient Hopewell and Adena remains from the Ohio Valley have genetic ties with contemporary Great Lakes populations (see Schurr 2004).

Controversial migration theories

Molecular genetics discoveries are contributing important new data to often heated debates surrounding less widely supported hypotheses of the settlement of the Americas. The restricted distribution of the X lineage in eastern North America was initially seen to support a "Solutrean hypothesis," that the ancestors of the Clovis people were

derived from an Upper Paleolithic population that migrated from Iberia (Bradley & Stanford 2004). According to this model, Solutrean people migrated from Western Europe during the LGM across an ice sheet skirting the Atlantic Ocean until they reached North America, where their descendants developed the Clovis lithic technology. However, the sum of the molecular evidence lends little support to this hypothesis. The coalescence ages of all five founding maternal lineages are very similar, suggesting they were all present in the Beringian founding population from which the Americas were colonized between 20 and 15kya (Fagundes et al. 2008; Perego et al. 2009). The failure to identify a closely related X lineage in Siberian populations is most likely due to the low frequency of this haplogroup in Asia.

There is currently little evidence that Native Americans migrated beyond the Americas into the Pacific or that Polynesians settled in large numbers in the Americas. A distinctive “Polynesian lineage” belonging to the mtDNA B haplogroup, which is shared by almost all extant Polynesians, has not been detected among Native American populations. There is currently no genetic evidence that peoples from Melanesia, Polynesia, Australia, Africa, Europe, China, or the Middle East contributed significantly to the pre-historic Native American gene pool. The molecular genetic data thus offer little support for settlement theories at the fringe of mainstream anthropology and archaeology. Interestingly, in response to molecular research, the Mormon Church (Church of Jesus Christ of Latter-day Saints) recently changed its belief that Israelites were the “principal ancestors” of Native Americans to a still overly hopeful qualification that they were “among the ancestors of the American Indians” (Moore 2007). However, the question of whether there could have been small admixtures from other parts of the world is frequently raised by journalists, maverick anthropologists, and revisionist historians.

Summary and outlook

Evidence from molecular research using mtDNA, NRY, and autosomal genetic systems suggests ancestral Native Americans entered the Americas after the LGM, between 20 and 15kya. Their entry predated the opening of the ice-free corridor and the appearance of the Clovis culture, and colonization was most likely to have occurred via a coastal route during their initial migration. The data suggest the ancestors of modern circumarctic populations, such as the Inuit, Aleuts, and Na-Dené entered the Americas during the middle Holocene. Further refinements of migration scenarios during the colonization of the Americas are likely to emerge as modern archaeological and genetics tools are used to address common questions. Research in more widely dispersed ancient populations will increase the accuracy of coalescence estimates.

SEE ALSO: 8 The human colonization of the Americas: archaeology; 41 Polynesia, East and South, including transpacific migration; 44 North America: Eskimo-Aleut linguistic history; 45 North America: Paleoeskimo and Inuit archaeology

References

- Bradley, B. & Stanford, D. (2004) The North Atlantic ice-edge corridor: a possible Palaeolithic route to the New World. *World Archaeology* 36, 459–478.
- Brown, M. D., Hosseini, S. H., Torroni, A., et al. (1998) MtDNA haplogroup X: An ancient link between Europe western Asia and North America? *American Journal of Human Genetics* 63, 1852–1861.
- Derenko, M., Malyarchuk, B., Grzybowski, T., et al. (2007) Phylogeographic analysis of mitochondrial DNA in Northern Asian Populations. *American Journal of Human Genetics* 81, 1025–1041.
- Fagundes, N. J. R., Kanitz, R., Eckert, A., et al. (2008) Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *American Journal of Human Genetics* 82, 583–592.
- Gilbert, M. T. P., Jenkins, D. L., Götherstrom, A., et al. (2008) DNA from pre-Clovis human coprolites in Oregon, North America. *Science* 320, 786–789.
- Goebel, T., Waters, M. R., & O'Rourke, D. H. (2008) The Late Pleistocene dispersal of modern humans in the Americas. *Science* 319, 1497–1502.
- Goebel, T., Slobodin, S. B., & Waters, M. R. (2010) New dates from Ushki-1, Kamchatka, confirm 13,000calBP age for earliest Paleolithic occupation. *Journal of Archaeological Science* 37, 2640–2649.
- Greenberg, J. H., Turner, C. G., & Zegura, S. L. (1986) The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Current Anthropology* 27, 477–497.
- Kaestle, F. A. & Smith, D. G. (2001) Ancient mitochondrial DNA evidence for prehistoric population movement: the Numic expansion. *American Journal of Physical Anthropology* 115(1), 1–12.
- Karafet, T. M., Zegura, S. L., Posukh, O., et al. (1999) Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. *American Journal of Human Genetics* 64, 817–831.
- Karafet, T. M., Zegura, S. L., & Hammer, M. F. (2008) Y Chromosomes. In D. Ubelaker (ed.), *Handbook of North American Indians*, vol. 3, *Environment, Origins, and Population*. Washington, DC: Smithsonian.
- Leonard, J. A., Wayne, R. K., Wheeler, J., et al. (2002) Ancient DNA evidence for Old World origin of New World dogs. *Science* 298, 1613–1616.
- Moore, C. A. (2007) Intro change in Book of Mormon spurs discussion. Deseret News, Salt Lake City. (<http://www.deseretnews.com/article/695226049/Intro-change-in-Book-of-Mormon-spurs-discussion.html>) accessed Jan 22, 2010.
- O'Rourke, D. H. & Raff, J. A. (2010) The human genetic history of the Americas: the final frontier. *Current Biology* 20, R202–R207.
- O'Rourke, D. H., Hayes, M. G., & Carlyle, S. W. (2000) Ancient DNA studies in physical anthropology. *Annual Review of Anthropology* 29, 217–242.
- Perego, U. A., Achilli, A., Angerhofer, N., et al. (2009) Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Current Biology* 19, 1–8.
- Ray, N., Wegmann, D., Fagundes, N. J. R., et al. (2010) A statistical evaluation of models for the initial settlement of the American continent emphasizes the importance of gene flow with Asia. *Molecular Biology and Evolution* 27, 337–345.
- Reidla, M., Kivisild, T., Metspalu, E., et al. (2003) Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics* 73, 1178–1190.
- Rothhammer, F. & Dillehay, T. D. (2009) The Late Pleistocene colonization of South America: an interdisciplinary perspective. *Annals of Human Genetics* 73, 540–549.

- Schurr, T. G. (2004) The peopling of the New World: perspectives from molecular anthropology. *Annual Review of Anthropology* 33, 551–583.
- Schurr, T. G., Ballinger, S. W., Gan, Y. Y., et al. (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they are derived from 4 primary maternal lineages. *American Journal of Human Genetics* 46, 613–623.
- Schurr, T. G. & Sherry, S. T. (2004) Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. *American Journal of Human Biology* 16, 420–439.
- Starikovskaya, E. B., Sukernik, R. I., Derbeneva, O. A., et al. (2005) Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. *Annals of Human Genetics* 69, 67–89.
- Stone, A. C. & Stoneking, M. (1998) MtDNA analysis of a prehistoric Oneota population: Implications for the peopling of the New World. *American Journal of Human Genetics* 62, 1153–1170.
- Tamm, E., Kivisild, T., Reidla, M., et al. (2007) Beringian standstill and spread of Native American founders. *Public Library of Science (PLoS) ONE* 2(9); online e829. doi:10.1371/journal.pone.0000829.
- Wang, S., Lewis, C. M., Jakobsson, M., et al. (2007) Genetic variation and population structure in Native Americans. *Public Library of Science (PLoS) Genetics* 3, 2049–2067.