PALEOGENOMICS

Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations

Marieke van de Loosdrecht,¹ Abdeljalil Bouzouggar,^{2,3*†} Louise Humphrey,⁴ Cosimo Posth,¹ Nick Barton,⁵ Ayinuer Aximu-Petri,⁶ Birgit Nickel,⁶ Sarah Nagel,⁶ El Hassan Talbi,⁷ Mohammed Abdeljalil El Hajraoui,² Saaïd Amzazi,⁸ Jean-Jacques Hublin,³ Svante Pääbo,⁶ Stephan Schiffels,¹ Matthias Meyer,⁶ Wolfgang Haak,¹† Choongwon Jeong,^{1*†} Johannes Krause^{1*†}

North Africa is a key region for understanding human history, but the genetic history of its people is largely unknown. We present genomic data from seven 15,000-year-old modern humans, attributed to the Iberomaurusian culture, from Morocco. We find a genetic affinity with early Holocene Near Easterners, best represented by Levantine Natufians, suggesting a pre-agricultural connection between Africa and the Near East. We do not find evidence for gene flow from Paleolithic Europeans to Late Pleistocene North Africans. The Taforalt individuals derive one-third of their ancestry from sub-Saharan Africans, best approximated by a mixture of genetic components preserved in present-day West and East Africans. Thus, we provide direct evidence for genetic interactions between modern humans across Africa and Eurasia in the Pleistocene.

nder typical conditions (i.e., aside from intermittent greening periods), the Sahara desert poses an ecogeographic barrier for human migration between North and sub-Saharan Africa (1). Sub-Saharan Africa is home to the most deeply divergent genetic lineages among present-day humans (2), and the general view is that all Eurasians mostly descend from a single group of humans that dispersed outside of sub-Saharan Africa around 50,000 to 100,000 years before the present (yr B.P.) (3). This group likely represented only a small fraction of the genetic diversity within Africa, most closely related to a Holocene East African group (4). Present-day North Africans share a majority of their ancestry with present-day Near Easterners but not with sub-Saharan Africans (5). Thus, from a genetic perspective, present-day North Africa is largely a part of Eurasia. However, the temporal depth of this genetic connection be-

*Corresponding author. Email: krause@shh.mpg.de (J.K.); jeong@shh.mpg.de (C.J.); bouzouggar@eva.mpg.de (A.B.) †These authors contributed equally to this work. tween the Near East and North Africa is poorly understood and has been estimated only indirectly from present-day mitochondrial DNA (mtDNA) variation (*6*, *7*).

Owing to challenging conditions for DNA preservation, relatively few ancient genomes have been recovered from Africa. Genome-wide data from 23 individuals have been reported from South and East Africa, with the oldest dating back to 8100 yr B.P. (4, 8, 9). In North Africa, a genomic study of Egyptian mummies from the first millennium BCE showed that the genetic connection between the Near East and North Africa was established by that time (5). However, the genetic affinity of North African populations at a greater time depth has remained unknown.

Here we present genome-wide data from seven individuals, directly dated between 15,100 and 13,900 calibrated years before present (cal. yr B.P.) (table S1), from Grotte des Pigeons near Taforalt in eastern Morocco (10). These genomic data provide a critical reference point to help explain the deep genetic history of North Africa and the broader Middle East (Fig. 1). The Taforalt individuals are associated with the Later Stone Age Iberomaurusian culture, whose origin is debated. These individuals may have descended either directly from the manufacturers of the preceding Middle Stone Age technologies (Aterian or local West African bladelet technologies) or from an exogenous population with ties to the Upper Paleolithic technocomplexes of the Near East or Southern Europe (10, 11).

For nine Taforalt individuals (table S2), we created double-indexed single-stranded DNA libraries (12) for next-generation sequencing of DNA isolated from petrous bones. We then used in-solution capture probes (13) to enrich libraries for the whole mitochondrial genome

and ~1.240.000 single-nucleotide polymorphisms (SNPs) in the nuclear genome (14). The DNA fragments obtained from seven individuals, six genetic males and one female, had postmortem degradation characteristics typical of ancient DNA (tables S3 to S5 and fig. S6). We reconstructed the mitochondrial genomes of all seven individuals (102× to 1701× coverage, unmerged libraries; table S4) while maintaining a low level of contamination from the DNA of modern humans (1 to 8%; table S4). For the nuclear data analysis, in which ancient DNA is more susceptible to contamination than in mitochondrial analyses, we analyzed five individuals (four males and one female) on the basis of coverage (table S3, merged libraries) and negligible modern human contamination for males (1.7 to 2.5%; table S5). For each individual, we randomly chose a single base per site as a haploid genotype. We intersected our new data with data from a panel of worldwide present-day populations, genotyped on the Affymetrix Human Origins array for ~600,000 markers, as well as ancient genomic data covering Europe, the Near East, and sub-Saharan Africa (4, 8, 15-17). The final data set includes 593,124 intersecting autosomal SNPs with 183,041 to 544,232 SNP positions covered for each of the five individuals (table S3). For group-based analyses involving other ancient individuals, we adopted the population labels from the original studies (4, 16). We found an overall high genetic relatedness between the Taforalt individuals, suggesting a strong population bottleneck (fig. S26).

We analyzed the genetic affinities of the Taforalt individuals by performing principal components analysis and model-based clustering of worldwide data (Fig. 2). When projected onto the top principal components of African and west Eurasian populations, the Taforalt individuals form a distinct cluster in an intermediate position between present-day North Africans [e.g., Amazighes (Berbers), Mozabites, and Saharawis] and East Africans (e.g., Afars, Oromos, and Somalis) (Fig. 2A). Consistently, we find that all males with sufficient nuclear DNA preservation carry Y haplogroup E1b1b1a1 (M-78; table S16). This haplogroup occurs most frequently in present-day North and East African populations (18). The closely related E1b1b1b (M-123) haplogroup has been reported for Epipaleolithic Natufians and Pre-Pottery Neolithic Levantines (Levant_N) (16). Unsupervised genetic clustering also suggests a connection of Taforalt to the Near East. The three major components that make up the Taforalt genomes are maximized in early Holocene Levantines, East African hunter-gatherer Hadza from north-central Tanzania, and West Africans (number of genetic clusters K = 10; Fig. 2B). In contrast, present-day North Africans have smaller sub-Saharan African components with minimal Hadza-related contribution (Fig. 2B).

We calculated outgroup f_3 statistics of the form f_3 (Taforalt, X; Mbuti) across worldwide ancient and present-day test populations. Consistent with previous analyses, we find that ancient Near Eastern populations, especially Epipaleolithic Naturfians and early Neolithic Levantines, show the highest

¹Department of Archaeogenetics, Max Planck Institute for the Science of Human History (MPI-SHH), Jena, Kahlaische Strasse 10, D-07745, Germany. ²Origin and Evolution of Homo sapiens in Morocco Research Group, Institut National des Sciences de l'Archéologie et du Patrimoine, Hay Riad, Madinat Al Irfane, Angle rues 5 et 7, Rabat-Instituts, 10 000 Rabat, Morocco. ³Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology (MPI-EVA), Leipzig, Deutscher Platz 6, D-04103, Germany. ⁴Department of Earth Sciences, The Natural History Museum, London SW7 5BD, UK. ⁵Institute of Archaeology, University of Oxford, 36 Beaumont Street, Oxford OX1 2PG, UK. 6 Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology (MPI-EVA), Leipzig, Deutscher Platz 6, D-04103, Germany. ⁷Faculté des Sciences, Campus d'Al Qods, Université Mohammed Premier, B.P. 717 Oujda, Morocco. ⁸Mohammed V University, Avenue Ibn Batouta, Rabat, Morocco.

outgroup f_3 values with Taforalt (Fig. 3A). This is confirmed by f_4 symmetry statistics of the form f_4 (Chimpanzee, Taforalt; NE₁, NE₂) that measure a relative affinity of a pair of Near Eastern (NE) groups to Taforalt. A positive value indicates that NE₂ is closer than NE₁ to Taforalt. We consistently find positive f_4 values when the NE₂ group is Natufian or Levant_N and the NE₁ group is representative of other populations [z score = 2.2 to 11.0 standard error (SE); table S6]. Congruent to the outgtoup- f_3 results, the Natufian population shows higher affinity to Taforalt than does the Levant_N group (z score = 2.2 SE; table S6). This indicates that the early Holocene Levantine populations, overlapping with or postdating our Taforalt individuals by up to 6000 years (16), are most closely related to the Taforalt group, among Near Eastern populations. Next, we evaluated whether the Taforalt individuals have sub-Saharan African ancestry by calculating f_4 (Chimpanzee, X; Natufian, Taforalt). We observe significant positive f_4 values for all sub-Saharan African groups and significant negative values for all Eurasian populations, supporting a substantial contribution from sub-Saharan Africa (Fig. 3B). West Africans, such as Mende and Yoruba, most strongly pull out the sub-Saharan African ancestry in Taforalt (Fig. 3B and figs. S15 and S16).

We investigated whether two first-hand proxies, Natufians and West Africans, are sufficient to explain the Taforalt gene pool or whether a more complex admixture model is required. We thus tested whether Natufians could be a sufficient proxy for the Eurasian ancestry in Taforalt without explicit modeling of its African ancestry (fig. S18). This line of investigation was inspired by proposed archaeological connections between the Iberomaurusian and Upper Paleolithic cultures in Southern Europe, either via the Strait of Gibraltar (*19*) or Sicily (*20*). If this connection is true, both the Upper Paleolithic European and Natufian ancestries will be required to explain the Taforalt gene pool. For our admixture modeling with the program qpAdm (16), we chose outgroups that can distinguish sub-Saharan African, Natufian, and Paleolithic European ancestries but are blind to differences between sub-Saharan African lineages (11). A two-way admixture model, comprising Natufian and sub-Saharan African populations, does not significantly deviate from our data ($\chi^2 P \ge 0.128$), with 63.5% Natufian and 36.5% sub-Saharan African ancestry, on average (table S8). Adding Paleolithic European lineages as a third source only marginally increased the model fit ($\chi^2 P = 0.019$ to 0.128; table S9). Consistently, by using the qpGraph package (21), we find that a mixture of Natufian and Yoruba reasonably fits the Taforalt gene pool ($|z| \le 3.7$; fig. S19 and table S10). Adding gene flow from Paleolithic Europeans does not improve the model fit and provides an ancestry contribution estimate of 0% (fig. S19). We thus find no evidence of gene flow from





95% confidence intervals of radiocarbon dates across all dated individuals (cal. yr B.P. on the *x* axis). Group labels are taken from previous studies reporting each ancient genome (*4*, *16*, *27*). N, Neolithic; WHG, Western European hunter-gatherers; EHG, Eastern European hunter-gatherers.

Paleolithic Europeans into Taforalt within the resolution of our data.

We further characterized the sub-Saharan African–related ancestry in the Taforalt individuals by using f_4 , statistics in the form f_4 (Chimpanzee,

Fig. 2. Summary of the genetic profile of the Taforalt individuals.

(A) The top two principal components (PCs) calculated from present-day African, Near Eastern, and Southern European individuals from 72 populations. The Taforalt individuals are projected thereon (red inverted triangles), and selected present-day populations are denoted by various colored symbols. Labels for other populations (denoted by small gray squares) are provided in fig. S8. (B) ADMIXTURE analysis results of chosen African and Middle Eastern populations (K = 10). Ancient individuals are labeled in red. Major ancestry components in Taforalt individuals are maximized in early Holocene Levantines (green), West Africans (purple), and East African Hadza (brown). The ancestry component prevalent in pre-Neolithic Europeans (beige) is absent in Taforalt.

A



Fig. 3. Geographic

African; Yoruba/Mende, Natufian). We find that Yoruba or Mende and Natufians are symmetrically related to two deeply divergent outgroups, an ancient South African group from 2000 yr B.P. (aSouthAfrica) and Mbuti Pygmy, respectively $||z|| \le 1.564$ SE; table S11). Because f_4 statistics are linear under admixture, we expect the Taforalt population not to be any closer to these outgroups than Yoruba or Natufians if the two-way admixture model is correct. However, we find



comparison to Natufians, as measured by f_4 statistics in the form f_4 (Chimpanzee, X; Natufian, Taforalt). Large diamonds and squares represent the 10 most positive and negative f_4 values, respectively. Sub-Saharan Africans show high positive values, with West African Yoruba and Mende having the highest values, supporting the presence of sub-Saharan African ancestry in Taforalt individuals. In contrast, all Eurasian populations are genetically closer to Natufians than to the Taforalt group. The statistics and their associated SEs for the top 30 signals are presented in fig. S16.



Fig. 4. Relative genetic affinity of representative sub-Saharan African groups to a mixture of Yoruba and Natufians in comparison to the **Taforalt group.** We measured f_4 statistics in the form f_4 (Chimpanzee, African; Yoruba+Natufian, Taforalt) by using (**A**) aSouthAfrica, (**B**) Mbuti, and (**C**) Hadza as the African group. The f_4 statistics were calculated for the

proportions of Natufian-related ancestry ranging from 0 to 100% in increments of 1%. The blue rectangle marks a plausible range of Natufian ancestry proportion, estimated by our qpAdm modeling [0.637 ± (2 × 0.069)]. Gray solid and dotted lines represent ±1 and –3 SE ranges, respectively. SEs were calculated by 5-centimorgan block jackknife method.

instead that the Taforalt group is significantly closer to both outgroups (aSouthAfrica and Mbuti) than any combination of Yoruba and Natufians $(z \ge 2.728 \text{ SE}; \text{Fig. 4})$. A similar pattern is observed for the East African outgroups Dinka, Mota, and Hadza (table S11 and fig. S20). These results can only be explained by Taforalt harboring an ancestry that contains additional affinity with South, East, and Central African outgroups. None of the present-day or ancient Holocene African groups serve as a good proxy for this unknown ancestry, because adding them as the third source is still insufficient to match the model to the Taforalt gene pool (table S12 and fig. S21). However, we can exclude any branch in human genetic diversity more basal than the deepest known one represented by aSouthAfrica (4) as the source of this signal: it would result in a negative affinity to aSouthAfrica. not a positive one as we find (Fig. 4). Both an unknown archaic hominin and the recently proposed deep West African lineage (4) belong to this category and therefore cannot explain the Taforalt gene pool.

Mitochondrial consensus sequences of the Taforalt individuals belong to the U6a (six individuals) and M1b (one individual) haplogroups (15), which are mostly confined to present-day populations in North and East Africa (7). U6 and M1 have been proposed as markers for autochthonous Maghreb ancestry, which might have been originally introduced into this region by a back-to-Africa migration from West Asia (6, 7). The occurrence of both haplogroups in the Taforalt individuals proves their pre-Holocene presence in the Maghreb. We used the BEAST v1.8.1 package (24) to analyze the seven ancient Taforalt individuals in combination with four Upper Paleolithic European mtDNA genomes (22, 23) and present-day individuals belonging to U6 and M1 (7). By using a human mtDNA mutation rate inferred from tip calibration of ancient mtDNA genomes (23), we obtained divergence estimates for U6 at 37,000 yr B.P. (40,000 to 34,000 yr B.P. for 95% highest posterior density, HPD) and M1 at 24,000 yr B.P. (95% HPD: 29,000 to 20,000 yr B.P.) (table S15). Our estimated dates are considerably more recent than those of a study using present-day data only (45,000 ± 7000 yr B.P. for U6 and 37,000 ± 7000 yr B.P. for M1) (7) but are similar to those of Pennarun et al. (25). Moreover, we observed an asynchronous increase in the effective population size for U6 and M1 (fig. S24), which suggests that the demographic histories of these North and East African haplogroups do not coincide and might have been influenced by multiple expansions in the Late Pleistocene (25). Notably, the diversification of haplogroups U6a and M1 found for Taforalt is dated to ~24,000 yr B.P. (fig. S23), which is close in time to the earliest known appearance of the Iberomaurusian culture in Northwest Africa [25,845 to 25,270 cal. yr B.P. at Tamar Hat (26)].

The relationships of the Iberomaurusian culture with those of the preceding Middle Stone Age, including the local backed bladelet technologies in Northeast Africa, and the Epigravettian in Southern Europe have been questioned (13). The genetic profile of Taforalt suggests substantial Natufian-related and sub-Saharan Africanrelated ancestries (63.5 and 36.5%, respectively) but not additional ancestry from Epigravettian or other Upper Paleolithic European populations. Therefore, we provide genomic evidence for a Late Pleistocene connection between North Africa and the Near East, predating the Neolithic transition by at least four millennia, while rejecting the hypothesis of a potential Epigravettian gene flow from Southern Europe into northern Africa, within the resolution of our data. Archaeogenetic studies on additional Iberomaurusian sites will be critical to evaluate the representativeness of Taforalt for the Iberomaurusian gene pool. We speculate that the Natufian-related ancestral population may have been widespread across North Africa and the Near East, associated with microlithic backed bladelet technologies that started to spread out in this area by at least 25,000 yr B.P. [(10) and references therein]. However, given the absence of ancient genomic data from a similar time frame for this broader area, the epicenter of expansion, if any, for this ancestral population remains unknown.

Although the oldest Iberomaurusian microlithic bladelet technologies are found earlier in the Maghreb than their equivalents in northeastern Africa (Cyrenaica) and the earliest Natufian in the Levant, the complex sub-Saharan ancestry in Taforalt makes our individuals an unlikely proxy for the ancestral population of later Natufians who do not harbor sub-Saharan ancestry. An epicenter in the Maghreb is plausible only if the sub-Saharan African admixture into Taforalt either postdated the expansion into the Levant or was a locally confined phenomenon. Alternatively, placing the epicenter in Cyrenaica or the Levant requires an additional explanation for the observed archaeological chronology.

REFERENCES AND NOTES

- I. S. Castañeda *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20159–20163 (2009).
- B. M. Henn et al., Proc. Natl. Acad. Sci. U.S.A. 108, 5154–5162 (2011).
- 3. S. Mallick et al., Nature 538, 201-206 (2016).
- 4. P. Skoglund et al., Cell 171, 59-71.e21 (2017).
- V. J. Schuenemann et al., Nat. Commun. 8, 15694–15704 (2017).
 A. M. González et al., BMC Genomics 8, 223–234 (2007).
- A. M. Gonzalez et al., Binc Genomics 8, 223–234 (2)
 A. Olivieri et al., Science 314, 1767–1770 (2006).
- A. Gallego Llorente et al., Science 350, 820–822 (2015).
- 9. C. M. Schlebusch et al., Science **358**, 652–655 (2017).
- 10. R. N. Barton et al., J. Hum. Evol. 65, 266–281 (2013).
- 11. See supplementary materials
- 12. M.-T. Gansauge et al., Nucleic Acids Res. 45, e79 (2017).
- 13. Q. Fu et al., Proc. Natl. Acad. Sci. U.S.A. 110, 2223-2227 (2013)
- 14. Q. Fu et al., Nature 524, 216–219 (2015).
- 15. E. R. Jones et al., Nat. Commun. 6, 8912-8919 (2015).
- 16. I. Lazaridis et al., Nature 536, 419-424 (2016).
- 17. M. Raghavan et al., Nature 505, 87-91 (2014).
- 18. F. Cruciani et al., Mol. Biol. Evol. 24, 1300-1311 (2007).
- P. Pallary, in Mémoires de la Société Historique Algérienne (Jourdan, 1909), vol. 3.

- G. Camps, Les Civilisations Préhistoriques de l'Afrique du Nord et du Sahara (Doin, 1974).
- 21. N. Patterson et al., Genetics 192, 1065–1093 (2012).
- 22. M. Hervella et al., Sci. Rep. 6, 25501-25505 (2016).
- 23. C. Posth et al., Curr. Biol. 26, 827–833 (2016).
- 24. A. J. Drummond, A. Rambaut, *BMC Evol. Biol.* 7, 214 (2007).
- E. Pennarun *et al.*, *BMC Evol. Biol.* **12**, 234–245 (2012).
 J. T. Hogue, R. Barton, *Quat. Int.* **413**, 62–75 (2016).
- 27. Q. Fu et al., Nature **534**, 200–205 (2016).

ç, ..., , ...

ACKNOWLEDGMENTS

We thank H. Temming and A. Le Cabec (MPI-EVA) for CT scanning and G. Brandt, A. Wissgott, F. Aron, M. Burri, C. Freund, and R. Stahl (MPI-SHH) for DNA sequencing. **Funding:** This work was supported by the Max Planck Society, Institut National des Sciences de l'Archéologie et du Patrimoine (Protars grant P32/09-CNRST), the Natural

Environment Research Council (grants EFCHED NER/T/S/2002/ 00700 and RESET NE/E015670/1), the Leverhulme Trust (grant F/ 08 735/F), the British Academy, Oxford University (Fell Fund, Boise and Meyerstein), the Natural History Museum (Human Origins Research Fund), and the Calleva Foundation. Author contributions: J.K., A.B., J.J.-H., and L.H. conceived of the study. A.B., L.H., N.B., and J.-J.H. provided archaeological material and input for the archaeological interpretation. M.v.d.L., B.N., and S.N. performed laboratory work with the help of A.A.-P. and M.M. M.v.d.L., C.J., C.P., and W.H. analyzed data. M.v.d.L., C.J., J.K., C.P., A.B., L.H., N.B., and M.M. wrote the manuscript with input from all coauthors. Competing interests: A.B. keeps an additional affiliation with the MPI-EVA; this institute supported his excavation efforts and worked with him on the site. This is also reflected in the coauthorship of J.J.-H., director of MPI-EVA. The authors declare no competing interests. Data and materials availability: Genomic data (BAM format) are available through the

Sequence Read Archive (accession number SRP132033) and consensus mitogenome sequences (FASTA format) in GenBank (accession numbers MG936619 to MG936625).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6388/548/suppl/DC1 Supplementary Text Figs. S1 to S26 Tables S1 to S16 References (28–114)

21 December 2017; accepted 28 February 2018 Published online 15 March 2018 10.1126/science.aar8380

Science

Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations

Marieke van de Loosdrecht, Abdeljalil Bouzouggar, Louise Humphrey, Cosimo Posth, Nick Barton, Ayinuer Aximu-Petri, Birgit Nickel, Sarah Nagel, El Hassan Talbi, Mohammed Abdeljalil El Hajraoui, Saaïd Amzazi, Jean-Jacques Hublin, Svante Pääbo, Stephan Schiffels, Matthias Meyer, Wolfgang Haak, Choongwon Jeong and Johannes Krause

Science **360** (6388), 548-552. DOI: 10.1126/science.aar8380originally published online March 15, 2018

Relationships among North Africans

The general view is that Eurasians mostly descend from a single group of humans that dispersed outside of sub-Saharan Africa around 50,000 to 100,000 years ago. Present-day North Africans share a majority of their ancestry with present-day Near Easterners, but not with sub-Saharan Africans. To investigate this conundrum, Van de Loosdrecht *et al.* sequenced high-quality DNA obtained from bone samples of seven individuals from Taforalt in eastern Morocco dating from the Later Stone Age, about 15,000 years ago. The Taforalt individuals were found to be most closely related to populations from the Near East (Natufians), with a third of their ancestry from sub-Saharan Africa. No evidence was found for introgression with western Europeans, despite attribution to the Iberomaurusian culture. None of the present-day or ancient Holocene African groups are a good proxy for the sub-Saharan genetic component. *Science*, this issue p. 548

ARTICLE TOOLS	http://science.sciencemag.org/content/360/6388/548
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2018/03/14/science.aar8380.DC1
REFERENCES	This article cites 102 articles, 25 of which you can access for free http://science.sciencemag.org/content/360/6388/548#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.