



Raghavan, M., Steinrücken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., ... Willerslev, E. (2015). Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*, 349(6250), [aab3884]. DOI: 10.1126/science.aab3884

Peer reviewed version

Link to published version (if available):
[10.1126/science.aab3884](https://doi.org/10.1126/science.aab3884)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via American Association for the Advancement of Science at DOI: 10.1126/science.aab3884 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>

Genomic evidence for the Pleistocene and recent population history of Native Americans.

Maanasa Raghavan^{1†}, Matthias Steinrücken^{2,3,4†}, Kelley Harris^{5†}, Stephan Schiffels^{6†}, Simon Rasmussen^{7†}, Anders Albrechtsen^{8†}, Cristina Valdiosera^{1,9†}, María C. Ávila-Arcos^{1,10†}, Anna-Sapfo Malaspinas^{1†}, Michael DeGiorgio^{11,12†}, Anders Eriksson^{13,14}, Ida Moltke⁸, Mait Metspalu^{15,16}, Julian R. Homburger¹⁰, Jeff Wall¹⁷, Omar E. Cornejo¹⁸, J. Víctor Moreno-Mayar¹, Thorfinn S. Korneliussen¹, Tracey Pierre¹, Morten Rasmussen^{1,10}, Paula F. Campos^{1,19}, Peter de Barros Damgaard¹, Morten E. Allentoft¹, John Lindo²⁰, Ene Metspalu^{15,16}, Ricardo Rodríguez-Varela²¹, Josefina Mansilla²², Celeste Henrickson²³, Andaine Seguin-Orlando¹, Helena Malmström^{1,24}, Thomas Stafford Jr.^{1,25}, Suyash S. Shringarpure¹⁰, Andrés Moreno-Estrada^{10,26}, Monika Karmin^{15,16}, Kristiina Tambets^{15,16}, Anders Bergström⁶, Yali Xue⁶, Vera Warmuth^{27,28}, Andrew Friend²⁹, Joy Singarayer³⁰, Paul Valdes³¹, Francois Balloux²⁷, Ilán Lebreiro²², Jose Luis Vera³², Hector Rangel-Villalobos³³, Davide Pettener³⁴, Donata Luiselli³⁴, Loren G. Davis³⁵, Evelyne Heyer³⁶, Christoph P. E. Zollikofer³⁷, Marcia S. Ponce de León³⁷, Colin I. Smith⁹, Vaughan Grimes^{38,39}, Kelly-Anne Pike³⁸, Michael Deal³⁸, Benjamin T. Fuller⁴⁰, Bernardo Arriaza⁴¹, Vivien Standen⁴², Maria F. Luz⁴³, Ludmila Osipova^{44,45}, Mikhail I. Voevoda^{44,46,47}, Oleg Balanovsky^{48,49}, Maria Lavryashina⁵⁰, Yuri Bogunov⁴⁸, Elza Khusnutdinova^{51,52}, Marina Gubina⁴⁴, Elena Balanovska⁴⁹, Sardana Fedorova^{53,54}, Sergey Litvinov^{15,51}, Boris Malyarchuk⁵⁵, Miroslava Derenko⁵⁵, M.J. Mosher⁵⁶, David Archer⁵⁷, Jerome Cybulski^{58,59,60}, Barbara Petzelt⁶¹, Joycelynn Mitchell⁶¹, Rosita Worl⁶², Paul J. Norman⁶³, Peter Parham⁶³, Brian Kemp^{64,65}, Toomas Kivisild^{15,66}, Chris Tyler-Smith⁶,

Manjinder S. Sandhu^{6,67}, Michael Crawford⁶⁸, Richard Villems^{15,16}, David Glenn Smith⁶⁹, Michael R. Waters^{70,71,72}, Ted Goebel⁷⁰, John R. Johnson⁷³, Ripan S. Malhi^{20,74}, Mattias Jakobsson²⁴, David J. Meltzer⁷⁵, Andrea Manica¹³, Richard Durbin⁶, Carlos D. Bustamante¹⁰, Yun S. Song^{2,3,5*}, Rasmus Nielsen^{5*}, Eske Willerslev^{1*}

1. Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen, Denmark.
2. Computer Science Division, University of California, Berkeley, CA 94720, USA.
3. Department of Statistics, University of California, Berkeley, CA 94720, USA.
4. Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, MA 01003, USA.
5. Department of Integrative Biology, University of California, 3060 Valley Life Sciences Bldg #3140, Berkeley, CA 94720, USA.
6. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK
7. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kemitorvet, Building 208, 2800 Kongens Lyngby, Denmark.
8. The Bioinformatics Centre, Department of Biology, University of Copenhagen, Ole Maaloes Vej 5, 2200 Copenhagen, Denmark.
9. Department of Archaeology and History, La Trobe University, Melbourne, Victoria 3086, Australia.

10. Department of Genetics, School of Medicine, Stanford University, 300 Pasteur Dr. Lane Bldg Room L331, Stanford, California 94305, USA.
11. Department of Biology and Statistics, Pennsylvania State University, 502 Wartik Laboratory, University Park, PA 16802, USA.
12. Department of Statistics, Pennsylvania State University, 502 Wartik Laboratory, University Park, PA 16802, USA.
13. Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.
14. Integrative Systems Biology Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Kingdom of Saudi Arabia.
15. Estonian Biocentre, Evolutionary Biology Group, Tartu 51010, Estonia.
16. Department of Evolutionary Biology, University of Tartu, Tartu 51010, Estonia.
17. Institute for Human Genetics, University of California San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143, USA.
18. School of Biological Sciences, Washington State University, PO Box 644236, Heald 429, Pullman, Washington 99164, USA.
19. CIMAR/CIIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Rua dos Bragas 289, 4050-123 Porto, Portugal.
20. Department of Anthropology, University of Illinois at Urbana-Champaign, 607 S. Mathews Ave, Urbana, IL 61801, USA.

21. Centro Mixto, Universidad Complutense de Madrid–Instituto de Salud Carlos III de Evolución y Comportamiento Humano, Madrid, Spain.
22. Instituto Nacional de Antropología e Historia, Moneda 13, Centro, Cuauhtémoc, 06060 Mexico Mexico City, Mexico.
23. University of Utah, Department of Anthropology, 270 S 1400 E, Salt Lake City, Utah 84112, USA.
24. Department of Evolutionary Biology and Science for Life Laboratory, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden.
25. AMS 14C Dating Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus, Denmark.
26. National Laboratory of Genomics for Biodiversity, Center for Research and Advanced Studies of the National Polytechnic Institute, Irapuato 36821, Mexico.
27. UCL Genetics Institute, Gower Street, London WC1E 6BT, UK.
28. Evolutionsbiologiskt Centrum, Norbyvägen 18D, 75236 Uppsala, Sweden.
29. Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, UK.
30. Centre for Past Climate Change and Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading, UK.
31. School of Geographical Sciences, University Road, Clifton, Bristol BS8 1SS, UK.
32. Escuela Nacional de Antropología e Historia, Periférico Sur y Zapote s/n. Colonia Isidro Fabela, Tlalpan, Isidro Fabela, 14030 Mexico City, Mexico.

33. Instituto de Investigación en Genética Molecular, Universidad de Guadalajara, Ocotlán, Mexico.
34. Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA), Università di Bologna, Via Selmi 3, 40126 Bologna, Italy.
35. Department of Anthropology, Oregon State University, 238 Waldo Hall, Corvallis, OR, 97331 USA.
36. Museum National d'Histoire Naturelle, CNRS, Université Paris 7 Diderot, Sorbonne Paris Cité, Sorbonne Universités, Unité Eco-Anthropologie et Ethnobiologie (UMR7206), Paris, France.
37. Anthropological Institute and Museum, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.
38. Department of Archaeology, Memorial University, Queen's College, 210 Prince Philip Drive, St. John's, Newfoundland, A1C 5S7, Canada.
39. Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany.
40. Department of Earth System Science, University of California, Irvine, Keck CCAMS Group, B321 Croul Hall, Irvine, California, 92697, USA.
41. Instituto de Alta Investigación, Universidad de Tarapacá, 18 de Septiembre 2222, Casilla 6-D Arica, Chile.
42. Departamento de Antropología, Universidad de Tarapacá, 18 de Septiembre 2222. Casilla 6-D Arica, Chile.
43. Fundação Museu do Homem Americano, Centro Cultural Sérgio Motta, Campestre, 64770-000 Sao Raimundo Nonato, Brazil.

44. Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Prospekt Lavrentyeva 10, 630090 Novosibirsk, Russia.
45. Novosibirsk State University, 2 Pirogova Str., 630090 Novosibirsk, Russia.
46. Institute of Internal Medicine, Siberian Branch of RAS, 175/1 ul. B. Bogatkova, Novosibirsk 630089, Russia.
47. Novosibirsk State University, Laboratory of Molecular Epidemiology and Bioinformatics, 630090 Novosibirsk, Russia.
48. Vavilov Institute of General Genetics, Gubkina 3, 119333 Moscow, Russia.
49. Research Centre for Medical Genetics, Moskvorechie 1, 115478 Moscow, Russia.
50. Kemerovo State University, Krasnaya 3, 650000 Kemerovo, Russia.
51. Institute of Biochemistry and Genetics, Ufa Scientific Center of RAS, Prospekt Oktyabrya 71, 450054 Ufa, Russia.
52. Department of Genetics and Fundamental Medicine, Bashkir State University, Zaki Validi 32, 450076 Ufa, Russia.
53. Department of Molecular Genetics, Yakut Scientific Centre of Complex Medical Problems, Sergelyahskoe Shosse 4, 677010 Yakutsk, Russia.
54. Laboratory of Molecular Biology, Institute of Natural Sciences, M.K. Ammosov North-Eastern Federal University, 677000 Yakutsk, Russia.
55. Institute of Biological Problems of the North, Russian Academy of Sciences, Portovaya Street 18, Magadan 685000, Russia.
56. Department of Anthropology, Western Washington University, Bellingham Washington 98225, USA.

57. Department of Anthropology, Northwest Community College, 353 Fifth Street, Prince Rupert, British Columbia V8J 3L6, Canada.
58. Canadian Museum of History, 100 Rue Laurier, Gatineau, Quebec K1A 0M8, Canada.
59. University of Western Ontario, London, Ontario N6A 3K7, Canada.
60. Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada.
61. Metlakatla Treaty Office, PO Box 224, Prince Rupert, BC, Canada V8J 3P6.
62. Sealaska Heritage Institute, 105 S. Seward Street, Juneau, Alaska 99801, USA.
63. Department of Structural Biology, Stanford University School of Medicine, D100 Fairchild Science Building, Stanford, California 94305-5126, USA.
64. Department of Anthropology, Washington State University, Pullman Washington 99163, USA.
65. School of Biological Sciences, Washington State University, Pullman Washington 99163, USA.
66. Division of Biological Anthropology, University of Cambridge, Henry Wellcome Building, Fitzwilliam Street, CB2 1QH, Cambridge, UK.
67. Dept of Medicine, University of Cambridge, MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK.
68. Laboratory of Biological Anthropology, University of Kansas, 1415 Jayhawk Blvd., 622 Fraser Hall, Lawrence, Kansas 66045, USA.

69. Molecular Anthropology Laboratory, 209 Young Hall, Department of Anthropology, University of California, One Shields Avenue, Davis, California 95616, USA.
70. Center for the Study of the First Americans, Texas A&M University, College Station, Texas 77843-4352, USA.
71. Department of Anthropology, Texas A&M University, College Station, Texas 77843-4352, USA.
72. Department of Geography, Texas A&M University, College Station, Texas 77843-4352, USA.
73. Santa Barbara Museum of Natural History, 2559 Puesta del Sol, Santa Barbara, CA 93105, USA.
74. Carle R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, 61801, USA.
75. Department of Anthropology, Southern Methodist University, Dallas, Texas 75275, USA.

† Equal contribution

*Corresponding authors

How and when the Americas were initially populated remains contentious. Using the largest ancient and modern genome-wide dataset to date, we find that Native American ancestors entered the Americas from Siberia no earlier than 23 thousand years ago (KYA), with subsequent gene flow between them and East Asians. We find that Native Americans diversified into two basal genetic branches about 13 KYA; one is currently represented in both North and South America while the other appears restricted to North America. We further find some evidence for a secondary gene flow into some Native Americans from populations sharing ancestry with Australo-Melanesians that we hypothesize may have been transmitted through populations represented by Aleutian islanders. Contrary to a number of recent claims, we find that there was no long-term isolation of early Americans in Beringia, that Amerindian and Athabascan ancestors entered the Americas at the same time, and that present-day South Americans are not related to modern Australo-Melanesians.

The peopling of the Americas has been a topic of research for more than a century (1). Despite this, questions regarding when and from where humans entered the Americas remain mired in controversy. The long-term view among archaeologists that humans first arrived in the Americas no earlier than *ca.* 13 KYA (the Clovis-first model) (2) has been seriously questioned in recent years with archaeological sites dating one-two thousand years older than Clovis (3-6), while claims of archaeological sites predating the Last Glacial Maximum (*ca.* 20 KYA) (*e.g.* 7-10) remain controversial. Molecular-based studies of mitochondrial DNA (mtDNA) and Y-chromosome markers have suggested that humans first entered the Americas between 15-23 KYA (11-16). However, these broad estimates are limited due to the use of single, uniparentally inherited loci.

Linguistic, dental and genome-wide data from contemporary individuals have suggested that Native Americans, represented by Amerindians and Athabascan speakers, have separate times of entrance into the Americas (17-19). By contrast, studies based on mtDNA and Y-chromosome data have argued for one founding migration for all Native Americans (*e.g.* 20-24). Similarly, some argue for an isolation period for Native American ancestors in Siberia prior to entering the Americas, possibly for as long as 15,000 years (Beringian Incubation Model) (11, 25, 26), while others have questioned this, arguing for continuity between Beringian and other Eurasian populations (15).

While the predominant view is that ancestors of Native Americans were the first humans to enter the Americas, some claim an early migration to the Americas from Europe (Solutrean model) (27) or from populations with recent shared ancestry with Australo-

Melanesians (Paleoamerican model) (28-33). Recent ancient genetic evidence has largely put the Solutrean model to rest (34, 35), while the Paleoamerican model based on skull morphology remains to be properly assessed using genomic approaches.

In this study, we address these unresolved paradigms relating to the early peopling of the Americas by using the largest genome-wide dataset to date from present-day and ancient individuals from North and South America, Siberia, East Asia, South East Asia and Oceania. We focus our study on the population history of the Native Americans, since a previous study has already addressed the early peopling of the American Arctic (36).

Dataset

We sequenced 31 present-day genomes from the Americas, Siberia and Oceania to an average depth of *ca.* 20X: Siberians – Altai (n=2), Buryat (n=2), Ket (n=2), Koryak (n=2), Yakut (n=2), Siberian Yupik (n=2); North American Native Americans – Tsimshian (n=1); South/Central American Native Pima (n=1), Huichol (n=1), Aymara (n=1), Yukpa (n=1); Oceanian – Papuan (n=14) (supplementary text S1, Table S1) (37). We additionally sequenced to low depth (0.003X to 1.7X), 23 ancient genomes dating between ~0.2-6 KYA from North and South America (supplementary text S1, Table S5). We further generated single nucleotide polymorphism (SNP) array data from 61 present-day individuals from 27 populations, including Native Americans and Siberians that have limited representation in the genomic literature (supplementary text S1, Tables S3, S4). These datasets were analyzed together with previously published genome and SNP chip

data (Tables S1, S3, S4). We masked for recent European admixture in some Native American populations prior to downstream analyses (supplementary text S5).

Time of entrance into the Americas

We first used *TreeMix* (38) to estimate the structure of the admixture tree relating American populations to worldwide populations (Figs. 1, S5, supplementary text S6). We found all Native Americans to be a monophyletic group, with further diversification into two basal branches, one representing Amerindians and the other representing Athabascan speakers. We found the Paleo-Eskimos and Inuit to be a separate clade to the Native Americans, as reported previously (36).

Next, we estimated divergence times between Native Americans and East Asians by applying *diCal2.0* (supplementary text S7), a new version of the method *diCal* (39) extended to handle complex demographic models involving multiple populations with migration (40), and an identity-by-state (IBS) tract method (41) (supplementary text S7) on high-coverage whole genome datasets. As we identified the Siberian Koryak to be one of the closest sampled East Asian populations to the Native Americans using *TreeMix* (Figs. 1, S5), we first determined the statistical estimates of divergence times between Native Americans and Koryak. We considered demographic models that reflected a clean split as well as gene flow post-dating the split between population pairs (Fig. 2A). Using both *diCal2.0* and IBS tracts we dated the divergence of Native Americans, including Amerindians and Athabascans, to Koryak to be close to 20 KYA, or perhaps slightly

older but no more than 23 KYA (Figure 2B, supplementary text S7). We also found that the models that include gene flow between the population pairs provided a better fit to the data (Fig. 2B, supplementary text S7). We used the Multiple Sequentially Markovian Coalescent (MSMC) method (42) (supplementary text S7) to estimate relative cross coalescence rates (CCR) between pairs of individuals. Overall, simulations based on the model estimates obtained using *diCal2.0* and real data show very similar relative CCR and IBS tract length distributions, which serves as an independent confirmation for *diCal*'s model estimates (Fig. 2C, D). We evaluated all methods using simulations under complex demographic models, and taking switch-errors in haplotype phasing into account (supplementary text S7).

We then proceeded to utilize *diCal2.0* with the model that allows for gene flow between populations after their split to estimate divergence times for Native Americans from more distant East Asians, namely the Siberian Nivkh and Han Chinese. The estimates for both Amerindians and Athabascans were again very similar to one another, *ca.* 23 KYA (supplementary text S7), suggesting that Amerindian and Athabascan belong to the same migration source. This was further supported by a spatially explicit population genetics model in which local demography is driven by paleo-climate and paleo-vegetation reconstructions (Fig. 3, supplementary text S8) (43). In all cases, we found that gene flow between Siberian and Native American populations is required to provide good fits of the models to the IBS tract distribution and relative CCR by MSMC. Moreover, we found strong evidence for gene flow between Athabascans and the Inuit (Table S11B), which is also supported by *TreeMix* (38) (supplementary text S6, Fig. S5), *D*-statistics employing

both sequencing and genotyping data (44-46) (supplementary text S6, Fig. S6A, Fig. S8) and, outgroup f_3 statistics (34, 46) (supplementary text S6, Fig. S12). We note that while we attempted to estimate the divergence times for Inuit-Siberians and Inuit-Native Americans (supplementary text S7, Table S11, Figs. S19, S25-27), our analyses are complicated by the extensive gene flow between Inuit and the Athabascans as well as complex admixture between Siberians and American Arctic groups (supplementary text S6, Figs. 1, S5). Additionally, all our estimates are based on assumptions regarding mutation and recombination rates. There is still some debate regarding human mutation rates and this uncertainty could affect our results (47).

Together, these results support a common Siberian origin for all Native Americans and their isolation and entrance into the Americas no earlier than 23 KYA, but with subsequent admixture with East Asian populations. This additionally suggests that Mal'ta-related admixture into the early Americans (34), representing ancestors of both Amerindians and Athabascans (Fig. 1), occurred sometime after 23 KYA, after the Native Americans split from East Asians. When we introduced stopping time of gene flow as a free parameter in the *diCal2.0* analysis, we still obtained the highest likelihood for a divergence time of 22 KYA for both Amerindians-Siberian and Athabascans-Siberian pairs, but the two pairs differed in the estimated gene flow rate and stopping time of the gene flow (supplementary text S7, Table S11C). Gene flow between Athabaskan and Siberian seems to have stopped *ca.* 12 KYA, while we found evidence for comparatively weaker gene flow until recently (*ca.* 2 KYA) between at least some East Asian populations and some Amerindian groups. Contributing factors may include

gene flow from Inuit into Siberians, although an indirect East Asian signal in South Americans cannot be discounted.

Diversification within the Americas

Using *diCal2.0*, we also estimated that Athabascans and Amerindians diversified *ca.* 13 KYA, which is also consistent with MSMC results (supplementary text S7, Table S11B), further supporting a common entrance of the two groups into the Americas after their split from Siberians after 23 KYA. We additionally employed outgroup f_3 statistics (34, 46) to determine the genetic affiliations of ancient and modern samples within the Americas (supplementary text S6, Figs. 4, S13, S14). We found that there was a marked differentiation in the Native American gene pool at least by the end of the Clovis tradition, *ca.* 12.6 KYA, with one branch representing Native Americans from both North and South America and the other being restricted to North America (Figs. 4, S13, S14). Moreover, using a large SNP chip panel masked for non-Native ancestry with increased representation of northern Amerindians, we observed a signal of gene flow between not only the Athabascans and the Inuit, as discussed before, but also between the Inuit and northwest Amerindians residing in the same region as the northern Athabascans (supplementary text S6, Fig. S8). This signal was not detected in north-central or northeastern Amerindian populations and provides a possible explanation for the southern Amerindians being genetically closer to the north-central and northeastern Amerindians than to the Athabascans and northwestern Native Americans.

In contrast to the Pleistocene human from Anzick-1, all the Holocene individuals from the Americas tend to be related to present-day populations from the same geographical regions (Figs. 4, S13, S14). As such, our results suggest that the diversification of Athabascans and Amerindian groups happened within the Americas. The data also point to a degree of geographic continuity of ancient and modern populations over, at least, the last 5 KYA.

Australo-Melanesian gene flow into the Americas

When testing for gene flow between Athabascans and Inuit using masked SNP chip based *D*-statistics (46), we observed a weak gradient in the *D*-statistic values across the Americas (Fig. S8). There is a tendency for the Inuit to be much closer to the Athabascans, than to certain Amerindians like the Arhuaco, Algonquin, Cree and Yaqui when compared to other Amerindians such as Palikur and Surui (Fig. S8). To further investigate this trend, we tested for additional gene flow from Eurasian populations into parts of the Americas using *D*-statistics (46). Using the masked SNP chip dataset with multiple Native American populations, we found that some American populations, including the Aleutian Islanders, Surui from Brazil and the Athabascans are closer to Australo-Melanesians than other Native Americans such as North American Ojibwa, Cree and Algonquin and South American Purepecha, Arhuaco and Wayuu (Fig. S10). Surui is one of closest Native American populations to Australo-Melanesians including Papuans, non-Papuan Melanesians, Solomon Islanders, and South East Asian hunter-gatherers such as Aeta. While we note that our results show only weak trends, they

nevertheless suggest a possible scenario leading to a Eurasian signal in some Native Americans that is related to Australo-Melanesians and East Asians.

Based on both archaeological evidence and mtDNA data from ancient and modern samples, the Aleutian Islands are hypothesized to have been colonized as early as *ca.* 9 KYA by ‘Paleo-Aleuts’ who were succeeded by the ‘Neo-Aleuts’, with present-day Aleutian Islanders potentially resulting from admixture between these two streams (48, 49). While Aleutian Islanders were previously found to be closely related to the Inuit (36), their increased affinity to East Asians and Oceanians compared to the rest of the Americas suggests a possible ‘northern’ route for the spread of an Oceanian signal into some Native Americans. Using enhanced D-statistics on whole genomes (50), we again found some evidence for more Papuan and Denisovan ancestry in the Aleutian Islander than in the Native American samples (supplementary text S6, Table S10). The differential affinity of Native Americans to Oceanians, ranging from a strong signal in the Surui to much weaker signal in northern Amerindians such as Ojibwa, points to this signal coming into the Americas after the initial peopling by Native American ancestors. This is also supported by the results from the *diCal2.0* and MSMC analyses showing a weak but recent gene flow into South Americans from populations related to contemporary Northeast Asians, which might be considered a proxy for the related Aleutian Islanders (Fig. 2). However, how this signal ultimately reached South America remains unclear, but one possible route could be through the Athabascan expansion southwards that might have brought with it an indirect Oceanian signal through the Aleutian Islanders. Additionally, we caution that all results are based on analyses of a small fraction of the

Aleutian Islanders datasets (masked SNP chip and whole genome data), and this may bias the results in an unforeseen manner.

The Paleoamerican model

To further test if the Australo-Melanesian gene flow could provide support for the Paleoamerican model, we used ancient genomes from groups (Pericúes from Mexico and the Fuego-Patagonians from Chile and Argentina) that, based on their distinctive skull morphologies, are claimed to be ancient relicts of Paleoamericans (28, 51-53). Using Principal Component Analysis and ADMIXTURE (54), we show that the ancient samples cluster with other Native American groups and outside the range of Oceanian genetic variation (supplementary text S10, Figs. S32, S33, 34). Similarly, outgroup f_3 statistics (34, 46) reveal low shared genetic ancestry between the ancient samples and Oceanians (supplementary text S10, Figs. S36, S37). Furthermore, from genome-based and masked SNP chip data-based D -statistics (44-46), we found no evidence for gene flow into the Pericúes from Oceanians, but observed the Fuego-Patagonians to have at least some signal from Oceanian groups (supplementary text S10, Fig. 39).

As the Paleoamerican model is based on cranial morphology (28, 51-53), we also measured craniometric data for the ancient samples (supplementary text S10). We assessed their phenotypic affinities to supposed Paleoamericans, Amerindians and worldwide populations using more recently accepted methodologies (supplementary text

S10, Fig. S41). The majority of the studied samples show close craniometric affinities to present-day populations from North America, the Arctic, and East Asia (Fig. S41), implying that previous reporting of the close similarity of Paleoamerican and Australo-Melanesian cranial morphologies could not be replicated (30). Admittedly, our morphometric analyses show that these ancient samples are not true Paleoamerican relicts as claimed, and hence do not support the Paleoamerican model. Regardless, our genomic data provide no support for an early migration of populations related to Australo-Melanesians to South America.

Discussion

Overall, our findings suggest that Native Americans had a high level of continuity with Eurasian populations, in particular the northeast Siberians, prior to entering the Americas, and that all present-day Native American populations originated from the same migration with subsequent gene flow from East Asia, including some that relates to Australo-Melanesians. Hence, our results contradict the long-term Beringian incubation model (11, 25, 26) and the dual migration model for the Amerindians and Athabascans (17-19). Moreover, Native Americans diverging no earlier than 23 KYA from East Asians provides an upper bound on the isolation, the entrance date of Native Americans, and the subsequent spread of Native American ancestors into the Americas, although our data are unable to provide exact geographical contexts for these processes. This is also in agreement with the pre-Clovis model suggesting initial occupation of the Americas a few thousand years prior to Clovis (3-6). However,

archaeological sites in the Americas significantly older than the Last Glacial Maximum (*ca.* 20 KYA) (7-10) should, in our view, be treated with caution. We further find a diversification of ancestral Native Americans leading to the formation of a ‘northern’ and ‘southern’ branch, which appears to have taken place within the Americas. As this event coincides roughly with the opening of habitable routes along the coastal and the interior corridors into unglaciated North America some 16 KYA and 14 KYA, respectively (53-56). This suggests a possible role of one or both these routes in the isolation and subsequent dispersal of Native Americans within the Americas.

References and Notes

1. D.J. Meltzer, *First Peoples in a New World: Colonizing Ice Age America* (Univ. of California Press, Berkeley, 2009).
2. S. J. Fiedel, The peopling of the New World: present evidence, new theories, and future directions. *J. Archaeol. Theory* **8**, 39–103 (2000).
3. T. D. Dillehay, The late Pleistocene cultures of South America. *Evol. Anthropol.* **7**, 206–16 (1999).
4. M. Waters *et al.*, The Buttermilk Creek Complex and the Origins of Clovis at the Debra L. Friedkin Site, Texas. *Science* **331**, 1599-1603 (2011).
5. M. T. Gilbert *et al.*, Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. *Science* **320**, 1787-1789 (2008).
6. D. L. Jenkins *et al.*, Clovis age western stemmed projectile points and human coprolites at the Paisley Caves. *Science* **337**, 223–228 (2012).

7. D. L. Lowery, M. A. O'Neal, J. S. Wah, D. P. Wagner, D. J. Stanford, Late Pleistocene upland stratigraphy of the western Delmarva Peninsula, USA. *Quat. Sci. Rev.* **29**, 1472-1480 (2010).
8. S. R. Holen, K. Holen, K., in *Paleoamerican Odyssey*, K. E. Graf, C. V. Ketron, M. R. Waters, Eds. (Texas A&M University Press, College Station, 2014), pp. 429-444.
9. E. Boëda *et al.*, A new late Pleistocene archaeological sequence in South America: the Vale da Pedra Furada (Piauí, Brazil). *Antiquity* **88**, 927-941 (2014).
10. G. M. Santos *et al.* A revised chronology of the lowest occupation layer of Pedra Furada Rock Shelter, Piauí, Brazil: the Pleistocene peopling of the Americas. *Quat. Sci. Rev.* **22**, 2303–2310 (2003).
11. E. Tamm *et al.*, Beringian stand- still and spread of Native American founders. *PLoS One* **2**, e829 (2007)
12. A. Achilli *et al.*, The phylogeny of the four pan-American mtDNA haplogroups: Implications for evolutionary and disease studies. *PLoS One* **3**, e1764 (2008).
13. U. A. Perego *et al.*, Distinctive paleo-indian migration routes from beringia marked by two rare mtDNA haplogroups. *Curr Biol* **19**, 1-8 (2009).
14. U. A. Perego *et al.*, The initial peopling of the Americas: A growing number of founding mitochondrial genomes from Beringia. *Genome Res.* **20**, 1174-1179 (2010).
15. M. C. Dulik, *et al.*, Mitochondrial DNA and Y Chromosome Variation Provides Evidence for a Recent Common Ancestry between Native American and Indigenous Altaians. *Am. J. Hum. Genet.* **90**, 229-246 (2012).
16. V. Battaglia *et al.*, The First Peopling of South America: New Evidence from Y-Chromosome Haplogroup Q. *PLoS One* **8**, e71390 (2013).

17. J. H. Greenberg, C. G. Turner, S. L. Zegura, The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. *Curr. Anthropol.* **27**, 477–497 (1986).
18. A. Torroni *et al.*, Asian Affinities and Continental Radiation of the Four Founding Native American mtDNAs. *Am. J. Hum. Genet.* **53**, 563-590 (1993).
19. D. Reich *et al.*, Reconstructing Native American population history, *Nature* **488**, 370-374 (2012).
20. C. J. Kolman, N. Sambuughin, E. Bermingham, Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. *Genetics* **142**:1321–34 (1996).
21. D. A. Merriwether, F. Rothhammer, R. E. Ferrell, 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 (1995).
22. T. M. Karafet *et al.*, Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. *Am. J. Hum. Genet.* **64**, 817–831 (1999).
23. S. L. Zegura, T. M. Karafet, L.A. Zhivotovsky, M. F. Hammer, High resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol. Biol. Evol.* **21**, 164-175 (2004).
24. N. J. R. Fagundes, R. Kanitz, S. L. Bonatto, A Reevaluation of the Native American MtDNA Genome Diversity and Its Bearing on the Models of Early Colonization of Beringia. *PLoS One* **3**, e3157 (2008).
25. A. Kitchen, M. M. Miyamoto, C. J. Mulligan, A Three-Stage Colonization Model for the Peopling of the Americas. *PLoS One* **3**, e1596 (2008).

26. C. J. Mulligan, A. Kitchen, M. M. Miyamoto, Updated Three-Stage Model for the Peopling of the Americas. *PLoS One* **3**, e3199 (2008).
27. S. Oppenheimer, B. Bradley, D. Stanford, Solutrean hypothesis: genetics, the mammoth in the room. *World Archaeology* **46**, 752-774 (2014).
28. R. González-José *et al.*, Craniometric evidence for Palaeoamerican survival in Baja California. *Nature* **425**, 62-65 (2003).
29. M. Hubbe, W. Neves, K. Harvati, Testing evolutionary and dispersion scenarios for the settlement of the new world. *PLoS One* **5**, e11105 (2010).
30. W. Neves, M. Hubbe, Cranial morphology of early Americans from Lagoa Santa, Brazil: implications for the settlement of the New World. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18309-18314 (2005).
31. W. A. Neves *et al.*, Early Holocene human skeletal remains from Cerca Grande, Lagoa Santa, Central Brazil, and the origins of the first Americans. *World Archaeology* **36**, 479-501 (2004).
32. J. Powell, W. Neves, Craniofacial morphology of the first Americans: Pattern and process in the peopling of the New World. *Am. J. Phys. Anthropol.* **29**, 153-188 (1999).
33. W. Neves *et al.*, in *Paleoamerican Odyssey*, K. E. Graf, C. V. Ketron, M. R. Waters, Eds. (Center for the Study of the First Americans, Santa Fe, 2013), pp. 397-412.
34. M. Raghavan *et al.*, Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87-91 (2014).
35. M. Rasmussen *et al.*, The genome of a late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**, 225-229 (2014).

36. M. Raghavan *et al.*, The genetic prehistory of the New World Arctic. *Science* **345**, 1255832–1255832 (2014).
37. Supplementary materials are available on Science Online.
38. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
39. S. Sheehan, K. Harris, and Y. S. Song, Estimating Variable Effective Population Sizes from Multiple Genomes: A Sequentially Markov Conditional Sampling Distribution Approach. *Genetics* **194**, 647-662 (2013).
40. M. Steinrücken, J. S. Paul, Y. S. Song, A sequentially Markov conditional sampling distribution for structured populations with migration and recombination. *Theor. Popul. Biol.* **87**, 51-61 (2013).
41. K. Harris, R. Nielsen, Inferring Demographic History from a Spectrum of Shared Haplotype Lengths. *PLoS Genet* **9**, e1003521 (2013).
42. S. Schiffels R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014).
43. A. Eriksson, A. Manica. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13956-1396046 (2012).
44. R. E. Green *et al.*, A draft sequence of the Neandertal genome. *Science* **328**, 710-722 (2010).
45. D. Reich *et al.*, Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053-1060 (2010).

46. N. Patterson *et al.*, Ancient admixture in human history, *Genetics* **192**, 1065-1093 (2012).
47. R. Scally, R. Durbin, Revising the human mutation rate: implications for understanding human evolution, *Nat Rev Genet.* **13**, 745-753
48. R. S. Davis, R.A. Knecht, Continuity and Change in the Eastern Aleutian Archaeological Sequence. *Hum. Biol.* **82**, 507-524 (2010).
49. M. H. Crawford, R. C. Rubicz, M. Zlojutro, Origins of Aleuts and the genetic structure of populations of the archipelago: molecular and archaeological perspectives. *Hum. Biol.* **82**, 695-717 (2010).
50. M. Meyer *et al.*, A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science* **338**, 222-226 (2012).
51. R. González-José, S. Dahinten, M. Luis, M. Hernández, H. Pucciarelli, Craniometric variation and the settlement of the Americas: testing hypotheses by means of R-matrix and matrix correlation analyses. *Am. J. Phys. Anthropol.* **116**, 154-165 (2001).
52. M. Hernández, C. Fox, C. García-Moro, Fuegian cranial morphology: the adaptation to a cold, harsh environment. *Am. J. Phys. Anthropol.* **103**, 103-117 (1997).
53. M. Lahr, Patterns of modern human diversification: Implications for Amerindian origins. *Am. J. Phys. Anthropol.* **38**, 163-198 (1995).
54. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655-1664 (2009).

55. C. A. S. Mandryk, H. Josenhans, D. W. Fedje, Late Quaternary paleoenvironments of Northwestern North America: implications for inland versus coastal migration routes. *Quat. Sci. Rev.*, **20**, 301-314 (2001).
56. A. S. Dyke, in *Quaternary Glaciations - Extent and Chronology, Part II*, J. Ehlers, P. L. Gibbard, Eds. (Elsevier, Amsterdam, 2004), pp. 373-424.
57. E. J. Dixon, Late Pleistocene colonization of North America from Northeast Asia: New insights from large-scale paleogeographic reconstructions. *Quaternary International* **285**, 57–67 (2013).
58. M. Sarnthein, T. Kiefer, P. M. Grootes, H. Elderfield, H. Erlenkeuser, Warmings in the far northwestern Pacific promoted pre-Clovis immigration to America during Heinrich event 1. *Geology* **34**, 141–144 (2006).
59. T. D. Dillehay, Ed., *Monte Verde: A Late Pleistocene Settlement in Chile*, vol. 2, *The Archaeological Context and Interpretation* (Smithsonian Institution Press, Washington, DC, 1997).

Acknowledgements

We acknowledge Jorge Gómez Valdés for providing craniometric measurements of the Pericúes at the National Museum of Anthropology in México, Armando Romero Monteverde from CINAH-Baja California Sur and Veronique Laborde at the Musée de l'Homme in Paris for providing documentation on Pericú and Fuego-Patagonian samples, respectively, Tom Gilbert, Mark McCoy, Clio Der Sarkissian, Martin Sikora, Ludovic Orlando for helpful discussions and input, Daniele Yang Yao and Chiara

Barbieri for helping with the collection of the Aymara population sample, Brenna Henn and Jeff Kidd for providing early access to the Mayan sequencing data, Canadian Museum of History, Metlakatla and Lax Kw'alaams First Nations, Listuguj Mi'gmaq Band Council, Alison Pye of TERRA Facility, Core Research Equipment & Instrument Training (CREAIT) Network at Memorial University, and the Danish National High-throughput DNA Sequencing Centre (Copenhagen) for help with sequencing. This study was supported by several funding bodies; Lundbeck Foundation and the Danish National Research (Centre for GeoGenetics members), Wellcome Trust grant 098051 (S.S., A.B., Y.X., C.T.-S., M.S.S., R.D.), Marie Curie Intra-European Fellowship-FP7-People-PIEF-GA-2009-255503 and the Transforming Human Societies RFA Fellowship from La Trobe University (C.V.), George Rosenkranz Prize for Health Care Research in Developing Countries and National Science Foundation award DMS-1201234 (M.C.A.A), Swiss National Science Foundation (PBSKP3_143529) (A.-S.M.), Ministerio de Ciencia e Innovación (MICINN) Project CGL2009-12703-C03-03 and MICINN (BES-2010-030127) (R.R.-V.), Consejo Nacional de Ciencia y Tecnología (Mexico) (J.V.M.M), European Research Council and Marie Curie Actions Grant 300554 (M.E.A.), Wenner-Gren Foundations and the Australian Research Council Future Fellowship FT0992258 (C.I.S.), European Research Council ERC-2011-AdG 295733 grant (Langelin) (D.P. and D.L.), Bernice Peltier Huber Charitable Trust (C.H., L.G.D.), Russian Foundation for Basic Research grant 13-06-00670 (E.B.), Russian Foundation for Basic Research (RFBR) grant 14-0400725 (E.K.), EU European Regional Development Fund through the Centre of Excellence in Genomics to Estonian Biocentre and Estonian

Institutional Research grant IUT24-1 (E.M., K.T., M.M., M.K., R.V.), Estonian Science Foundation grant 8973 (M.M.), Stanford Graduate Fellowship (J.R.H.), National Science Foundation (NSF) BCS-1025139 (R.S.M.), Social Science Research Council of Canada (K.-A.P., V.G.), National Institutes of Health (NIH) grant R01-GM094402 (M.S., Y.S.S.), Packard Fellowship for Science and Engineering (Y.S.S.), Russian Science Fund (RNF) grant 14-04-00827 and Presidium of Russian Academy of Sciences (RAS) MCB Programme (O.B.), National Institutes of Health (NIH) grant R01 - AI17892 (P.J.N., P.P.), and Russian Foundation for Basic Research (RFBR) grant 14-06-00384 (Y.B.). Informed consent was obtained for the sequencing of the modern individuals, with ethical approval from The National Committee on Health Research Ethics, Denmark (H-3-2012-FSP21). Data for some of the present-day samples are only available for demographic research under data access agreement with E.W. and R.S.M. (Tsimshian genome) (see Table S1 for a list of these samples). Raw reads from the ancient and the remainder of the present-day genomes are available for download through European Nucleotide Archive (ENA) accession no. XX, and the corresponding alignment files (BAMs) are available at <http://www.cbs.dtu.dk/suppl/NativeAmerican/>. The remainder of the Illumina genotype data can be accessed through Gene Expression Omnibus (GEO) series accession no. XX and at www.ebc.ee/free_data. The authors declare no competing financial interests.

Figure legend

Fig. 1. Origins and pre-Columbian admixture history of present-day Native

Americans. A. Our results support a single migration for the ancestors of all present-day Native Americans, including Amerindian and Athabascans speakers (purple). The gradient in colour provides an estimate of the initial divergence of Native American ancestors from the ancestors of present-day East Asians (~23 KYA), northeast Siberians (~20 KYA), and finally their appearance in the Americas ~15 KYA, which is supported by archaeological evidence from sites like Monte Verde (59). Subsequently, ~13 KYA (red), there was a split between the northern and southern branches of Native Americans, with the former comprising Athabascans and northern Amerindians and the latter consisting of Amerindians speakers in Central and South America including the ~12,600-year-old Clovis (35). There is an admixture signal between Inuit and Athabascans and northwestern Amerindians, and between Inuit and Siberians (yellow arrows); however, the gene flow direction is unresolved due to the complexity of the admixture events (SOM). Additionally, we see a weak signal of gene flow from Papuan-related populations into some Native Americans, which may have been mediated through East Asians and Aleutian Islanders. Also shown is the Mal'ta gene flow into Native American ancestors some 23 KYA (Raghavan et al. 2014). We note that it is currently not possible for us to ascertain the exact geographical locations of the depicted events; hence, the positioning of the arrows should not be considered a reflection of these. **B.** Admixture plot, based on *TreeMix* (38) results and supported by *D*-statistics (SOM), shows that all Native Americans form a clade, separate from the Inuit, with admixture between some Native Americans and the North American Arctic.

Fig. 2. Divergence estimates between Native Americans and Siberian Koryak. A.

The demographic model used allows for continuous gene flow between populations 1 and 2, starting from the time T_{DIV} of divergence and ending at T_M . The backward probability of migration per individual per generation is denoted by m . The bottleneck at T_B captures the out-of-Africa event. **B.** The red curves depict empirical distributions of IBS tracts shared between Karitiana-Koryak (solid) and Athabascan-Koryak (dashed). The blue and orange curves depict IBS tracts shared between the two population pairs, simulated under two demographic models based on results from *diCal2.0*. Overall, the migration scenarios (orange) match the empirical curves (red) better than the clean split scenarios (blue), with more long IBS tracts showing evidence of recent common ancestry between Koryaks and Native Americans. **C and D.** Relative CCR for the Karitiana-Koryak and Athabascan-Koryak divergence (red), respectively, including data simulated under the two demographic models in panel B. In both cases, the model with gene flow (orange) fits the data (red) better than the clean split model (blue). The migration model explains a broader CCR tail in the case of Karitiana-Koryak and the relatively late onset of the CCR decay for Athabascan-Koryak.

Fig. 3. Testing migrations into the Americas using a spatially explicit population genetics model. Estimates of differentiation between Central and South American populations versus present-day Athabascan and Greenlandic Inuit, and the ancient Saqqaq and Anzick-1 genomes (black vertical lines), compared to predictions from a Climate-Informed Spatial Genetic model reconstructing a single wave into the Americas (curves, the solid part represents the 95% credibility interval). Both Anzick-1 and the Athabascans were part of the same main wave into the Americas to which other Central

and South American populations belonged to, while the Inuit and Saqqaq are the descendants of different waves (observed values outside the 95% credibility interval).

Fig. 4. Diversification within the Americas. We used SNP chip data-based f_3 -statistics (34,45) of the form $f_3(X, \text{Ancient}; \text{Yoruba})$ to estimate the shared ancestry between Anzick-1 and low coverage Holocene-aged samples from the Americas to a large panel of worldwide populations (X), including previously unrepresented Athabascan and Amerindian groups from North America (Table SX). Individuals from the Americas were masked for non-Native ancestry prior to the analysis. The f_3 statistics are depicted as heat maps with the sampling location of the ancient sample marked by the dotted lines, and corresponding ranked plots with error bars are shown in Fig. SY. We find the Anzick-1 sample to share most ancestry with the ‘southern’ branch of Native Americans when using multiple northern Native Americans sequenced in this study, consistent with (35). The rest of the ancient samples share most ancestry with Native Americans, as expected, with a general tendency to be genetically closer to present-day populations from the same geographical region.

Figure 1

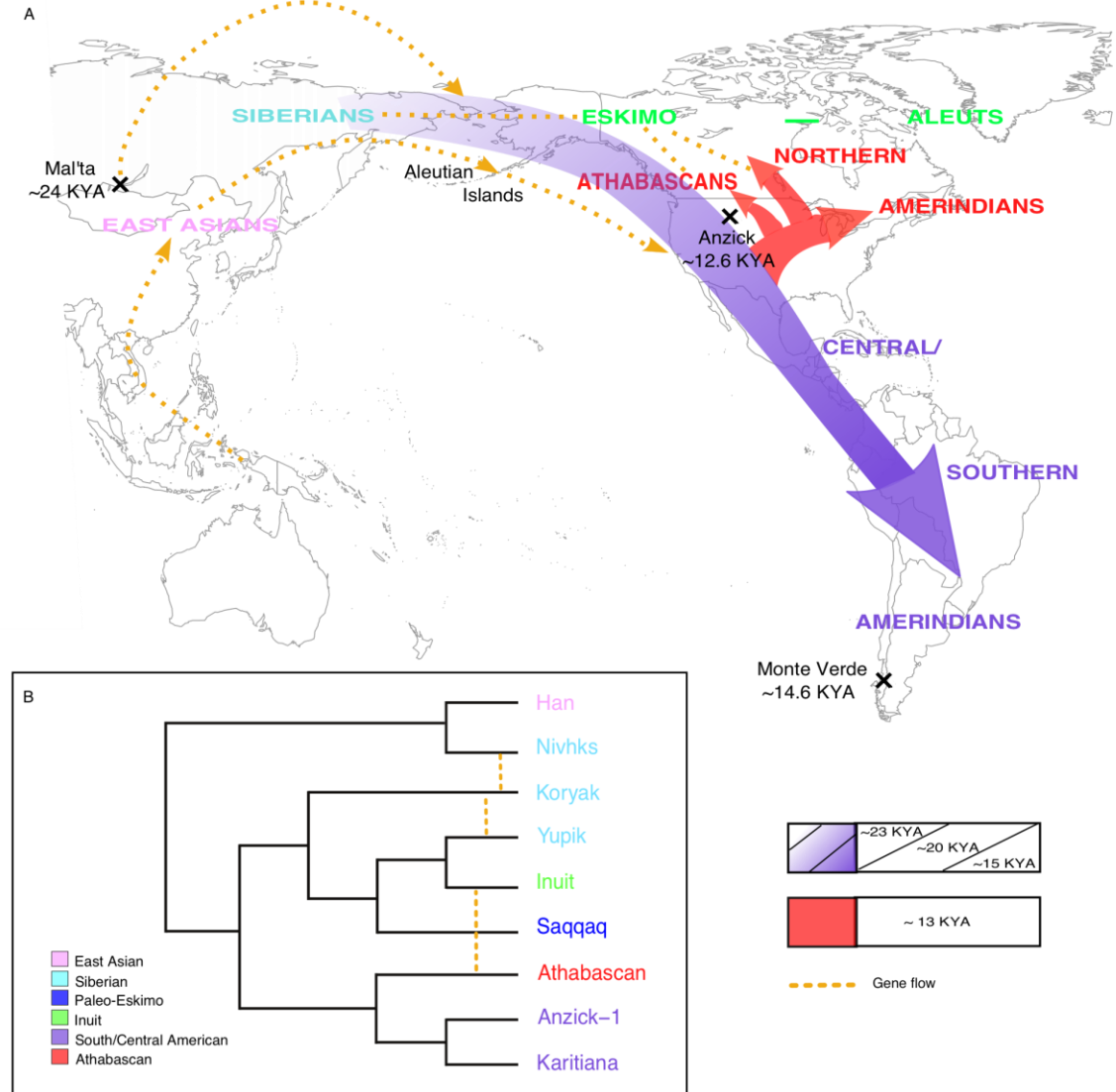


Figure 2

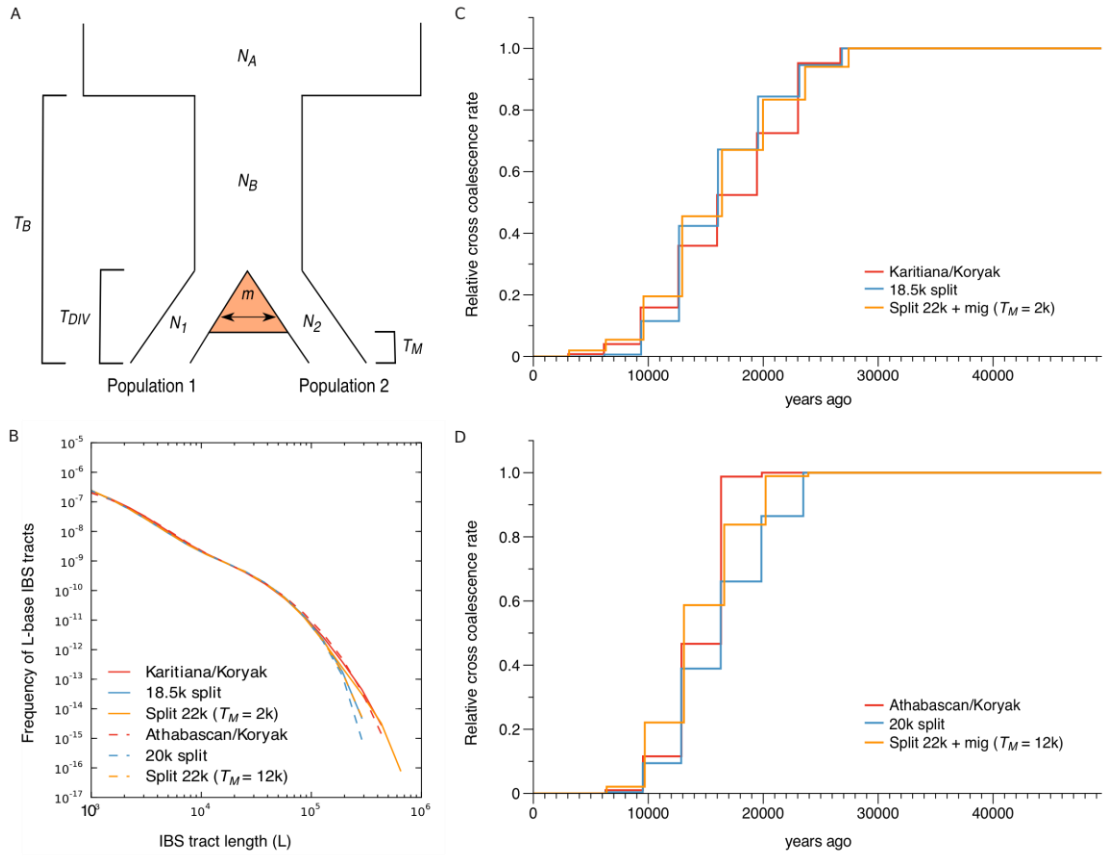


Figure 3

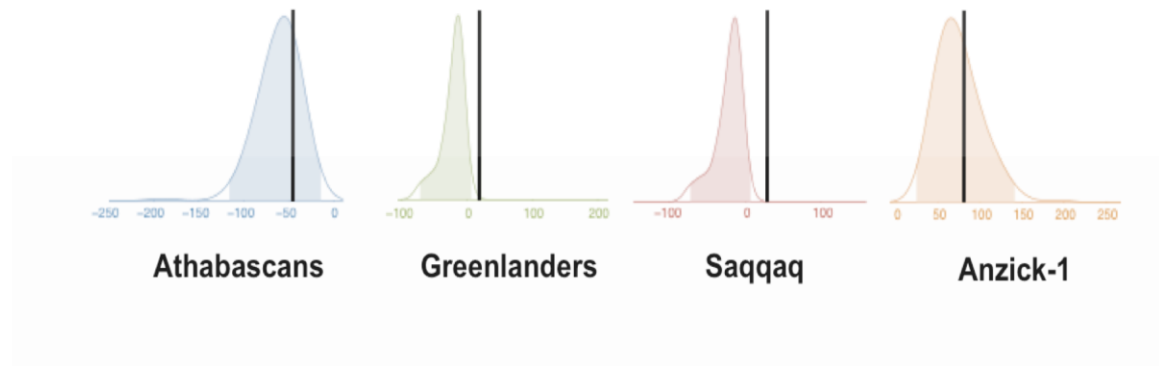


Figure 4

