



# Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia

Constanza de la Fuente<sup>a,b</sup>, María C. Ávila-Arcos<sup>c</sup>, Jacqueline Galimany<sup>a</sup>, Meredith L. Carpenter<sup>d,e</sup>, Julian R. Homburger<sup>f</sup>, Alejandro Blanco<sup>a</sup>, Paloma Contreras<sup>a</sup>, Diana Cruz Dávalos<sup>g</sup>, Omar Reyes<sup>h</sup>, Manuel San Roman<sup>h</sup>, Andrés Moreno-Estrada<sup>i</sup>, Paula F. Campos<sup>b,j</sup>, Celeste Eng<sup>k</sup>, Scott Huntsman<sup>k</sup>, Esteban G. Burchard<sup>l</sup>, Anna-Sapfo Malaspinas<sup>m</sup>, Carlos D. Bustamante<sup>d,f</sup>, Eske Willerslev<sup>b,n,o</sup>, Elena Llop<sup>a</sup>, Ricardo A. Verdugo<sup>a,1</sup>, and Mauricio Moraga<sup>a,1</sup>

<sup>a</sup>Human Genetics Program, Institute of Biomedical Sciences, Faculty of Medicine, University of Chile, Santiago 8380453, Chile; <sup>b</sup>Centre for GeoGenetics, University of Copenhagen, 1350 Copenhagen, Denmark; <sup>c</sup>International Laboratory for Human Genome Research, National Autonomous University of Mexico, Juriquilla 76230, Santiago de Querétaro, Mexico; <sup>d</sup>Center for Computational, Evolutionary and Human Genomics, Stanford University, Stanford, CA 94305; <sup>e</sup>Arc Bio, LLC, Menlo Park, CA 94025; <sup>f</sup>Department of Genetics, Stanford University, Stanford, CA 94305; <sup>g</sup>Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland; <sup>h</sup>Centro de Estudios del Hombre Austral, Instituto de la Patagonia, Universidad de Magallanes, Punta Arenas 6213029, Chile; <sup>i</sup>National Laboratory of Genomics for Biodiversity, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, 36821 Guanajuato, Mexico; <sup>j</sup>Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Novo Edifício do Terminal de Cruzeiros do Porto de Leixões, 4450-208 Matosinhos, Portugal; <sup>k</sup>Department of Medicine, University of California, San Francisco, CA 94131; <sup>l</sup>Department of Bioengineering and Therapeutic Science, University of California, San Francisco, CA 94158; <sup>m</sup>Department of Computational Biology, University of Lausanne, 1015 Lausanne, Switzerland; <sup>n</sup>Department of Zoology, University of Cambridge, Cambridge CB2 1TN, United Kingdom; and <sup>o</sup>Wellcome Genome Campus, Wellcome Trust Sanger Institute, Hinxton CB10 1SA, United Kingdom

Edited by Tony Tosi, Kent State, and accepted by Editorial Board Member C. O. Lovejoy February 5, 2018 (received for review September 30, 2017)

Patagonia was the last region of the Americas reached by humans who entered the continent from Siberia ~15,000–20,000 y ago. Despite recent genomic approaches to reconstruct the continental evolutionary history, regional characterization of ancient and modern genomes remains understudied. Exploring the genomic diversity within Patagonia is not just a valuable strategy to gain a better understanding of the history and diversification of human populations in the southernmost tip of the Americas, but it would also improve the representation of Native American diversity in global databases of human variation. Here, we present genome data from four modern populations from Central Southern Chile and Patagonia ( $n = 61$ ) and four ancient maritime individuals from Patagonia (~1,000 y old). Both the modern and ancient individuals studied in this work have a greater genetic affinity with other modern Native Americans than to any non-American population, showing within South America a clear structure between major geographical regions. Native Patagonian Kawéskar and Yámana showed the highest genetic affinity with the ancient individuals, indicating genetic continuity in the region during the past 1,000 y before present, together with an important agreement between the ethnic affiliation and historical distribution of both groups. Lastly, the ancient maritime individuals were genetically equidistant to a ~200-y-old terrestrial hunter-gatherer from Tierra del Fuego, which supports a model with an initial separation of a common ancestral group to both maritime populations from a terrestrial population, with a later diversification of the maritime groups.

resources. A systematic and specialized exploitation of the marine environment began only ~6,000 YBP (3). Evidence of this adaptation is found in three areas: (i) Chiloé (north of the Patagonian archipelago), (ii) Strait of Magellan–Otway Sound (western shore), and (iii) the Beagle Channel (4, 5). The latter two areas are located south of the Patagonian archipelago (Fig. 1) and show the earliest evidence of a maritime tradition, although it is

## Significance

Recent genomic studies of ancient and modern humans from the Americas have given a comprehensive view of the peopling of the continent. However, regional characterization of ancient and modern individuals is lacking, being key to unveiling fine-scale differences within the continent. We present genome-wide analyses of ancient and modern individuals from South America from Western Patagonia. We found a strong affinity between modern and ancient individuals from the region, providing evidence of continuity in the region for the last ~1,000 years and regional genetic structure within Southern South America. In particular, the analysis of these ancient genomes helps address questions related to the maritime tradition in the region and its diversification posterior to the split from terrestrial hunter-gatherers.

paleogenomics | Patagonia | maritime hunter-gatherers

The Southern part of South America, known as Patagonia, is a highly diverse territory with two contrasting environments: the Andean mountains to the Southwest—encompassing an irregular coast, numerous islands and archipelagos, and plateaus—and lowlands in the Northeast. The earliest evidence of human occupation is found in North Patagonia in the Monte Verde site, dated to 14,500 y before present (YBP) (1). During this time, Tierra del Fuego island was still connected to the mainland, and it was not until ~8,000 YBP that the Strait of Magellan was formed, isolating some human groups on the island (2). Different disciplines have contributed to unraveling the first human occupation and dispersion in Patagonia. However, from a genetic point of view, little is known about their affinities at a regional scale.

The first inhabitants of Patagonia were broad-spectrum hunter-gatherers with a nonspecialized exploitation of marine

Author contributions: C.d.l.F., R.A.V., and M.M. designed research; C.d.l.F., M.C.Á.-A., M.L.C., A.M.-E., P.F.C., R.A.V., and M.M. performed research; O.R., M.S.R., A.M.-E., C.E., S.H., E.G.B., C.D.B., E.W., and E.L. contributed new reagents/analytic tools; C.d.l.F., M.C.Á.-A., J.G., J.R.H., A.B., P.C., D.C.D., and A.-S.M. analyzed data; and C.d.l.F., M.C.Á.-A., O.R., M.S.R., P.F.C., R.A.V., and M.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. T.T. is a guest editor invited by the Editorial Board.

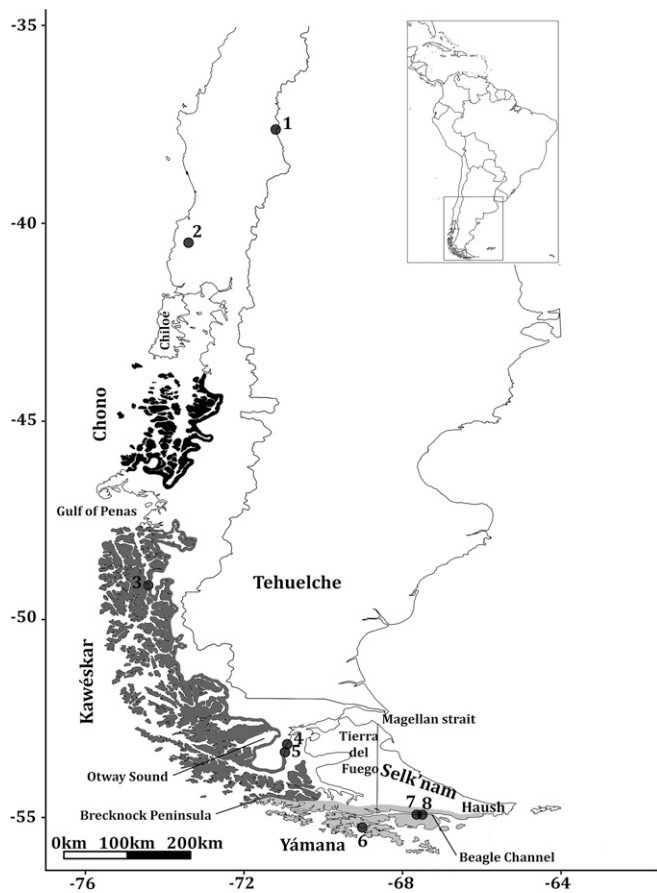
Published under the PNAS license.

Data deposition: All the modern data from this paper can be accessed through request to the authors. Researchers interested in accessing individual-level data must send a signed letter stating: "With regards to the non-public samples from the PatagoniaDNA Project, I agree that: (a) I will not secondarily distribute the data to anyone, (b) I will not post it publicly, (c) I will make no attempt to connect the genetic data to personal identifiers for the samples, (d) I will not use the data for any commercial purposes." The ancient data are available for download through the European Nucleotide Archive (ENA) (accession no. PRJEB24629).

<sup>1</sup>To whom correspondence may be addressed. Email: raverdugo@u.uchile.cl or mmoraga@med.uchile.cl.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1715688115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1715688115/-DCSupplemental).

Published online April 9, 2018.



**Fig. 1.** Approximate distribution of Native populations from Patagonia. The distribution of the maritime groups is shown in gray, and the associated labels are adjacent to each area. The distribution of the terrestrial groups is displayed with their name at the center of their range. Numbers display the location of the ancient individuals and modern population: 1, Pehuenche; 2, Huilliche; 3, IPK13; 4, Kawéskar; 5, IPK12; 6, IPY08; 7, Yámana; and 8, IPY10.

not clear whether these early adaptations developed into distinct cultural groups and how they are connected to present-day ethnic groups from Patagonia (2, 4, 6, 7).

After the arrival of Europeans in the 16th century, six major ethnic groups were recognized in the area: three maritime hunter-gatherer groups (from north to south): Chono, Kawéskar or Alacalufe, and Yámana or Yaghan; and three terrestrial hunter-gatherer groups, known as Tehuelche or Aonikenk, Selk'nam or Ona, and Haush (Fig. 1). In addition to different subsistence strategies, these groups were identified based on their language, physical appearance, and geographical location. The maritime groups occupied the archipelagos to the west of the Andes, from Reloncaví Sea to the Cape Horn (~40°S to 56°S), while the terrestrial groups inhabited the plateaus and lowlands from the Argentinean Patagonia (Tehuelches) to Tierra del Fuego (Selk'nam and Haush) (8–11).

The biological variation of the native populations from Patagonia has been described by using morphological and genetic data. Only a few human remains have been found dating to between 12,000 and 3,000 YBP. Consequently, most of the morphological and genetic studies have focused on modern populations and late historical groups, such as Selk'nam, Kawéskar, and Yámana. Both approaches have suggested a close relationship between these groups (12–18). Studies based on mitochondrial DNA (mtDNA) variation have identified only two main lineages (C and D) in the ancient Patagonian populations (12, 14). A more recent study showed a high differentiation at the D-loop region of the mtDNA between ancient Kawéskar and Yámana, with a high frequency of D lineages, particularly D4h3a5, in the former and a higher frequency of C in the latter (17). Furthermore, lineages B2i2, C1b13, and D1g have been identified in present-day individuals from Southern Chile and Argentina. The coalescent time and distribution of these lineages suggest a rapid coastal migration through America and, subsequently, extensive trans-Andean migration in Southern South America (18–20).

Genome-wide data of modern individuals from Patagonia have been analyzed in the context of the peopling of the Americas, strongly suggesting that the populations from the area descend from a single major migration wave into the continent (21, 22). Additionally, the study of complete genomes from ancient Patagonian individuals shows no closer affinity between these and modern Australo-Melanesian populations (21), as had been suggested by craniometrical analysis (23). However, the information obtained from ancient or present-day genomes from peoples in Patagonia has not been thoroughly analyzed to address critical questions about population dynamics within the region.

The analysis of nuclear data and, in particular, complete genomes from Patagonian individuals from pre- and post-European contact is a valuable strategy to explore further and discuss the microevolutionary processes in the area, usually disregarded in

**Table 1.** Description and sequence data statistics for the four ancient Patagonian hunter-gatherers analyzed in this study

Sample ID	IPK12	IPK13	IPY08	IPY10
Territory	Kawéskar	Kawéskar	Yámana	Yámana
Coordinates	53°21'0"S 70°58'0"W	49°8'23"S 74°27'7"W	55°15'S 69°01'W	54°56'00"S 67°37'00"W
Date*	1,000 ± 30 BP† (Cal BP 930–795)	1,320 ± 30 BP‡ (Cal BP 1,275–1,110)	No date	910 ± 30 BP§ (Cal BP 895–725)
δ <sub>13</sub> C	−15.8	−10.8	No data	−12.4
δ <sub>15</sub> N	17.5	14.9	No Data	16.7
Total reads	636,310,871	1,047,656,357	911,455,297	845,604,155
Endogenous content	63.8%	24.6%	10.2%	59.6%
Average depth	7.8	3.5	1.7	9.1
Average depth mtDNA	151.8	148.9	164.1	164.7
Morphological sex estimation	Female	No data	No data	Male
Molecular sex estimation	Female	Male	Male	Male
Contamination percentage mtDNA	1.6	0.6	0.3	0.6
Contamination percentage X-chr.	—	0.7	1.4	0.8

\*Calibrated radiocarbon years (cal BP) with two-sigma calibration on parenthesis.

†Beta-413911.

‡Beta-413912.

§Beta-413909.

**Table 2. New modern populations genotyped in this study**

Population	Individuals	Location
Pehuenche	17	Trapa-Trapa, Alto Bío Bío. Region Biobío
Huilliche	20	La Misión, San Juan de la Costa. Region Los Lagos
Kawéskar	4	Punta Arenas. Region of Magallanes and Antártica Chilena
Yámana	20	Puerto Williams, Isla Navarino. Region of Magallanes and Antártica Chilena
Total	61	

favor of a continental picture. In this work, we present the analysis of four ancient genomes from maritime hunter-gatherer individuals dated between  $910 \pm 30$  and  $1,320 \pm 30$  YBP and genotyped data from 61 modern individuals from Central-Southern Chile and Patagonia. The joint analysis of these datasets reveals important aspects of the population structure in the area, the continuity of the native populations in Patagonia, and the relationship between the maritime hunter-gatherer groups and Selk'nam, a terrestrial hunter-gatherer group from Tierra del Fuego. In addition, we discuss the contribution of the ancient genomic data to elucidate the origin of late maritime hunter-gatherer groups.

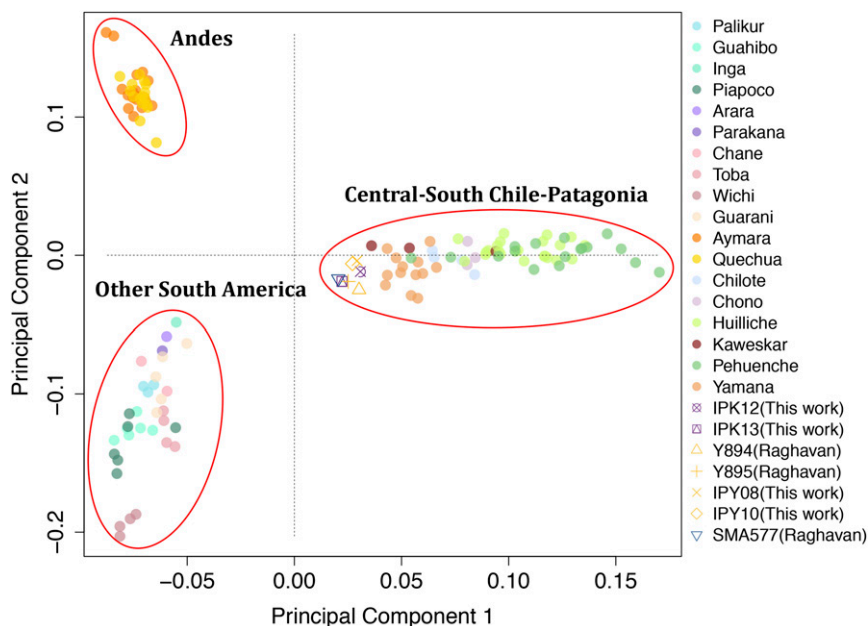
## Results

We sequenced the genomes of four ancient maritime hunter-gatherers from Patagonia using Illumina HiSeq 2000 in single-read mode (100 cycles). The geographical locations of these individuals agreed with the historical distribution of Kawéskar and Yámana and were, therefore, associated to ancient Kawéskar ( $n = 2$ ) and ancient Yámana ( $n = 2$ ) groups (Fig. 1). Isotopic results for three individuals were available, showing high nitrogen values, in accordance with a marine diet. Between 10% and 63.8% of the reads from the sequenced libraries mapped to the reference human genome. The assembled genomes reached depths of coverage between 1.7- and 9.1-fold (Table 1). The DNA fragmentation and misincorporation patterns were consistent with postmortem damage (SI Appendix, Fig. S1), and the contamination estimates from mtDNA (24) and X chromosome (25) were all <1.6% and 1.4%, respectively, attesting to the endogenous nature of the data (Table 1). Morphological sex estimation was available for two individuals (IPK12 and IPY10; male and female, respectively), and it was consistent with the sex

estimation based on the ratio of reads mapping to chromosome X and Y (26). The other samples (IPK13 and IPY08), which lacked a morphological assignment, were found to be male in both cases (Table 1). In addition, we obtained complete mitochondrial genomes for all of the samples, allowing the identification of the lineages C1b (IPK12 and IPY10), D1g (IPY08), and D4h3a5 (IPK13)—all common in the region (SI Appendix, Fig. S2). The Y chromosome haplogroups were identified as Q1a2a1a1-M3 for the three male individuals (Samples and Methods).

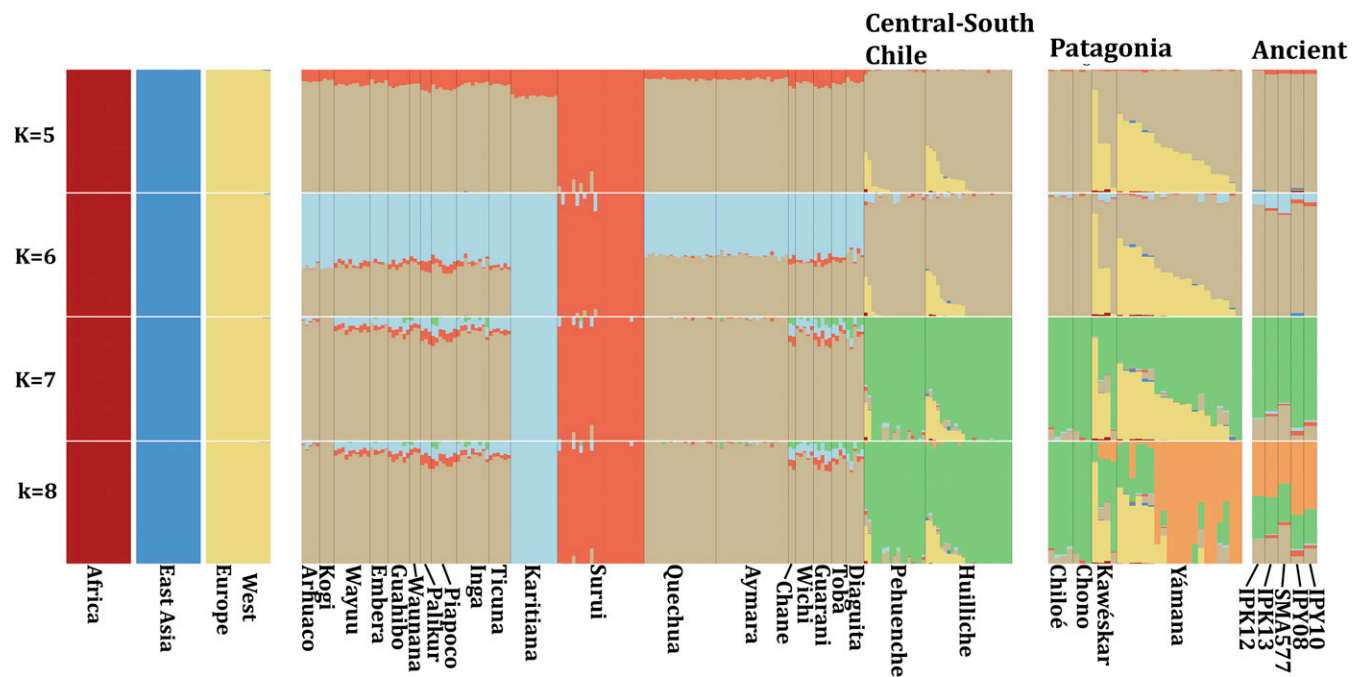
We evaluated the genetic affinities and population structure of the samples through principal components analysis (PCA). This analysis was performed by using a modern reference panel that included published data from 2,315 individuals (compiled by 22) and 61 newly genotyped individuals from Central-Southern Chile and Patagonia (Table 2). After merging the data, 93,465 common positions were kept. In addition, we included genomic data from three historical individuals from Patagonia (21): one Selk'nam (SMA577) and two Yámana (Y894 and Y895). As expected, the ancient samples from Patagonia clustered together with other South American populations (SI Appendix, Fig. S3). When only South American populations were considered (Fig. 2), the first two PCs showed three main clusters represented by: (i) Andean populations (Quechua and Aymara), (ii) Central-Southern Chile and Patagonia, and (iii) the remaining South American groups. Interestingly, the ancient Patagonians clustered close to each other and all fell in cluster two, specifically in proximity to modern Yámana and Kawéskar individuals.

We used the maximum-likelihood-based algorithm ADMIXTURE (27) to further characterize the genetic ancestry of the ancient and modern samples (SI Appendix, Figs. S4 and S5). Assuming four ancestral populations ( $K = 4$ ), we observed the



**Fig. 2.** PCA of ancient and present-day South American populations. All of the ancient individuals were projected onto the first two PCs by using the lsq option from smartpca.





**Fig. 3.** Ancestry proportions of ancient individuals ( $n = 5$ ; right bars) and selected modern individuals from African, Eurasian, and South American populations at  $K = 5-8$ . The analysis was performed by using the ADMIXTURE program (26). Individuals are displayed by population, with the South Americans sorted by latitude (north to south).

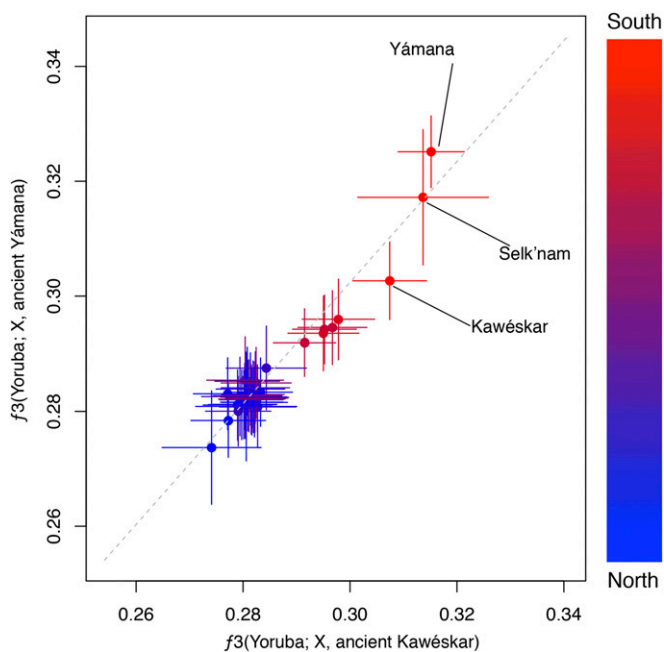
expected structure within South America, where there is a common ancestral component for all of the Native American populations (light brown). The lowest cross-validation error (*SI Appendix, Fig. S5*) was observed at  $K = 5$ , representing main continental groups and Surui. Further structure in South America and particularly in the Southern part was seen from  $K = 7$ , where one of the components (light green) was represented mainly in Central-Southern Chile (Pehuenche and Huilliche) and Patagonia (Chiloé, Kawéskar, and Yámana) (Fig. 3). At  $K = 8$ , a new component (light orange) was identified with a higher proportion in modern Yámana. Interestingly, approximately half the ancestry of the ancient individuals belonged to this component (higher in IPY08 and IPY10) and  $\sim 30\%$  to the one in Central-Southern Chile. Together, these components account for  $\sim 65-84\%$  of the total ancestry of the ancient individuals. Other ancestries in the ancient individuals included a major component found in South American populations (light brown) and the one present in the Andes (Aymara and Quechua) at  $K = 9$  (light purple), the latter found only in ancient Kawéskar (IPK12 and IPK13) and Selk'nam (SMA577). This component was also found in some modern individuals from Patagonia, reaching a maximum of  $\sim 9\%$ .

To measure genetic affinity, we calculated an outgroup- $f_3$  statistic (28) of the form (outgroup; X, ancient) to evaluate the shared genetic drift of the ancient individuals with modern populations from the Americas. The analysis performed grouping the ancient samples by territory (Kawéskar and Yámana) showed a clear latitudinal cline pattern of affinities increasing from north to south (Fig. 4). The closest relationships were observed between the ancient maritime groups with Selk'nam and the modern population from Southern Chile and Patagonia. Both ancient maritime groups had similar affinities with modern populations from South America and Selk'nam. However, modern Kawéskar and Yámana individuals seemed closer to ancient Kawéskar and Yámana, respectively. The analysis by individual did not affect these results, showing a similar pattern (*SI Appendix, Fig. S6*).

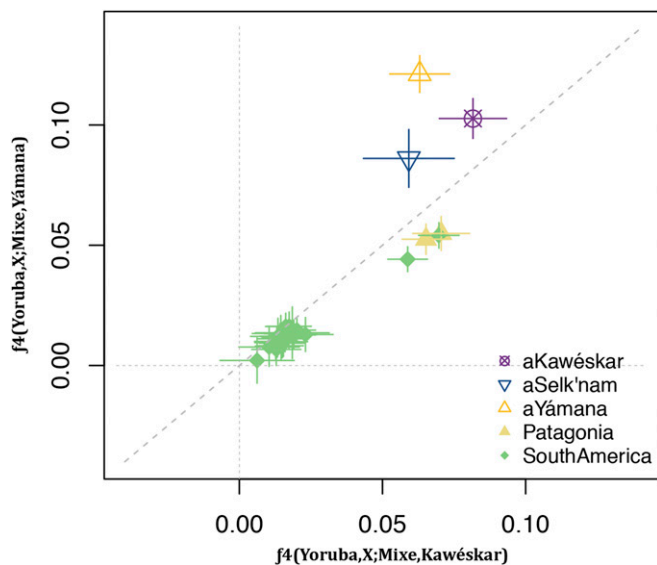
Likewise, a  $f_4$  analysis of the form (Yoruba, X; modern Patagonian, Mixe) confirmed that the ancient individuals from Patagonia were closer to modern Yámana and Kawéskar than to

any other modern population (Fig. 5). However, the closest proximity was found among all of the ancient individuals and modern Yámana, while modern Kawéskar seemed closer to other modern populations from Chiloé and Central-South Chile than modern Yámana.

To further explore the relationship between the ancient individuals, we calculated D-statistics using data from all of the available ancient complete genomes from Patagonia. These results



**Fig. 4.** Shared genetic drift between ancient Patagonian groups and modern population from South America. Bars indicate 95% confidence intervals. Color code represents the distance from the ancient groups.



**Fig. 5.**  $f_4$  analysis of the form (Yoruba, X; Mixe, modern Patagonian). Bars indicate 95% confidence intervals.

showed more clearly the affinity between the ancient individuals within the Kawéskar and Yámana geographical distributions, respectively. Furthermore, low coverage genomes from historical Kawéskar (K74) and Yámana (Y894 and Y895) were consistently closer to the  $\sim 1,000$ -y-old Kawéskar and Yámana individuals, respectively (*SI Appendix, Fig. S7*).

To further explore the population history of Patagonia, we estimated the fraction of the genomes under runs of homozygosity (ROHs). The distribution of the ROHs are reflective of past demographic events, where long contiguous homozygous segments are indicative of recent endogamy, and a short ROH is the product of old manifestations of small ancestral population size (29, 30). To ensure the quality of diploid genotype calling, only the ancient genomes with the highest coverage (IPK12 and IPY10) were included in the analysis. These were combined with 11 modern genomes from the Americas (21) and 90 modern individuals from the 1000 Genomes Project (31), keeping only common positions and transversions (*SI Appendix, Table S1*). Consistent with previous reports (32, 33), Native Americans showed a clear excess of ROH (Fig. 6), compared with non-American populations. This excess was mostly represented by short ROHs measuring hundreds of kilobases to several megabases (*SI Appendix, Fig. S8*). Longer ROHs were also found in some modern individuals from America, suggesting recent parental relatedness as a consequence of the bottleneck post-European contact. The variation between American populations has been associated to the different demographic histories within the continent precolonization and postcolonization period. In particular, the population size varies widely along America, with regions such as the Andes and current Mexico displaying higher population density than Patagonia, which could explain the variation we saw in our data (34–36). Both modern and ancient Patagonian individuals showed values slightly lower in short categories, with no representation in the longest ROH classes ( $>5$  Mb).

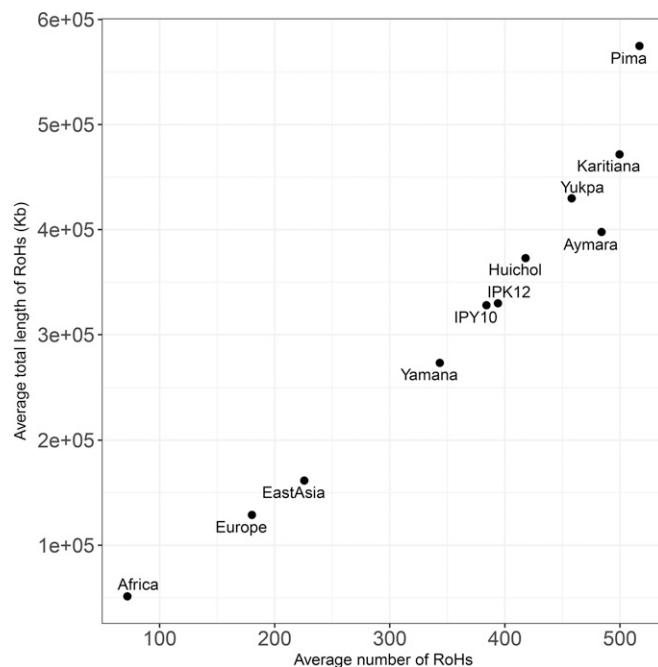
## Discussion

The native populations from Patagonia were an early focus of studies in the context of the peopling of the Americas because of their old presence in the area and their biological diversity. Initially, the skull morphology from Patagonians suggested a closer affinity with an early migration wave into the Americas who shared some attributes with present-day Australo-Melanesians (21, 37). However, genomic analysis of modern and ancient individuals from Patagonia so far do not reveal any affinity with Australo-Melanesian groups, instead showing a closer genetic proximity with other Native Ameri-

can populations (21). The analysis of ancient and modern genome-wide data generated in this work is in agreement with this observation, reflecting a closer relationship with other Native American populations, particularly Southern South Americans. Both late maritime (this work) and historical individuals (21) consistently showed a closer affinity between them and with modern native populations from Central-Southern Chile and Patagonia.

Furthermore, the analysis of the genetic structure of both ancient and modern Patagonians allowed us to identify a clear structure within South America and differentiation between some major geographical areas (e.g., Southern South America and the Andes). Populations from Central-Southern Chile (Pehuenche and Huilliche) and Patagonia clustered together and suggested the presence of an ancestral genetic component that is not shared with or has little contribution in other South American populations, reflecting the differentiation processes along South America after its occupation. Similarly, an additional ancestral component associated almost exclusively to modern Yámana ( $K = 8$ ) was most likely the result of the isolation of this population in the area, as seen in other groups in America (e.g., Surui and Karitiana). Different sources of evidence suggest a demographic history of Patagonia that is consistent with the isolation and small population sizes of these groups, where there is a greater effect of genetic drift. The analysis of ROHs from modern Yámana and ancient individuals suggested a small ancestral population size through the excess of short ROH categories compared with reference populations outside of America (*SI Appendix, Fig. S8*).

Within Patagonia, the ancestry and genetic affinities of modern Kawéskar may suggest some differences compared with Yámana. When analyzing only the Native American genetic ancestry, Kawéskar individuals showed a greater affinity with northern groups from the area of Chiloé and Central-Southern Chile (outgroup- $f_3$  and  $-f_4$  analyses). The region traditionally associated to Kawéskar encompassed a wide area from the Gulf of Penas to the Brecknock Peninsula (Fig. 1). From a linguistic perspective, at least three different dialects were identified within this region, suggesting some differentiation at the population level (38). The two ancient Kawéskar individuals, IPK12 and IPK13, included in this work were geographically associated to the north and central groups, respectively. Our analysis suggested a slight difference between them, with IPK13 showing a greater affinity to ancient Yámana than



**Fig. 6.** Average length and number of ROHs in selected individuals.

IPK12. On the other hand, the ancestry of the modern Kawéskar individuals suggested an interesting link with populations from Chiloé. Further structure within Kawéskar or later migrations from the area of Chiloé could explain this relationship, considering the intense movement and relocation of native populations post-European contact (39). In addition, earlier contact between Kawéskar with northern groups such as Chono is also a possibility.

We demonstrated a clear affinity between the ancient maritime individuals and modern populations from Central-Southern Chile and Patagonia. In particular, their ancestral composition is mainly associated to both of the components identified in Southern South America (from  $K = 4$  and  $K = 7$ ), strongly reflecting the continuity of native populations in the area. Recently, Llamas et al. (40) argued in favor of a lack of continuity in mtDNA lineages in America due to the extinction of lineages and the sharp decrease in population size during the colonization period. The analysis of nuclear data from ancient and modern Patagonians carried out here showed a marked continuity in the region, confirming previous results by us using partial data of the mtDNA data in modern and ancient individuals (17, 18). A decrease in the genetic diversity of Native American populations is expected, given the reduction of population size, relocation, and changes in the traditional lifestyles. However, using a genomic approach, we did not see evidence of discontinuity in Patagonia.

By exploring the shared genetic drift between the ancient genomes and other populations through  $f$ -statistics, we identified a higher similarity between individuals belonging to each maritime group, reflecting some degree of genetic homogeneity within each area and differentiation between both maritime groups, in agreement with their linguistic differences (38). In addition, the observation that both maritime groups are equally related to the Selk'nam individual might be consistent with an early separation of this group. Historical and archaeological evidence support an early separation of the populations from Tierra del Fuego (terrestrial groups) after the formation of the Magellan Strait  $\sim 8,000$  YBP (2). Later,  $\sim 6,000$  YBP, the first evidence of a maritime adaptation were found in the Beagle Channel and on the western shore around the Otway Sea and Magellan Strait (4). The close relationship between both maritime groups analyzed in this work and their differentiation with Selk'nam is in accordance with this scenario, with an earlier separation of an ancestral population of Selk'nam and an ancestral maritime hunter-gatherer group, followed by the diversification of the maritime populations Kawéskar and Yámana.

Different disciplines have contributed to our understanding of the initial human occupation and dispersion in Patagonia. Our work is a comprehensive genomic characterization focus on this region, aiming to investigate the population structure and affinity of local populations. The analysis of genomics data to reconstruct human past history of restricted areas such as Patagonia is a valuable strategy to improve the representation of Native American diversity worldwide and to gain a better insight into the microevolutionary processes along America.

## Samples and Methods

**DNA Extraction and Library Preparation.** Ancient samples come from different archaeological context and were obtained from the Instituto de la Patagonia (SI Appendix, SI Material and Methods). The DNA was extracted with a modified silica spin-column protocol implemented by ref. 28. Double-stranded libraries were built following Meyer and Kircher (41). All these procedures were performed in a dedicated ancient DNA laboratory at the Centre for Geogenetics, University of Copenhagen (SI Appendix, SI Materials and Methods).

1. Dillehay TD (2009) Probing deeper into first American studies. *Proc Natl Acad Sci USA* 106:971–978.
2. Morello F, et al. (2012) Hunter-gatherers, biogeographic barriers and the development of human settlement in Tierra del Fuego. *Antiquity* 86:71–87.
3. Piana EL, Orquera LA (2009) The southern top of the world: The first peopling of Patagonia and Tierra del Fuego and the cultural endurance of the Fuegian Sea-Nomads. *Arctic Anthropol* 46:103–117.
4. Orquera LA, Legoupil D, Piana EL (2011) Littoral adaptation at the southern end of South America. *Quat Int* 239:61–69.

**Sequencing and Data Analysis.** The sequencing was done at the Stanford High-Throughput Sequencing Facility on a HiSeq 2000. Reads were processed with AdapterRemoval (Version 1.5.2) (42) before being mapped to the human reference genome build hg19 (CRCh37) by using BWA (Version 0.7.10) (43). DNA damage patterns were estimated and plotted with MapDamage2 (44). The biological sex of all four samples was estimated with the method described in ref. 26. Contamination was evaluated by using the X chromosome and mtDNA as described in refs. 25 and 24, respectively. Mitochondrial haplogroups were determined by using Haplogrep, Phylotree build (Version 17) (45, 46). Y-chromosome lineages were identified by comparison with the 1000-Y phylogeny dataset from ref. 47.

Modern data were generated with the Axiom LAT1 platform (Affymetrix) on historical collection of modern DNA samples at the Program of Human Genetics of the University of Chile. All of the participants were healthy individuals,  $>18$  y old, and gave informed consent. The study was approved by the Ethics Committee of the Faculty of Medicine, University of Chile. Considering the variable levels of nonnative ancestry, a version of this array masking the nonnative genotypes was produced by using Shapelt2 to phase (48) and then RFMix (49) to estimate local ancestry. Low-coverage genomes from two Yámana, included in the previous Axiom array, were sequenced and processed in the sequencing facility of the Faculty of Medicine, University of Chile (SI Appendix, SI Materials and Methods).

**Genetic Structure and Affinities.** The data were merged with published data from ref. 22 by using PLINK1.7 (50). In addition, ancient genomes from America were included: one individual associated to Clovis culture (51) and four historical individuals from Patagonia: one Selk'nam (SMA577), one Kawéskar (K74), and two Yámana (Y894 and Y895) (21). We performed a PCA using smartpca implemented in the package EIGENSOFT (52) and projecting the ancient individuals on the PC inferred from the modern populations. To characterize the ancestry composition of the samples, we used ADMIXTURE (Version 1.3) (27). Unmasked genotypes of the modern populations were used for the admixture analysis to explicitly assess the level of European admixture and inform the ancestral component of the individuals without bias. The software was run assuming 3–11 “ancestral” populations ( $K = 3$  to  $K = 11$ ) in 10 replicates, starting from a random seed and keeping the run with the highest likelihood. We used the default settings to perform a cross-validation procedure to evaluate the fit of each  $K$ . We performed an outgroup  $f_3$  analysis (28) in the qp3pop program from ADMIXTOOLS (53) using Yoruba as outgroup. The same software was used to run D-statistic analyses (qpDstat). In addition, full genomes from all of the ancient samples, including four historical samples of Kawéskar ( $n = 1$ ), Yámana ( $n = 2$ ), and Selk'nam ( $n = 1$ ) (21), were used to estimate the D-statistic, as implemented in ANGSD (25). ROHs were estimated for the 2 samples with highest coverage (IPK12 and IPY10) together with 11 modern genomes from the Americas (SI Appendix, Table S1 and ref. 21) and 90 individuals with European ( $n = 30$ ), East Asian ( $n = 30$ ), and African ( $n = 30$ ) ancestry from the 1000 Genomes Project (31). Two of the modern American individuals were low-coverage genomes of Yámana sequenced in this project (SI Appendix, Table S1). This analysis was implemented in PLINK1.7 (50) as described in ref. 30.

**ACKNOWLEDGMENTS.** We thank the contribution of the “Instituto de la Patagonia” and Stanford Sequencing Center. We thank Consuelo Quinto for helpful comments on earlier versions of this manuscript. This work was supported by Comisión Nacional de Investigación Científica y Tecnológica Grant USA2013-0015 and Fondo Nacional de Ciencia y Tecnología Grants 1140544 and 1170726. M.C.Á.-A.'s laboratory is supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica–Universidad Nacional Autónoma de México Grant IA206817. A.-S.M. and D.C.D. were supported by Swiss National Science Foundation Grant PZ00P3\_154717 and European Research Council Starting Grant 679330. A.M.-E. was supported by International Center for Genetic Engineering and Biotechnology Grant CRP/MEX15-04\_EC. Sequencing of modern DNA was partially funded by Fund of Scientific and Technological Equipment (FONDEQUIP) Grant EQM140157.

5. San Román M (2014) Sea-level changes and coastal peopling in southernmost Pacific South America: Marine hunters from Patagonia. *Encyclopedia of Global Archaeology*, ed Smith C (Springer, New York), pp 6515–6525.
6. Morello F, et al. (2012) Arqueología de la Punta Santa Ana: Reconstrucción de secuencias de ocupación de cazadores-recolectores marinos del Estrecho de Magallanes, Patagonia. *Magallania (Punta Arenas)* 40:129–149.
7. Morello F, Stern C, San Román M (2015) Obsidiana verde en Tierra del Fuego y Patagonia: Caracterización, distribución y problemáticas culturales a lo largo del Holoceno. *Intersecciones Antropol* 6:139–153.



8. Chapman A (1986) *Los Selk'nam. La Vida de los Onas* (Emecé, Buenos Aires).
9. Gusinde M (1986) *Los Indios de Tierra del Fuego. Tomo II (1,2 y3) (Yámanas)* (CAEA, Buenos Aires).
10. Gusinde M (1991) *Los Indios de Tierra del Fuego. Tomo III (1 y 2) (Los Halakwulup)* (CAEA, Buenos Aires).
11. Lothrop SK (1928) *The Indians of Tierra del Fuego. Contribution of the Museum of the American* (Heye Foundation, Washington, DC).
12. Lalueza C, Pérez-Pérez A, Prats E, Cornudella L, Turbón D (1997) Lack of founding Amerindian mitochondrial DNA lineages in extinct aborigines from Tierra del Fuego-Patagonia. *Hum Mol Genet* 6:41–46.
13. González José R, García-Moro C, Dahinten S, Hernández M (2002) Origin of Fuegian-Patagonians: An approach to population history and structure using R matrix and matrix permutation methods. *Am J Hum Biol* 14:308–320.
14. García-Bour J, et al. (2004) Early population differentiation in extinct aborigines from Tierra del Fuego-Patagonia: Ancient mtDNA sequences and Y-chromosome STR characterization. *Am J Phys Anthropol* 123:361–370.
15. Moraga M, De Saint Pierre M, Torres F, Ríos J (2010) Vínculos de parentesco por vía materna entre los últimos descendientes de la etnia Kawésqar y algunos entierros en los canales patagónicos: Evidencias desde el estudio de linajes mitocondriales. *Magallania (Punta Arenas)* 38:103–114.
16. Perez SI, Bernal V, Gonzalez PN (2007) Morphological differentiation of aboriginal human populations from Tierra del Fuego (Patagonia): Implications for South American peopling. *Am J Phys Anthropol* 133:1067–1079.
17. de la Fuente C, et al. (2015) Ancient marine hunter-gatherers from Patagonia and Tierra del Fuego: Diversity and differentiation using uniparentally inherited genetic markers. *Am J Phys Anthropol* 158:719–729.
18. de Saint Pierre M, et al. (2012) An alternative model for the early peopling of southern South America revealed by analyses of three mitochondrial DNA haplogroups. *PLoS One* 7:e43486.
19. de Saint Pierre M, et al. (2012) Arrival of Paleo-Indians to the southern cone of South America: New clues from mitogenomes. *PLoS One* 7:e51311.
20. Bodner M, et al. (2012) Rapid coastal spread of first Americans: Novel insights from South America's southern cone mitochondrial genomes. *Genome Res* 22:811–820.
21. Raghavan M, Steinrücken M, Harris K, et al. (2015) Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349:aab3884.
22. Reich D, et al. (2012) Reconstructing native American population history. *Nature* 488:370–374.
23. Pucciarelli HM (2004) Migraciones y variación craneofacial humana en América. *Complutum* 15:225–247.
24. Fu Q, et al. (2013) A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* 23:553–559.
25. Korneliusen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics* 15:356.
26. Skoglund P, Storå J, Götherström A, Jakobsson M (2013) Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J Archaeol Sci* 40:4477–4482.
27. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19:1655–1664.
28. Raghavan M, et al. (2014) Upper Palaeolithic Siberian genome reveals dual ancestry of native Americans. *Nature* 505:87–91.
29. Cassidy LM, et al. (2016) Neolithic and bronze age migration to Ireland and establishment of the insular Atlantic genome. *Proc Natl Acad Sci USA* 113:368–373.
30. Gamba C, et al. (2014) Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun* 5:5257.
31. Auton A, et al.; 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526:68–74.
32. Pemberton TJ, et al. (2012) Genomic patterns of homozygosity in worldwide human populations. *Am J Hum Genet* 91:275–292.
33. Broushaki F, et al. (2016) Early Neolithic genomes from the eastern Fertile Crescent. *Science* 353:499–503.
34. Kirin M, et al. (2010) Genomic runs of homozygosity record population history and consanguinity. *PLoS One* 5:e13996.
35. Moreno-Estrada A, et al. (2014) The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* 344:1280–1285.
36. Goldberg A, Mychajliw AM, Hadly EA (2016) Post-invasion demography of prehistoric humans in South America. *Nature* 532:232–235.
37. Skoglund P, Reich D (2016) A genomic view of the peopling of the Americas. *Curr Opin Genet Dev* 41:27–35.
38. Viegas Barros JP (2005) *Voces en el Viento: Raíces Lingüísticas de la Patagonia* (Mondragón Ediciones, Buenos Aires).
39. Gusinde M (1951) *Hombres Primitivos de la Tierra del Fuego* (Escuela de Estudios Hispano-Americanos de Sevilla, Sevilla, Spain).
40. Llamas B, et al. (2016) Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci Adv* 2:e1501385.
41. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010:pbp05448.
42. Lindgreen S (2012) AdapterRemoval: Easy cleaning of next-generation sequencing reads. *BMC Res Notes* 5:337.
43. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.
44. Jónsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29:1682–1684.
45. Weissensteiner H, et al. (2016) HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res* 44:W58–W63.
46. van Oven M (2015) PhyloTree Build 17: Growing the human mitochondrial DNA tree. *Forensic Sci Int Genet Suppl Ser* 5:e392–e394.
47. Poznik GD, et al.; 1000 Genomes Project Consortium (2016) Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nat Genet* 48:593–599.
48. Delaneau O, Marchini J, Zagury J-F (2011) A linear complexity phasing method for thousands of genomes. *Nat Methods* 9:179–181.
49. Maples BK, Gravel S, Kenny EE, Bustamante CD (2013) RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet* 93:278–288.
50. Purcell S, et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575.
51. Rasmussen M, et al. (2014) The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* 506:225–229.
52. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2:e190.
53. Patterson N, et al. (2012) Ancient admixture in human history. *Genetics* 192:1065–1093.