

# SCIENTIFIC REPORTS

OPEN

## First complete mitochondrial genome data from ancient South American camelids - The mystery of the *chilihueques* from Isla Mocha (Chile)

Michael Westbury<sup>1</sup>, Stefan Prost<sup>2,3</sup>, Andrea Seelenfreund<sup>4</sup>, José-Miguel Ramírez<sup>5</sup>, Elizabeth A. Matisoo-Smith<sup>6</sup> & Michael Knapp<sup>6</sup>

Received: 27 July 2016  
Accepted: 10 November 2016  
Published: 08 December 2016

In South American societies, domesticated camelids were of great cultural importance and subject to trade and translocation. South American camelids were even found on remote and hard to reach islands, emphasizing their importance to historic and pre-historic South American populations. Isla Mocha, a volcanic island 35 km offshore of Central-South Chile, is an example of such an island. When Dutch and Spanish explorers reached the island in the early 17th century, they found that domesticated camelids called "*chilihueque*" played a major role in the island's society. The origin and taxonomy of these enigmatic camelids is unclear and controversial. This study aims to resolve this controversy through genetic analyses of Isla Mocha camelid remains dating from pre-Columbian to early historic times. A recent archaeological excavation of site P21-3 on Isla Mocha yielded a number of camelid remains. Three complete mitochondrial genomes were successfully recovered and analysed. Phylogenetic analyses suggest that "*chilihueque*" was a local term for a domesticated guanaco. Results from phylogeographic analyses are consistent with Isla Mocha camelids being sourced from Southern Chilean guanaco populations. Our data highlights the capability of ancient DNA to answer questions about extinct populations which includes species identity, potential translocation events and origins of founding individuals.

The translocation of animals to islands has been an important human adaptive strategy even before the development of domestication<sup>1</sup>. Live animals are transported for numerous reasons including as future food resources, for their by-products, which can include bones, teeth, fibre, or dung, or for their cultural/spiritual or ritual importance. Often sizable challenges and obstacles had to be overcome to transport larger animals across considerable water gaps, emphasizing the cultural importance of these translocated animals. The enigmatic camelids of Isla Mocha, an island located approximately 35 km off the coast in the Arauco region of South Central Chile, are an example of such a challenging translocation of animals.

Isla Mocha has not been connected to the mainland since it rose above sea level during the Pleistocene, and it has no native large land mammals<sup>2,3</sup>. The earliest human occupation of the island is dated to about 3500 years ago and evidence of permanent occupation dates from about 1500 years ago<sup>4,5</sup>. Archaeological excavations on Isla Mocha have provided a large number of bone artifacts made of large mammal bones such as cetaceans, pinnipeds and camelids. Camelid bones make up the majority of the mammal remains in archaeological sites on the island.

<sup>1</sup>Department of Mathematics and Natural Sciences, Evolutionary and Adaptive Genomics, Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany. <sup>2</sup>Department of Integrative Biology, University of California Berkeley, 3040 Valley Life Sciences Building, Berkeley, CA 94720-3140, USA. <sup>3</sup>Department of Biology, Stanford University, 371 Serra Street, Palo Alto, CA 94305-5020, USA. <sup>4</sup>School of Anthropology, Faculty of Social Sciences, Universidad Academia de Humanismo Cristiano, Condell 506, Santiago 7500828, Chile. <sup>5</sup>Centro de Estudios Avanzados, Universidad de Playa Ancha, Traslaviña 450, Viña del Mar, Chile. <sup>6</sup>Department of Anatomy, University of Otago, 270 Great King Street, Dunedin 9016, New Zealand. Correspondence and requests for materials should be addressed to M.K. (email: michael.knapp@otago.ac.nz)

Sample	Site	Unit	Level	Average coverage per base site	Number of mapping reads	Percent of genome covered	GenBank reference number
Mocha04	P21-3	100	A	85.80x	12549	98.64%	KX388532
Mocha05	P21-3	100	150–160 cm	3.01x	528	89.56%	KX388533
Mocha06	P21-3	100a	190–200 cm	28.50x	3970	98.43%	KX388534

**Table 1.** Excavation site and sequencing details for each successfully sequenced sample from Isla Mocha.

The large numbers of camelid remains present in archaeological sites on Isla Mocha suggest animals were likely living there from the time of permanent occupation. It is generally accepted that they were culturally important, used for ritual sacrifice, with the meat only consumed on special occasions<sup>2</sup>. Thus, live animals had to be translocated from the mainland to the island. The 35 km crossing from the mainland to Isla Mocha is a dangerous one, particularly in the winter time. Modern kayakers claim that the crossing from Tirua, the closest launching site on the mainland, to the island takes approximately 7 hours<sup>2</sup>. Transporting live camelids to Isla Mocha would have been particularly difficult and thus may have been a rare occurrence.

The origin and taxonomy of the Mocha camelids has been a matter of debate<sup>2</sup>. Early European descriptions of the island's resources, including that of van Speilbergen in 1615, describe the presence of sheep or odd looking camel like creatures: "sheep of a very wonderful shape, having a very long neck and a hump like a camel, a hare lip and very long legs"<sup>6</sup>. These animals were referred to as "*chilihueque*"<sup>6</sup>.

Today, South American camelids are classified as belonging to four species within two genera, the wild forms, guanaco (*Lama guanicoe*) and vicuña (*Vicugna vicugna*), and the domestic species, llama (*L. glama*), and alpaca (*V. pacos*)<sup>7</sup>. How "*chilihueques*" fit into this taxonomy is unclear. Bonacic (1991) described them as a group of camelids that existed during Pre-Hispanic and early colonial times in central and south-central Chile, which went extinct during the 16th or 17th century, likely due to the arrival of the European horse and sheep<sup>8</sup>. The author suggested that "*chilihueque*" was either a local name for a different form of the llama that was brought down from Northern Chile or that "*chilihueques*" were a sub-species of locally domesticated guanaco<sup>8</sup>. Gay<sup>9</sup>, on the other hand, referred to "*chilihueque*" simply as the Mapuche name for a domesticated guanaco without describing them as a separate taxonomic unit.

Here we use next generation sequencing technology coupled with DNA bait capture enrichment to obtain complete mitochondrial genomes of camelid remains recovered from archaeological excavations on Isla Mocha in order to identify the camelid species translocated to Isla Mocha and locate the likely origin of the potential source population.

## Results

**DNA preservation, recovery and authenticity.** A total of 14 samples were excavated and DNA was extracted from all samples. Ion Torrent sequencing libraries were prepared from all sample DNA extracts alongside three extraction blanks and a library preparation blank. All libraries including blank libraries were quantified using quantitative PCR with Ion Torrent adapter specific primers and checked for the presence of adapter dimers on an agarose gel. Of the 14 excavated samples from Isla Mocha, six samples yielded sequenceable libraries free of adapter dimers and of those, three libraries successfully recovered nearly complete camelid mitochondrial genome data (Table 1). All blanks produced alongside the successfully sequenced samples Mocha04, Mocha05, Mocha06 showed no detectable PCR products when amplified with a camelid specific primer test set targeting a 60 bp fragment of the camelid mitochondrial control region.

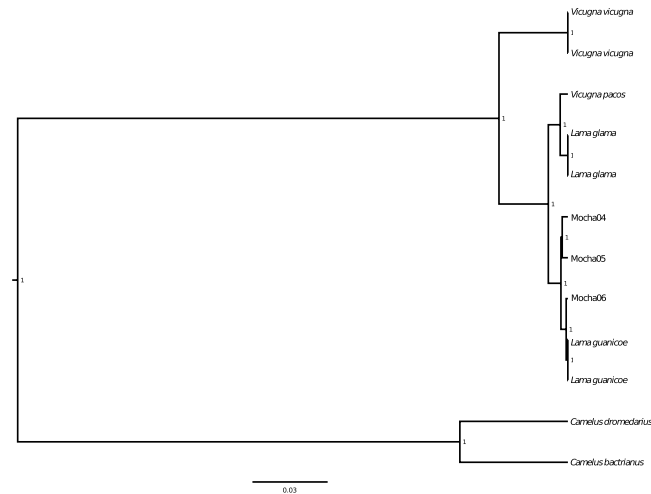
To further evaluate the authenticity of the obtained DNA sequence data, we inspected sequencing reads for characteristic deamination damage. The analysis with mapDamage showed damage patterns indicative of ancient DNA with the presence of increased C–T and G–A mutations at the respective 5' and 3' ends of sequenced molecules (see Supplementary Fig. S1). The average fragment length was 85–100 bp and therefore rather short, also consistent with ancient DNA (see Supplementary Fig. S2).

Ion Torrent sequencing reads from each sequencing library were mapped against both a *L. guanicoe* and a *V. vicugna* reference mitochondrial genome to identify potential ascertainment bias. In each case a higher coverage was achieved when mapping against *L. guanicoe* but both consensus sequences from the three samples respectively were identical independent of whether the sequencing reads were mapped against guanaco or vicuña. Phylogenetic analyses were conducted with the consensus sequences mapped to the guanaco due to the higher coverage mappings.

**Phylogenetic and phylogeographic analyses.** Bayesian phylogenetic reconstructions yielded a well-resolved tree topology with high posterior probabilities on all branches. Complete mitochondrial genome sequences of all three sequenced camelid samples from Isla Mocha grouped unambiguously with *L. guanicoe* (Fig. 1).

Phylogeographic analyses were conducted on control region data only, with the 176 guanaco control regions published by Marin *et al.*<sup>10</sup> being used as reference data. All three Isla Mocha camelids had different control region haplotypes, and all Mocha haplotypes differed from the published reference data (Fig. 2).

Mocha04 was most closely related to and one mutation removed from a haplotype shared between guanaco found in 2 different regions of Southern South America: in Punta Choros of the Cuarta Region (now Coquimbo region), Chile and Parque Nacional Pan de Azúcar of the Tercera Region (now Atacama Region), Chile.



**Figure 1. Bayesian reconstruction of phylogenetic relationships of Isla Mocha camelids.** Tree reconstructed with complete mitochondrial genome data from all extant South American camelids with two *Camelus* species as outgroup. Numbers at the branch nodes indicate posterior probability support values. All three Mocha samples are shown to group together with the *Lama guanicoe* mitogenome sequences with high support values.

Isla Mocha sample, Mocha05, was just one substitution removed from a common and widespread control region haplotype that is shared between 13 different guanaco populations from 13 different areas of South America. These populations were those from “Sureste del Chaco Boliviano, Bolivia”, “Sector Bajada del Diablo, Trelew, Provincia de Chubut, Argentina”, “Minera Pelambres, Cuarta Región, Chile” (now Coquimbo region), “INTA Bariloche, Provincia de Neuquén, Argentina”, “Reserva Provincial La Payunia, Provincia de Mendoza, Argentina”, “Valle Chacabuco, Undécima Región (now Aysén Region), Chile”, “Reserva Nacional Río Cipreses, Sexta Región (now General L. Bernardo O’Higgins Region), Chile”, “Uspallata, Provincia de Mendoza, Argentina”, “Ovalle, Cuarta Región, Chile” (now Coquimbo region), “Isla Navarino, Duodécima Región, Chile” (now Magallanes and Chilean Antartica Region), “Parvenir, Tierra del Fuego, Duodécima Región, Chile” (now Magallanes and Chilean Antartica Region), “Sector Calafate, Tierra del Fuego, Duodécima Región, Chile” and “Parque Nacional Torres del Paine, Duodécima Región, Chile” (now Magallanes and Chilean Antartica Region).

Mocha06 was one substitution away from three different regional populations. These regional populations coming from “Valle Chacabuco, Undécima Región, Chile” (now Aysén Region), “INTA Bariloche, Provincia de Neuquén, Argentina”, “Parque Nacional Llanos de Challe, Tercera Región (now Atacama Region), Chile”.

All the above populations that are closely related to the Mocha camelids, with the exception of “Sureste del Chaco Boliviano”, are found south of 26°S (approximately the latitude of the border between Argentina and Bolivia) (Fig. 3). All guanaco populations from north of 20°S sequenced by Marin *et al.*<sup>10</sup> are more distantly related to the Mocha camelids and at least three mutations removed from any of the Mocha samples.

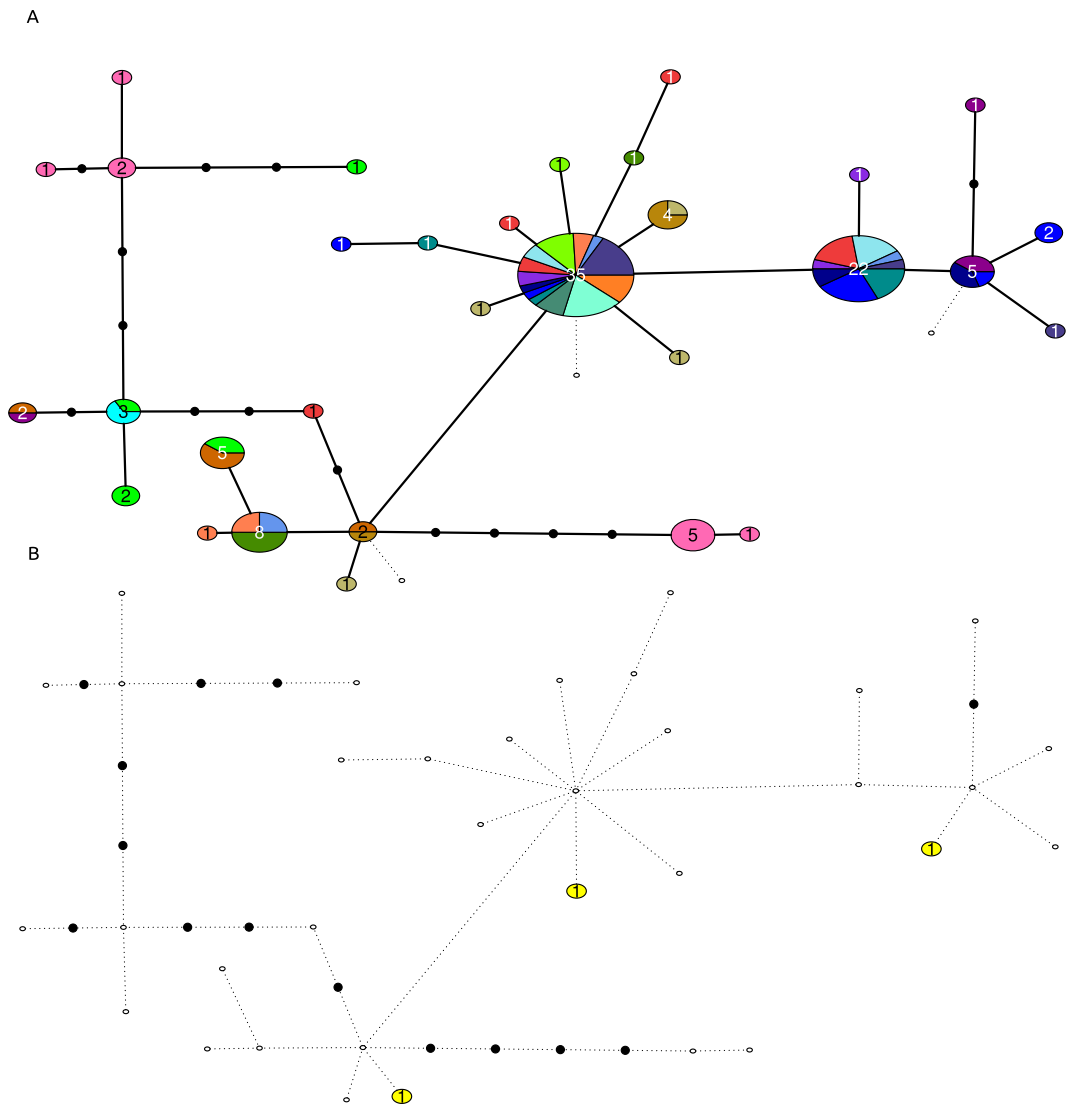
## Discussion

Our study helps to reconstruct the identity of the enigmatic camelids from Isla Mocha. Bayesian phylogenetic analyses, using the complete mitochondrial genomes of three ancient camelids grouped these individuals with modern *L. guanicoe*, to the exclusion of *Vicugna vicugna*, *Lama glama* and *Vicugna pacos* (Fig. 1). These results are consistent with previous analyses based on morphological data, which also identified the Isla Mocha camelids as guanacos<sup>2</sup>.

Identifying the exact geographic location of the source population of the Isla Mocha guanacos is challenging. Very limited complete mitochondrial data from guanaco populations from across South America is available to date. The only large-scale phylogeographic information for guanacos was based on the mitochondrial control region only<sup>10</sup>. Marín *et al.* discovered limited phylogeographic structure of guanacos across their range, but did find a well-defined separation between northern populations from Peru, Bolivia and Northern Chile and southern populations from Argentina and southern Chile. The only exception from this structure is one isolated population in the highlands of Bolivia, which groups with the southern haplotypes. The three Isla Mocha individuals sequenced in this study had unique and previously unsequenced haplotypes, but did group most closely with the southern populations.

These results suggest that the Isla Mocha camelids were most likely derived from southern South American guanaco populations rather than llamas brought in from the North, as suggested by the alternative hypothesis. The very limited phylogeographic structure of southern South American guanacos prevents a more precise determination of the geographic origins of the Isla Mocha camelids.

Together with morphological studies, our data provides some insights into the taxonomic status of the enigmatic “chilihueques” from Isla Mocha. All pre-historic and early historic camelid remains from Isla Mocha studied so far have been described as guanacos based on morphological evidence<sup>2,4,11</sup>. Similarly, our genetic data group the three individuals studied by us with guanacos. While the haplotypes of our samples are unique, they do not



### Key:

<span style="color: yellow;">■</span> - Mocha Island, Chile	<span style="color: lightgreen;">■</span> - Uspallata, Provincia de Mendoza, Argentina
<span style="color: cyan;">■</span> - <i>Lama glama</i>	<span style="color: darkgreen;">■</span> - Putaendo, Quinta Region, Chile
<span style="color: orange;">■</span> - Isla Navarino, Duodécima Region, Chile	<span style="color: red;">■</span> - Minera Pelambres, Cuarta Region, Chile
<span style="color: lightblue;">■</span> - Porvenir, Tierra del Fuego, Duodécima Region, Chile	<span style="color: blue;">■</span> - Ovalle, Cuarta Region, Chile
<span style="color: darkblue;">■</span> - Sector Calafate, Tierra del Fuego, Duodécima Region, Chile	<span style="color: brown;">■</span> - Punta Choros, Cuarta Region, Chile
<span style="color: teal;">■</span> - Parque Nacional Torres del Paine, Duodécima Region, Chile	<span style="color: olive;">■</span> - Vallenar, Tercera Region, Chile
<span style="color: darkred;">■</span> - Valle Chacabuco, Undécima Region, Chile	<span style="color: purple;">■</span> - Parque Nacional Llanos de Challe, Tercera Region, Chile
<span style="color: darkblue;">■</span> - INTA Bariloche, Provincia de Neuquen, Argentina	<span style="color: brown;">■</span> - Parque Nacional Pan de Azúcar, Tercera Region, Chile
<span style="color: purple;">■</span> - Sector Bajada del Diablo, Trelew, Provincia de Chubut, Argentina	<span style="color: darkblue;">■</span> - Sureste del Chaco, Bolivia
<span style="color: red;">■</span> - Reserva Provincial La Payunia, Provincia de Mendoza, Argentina	<span style="color: lightgreen;">■</span> - Putre, Primera Region, Chile
<span style="color: lightblue;">■</span> - Reserva Nacional Río Cipreses, Sexta Region, Chile	<span style="color: pink;">■</span> - Huayhua, Departamento de Ayacucho, Peru

**Figure 2. Haplotype network constructed using mitochondrial control regions of extant guanaco, llama and Isla Mocha camelids.** Each node represents a single mutation; colours correspond to regions of origin. (A) Present day camelids sequenced by Marín *et al.*<sup>10</sup>. (B) Ancient Isla Mocha camelids from this study in relation to Marín *et al.*<sup>10</sup> camelids. Isla Mocha samples are shown to fall into three separate parts of the network while all being distinct from any living guanaco or llama control region sequence.

form a monophyletic clade within the guanaco diversity as could for example be expected from a reproductively isolated subspecies. We therefore suggest that “*chilihueque*” was merely a local term for a guanaco. Whether or not the term “*chilihueque*” was specifically referring to domesticated guanacos, as suggested by Gay<sup>9</sup>, cannot



**Figure 3. Map of the genetically closest guanaco haplotype to our Isla Mocha camelids' sequences.** The geographic location of the genetically closest guanaco haplotype to Mocha04 is indicated in yellow, to Mocha05 is indicated in blue and to Mocha06 is indicated in red. Map generated using QGIS 2.0.1-Dufour (QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://www.qgis.org/>).

be inferred from our limited genetic data or the island's archaeology and remains a contentious question<sup>2</sup>. It is conceivable that the animals were transported to the island and released for hunting or that they were casually managed rather than fully domesticated.

Our data provide the first glimpse into the origin and taxonomy of the enigmatic Isla Mocha camelids or "*chili-hueques*", reinforcing the hypothesis that these camelids were guanacos derived from local mainland populations rather than domesticated llama populations brought in from further north. This result suggests a local sourcing of animals for translocation to the isolated Isla Mocha.

## Methods

In 2011, an archaeological excavation was undertaken on Mocha at site P21-3 located on Parcela 21 on the south-west site of the island. The site is located on a marine terrace approximately 20–30 m above sea level. Two units, 100 and 100a were identified. A total of 10 m<sup>3</sup> was excavated in unit 100a and total of 6.4 m<sup>3</sup> in unit 100. Excavations followed the natural stratigraphy of the sites. An area of 2 × 4 metres was excavated to a depth of 2 meters, in approximately 10 cm spits. P21-3 is believed to be an occupation site and midden linked to P21-1, a burial site located approximately 20 metres away and excavated in 2001 and 2003 by Daniel Quiroz and colleagues<sup>12</sup>. The P21-3 site contains evidence of continuous human occupation for at least 1,000 years. A number of intact cultural deposits containing a large amount of faunal, ceramic and lithic material from throughout the El Vergel complex (dating to 1,000–1,500 AD) and into the Pitren occupation period (0–600 AD) provide evidence for occupation over this time. Faunal remains recovered included fish, marine mammals, camelids, rodents, canids, and bird bones (unpublished data). The sequenced camelid remains were recovered from archaeological contexts at P21-3, with associated material dating from 689 ± 31 BP to 134 ± 20 BP (Wk 38379–38395). A total of 14 camelid samples were considered appropriate for ancient DNA analysis and were bagged and labeled for such purposes at the site.

**DNA extractions.** All ancient DNA extractions and sequencing library preparations were undertaken in a dedicated ancient DNA laboratory, at the University of Otago, Dunedin. 250 mg of bone material per sample was ground to a powder using a clean mortar and pestle then subsequently extracted using a silica based extraction protocol modified from Rohland and Hofreiter<sup>13</sup>. For every five samples extracted, one extraction blank was also processed under the same conditions. Three microliter aliquots of extraction blanks were amplified with a primer set (Supplementary Table S1), based on sequences from Marín *et al.*<sup>10</sup>, targeting a 60 bp fragment of the camelid mitochondrial control region to detect the presence of camelid DNA in extractions and blanks. The following thermal profile was used for this amplification: Initial denaturation step of 94 °C for nine minutes followed by 40



steps of 94 °C denaturation for 20 seconds, 55 °C annealing for 30 seconds, and 72 °C elongation for 30 seconds then a final elongation of 72 °C for 4 minutes and storage at 4 °C. PCR products were visually checked for amplification bands via agarose gel electrophoresis on a 2.5% agarose gel.

**Ion torrent library preparation.** Barcoded Ion torrent libraries were produced from all DNA extracts, extraction blanks and no-template controls by ligating ion torrent specific sequencing adapters to the ancient DNA extracts using a modified version of the ancient DNA library build protocol described by Knapp *et al.*<sup>14</sup>. Subsequent to library preparation, barcoded libraries were PCR amplified to produce easy to store stock libraries while at the same time retaining maximum library complexity<sup>14</sup>. To reduce amplification artifacts, the PCR was stopped after amplification plateau was reached, as determined by quantitative PCR.

Agarose gel electrophoresis (2.5% gel) of amplified libraries was used to visually check the success of the library preparation and to identify the presence of adapter dimers. Libraries with strong adapter dimers were excluded from sequencing as they strongly reduce the reads on target in a sequencing run. Only libraries that showed a characteristic smear of expected size (c. 150–300 bp) and no adapter dimer bands were used for downstream analyses. To check for potential cross contamination, no-template control libraries were PCR amplified with the same 60 bp camelid control region specific primers and the same thermal profile used to test for contamination of the extraction blanks (Supplementary Table S1). Three sequencing libraries passed all quality checks and were used for downstream analyses. Amplified libraries were purified using a Qiagen MinElute kit following the manufacturers' protocol.

**Hybridisation capture.** Sequencing libraries were enriched for mitochondrial DNA using hybridisation capture enrichment<sup>15</sup>. Two micrograms of each library were required for hybridisation capture. To obtain this amount of template library, four 1 µl aliquots of each stock library were PCR amplified in 100 µl reaction volume each with reactions again being terminated after reaching PCR amplification plateau. The four amplicons from each library were then pooled, purified using a Qiagen Minelute kit, and eluted in 20 µl of 0.1xTE. DNA extracted from modern *Vicugna pacos* (alpaca) samples was used to produce the capture bait as described in Maricic *et al.*<sup>15</sup>. In brief, modern samples were extracted using a Qiagen DNeasy extraction kit, complete mitochondrial genomes were amplified by long range PCR as two separate fragments using primers based on the *Vicugna pacos* mitochondrial genome NCBI Reference Sequence: NC\_002504.1 (Supplementary Table S2), purified using Qiagen QIAquick DNA purification kit and quantified using a thermo fisher scientific nanodrop 2000c spectrophotometer. Long range PCR products were pooled in 750ng equimolar amounts and fragmented through sonication. Biotinylated adapters were ligated to the sonicated mitochondrial genome fragments. Adapter ligated fragments were then bound to streptavidin coated beads to be used as bait in the hybridisation capture.

**Sequencing.** Capture enriched, barcoded libraries were pooled in equimolar ratios, depending on quantities determined by qPCR, and sequenced on an Ion Torrent sequencer at the University of Otago, Dunedin.

**Raw data processing and mitochondrial genome mapping.** Ion torrent sequencing reads were separated according to their barcode and processed following Clarke *et al.*<sup>16</sup>. Any adapter sequences and reads shorter than 25 bp were trimmed using cutadapt<sup>17</sup>. Processed reads were mapped against the guanaco and vicuña reference mitochondrial genomes NC\_011822.1 (guanaco) and FJ456892.1 (vicuña) using the Burrows-Wheeler Aligner (BWA)<sup>18</sup>. Duplicate read removal and quality control were performed using SAMtools (a tool for interacting and post-processing short DNA read alignments) while consensus sequences (bcftools and mpileup) and coverage plots were constructed using SAMtools mpileup<sup>19</sup>. Consensus sequences were submitted to GenBank (accession numbers: KX388532 - KX388534, see Table 1). The perl script make\_coverage\_plots.pl was used to visualize the coverage plots (see Supplementary Fig. S3).

The commonly used Python and R computational framework known as mapDamage was implemented to track and quantify ancient DNA damage patterns in order to assess for ancient DNA authenticity<sup>20</sup>.

**Phylogenetic analysis.** A multiple sequence alignment (MSA) was performed using the three consensus Mocha camelid mitochondrial genomes alongside guanaco, llama, alpaca, vicuña and outgroup sequences obtained from NCBI Genbank (Supplementary Table S3) using Mafft with parameters specified for accuracy over speed<sup>21</sup>. The MSA was inspected by eye using Jalview<sup>22</sup>, all sites with an unknown base (n) were removed. BEAUti was used to prepare alignments for phylogenetic tree reconstruction, which was carried out using BEASTv1.8.1. Parameters included a HKY+G substitution as determined using jmodeltest2.1.7<sup>23</sup> and a Speciation:Birth-Death process tree prior<sup>24</sup>. The Markov Chain Monte Carlo (MCMC) was run from a random starting tree for 10,000,000 iterations, sampling every 1,000th tree with a burn in of 100,000 states. Effective sample size for estimated parameters was checked using Tracer v. 1.6<sup>25</sup>. All parameters were estimated from at least 200 fully independent trees sampled by the MCMC. The MCMC was run three times and results were compared to check for convergence of the algorithm. Posterior probabilities were annotated onto the BEAST output tree using TreeAnnotator.

**Network analysis.** As phylogenetic analyses indicated that Isla Mocha camelids were guanacos, a network analysis against a guanaco dataset was undertaken in order to fit the Isla Mocha camelids into the phylogeographic framework of wild guanacos and identify their potential region of origin. Due to lack of sufficient complete mitochondrial genomes available for comparison, the dataset was limited to the control region of the mitochondrial genome, for which an extensive phylogeographic reference database is available<sup>10</sup>. These control regions were aligned with the Isla Mocha camelid samples using Mafft, then analysis and network construction was performed using the freely available R script, TempNet<sup>26</sup>.

## References

1. Flannery, T. F. & White J. P. Animal translocation. *Research & Exploration* **7**, 96–113 (1991).
2. Campbell, R. Socioeconomic differentiation, leadership, and residential patterning at an araucanian chiefly center (Isla Mocha, AD 1000–1700). Doctoral dissertation, Dietrich School of Arts and Sciences. University of Pittsburgh (2011).
3. Pefaur, J. & Yáñez, J. Ecología descriptiva de la Isla Mocha (Chile), en relación al poblamiento de vertebrados. *Boletín del Museo Nacional de Historia Natural* **37**, 103–112 (1980).
4. Quiroz, D. & Sánchez, M. Poblaciones tempranas en Isla Mocha (Siglo XIV AC). *Museos* **15**, 9–11 (1993).
5. Quiroz, D. & Sánchez, M. Poblamientos iniciales en la costa septentrional de la Araucanía (6.500–2.000 a.p.). *Chungara (Arica)* **36**, 289–302 (2004).
6. van Speilbergen, J., Maire, J. L. & DeVilliers, J. A. J. Spiegel der Australische navigatie. English. *Volume 18 of The East and West Indian Mirror: Being an Account of Joris Van Speilbergen's Voyage Round the World (1614–1617), and the Australian Navigations of Jacob Le Maire*. (Hakluyt Society, 1906).
7. Mengoni-Gonalons, G. L. & Yacobaccio, H. D. The Domestication of South American Camelids: A view from the South-Central Andes in *Documenting Domestication: New genetic and archaeological paradigms* (Eds Zeder, M. A., Bradley, D. G., Emshwiller, E., Smith, B. D.) 228–244 (University of California Press, Berkeley, 2006).
8. Bonacic, C. Características biológicas y productivas de los camélidos sudamericanos. *Avances en Ciencias Veterinarias* **6**, 3–15 (1991).
9. Gay, C. Historia física y política de Chile. Vertebrados, Mamíferos. *Zoología* **1**, 19–182 (1847).
10. Marín, J. C. *et al.* Mitochondrial DNA variation, phylogeography and systematics of guanaco (*Lama guanicoe*, artiodactyla: camelidae). *Journal of Mammalogy* **89**(2), 269–281 (2008).
11. Quiroz, D. & Vásquez, M. La presencia del arcaico tardío en Isla Mocha: excavaciones preliminares del sitio P27-1. *Museos* **21**, 21–26 (1996).
12. Goicovich, F. & Quiroz, D. De insulares a continentales: la historia de los mochanos, desde los orígenes hasta su desintegración social en la Misión de San José de la Mocha. Facultad de Filosofía y Humanidades, Universidad de Chile, Santiago (2008).
13. Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth. *Nature Protocols* **2**, 1756–1762 (2007).
14. Knapp, M., Stiller, M. & Meyer, M. Generating barcoded libraries for multiplex high-throughput sequencing in: *Ancient DNA: Methods and Protocols* Shapiro, B., Hofreiter, M. (Eds). Humana Press, Springer, New York, 155–170 (2012).
15. Maricic, T., Whitten, M. & Pääbo, S. Multiplexed DNA sequence capture of mitochondrial genomes Using PCR Products. *PLoS One* **5**, e14004 (2010).
16. Clarke, Andrew C. *et al.* “From cheek swabs to consensus sequences: an A to Z protocol for high-throughput DNA sequencing of complete human mitochondrial genomes”. *BMC genomics* **15**, 1 (2014).
17. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal* **17**(1), 10 (2011).
18. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
19. Li, H. *et al.* The Sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
20. Ginolhac, A., Rasmussen, M., Gilbert, M. T. P., Willerslev, E. & Orlando, L. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* **27**, 2153–2155 (2011).
21. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**(4), 772–780 (2013).
22. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**(9), 1189–1191 (2009).
23. Darriba, D., Taboada G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**(8), 772–772 (2012).
24. Gernhard, T. The conditioned reconstructed process. *Journal of Theoretical Biology*. **253**, 769778 (2008).
25. Rambaut, A., Suchard, M. A., Xie, D. & Drummond, A. J. Tracer v1. 6. URL <http://beast.bio.ed.ac.uk/tracer> (2014).
26. Prost, S. & Anderson, C. N. K. TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. *Methods in Ecology and Evolution* **2**, 663–667 (2011).

## Acknowledgements

The project was supported by a Royal Society of New Zealand Marsden grant (UOO 0926) to EM-S. MK was supported by the Allan Wilson Centre for Molecular Ecology and Evolution.

## Author Contributions

Westbury M. contributed to labwork, data analyses and manuscript writing. Prost S. contributed to data analyses. Seelenfreund A. conceived the research question and contributed samples. Ramírez J.M. contributed to sample collection. Matisoo-Smith E. conceived the research question and experiments, contributed to manuscript writing and sample collection. Knapp M. conceived research question and experiments, contributed to labwork, data analyses and manuscript writing. All authors reviewed the manuscript

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Westbury, M. *et al.* First complete mitochondrial genome data from ancient South American camelids - The mystery of the *chilihueques* from Isla Mocha (Chile). *Sci. Rep.* **6**, 38708; doi: 10.1038/srep38708 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016