

SCIENTIFIC REPORTS



OPEN

Whole mitochondrial DNA sequencing in Alpine populations and the genetic history of the Neolithic Tyrolean Iceman

V. Coia¹, G. Cipollini¹, P. Anagnostou², F. Maixner¹, C. Battaglia², F. Brisighelli^{3,†}, A. Gómez-Carballa⁴, G. Destro Bisol^{2,5}, A. Salas⁴ & A. Zink¹

Received: 07 July 2015
Accepted: 01 December 2015
Published: 14 January 2016

The Tyrolean Iceman is an extraordinarily well-preserved natural mummy that lived south of the Alpine ridge ~5,200 years before present (ybp), during the Copper Age. Despite studies that have investigated his genetic profile, the relation of the Iceman's maternal lineage with present-day mitochondrial variation remains elusive. Studies of the Iceman have shown that his mitochondrial DNA (mtDNA) belongs to a novel lineage of haplogroup K1 (K1f) not found in extant populations. We analyzed the complete mtDNA sequences of 42 haplogroup K bearing individuals from populations of the Eastern Italian Alps – putatively in genetic continuity with the Tyrolean Iceman—and compared his mitogenome with a large dataset of worldwide K1 sequences. Our results allow a re-definition of the K1 phylogeny, and indicate that the K1f haplogroup is absent or rare in present-day populations. We suggest that mtDNA Iceman's lineage could have disappeared during demographic events starting in Europe from ~5,000 ybp. Based on the comparison of our results with published data, we propose a scenario that could explain the apparent contrast between the phylogeographic features of maternal and paternal lineages of the Tyrolean Iceman within the context of the demographic dynamics happening in Europe from 8,000 ybp.

The Tyrolean Iceman (also known as Ötzi or Similaun Man) was discovered in 1991 on the *Tisenjoch* Pass in the Italian part of the Ötztal Alps, close to the Austrian border (3210 m s.l.m.). The Iceman is one of the oldest natural European mummies (dated to 5350–5100 ybp), and an extraordinarily well-preserved individual due to a spontaneous freeze-drying process. Scientific interest is further increased by the preservation of parts of his clothes and equipment (e.g. loincloth, shoes, leggings, axe, bow).

This important finding provided unprecedented insights into aspects of the daily life during the Copper Age in central Europe^{1–3}. Both stable isotopes analyses of Iceman's teeth and bones and the identification of pollen and moss fragments in his intestines have pinpointed his origin to south of the Alpine ridge to a few valleys within ~60 kilometers South-East of the discovery site (*Vinschgau* and *Schnals* valleys in actual South-Tyrol, Eastern Italian Alps)^{4,5}.

The Tyrolean Iceman has also been extensively characterized genetically. Whole genome sequencing data indicate that the Iceman was blood group 0, lactose intolerant and genetically predisposed to cardiovascular diseases⁶. Moreover, Keller and colleagues⁶ showed that the Iceman belongs to the Y-chromosome haplogroup G2a-L91. Comparing the Iceman's genomic data to the human reference genome (UCSC hg19) specified the Iceman's lineage into one of the four sub-branches of G2a L-91 (G2a2a1a2a1a-FGC5672) following the latest version of the Y-DNA Haplogroup Tree⁷. Comparative analysis based on informative autosomal SNPs and Y-chromosome data showed a clear link with genetic diversity of extant populations further supported by recent whole genome-data analysis⁸. However, an important question regarding the Iceman genetic heritage remains unanswered. In fact,

¹Accademia Europea di Bolzano (EURAC-Research), Istituto per le mummie e l'Iceman, Bolzano, Italy. ²Dipartimento Biologia Ambientale, Università La Sapienza, Roma, Italy. ³Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK. ⁴Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, Galicia, Spain. ⁵Istituto Italiano di Antropologia, Roma, Italy. [†]Present address: Sezione di Medicina Legale-Istituto di Sanità Pubblica, Università Cattolica del Sacro Cuore, Roma, Italy. Correspondence and requests for materials should be addressed to V.C. (email: valentina.coia@eurac.edu)

despite the mitochondrial DNA (mtDNA) of the Iceman being well characterised^{9–12}, the relation of his lineage with present-day mitochondrial diversity remains elusive. The first study on the complete mtDNA of the Iceman¹¹ showed that the ancient mitogenome was a novel lineage of mtDNA haplogroup K1 (named K1f), not yet detected in extant populations. Ermini *et al.*¹¹ proposed that the Iceman's haplogroup may remain undetected simply due to inadequate sampling. This hypothesis should be considered due to the paucity of data regarding K1 haplogroups available at the time of the study, especially from alpine populations, which are putatively in genetic continuity with the Iceman. Alternatively, the K1f haplogroup and its closest neighbors may have been lost during repeated demographic and population events occurring in the Alps since the Neolithic.

To shed new light on these issues, we analyzed the mitogenomes of 42 K-haplogroup samples from Eastern Italian Alps and compared the Iceman's sequence with a large dataset of worldwide K1 mitogenomes. The results are discussed with currently available ancient mtDNA data from European Neolithic specimens. Finally, we combine mtDNA and Y-chromosomal data to propose a scenario that could explain the complex genetic history of the Tyrolean Iceman.

Results and Discussion

The Iceman mitochondrial genetic heritage. The mitogenomes of 42 haplogroup K individuals from Eastern Italian Alps were analyzed (Table S1). The most frequently observed K1 Alpine haplogroups were K1a4 and K1a1, similar to the worldwide K1 distribution (Fig. 1) and K1a + 195 (Table S2). Rare haplogroups were also observed in this study such as K1a19 and K1a24, and K1e1, which was represented in our dataset by a novel haplotype defined by a specific transition in the coding region (transition A15244G; sample #111Gm in the Fig. S2). Neither the Iceman's lineage nor any other intermediate haplotypes of K1f were detected.

Next, we analyzed all K1 mitogenomes available in the literature, scanning databases, in combination with our dataset ($n = 1077$; Table S2). The whole K1 tree was updated and the Iceman's sequence was added to the phylogeny (Fig. 1). This tree was compared with the one presented by Ermini and collaborators¹¹ that used in total 115 K sequences (85 of which were K1 mitogenomes). The Iceman's mitogenome falls within a cluster defined by transition T16362C (K1 + 16362) that includes K1d (5.6 thousand years [kya]; standard error: 0.6–10.7) and K1e (13.8 kya; standard error: 6.5–21.4) haplogroups. The collected data created a clearer and more robust representation of K1 + 16362 cluster despite being defined exclusively by an unstable mutation on the rapidly mutating mtDNA control region. Indeed, only one sequence of unknown origin (EU073969) shared the mutation T16362C with the Iceman's mitogenome in the tree built by Ermini and colleagues¹¹. Currently, this cluster encompasses 13 other modern sequences mostly derived from individuals of known European ancestry (five sampled in Denmark, one in Scotland, one in the Alps, and six of unknown origin; see Fig. 2). The closest observed haplotype (JQ704724) to the Iceman's mitogenome is, however, quite distant from K1f, being separated by six mutational steps (five of which are in the coding region).

Overall, the results indicate that K1f has been lost from present-day populations, rather than being a rare haplotype undetected due to inadequate sampling. With a binomial distribution^{13,14} given the size of our dataset, we estimated (with a 95% of probability) that the K1f lineage should not occur with a frequency above 0.3%. In addition, the most parsimonious phylogenetic tree showed that there exist no intermediate haplotypes connecting the Tyrolean Iceman's lineage to the main root of K1 (Fig. 2).

Recently, considerable knowledge has been gained on the mtDNA diversity of ancient Europeans¹⁵. It appears that the Iceman's haplogroup may not be the only case of a described mtDNA lineages that has apparently been lost. Two examples include lineages from haplogroup H (H46b and H88) that have been detected in two Linear Pottery culture individuals (LBK, 5450–4775 BC) in Germany¹⁶. However, H46b and H88 haplogroups are only one-step mutation away from the extant ancestral haplogroups (H46 and H*, respectively). This is in contrast to the described phylogeographic features of the K1f haplogroup, suggesting that the entire branch of K1f was lost.

Reconstructing the genetic history of the Tyrolean Iceman. The focused analysis of the Tyrolean Iceman has produced important information on his origin and genetic relations between Iceman and extant European populations. The availability of data on genetic markers with different modes of inheritance and rates of evolution allows to perform comparative analysis that shed further light on the genetic history of the Tyrolean Iceman. Interestingly, there is a contrast between the Iceman's maternal and paternal genetic heritage. While the branch of mtDNA K1f lineage has probably been lost from present-day populations according to this study, the Y-chromosome branch G2a-L91 is still observed in Europe, and reaches remarkable frequencies (> 10%) in groups from the Mediterranean area, i.e. Sardinians and Corsicans^{6,17}. How can this pattern be explained?

The direct comparison of the evolutionary history of specific female and male lineages can be complicated by differences in their phylogenetic structure and resolution. However, it is likely that the dissimilarity observed in the diachronic patterns of the Iceman's genetic lineages are due mainly to the distributions of K1f and G2a in Europe prior to 5,000 ybp and later demographic events that shaped the European genetic structure.

Recent ancient DNA studies suggest that both mtDNA haplogroup K1 and Y-chromosome G2a reached the European continent around 8,000 ybp by migrations of Early Neolithic people from the Near East through continental and Mediterranean routes^{15,18–20} (see Fig. 3A). Subsequently, these haplogroups spread across the continent, as shown by the distribution of K1 and G2a from several Europeans specimens dated > 5,000 ybp (see Fig. 3B and its legend for references and details). At that time, the K1 haplogroup was present in West, South, East, and North Europe with its main branches (K1a, K1b) and rarer branch like K1e, whereas K1f was virtually restricted to the Alps since it has not been observed outside this region. On the other hand, G2a was the predominant Y-chromosome lineage observed in the European Neolithic samples analyzed thus far^{15,19–23}.

While the Iceman's Y-chromosome lineage was part of the Early Neolithic genetic substrate, as suggested from current distributions of G2a-L91 (Fig. 3C), the mtDNA K1f lineage developed locally in the Eastern Alps at least from ~5,200 ybp (the time of the Iceman).

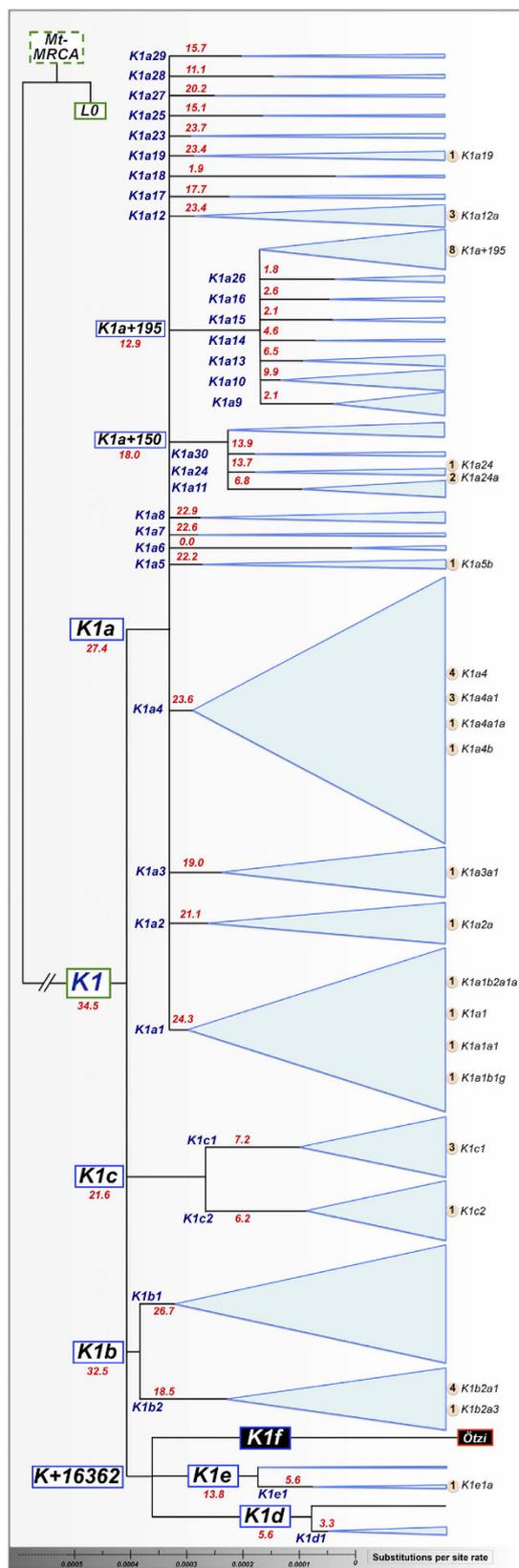


Figure 1. Main skeleton of haplogroup K1 and its branches based on all of the mitogenomes. The area of the triangles is proportional to the occurrence of K1 haplogroups in the total dataset (see also Table S2). The haplogroup assignment of K1 Alpine mitogenomes (thirty-six from the present study and six from³⁷) are listed on the right-hand side of the tree. Age estimates of haplogroups are reported close to each branch (see also Table S3).

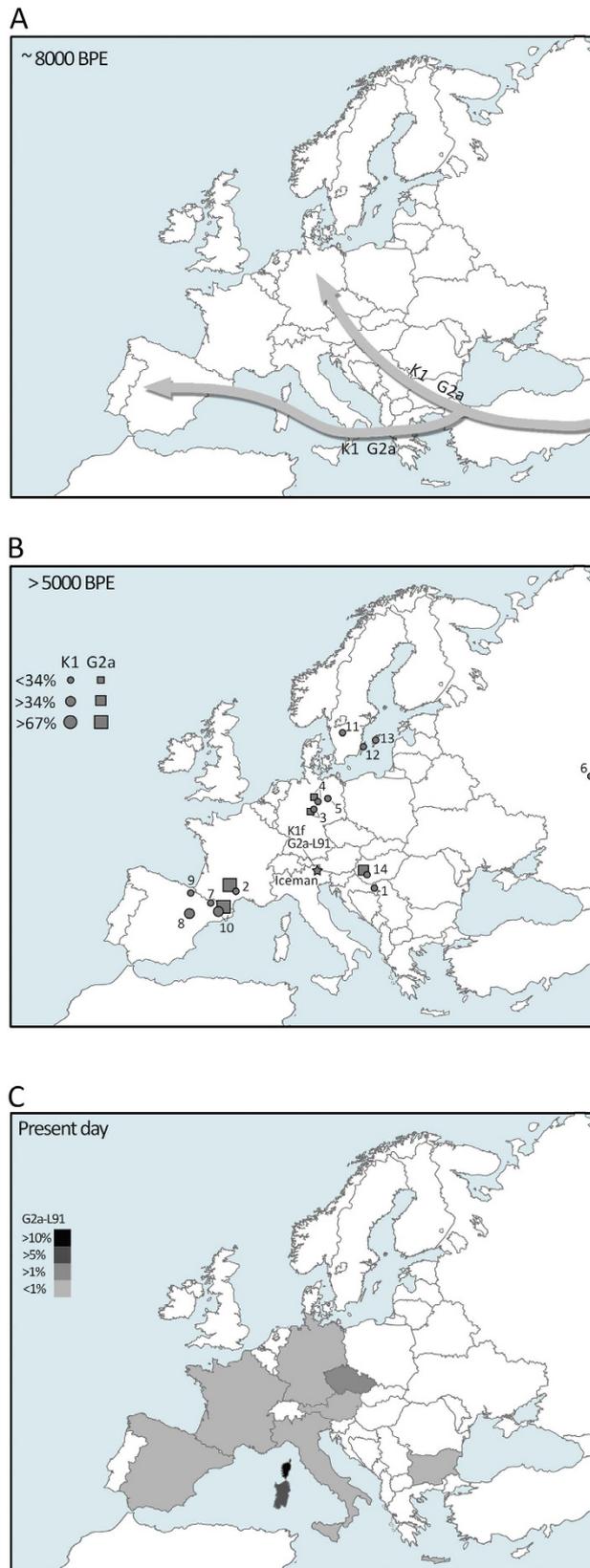


Figure 3. (A) Migrations occurred around 8,000 ybp of Early Neolithic people from the Near East to Europe carrying the main mtDNA (K1) and Y chromosome (G2a) haplogroups according to recent data on ancient DNA studies^{15,18–20}. (B) Distribution of mtDNA K1 (circles) and Y chromosome G2a (rectangles) haplogroups in ancient samples dated > 5,000 ybp and their absolute frequencies. Site 1: *Vinkovci Nama*, Croatia²³; site 2: *Treilles at Saint-Jean-et-Saint-Paul*, France¹⁹; site 3: *Derenburg-Meerensstieg II*, Germany²²; site 4: *Halberstadt-Sonntagsfeld*, Germany²¹; site 5: South of *Saxony-Anhalt*, Germany¹⁸; site 6: *Yamnaya*, Russia²¹; site 7: *Els Trocs*,

Spain²¹; site 8: *La Mina*, Spain²¹; site 9: *Navarre*, Spain³⁸; site 10: *Avellaner*, Catalonia, Spain²⁰; site 11: *Gökhem*, Sweden^{39,40}; site 12: *Köpingsvik*, Sweden⁴⁰; site 13: *Gotland*, Sweden⁴⁰; site 14: *Alsónyék-Bátaszék, Mérnöki telep, Lánycsók, Gata-Csotola, Bölcske-Gyűrűsvölgy, Budakeszi, Szőlőskert-Tangazdaság*, Hungary²³. (C). Distribution and approximate frequencies of haplogroup G2a-L91 in Europe based on the data from Rootsi *et al.*¹⁷. Maps of Europe available from Wikipedia Common web page (https://commons.wikimedia.org/wiki/File:Blank_political_map_Europe_in_2006_WF.svg?uselang=it#filelinks) were modified using Adobe Photoshop CS6 software.

PhyloTree Build 16³¹ was used as a reference to draft the skeleton of K1 haplogroup and its branches; some branches were updated according to new available data (Fig. 1). For the most closely related sequences of the Iceman haplotype we reconstructed the phylogeny using the reduced-median algorithm³³ (Fig. 2). Reticulations were resolved using the relative mutation rates at each position³⁴. The phylogenetic tree was drafted using mtPhyl v3.520 (<http://eltsov.org>) and adjusted manually according to PhyloTree Build 16³¹ and the data added in the present study.

We estimated the Time of the Most Recent Common Ancestor (TMRCA) of clades utilizing a time-dependent clock incorporating a correction for purifying selection³⁴ and using the maximum likelihood (ML) procedure with PAML 3.13³⁵. A skeleton of the tree based on a well known mtDNA phylogeny (PhyloTree Build 16) was initially provided to the software in newick format. Then, TMRCA was obtained assuming the HKY85 mutation model (excluding indels, and hotspot mutations such as 16182C, 16183C and 16519C), and used the gamma-distributed rates (approximated by a discrete distribution with 32 categories) and three partitions: Hypervariable region I (HVS-I: positions 16051–16400), Hypervariable region II (HVS-II: positions 68–263), and the coding region. A set of outgroups from haplogroups were used that represent most of the main branches of the worldwide phylogeny, from the rCRS to the oldest African nodes³¹.

The BSP showing changes of the effective population size of haplogroup K1 through time was performed with the Bayesian Evolutionary Analysis Sampling Trees (BEAST)³⁶ software. The analysis was made using only the mtDNA coding region. MCMC samples were based on 40,000,000 generations with the first 4,000,000 generations discarded as burn-in. We used a HKY substitution model and a strict clock with a substitution rate of 1.69×10^{-8} as reported in³⁴. BEAST outputs were analyzed with the software Tracer (<http://tree.bio.ed.ac.uk/software/tracer/>).

References

- Spindler, K., Rastbichler-Zissernig, E., Wilfing, H., ZurNedden, D., Nothdurfter, H. (eds). *Der Mann im Eis: neue Funde und Ergebnisse* (Springer, 1995).
- Spindler, K., Wilfing, H., Rastbichler-Zissernig, E., ZurNedden, D., Nothdurfter, H. (Eds). *Human Mummies: A Global Survey of their Status and the Techniques of Conservation in The Man in the Ice*. vol 3, (Springer, 1996).
- Bortenschlager, S., Oeggl, K. (eds). *The Iceman and his Natural Environment: Palaeobotanical Results in The man in the ice*. vol 4, (Springer, 2000).
- Müller, W., Fricke, H., Halliday, A. N., McCulloch, M. T. & Wartho, J. A. Origin and migration of the Alpine Iceman. *Science*. **302**, 862–866 (2003).
- Oeggl, K. *et al.* The reconstruction of the last itinerary of “Ötzi”, the Neolithic Iceman, by pollen analyses from sequentially sampled gut extracts. *Quat. Sci. Rev.* **26**, 853–861 (2007).
- Keller, A. *et al.* New insights into the Tyrolean Iceman’s origin and phenotype as inferred by whole-genome sequencing. *Nat. Commun.* **3**, 698 (2012).
- International Society of Genetic Genealogy 2015, Y-DNA Haplogroup Tree 2015, Version: 10.44 (<http://www.isogg.org/tree/>).
- Sikora, M. *et al.* Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe. *PLoS Genet.* **10**, e1004353 (2014).
- Handt, O. *et al.* Molecular genetic analyses of the Tyrolean Ice Man. *Science*. **264**, 1775–1778 (1994).
- Rollo, F. *et al.* Fine characterization of the Iceman’s mtDNA haplogroup. *Am. J. Phys. Anthropol.* **130**, 557–564 (2006).
- Ermini, L. *et al.* Complete mitochondrial genome sequence of the Tyrolean Iceman. *Current Biol.* **18**, 1–7 (2008).
- Endicott, P. *et al.* Genotyping human ancient mtDNA control and coding region polymorphisms with a multiplexed Single-Base-Extension assay: the singular maternal history of the Tyrolean Iceman. *BMC Genet.* **10**, 29 (2009).
- Coia, V. *et al.* Brief communication: mtDNA variation in North Cameroon: lack of Asian lineages and implications for back migration from Asia to sub-Saharan Africa. *Am. J. Phys. Anthropol.* **128**, 678–861 (2005).
- Plaza, S. *et al.* Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. *Hum Genet.* **115**, 439–447 (2004).
- Brandt, G. *et al.* Human paleogenetics of Europe—the known knowns and the known unknowns. *J. Hum. Evol.* **79**, 73–92 (2015).
- Brotherton, P. *et al.* Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat. Commun.* **4**, 1764 (2013).
- Rootsi, S. *et al.* Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. *Eur. J. Hum. Genet.* **20**, 1275–1282 (2012).
- Brandt, G. *et al.* Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. *Science*. **342**, 257–261 (2013).
- Lacan, M. *et al.* Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *Proc. Natl. Acad. Sci.* **108**, 9788–9791 (2011).
- Lacan, M. *et al.* Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proc. Natl. Acad. Sci.* **8**, 18255–18259 (2011).
- Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. **522**, 207–211 (2015).
- Haak, W. *et al.* Ancient DNA from European early neolithic farmers reveals their near eastern affinities. *PLoS Biol.* **8**, e1000536 (2010).
- Szécényi-Nagy, A. *et al.* Tracing the genetic origin of Europe’s first farmers reveals insights into their social organization. *Proc. Biol. Sci.* **282**, doi: 10.1098/rspb.2015.0339 (2015).
- Festi, D., Putzer, A. & Oeggl, K. Mid and late Holocene land-use changes in the Ötztal Alps, territory of the Neolithic Iceman “Ötzi”. *Quat. Int.* **353**, 17–33 (2014).
- Coia, V. *et al.* Evidence of high genetic variation among linguistically diverse populations on a micro-geographic scale: a case study of the Italian Alps. *J. Hum. Genet.* **57**, 254–60 (2012).
- Capocasa, M. *et al.* Detecting genetic isolation in human populations: a study of European language minorities. *PLoS One.* **8**, e56371 (2013).

27. Capocasa, M. *et al.* Linguistic, geographic and genetic isolation: a collaborative study of Italian populations. *J. Anthropol. Sci.* **92**, 201–231 (2014).
28. Pichler, I. *et al.* Genetic structure in contemporary south Tyrolean isolated populations revealed by analysis of Y-chromosome, mtDNA, and Alu polymorphisms. *Hum. Biol.* **78**, 441–464 (2006).
29. Torroni, A. *et al.* Do the four clades of the mtDNA haplogroup L2 evolve at different rates? *Am. J. Hum. Genet.* **69**, 1348–1356 (2001).
30. Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. & Howell, N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147 (1999).
31. Van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, 386–394 (2009).
32. Gómez-Carballa, A., Pardo-Seco, J., Amigo, J., Martín-Torres, F. & Salas, A. Mitogenomes from The 1000 Genome Project reveal new Near Eastern features in present-day Tuscans. *PLoS One.* **18**, e0119242 (2015).
33. Bandelt, H. J. *et al.* Mitochondrial portraits of human populations using median networks. *Genetics.* **141**, 743–753 (1995).
34. Soares, P. *et al.* Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* **84**, 740–759 (2009).
35. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
36. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **8**, 214 (2007).
37. Pichler, I. *et al.* Drawing the history of the Hutterite population on a genetic landscape: inference from Y-chromosome and mtDNA genotypes. *Eur. J. Hum. Genet.* **18**, 463–470 (2010).
38. Hervella, M. *et al.* Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the Neolithic expansion into Europe. *PLoS One.* **7**, e34417 (2012).
39. Skoglund, P. *et al.* Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science.* **344**, 747–750 (2014).
40. Malmström, H. *et al.* Ancient mitochondrial DNA from the northern fringe of the Neolithic farming expansion in Europe sheds light on the dispersion process. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 20130373 (2015).

Acknowledgements

We are greatly indebted to the blood donors and the nursing staff of the *Laboratorio di Patologia Clinica* at the Brunico Hospital, especially to Dr. Peter Santer, for helping us in collecting samples from South Tyrol. We are also grateful to Peter Pramstaller and Irene Pichler (EURAC-Research) for providing us with South Tyrolean DNA samples; to Donata Luiselli (University of Bologna) for providing us with three unpublished sequences from South Tyrol; to Antonio Torroni and Francesca Gandini (University of Pavia) for their methodological advices; to Cristian Capelli (Oxford University) for his contribution in the initial phase of the study design; to Valentina Dominici (University of Rome La Sapienza) for helping us with the Beast analysis, to Kathrin Renner (EURAC-Research) for providing us with the map of South Tyrol and, finally, to Niall O'Sullivan (EURAC-Research) for kindly helping us with the re-analysis of the Iceman's Y-chromosome genome. This study was financed by the *Provincia Autonoma di Bolzano – Alto Adige, Ripartizione Diritto allo studio, università e ricerca scientifica*, funds to VC.

Author Contributions

Designed the research V.C. and A.Z. collected the South Tyrolean samples V.C. provided alpine samples from Veneto and Friuli G.D.B. performed the experiments V.C. and G.C. contributed to the experimental steps F.M., F.B. and C.B. analyzed the molecular data A.S. and A.G.C. together with P.A. and V.C. Wrote the paper V.C. with collaboration of A.Z. and G.D.B. All authors have read and approved the final 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: Coia, V. *et al.* Whole mitochondrial DNA sequencing in Alpine populations and the genetic history of the Neolithic Tyrolean Iceman. *Sci. Rep.* **6**, 18932; doi: 10.1038/srep18932 (2016).



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/>