



A genomic view of the peopling of the Americas

Pontus Skoglund^{1,2,3} and David Reich^{1,2,4}

Whole-genome studies have documented that most Native American ancestry stems from a single population that diversified within the continent more than twelve thousand years ago. However, this shared ancestry hides a more complex history whereby at least four distinct streams of Eurasian migration have contributed to present-day and prehistoric Native American populations. Whole genome studies enhanced by technological breakthroughs in ancient DNA now provide evidence of a sequence of events involving initial migrations from a structured Northeast Asian source population with differential relatedness to present-day Australasian populations, followed by a divergence into northern and southern Native American lineages. During the Holocene, new migrations from Asia introduced the Saqqaq/Dorset Paleo Eskimo population to the North American Arctic ~4500 years ago, ancestry that is potentially connected with ancestry found in Athabaskan-speakers today. This was then followed by a major new population turnover in the high Arctic involving Thule-related peoples who are the ancestors of present-day Inuit. We highlight several open questions that could be addressed through future genomic research.

Addresses

¹ Harvard Medical School, Boston, MA, USA

² Broad Institute of MIT and Harvard, Cambridge, MA, USA

³ Department of Archaeology and Classical History, Stockholm, Sweden

⁴ Howard Hughes Medical Institute, Boston, MA, USA

Corresponding author: Skoglund, Pontus (pontus.skoglund@gmail.com)

Current Opinion in Genetics & Development 2016, 41:27–35

This review comes from a themed issue on **Genetics of human origin**

Edited by **Joshua Akey** and **Anna Di Rienzo**

<http://dx.doi.org/10.1016/j.gde.2016.06.016>

0959-4377/© 2016 Elsevier Ltd. All rights reserved.

Introduction

The peopling of the Americas represented the culmination of a Late Pleistocene expansion of anatomically modern humans out of Africa. Archaeological evidence indicates that groups subsisting on hunting lived in extreme northeast Siberia (71° N) by at least 28,000 years ago [1]. Human groups adapted to the mammoth steppe habitat were thus poised to enter Beringia — the land-mass between Alaska and Eurasia that is now submerged — by this time [2,3]. The path from Beringia

to the more temperate parts of the American continents, however, was blocked by the merged Laurentide and Cordilleran ice sheets that covered northern North America until the end of the Last Glacial Maximum. The ice retreated from parts of the Pacific coast ~16,000 years ago, raising the possibility of a coastal migration after this time, and within a few thousand years a habitable corridor through the center of the continent opened between the two ice sheets [4]. The first unambiguous evidence of modern humans in the Americas dates to between 14,000 and 15,000 years ago [5–8], and was likely the consequence of migration from Beringia.

Major debates about the peopling of the Americas have focused on the question of whether the first early human populations in the Americas are directly ancestral to present-day Native Americans, as well as on the number, mode, and timing of secondary migrations to the Americas. Advances in genomics have, within the last five years, enabled the collection of far more data from present-day Native Americans than were available previously, and have also made it possible for the first time to access DNA from ancient Native American remains. Analysis of these data has highlighted a minimum of four distinct streams of Asian ancestry, some of which were not clear from the archaeological evidence. We review the contributions of genomic data to understanding the prehistory of the Americas, and outline outstanding questions where it may be able to provide additional insight.

The power of the whole genome

The first meaningful genetic insights about Native American population history came from mitochondrial DNA, a segment of about 16,500 base pairs (approximately 1/200,000th of the genome) that is passed exclusively along the maternal line. Mitochondrial DNA was one of the first parts of the genome to be heavily investigated to learn about human population history for several reasons. First, it is highly variable on a per-nucleotide level and thus sequencing only a short stretch can detect non-trivial amounts of human variation. Mitochondrial DNA's high variability and short length also meant that it was practical to sequence in large numbers of samples at a time when it was prohibitively expensive to generate genome scale data. For ancient DNA studies, mitochondrial DNA had the further advantage that it exists in about a thousand-fold higher copy number than any other single place in the genome. Since one of the main challenges of ancient DNA is obtaining sufficient amount of material from any position in the genome to be able to analyze, beginning with more material can be an advantage [9].

The greatest contribution of mitochondrial DNA studies to the understanding of Native American prehistory has been in the area of reconstruction of population size history. Mitochondrial DNA analyses were the first to document that the ancestry of most Native Americans derives from a population that experienced a profound founder event [10], with a relatively small number of individuals giving rise to a large number of descendants today. The evidence for this is that all Native American mitochondrial DNA lineages today descend from just five founding maternal lines [11–13] that each had a common ancestor around 18,000 to 15,000 years ago, implying a population size bottleneck around this time [14–18]. The evidence for a profound population bottleneck has since been confirmed and its intensity measured more accurately with genome scale data [19–22,23**], but it is important to note that there are still challenges with disentangling the number of founder individuals from the duration of the population size reduction using all the reported methods.

A second finding about Native American population history based on mitochondrial DNA data is that the founder event may have been preceded by an extended period (many thousands of years) of little or no shared ancestry with non-Native American mitochondrial DNA lineages. This suggested to some researchers the hypothesis of a ‘Beringian standstill’, whereby the first founding population of the Americas was isolated from Eurasian populations before its radiation into a multitude of sub-populations in America [15].

At the same time, some observations from mitochondrial DNA studies of the Americas have been deceptive. For example, the mitochondrial DNA subtype called D4h3a is today almost entirely restricted to Pacific coastal populations, both in North and South America. This observation led to the hypothesis that D4h3a was carried south of the ice sheets along a coastal route, in a migratory movement that was distinctive from the one that led to many other Native American populations [24]. However, ancient DNA studies have since found the same mitochondrial DNA type in a ~12,600 year old individual from present-day Montana, which based on its genome-wide data is unambiguously from the main ancestral lineage leading to most Native Americans [25**].

It is now clear that so many founder events and fluctuations in population size have occurred before, during, and after the peopling of the Americas that the evidence from one position in the genome — mitochondrial DNA, the Y chromosome, or any other location — is too subject to random changes in frequency (genetic drift) to provide a complete picture by itself. Only by taking the independent testimony of many locations in the genome simultaneously can we obtain a high-resolution picture of the deep past. The remainder of this article focuses on

insights from whole genome studies of Native American population history. While these studies are still in their early days, they have already upended our understanding of key events. Application of ancient DNA technology promises further insights in years to come.

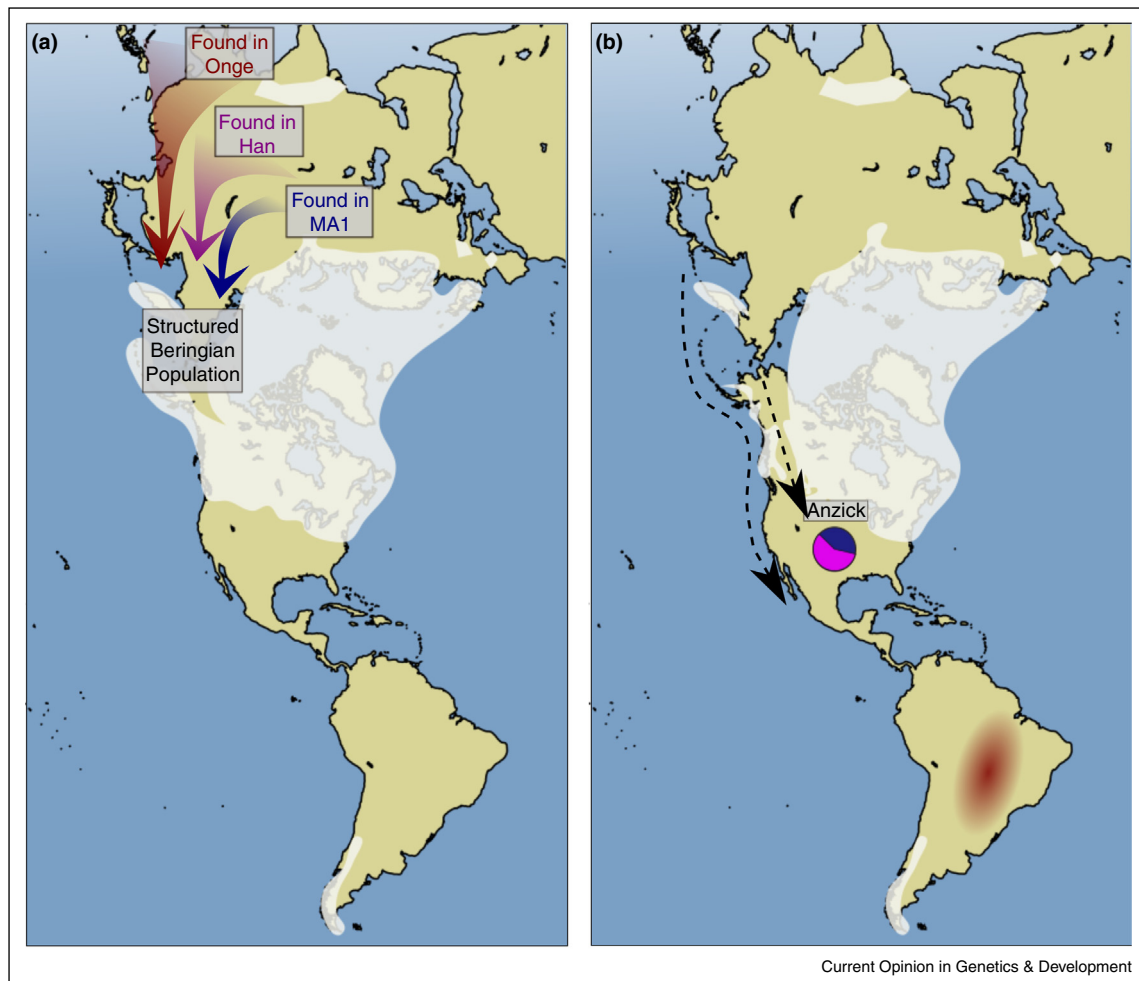
Sources of Native American ancestry

Under the hypothesis that Native American ancestry stems from a single founder population that separated earlier from Eurasian populations, differences in allele frequencies between Native American groups should have developed independently from Eurasian allele frequencies. This simple null hypothesis makes it possible to explicitly test hypotheses about the number of American founder populations. Reich *et al.* [23**] applied this idea to the first comprehensive genome-wide data from Native American populations (52 populations, but none from the continental United States), and concluded that at least three ancestral populations were required to explain the similarities between Native Americans and East Asians. According to the initial study [23**], all Native American groups from Central and South America fit a model of a single founder population. An additional source of ancestry was necessary to account for genetic variation in Eskimo-Aleut speakers. In addition, analysis of the Athabaskan-speaking Chipewyan revealed that they could not solely have their ancestry from the same founding population as other Northern-, Meso- and South American populations.

The main ancestral stream giving rise to Native American ancestry

One of the most important pieces of genetic evidence relevant to the peopling of the Americas was the sequencing of a genome from the remains of a child (‘Anzick-1’) buried with Clovis artifacts in western Montana and directly dated to 12,600 before present (BP) [25**]. This child was consistent with deriving all of his ancestry from the same founding population as Central and South Americans (Figure 1), contradicting the ‘Solutrean hypothesis’ of transatlantic migration from Upper Paleolithic Europe giving rise to the Clovis industry [26] which thus posits genetic discontinuity between the makers of the Clovis industry and present-day Native Americans [27]. The most surprising finding was that the Anzick individual is from a population more closely related to Central Americans and South Americans than to some northern North Americans (including all speakers of Algonquian languages studied to date), despite the apparent common ancestral origin of Native Americans across the continents. This suggests that the present-day structuring of the main ancestry in Native Americans into northern and southern lineages [23**] dates back to more than 12,600 years ago [25**]. This event divided the ancestry of present-day Native Americans into two main streams, one of which includes the ancestors of present-day Northern Native Americans

Figure 1



Eurasian source populations of the earliest Native Americans. **(a)** Diverse population lineages that genetic evidence suggest were present in Northeast Asia and contributed to the ancestry of present-day Native Americans. These comprise people with ancestry related to present-day Andamanese and Australo-Melanesians (red), people related to present-day East Asians such as Han Chinese (pink), and people related to the 24,000 year old Mal'ta child buried near Lake Baikal (dark blue). By 12,600 years ago, at least the Mal'ta-related and east Asian-related ancestries were already present in the proportions found in Native Americans today in the Anzick child from Western Montana **(b)**. Today, the Andamanese-related ancestry can be detected as an excess affinity to present-day Amazonians **(b)**.

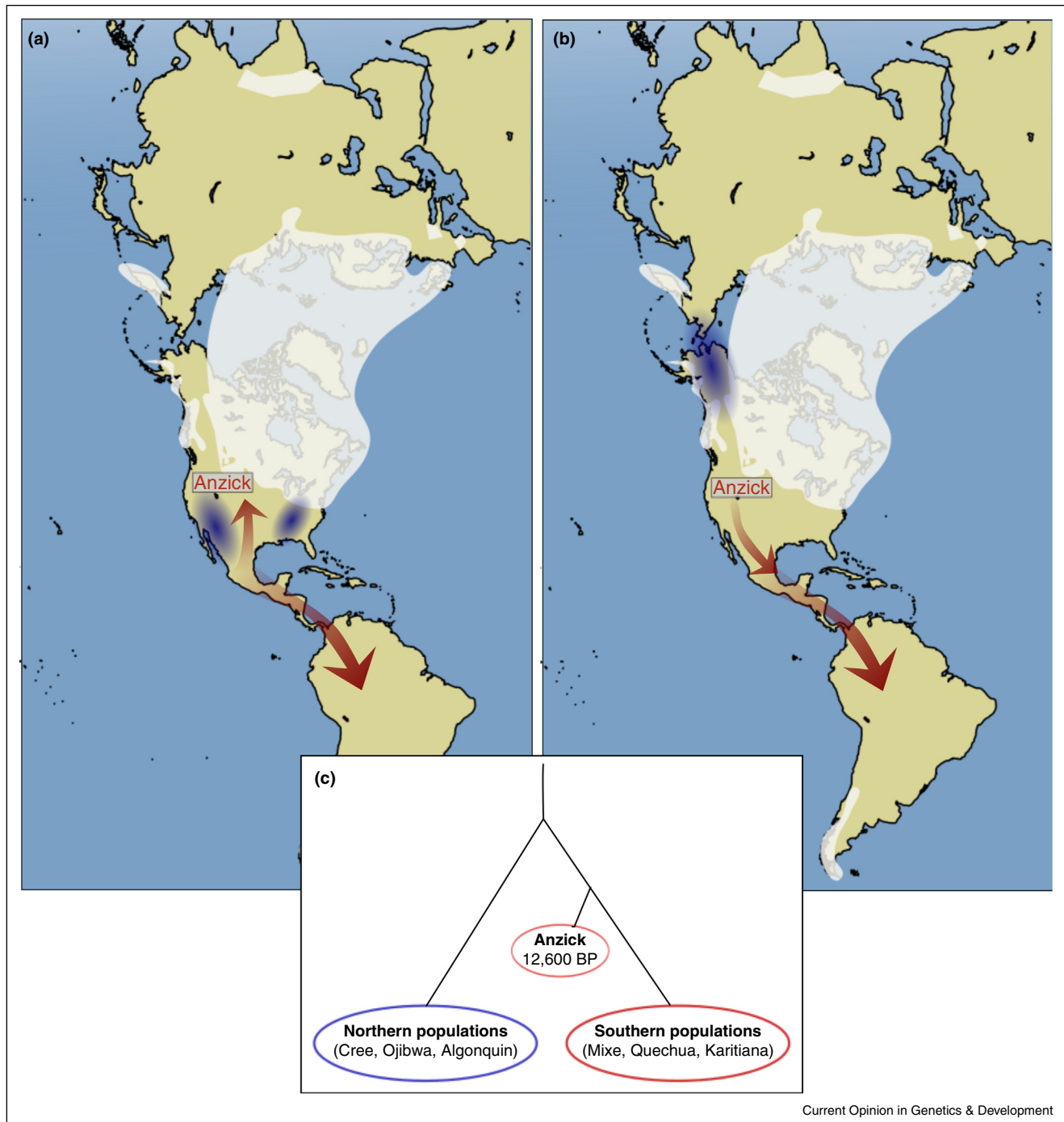
analyzed ('NNA': including Cree, Ojibwa, and Algonquin), and the other of which includes the Anzick individual and present-day Central American groups and South American groups ('SA': including Mixe, Quechua, and Yaghan).

While we thus have evidence for SA ancestry in the Late Pleistocene in the form of the Anzick genome, an outstanding question is where the ancestors of NNA were localized. One possibility is that they were confined to ice-free regions north of the corridor between the Cordilleran and Laurentide ice sheets, and only expanded south after 12,600 BP, where this ancestry would eventually displace the SA populations represented by Anzick (Figure 2). Another possibility might be an expansion of

the Clovis industry from a southern origin that also represented a population expansion by SA populations into regions where NNA populations were located (Figure 2). Thus SA populations might have brought Clovis technology to some regions in northern North America only to later be displaced by an NNA resurgence.

More recent genome sequencing [28] of the ~9000 year-old Kennewick individual did not provide resolution about this issue. While the genome analysis convincingly showed that the Kennewick individual had the great majority of its ancestry from the same deep source as other present-day Native Americans, his affinity to the SA and NNA lineages remains ambiguous.

Figure 2



Two scenarios explaining why the 12,600 year old Anzick individual is more closely related to South and Central-Americans than Northern Native Americans. **(a)** Clovis technology appears first in the southern parts of North America. It is thus possible that Anzick, being associated with Clovis artifacts and dating to the end of the Clovis period, represents a northwards expansion into a region where the ancestors of present-day northern Native Americans were already established. **(b)** An alternative is that the southern lineage associated with the Anzick individual represents the first migration south of the ice sheets, whereas the ancestors of present-day Northern Native Americans resided north of the ice sheets at 12,600 years ago and only later migrated southwards, replacing the Anzick-like population. **(c)** Both these scenarios could explain the observation that Anzick is more closely related to Southern Native American populations.

Major admixture during the formation of the ancestral population of Native Americans

Until 2014, efforts to infer demographic parameters for the peopling of the New World based on genetic data focused on modeling Native Americans as an offshoot of East Asian ancestry [17,20,23^{••},29–33,34^{*}]. However, the analysis of the genome of a 24,000 year-old individual from the Mal'ta site near Lake Baikal in Central Siberia revealed that this model is untenable. The Mal'ta individual shared genetic affinities to both European (West Eurasian) and Native American populations [35^{••}]. Analyses showed that a scenario in which Mal'ta descended from an admixture between a West Eurasian population and an ancient population that was also ancestral to Native Americans could not explain all features of the data [35^{••}]. However, the data was consistent with a model in which Native Americans are admixed between lineages related to Mal'ta on the one hand (between ¼ and ½ of Native American ancestry) and East Asians on the other hand (Figure 1). Thus, Native Americans and East Asians do not in fact descend from a common ancestral population that separated earlier from a lineage leading to Mal'ta and to West Eurasians [35^{••}].

The finding of ancient mixture in the ancestry of Native American — before diversification within the Americas — also has consequences for modeling other features of Native American population history. Genetic debate about the date of the first migrants into the Americas has had, as one of its important themes, estimation of the date of genetic divergence of the lineages giving rise to Native American and East Asian populations [31]. The major admixture related to the Mal'ta lineage in the ancestry of all Native Americans is inconsistent with the assumption of a simple population split between Native Americans and East Asians that has been the basis for the most attempts to date to infer the population split times, and which have suggested dates of around 23,000 years ago [33,34^{*}]. The substantial contribution of the Mal'ta lineage to Native Americans may have the effect of upwardly biasing estimates of the time of divergence of Native Americans and East Asians. Future models estimating parameters for the founding populations of the Americas will need to consider this admixture explicitly.

An Australasian connection

Recently, we carried out a stringent test of the null hypothesis of a single founding population of Central and South Americans using genome-wide data from diverse Native Americans [36^{••}]. We detected a statistically clear signal linking Native Americans in the Amazonian region of Brazil to present-day Australo-Melanesians and Andaman Islanders ('Australasians'). Specifically, we found that Australasians share significantly more genetic variants with some Amazonian populations — including ones speaking Tupi languages — than they do with other Native Americans. We called this putative ancient Native

American lineage 'Population Y' after *Ypykuéra*, which means 'ancestor' in the Tupi language family.

To learn more about the Population Y ancestry present in the Americas, we carried out a series of statistical modeling analyses. We found that the patterns of genomic variation of present-day Amazonians could be explained by as little as 2% admixture from an Australasian-related population, that would thus have penetrated deep inside the Americas without mixing with the main ancestral lineage of present-day Native Americans. Alternatively, the patterns could be explained by a larger proportion of ancestry (2–85%) from a population that existed in a substructured Northeast Asia, and was similar to the main lineage that gave rise to other Native Americans while retaining more Australasian affinity.

We considered the possibility that these genetic patterns relate to the claims based on skeletal morphology that there was a wave of migration related to Australasians that contributed to early Native Americans, followed by later large-scale population replacement by the primary ancestors of present-day Native Americans [37,38]. While this evidence has been contested on morphological grounds — with the most important critique being that the analyses are not statistically compelling [2,39] — the fact that the morphological evidence is strongest in Brazil where Population Y ancestry is prevalent justifies investigation into possible connections. However, new genetic findings reject one set of the arguments based on morphology: that the morphology of groups such as historic-period Native Americans in the Baja California region of Mexico and the Tierra del Fuego region in the southern tip of South America suggests that they are related to the hypothesized earlier Australasian-related migration. Genomic DNA from these populations do not have any evidence of affinity to Australasian populations [34^{*}].

What is the history behind the Population Y ancestry in Amazonian Native Americans today? Little can be said at present, as we do not have ancient DNA from individuals carrying detectable amounts of this ancestry. The implication that the Siberian populations that gave rise to the ancestors of Native Americans was substructured, however, is not surprising given that 24,000–18,000 year-old individuals from the Lake Baikal region of Siberia (the Mal'ta and Afontova Gora sites) were genetically very divergent from present-day East Asians [35^{••},40]. The Population Y results suggest that a population that has not yet been sampled with ancient DNA data — one with more Australasian-related ancestry — may also have been present in the broad geographic area in order to contribute to the founders of Native American founders. Notably, Andaman Islanders, the population with the single strongest affinity to Amazonians, are not as good match for the non-Mal'ta like ancestry in Central Americans as are Chinese populations [36^{••}]. These strands of evidence suggest a minimum

three-part ancestry for the Beringian populations that came to populate the Americas (Figure 1). Two of these strands were fully braided together to form the main ancestral lineage of Native Americans by the time of the Beringian bottleneck. However, the third strand, with an affinity to Australasians, was not.

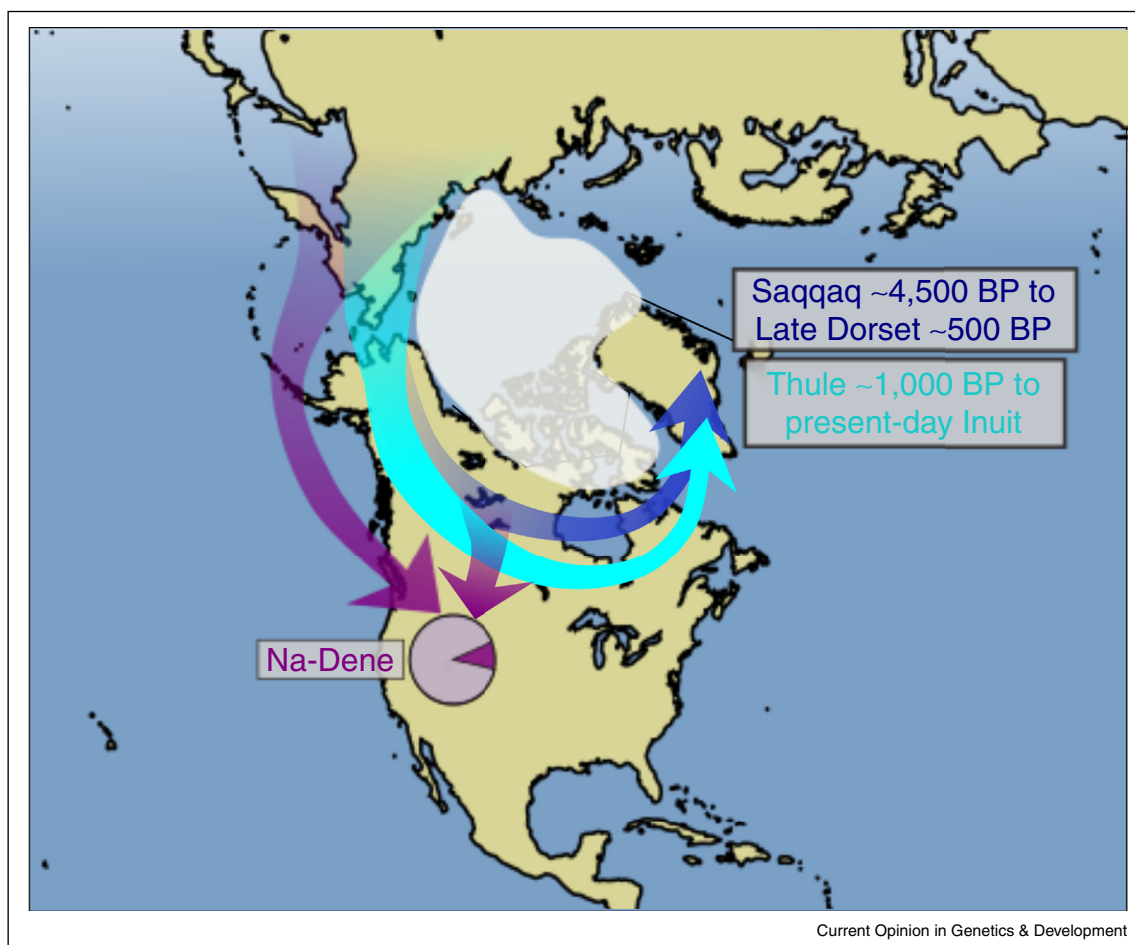
It has been suggested that Native American ancestors may have entered the more temperate parts of North America both by an early coastal route and the later ice free corridor [2]. One possibility is that the different sources of deep ancestry that are inferred by these genetic patterns may reflect movements of a substructured Beringian population through these different routes. Alternatively, the patterns could reflect pulses of migration from Beringia occurring at different times (e.g. multiple pulses through the ice free corridor).

Restricting study to South America, a related question is the explanation for the deeply structured population lineages East and West of the Andes in South America that have been documented based on genomic data [23**]. Population Y ancestry may be limited entirely to the eastern populations, raising the possibility that this split was extremely ancient.

Post-glacial migrations into North America

The archaeological record of the North American Arctic documents a transformation beginning around a thousand years ago and associated with the Thule cultural complex. The Thule culture advanced into a North American Arctic occupied by the descendants of previous migrations (Figure 3), referred to as Paleoeskimos and culturally represented by the Saqqaq and Dorset technological complexes [41]. Genome-wide data from a ~4000 year old

Figure 3



Holocene migrations to the Americas. There is a current consensus supporting a Paleoeskimo migration to Greenland ~4500 years ago, giving rise to both the Saqqaq culture and the Early, Middle and Late Dorset cultures. This Paleoeskimo population was displaced by Neo-Eskimos associated with the Thule culture migrating through the Arctic ~1000 years ago, although there remains a possibility of some admixture. Two scenarios can plausibly explain the Asian ancestry in present-day speakers of Athabaskan (Na-Dene) languages. (1) The Na-Dene have ancestry from a distinct migration to the Americas from that seen in any other ancient samples. (2) The Na-Dene-specific ancestry comes from the same migration as the Paleoeskimos [23**]. Note that these arrows are not intended to show exact geographic migration routes.

Saqqaq culture individual in Greenland — the first genome-wide dataset from an ancient human — was interpreted as showing that the Saqqaq population went extinct in North America, since the Saqqaq genome was most similar to present-day Koryak and Chukchi on the eastern side of the Bering strait in clustering analyses [42^{*}]. Genome-wide analysis of additional samples from the Early Dorset and Middle Dorset cultures — along with Thule individuals — confirmed the picture of distinct ancestry in Paleoescimos, and a population break associated with the Thule that led directly to present-day Inuit populations [43^{*}].

A major outstanding issue in the interpretation of the peopling of the Arctic relates to the geographic and temporal origins of the distinct Asian ancestry present in speakers of Athabaskan (Na-Dene) languages [23^{**},43^{*},44]. Two distinct models for Paleoescimo and Athabaskan-speaking ancestry have been proposed based on the genetic data. Raghavan *et al.* [43^{*}] proposed that Athabaskan speakers share all of their ancestry with Native Americans in Central and South America. However, Reich *et al.* [23^{**}] showed that ancestry in Athabaskan speakers cannot be modeled as a mixture of the lineages that have given rise to other Native American and Inuit population living today. Instead, a model in which Athabaskan-speakers harbor about 10% of their ancestry from a source that is distinct from both the Inuit and the main ancestry in other Native Americans — possibly related to the Saqqaq Paleoescimo — may be consistent with the data [23^{**}]. A caveat is that several of these inferences may be complicated if Siberian populations today have substantial amounts of ancestry from back-migrations out of America. An important direction for future research is to determine the origin of the distinct Asian ancestry present in Athabaskans using higher resolution methods and additional data.

Prospects

Far more remains to learn about the genomic history of Native Americans than has already been discovered. Major directions for future data collection include more ancient DNA data, and filling in sampling gaps of present-day Native American populations, especially in the United States. We conclude by highlighting five major open questions about Native American population history that we believe may be meaningfully addressed using genomic data in the coming years.

- (1) What is the origin of Population Y ancestry in Amazonia — can it be found in ancient DNA?
- (2) What is the origin of the lineage found in many present-day northern North Americans but not in the 12,600 year old Anzick genome?
- (3) What is the history of peopling and migration in South America? Was there an early population split east- and

west of the Andes, and how much major migration occurred after the first people arrived?

- (4) What is the genetic legacy of the Paleoescimos in the North American Arctic? Can this widespread population have disappeared completely after the arrival of the Thule and Aleuts or have they left some descendants, perhaps in admixed form in Athabaskan speakers?
- (5) What was the structure of Native American populations in North America before the disruption of the last 500 years?

We conclude by emphasizing that true understanding of the population history of any group or region cannot be achieved through genomic studies alone, but requires a synthesis of insights from genomics with information from anthropology, linguistics, archaeology, and sociology. It will also be important to involve Native American organizations and communities in dialogues about these studies, as their perspectives have been underrepresented in these studies in the past.

Acknowledgements

We thank Lars Fehren-Schmitz, David Meltzer and Jennifer Raff for critical comments. PS was supported by the Wenner-Gren foundation and the Swedish Research Council (VR grant 2014-453). DR was supported by NIH grant GM100233, by NSF HOMINID BCS-1032255, and is a Howard Hughes Medical Institute investigator.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Pitulko V, Nikolsky P, Giryay E, Basilyan A, Tumskey V, Koulakov S, Astakhov S, Pavlova E, Anisimov M: **The Yana RHS site: humans in the Arctic before the last glacial maximum.** *Science* 2004, **303**:52-56.
 2. Meltzer DJ: *First Peoples in a New World: Colonizing Ice Age America.* Univ of California Press; 2009.
 3. Goebel T, Waters MR, O'Rourke DH: **The late Pleistocene dispersal of modern humans in the Americas.** *Science* 2008, **319**:1497-1502.
 4. Heintzman PD, Froese D, Ives JW, Soares AER, Zazula GD, Letts B, Andrews TD, Driver JC, Hall E, Hare PG *et al.*: **Bison phylogeography constrains dispersal and viability of the Ice Free Corridor in western Canada.** *Proc Natl Acad Sci.* 2016.
 5. Gilbert MTP, Jenkins DL, Gotherstrom A, Naveran N, Sanchez JJ, Hofreiter M, Thomsen PF, Binladen J, Higham TFG, Yohe RM *et al.*: **DNA from pre-Clovis human coprolites in Oregon, North America.** *Science* 2008, **320**:786-789.
 6. Dillehay TD, Ramirez C, Pino M, Collins M, Rossen J, Pino-Navarro J: **Monte Verde: seaweed, food, medicine, and the peopling of South America.** *Science* 2008, **320**:784-786.
 7. Waters MR, Stafford TW, McDonald HG, Gustafson C, Rasmussen M, Cappellini E, Olsen JV, Szklarczyk D, Jensen LJ, Gilbert MTP *et al.*: **Pre-Clovis mastodon hunting 13,800 years ago at the Manis Site, Washington.** *Science* 2011, **334**:351-353.
 8. Jenkins DL, Davis LG, Stafford TW, Campos PF, Hockett B, Jones GT, Cummings LS, Yost C, Connolly TJ, Yohe RM *et al.*: **Clovis age Western Stemmed projectile points and human coprolites at the Paisley Caves.** *Science* 2012, **337**:223-228.

9. Pääbo S: **Ancient DNA: extraction, characterization, molecular cloning and enzymatic amplification.** *Proc Natl Acad Sci* 1989;86.
 10. Wallace DC, Garrison K, Knowler WC: **Dramatic founder effects in Amerindian mitochondrial DNAs.** *Am J Phys Anthropol* 1985, **68**:149-155.
 11. Torroni A, Schurr TG, Yang C-C, Szathmary EJ, Williams RC, Schanfield MS, Troup GA, Knowler WC, Lawrence DN, Weiss KM: **Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations.** *Genetics* 1992, **130**:153-162.
 12. Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K: **Peopling of the Americas, founded by four major lineages of mitochondrial DNA.** *Mol Biol Evol* 1993, **10**:23-47.
 13. Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC: **Asian affinities and continental radiation of the four founding Native American mtDNAs.** *Am J Hum Genet* 1993, **53**:563.
 14. Mulligan CJ, Kitchen A, Miyamoto MM: **Updated three-stage model for the peopling of the Americas.** *PLoS ONE* 2008, **3**:e3199.
 15. Tamm E, Kivisild T, Reidla M, Metspalu M, Smith D, Mulligan C, Bravi C, Rickards O, Martinez-Labarga C, Khusnutdinova E *et al.*: **Beringian standstill and spread of Native American founders.** *PLoS ONE* 2007, **2**:e829.
 16. Llamas B, Fehren-Schmitz L, Valverde G, Soubrier J, Mallick S, Rohland N, Nordenfellt S, Valdiosera C, Richards SM, Rohrlach A *et al.*: **Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas.** *Sci Adv* 2016;2.
 17. Hey J: **On the number of new world founders: a population genetic portrait of the peopling of the Americas.** *PLoS Biol* 2005, **3**:e193.
 18. Tackney JC, Potter BA, Raff J, Powers M, Watkins WS, Warner D, Reuther JD, Irish JD, O'Rourke DH: **Two contemporaneous mitogenomes from terminal Pleistocene burials in eastern Beringia.** *Proc Natl Acad Sci* 2015, **112**:13833-13838.
 19. Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL: **Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa.** *Proc Natl Acad Sci U S A* 2005, **102**:15942-15947.
 20. Wang S, Lewis CM Jr, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C *et al.*: **Genetic variation and population structure in Native Americans.** *PLoS Genet* 2007, **3**:e185.
 21. Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung H-C, Szpiech ZA, Degnan JH, Wang K, Guerreiro R: **Genotype, haplotype and copy-number variation in worldwide human populations.** *Nature* 2008, **451**:998-1003.
 22. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL *et al.*: **Worldwide human relationships inferred from genome-wide patterns of variation.** *Science* 2008, **319**:1100-1104.
 23. Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, ●● Parra M, Rojas W, Duque C, Mesa N *et al.*: **Reconstructing Native American population history.** *Nature* 2012, **488**:370-374.
- This paper reports the first comprehensive genome-wide data set of Native Americans. Analysis of these data documented evidence for at least three streams of ancestry into the Americas: the lineage giving rise to most Native American ancestry, one giving rise to the Inuit, and one that contributed some of the ancestry of present-day Athabaskan-speakers.
24. Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, Olivieri A, Kashani BH, Ritchie KH, Scozzari R, Kong Q-P *et al.*: **Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups.** *Curr Biol* 2009, **19**:1-8.
 25. Rasmussen M, Anzick SL, Waters MR, Skoglund P, DeGiorgio M, ●● Stafford TW Jr, Rasmussen S, Moltke I, Albrechtsen A, Doyle SM *et al.*: **The genome of a Late Pleistocene human from a Clovis burial site in western Montana.** *Nature* 2014, **506**:225-229.

This paper presents the oldest genome-wide data from the Americas reported to date, documenting the presence of the main source of ancestry in present-day Native Americans by at least around 12,600 years ago. In addition, it demonstrated structure between Northern and Southern groups persisting since the Late Pleistocene until today.

26. Bradley B, Stanford D: **The North Atlantic ice-edge corridor: a possible Palaeolithic route to the New World.** *World Archaeol* 2004, **36**:459-478.
 27. Raff JA, Bolnick DA: **Does mitochondrial haplogroup X indicate ancient trans-atlantic migration to the Americas? A critical re-evaluation.** *PaleoAmerica* 2015, **1**:297-304.
 28. Rasmussen M, Sikora M, Albrechtsen A, Korneliussen TS, Moreno-Mayar JV, Poznik GD, Zollikofer CPE, Ponce de Leon MS, Allentoft ME, Moltke I *et al.*: **The ancestry and affiliations of Kennewick Man.** *Nature* 2015, **523**:455-458.
 29. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD: **Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data.** *PLoS Genet* 2009, **5**:e1000695.
 30. Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, Conrad DF, Tito RY, Osipova LP, Tarskaia LA, Zhadanov SI *et al.*: **Haplotype background of a private allele at high frequency in the Americas.** *Mol Biol Evol* 2009, **26**:995-1016.
 31. Ray N, Wegmann D, Fagundes NJR, Wang S, Ruiz-Linares A, Excoffier L: **A statistical evaluation of models for the initial settlement of the american continent emphasizes the importance of gene flow with Asia.** *Mol Biol Evol* 2010, **27**:337-345.
 32. Lipson M, Loh P-R, Levin A, Reich D, Patterson N, Berger B: **Efficient moment-based inference of admixture parameters and sources of gene flow.** *Mol Biol Evol* 2013, **30**:1788-1802.
 33. Schiffels S, Durbin R: **Inferring human population size and separation history from multiple genome sequences.** *Nat Genet* 2014, **46**:919-925.
 34. Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, ● DeGiorgio M, Albrechtsen A, Valdiosera C, Ávila-Arcos MC, Malaspina A-S: **Genomic evidence for the Pleistocene and recent population history of Native Americans.** *Science* 2015, **aab3884**.
- This paper analyzed genome-wide data from historical Pericu and Fuegians, and rejected hypotheses based on morphology that these populations carried ancestry radically different from present-day Native Americans. The paper also reports analyses of additional ancient and present-day individuals from across the Americas.
35. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, ●● Moltke I, Rasmussen S, Stafford TW Jr, Orlando L, Metspalu E *et al.*: **Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans.** *Nature* 2014, **505**:87-91.
- This paper reports the draft genome sequence of a ~24,000 year old individual from Central Siberia, showing that Native Americans are not just a branch of East Asian lineages, but are an admixture of sources related to East and West Eurasians.
36. Skoglund P, Mallick S, Bortolini MC, Chennagiri N, Hünemeier T, ●● Petzl-Erler ML, Salzano FM, Patterson N, Reich D: **Genetic evidence for two founding populations of the Americas.** *Nature* 2015.
- This paper provides evidence of ancestry related to Australo-Melanesians and Andaman Islanders in present-day populations of Amazonia, documenting a previously unknown stream of migration into the Americas.
37. Neves W, Pucciarelli H: **The origin of the first Americans-an analysis based on the cranial morphology of early South American human remains.** *Am J Phys Anthropol* 1990.
 38. Neves WA, Prous A, González-José R, Kipnis R, Powell J: **Early Holocene human skeletal remains from Santana do Riacho, Brazil: implications for the settlement of the New World.** *J Hum Evol* 2003, **45**:19-42.
 39. González-José R, Bortolini MC, Santos FR, Bonatto SL: **The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view.** *Am J Phys Anthropol* 2008, **137**:175-187.

40. Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A *et al.*: **The genetic history of Ice Age Europe**. *Nature* 2016, **534**:200-205.
41. Bellwood P: *First Migrants: Ancient Migration in Global Perspective*. John Wiley & Sons; 2014.
42. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R *et al.*: **Ancient human genome sequence of an extinct Palaeo-Eskimo**. *Nature* 2010, **463**:757-762.

This paper presented the first ancient human genome, from a ~4000 year old Saqqaq Paleo-Eskimo, and showed that most of its ancestry derived from a migration from Asia distinct from the one that gave rise to most of the ancestry of present-day Native American and Inuit populations.

43. Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliussen TS, Grønnow B, Appelt M, Gulløv HC, Friesen TM *et al.*: **The genetic prehistory of the New World Arctic**. *Science* 2014:345.

This paper presents low-coverage genomic sequence data from Neo-Eskimos and Paleo-Eskimos, providing support for the hypothesis these two cultural groups were biologically distinct, and that Neo-Eskimos are the ancestors of present-day Inuit.

44. Verdu P, Pemberton TJ, Laurent R, Kemp BM, Gonzalez-Oliver A, Gorodezky C, Hughes CE, Shattuck MR, Petzelt B, Mitchell J *et al.*: **Patterns of admixture and population structure in native populations of northwest North America**. *PLoS Genet* 2014, **10**:e1004530.