

HUMAN GENETICS

Early Neolithic genomes from the eastern Fertile Crescent

Farnaz Broushaki,¹ Mark G. Thomas,² Vivian Link,^{3,4} Saioa López,² Lucy van Dorp,² Karola Kiršanow,¹ Zuzana Hofmanová,¹ Yoan Diekmann,² Lara M. Cassidy,⁵ David Díez-del-Molino,^{2,6} Athanasios Kousathanas,^{3,4,7} Christian Sell,¹ Harry K. Robson,⁸ Rui Martiniano,⁵ Jens Blöcher,¹ Amelie Scheu,^{1,5} Susanne Kreutzer,¹ Ruth Bollongino,¹ Dean Bobo,⁹ Hossein Davoudi,¹⁰ Olivia Munoz,¹¹ Mathias Currat,¹² Kamyar Abdi,¹³ Fereidoun Biglari,¹⁴ Oliver E. Craig,⁸ Daniel G. Bradley,⁵ Stephen Shennan,¹⁵ Krishna R. Veeramah,⁹ Marjan Mashkour,¹⁶ Daniel Wegmann,^{3,4*†} Garrett Hellenthal,^{2*†} Joachim Burger^{1*†}

We sequenced Early Neolithic genomes from the Zagros region of Iran (eastern Fertile Crescent), where some of the earliest evidence for farming is found, and identify a previously uncharacterized population that is neither ancestral to the first European farmers nor has contributed substantially to the ancestry of modern Europeans. These people are estimated to have separated from Early Neolithic farmers in Anatolia some 46,000 to 77,000 years ago and show affinities to modern-day Pakistani and Afghan populations, but particularly to Iranian Zoroastrians. We conclude that multiple, genetically differentiated hunter-gatherer populations adopted farming in southwestern Asia, that components of pre-Neolithic population structure were preserved as farming spread into neighboring regions, and that the Zagros region was the cradle of eastward expansion.

The earliest evidence for cultivation and stock-keeping is found in the Neolithic core zone of the Fertile Crescent (1, 2); a region stretching north from the southern Levant through eastern Anatolia and northern Mesopotamia, then east into the Zagros Mountains on the border of modern-day Iran and Iraq (Fig. 1). From there, farming spread into surrounding regions, including Anatolia and, later, Europe, southern Asia, and parts of Arabia and North Africa. Whether the transition to agriculture was

a homogeneous process across the core zone, or a mosaic of localized domestications, is unknown. Likewise, the extent to which core zone farming populations were genetically homogeneous, or exhibited structure that may have been preserved as agriculture spread into surrounding regions, is undetermined.

Ancient DNA (aDNA) studies indicate that early Aegean farmers dating to ~6500 to 6000 BCE are the main ancestors of early European farmers (3, 4), although it is not known if they were predominantly descended from core zone farming populations. We sequenced four Early Neolithic (EN) genomes from Zagros, Iran, including one to 10× mean coverage from a well-preserved male sample from the central Zagros site of Wezmeh Cave [WC1, 7455 to 7082 calibrated years (cal) BCE]. The three other individuals were from Tepe Abdul Hosein and were less well preserved (genome coverage between 0.6 and 1.2×) but are around 10,000 years old, and therefore are among the earliest Neolithic human remains in the world (tables S1 and S3).

Despite a lack of a clear Neolithic context, the radiocarbon-inferred chronological age and palaeodietary data support WC1 being an early farmer (tables S1 to S3 and fig. S7). WC1 bone collagen $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values are indistinguishable from those of a securely assigned Neolithic individual from Abdul Hosein and consistent with a diet rich in cultivated C_3 cereals rather than animal protein. Specifically, collagen from WC1 and Abdul Hosein is ^{13}C depleted compared to those from contemporaneous wild and domestic fauna from this region (5), which consumed C_4 plants. Crucially, WC1 and the Abdul Hosein farmers exhibit very similar genomic signatures.

The four EN Zagros genomes form a distinct cluster in the first two dimensions of a principal components analysis (PCA; Fig. 2); they plot closest to modern-day Pakistanis and Afghans and are well separated from European hunter-gatherers (HG) and other Neolithic farmers. In an outgroup f_3 -test (6, 7) (figs. S17 to S20), all four Neolithic Iranian individuals are genetically more similar to each other than to any other pre-historic genome except a Chalcolithic genome from northwestern Anatolia (see below). Despite ^{14}C dates spanning around 1200 years, these data are consistent with all four genomes being sampled from a single eastern Fertile Crescent EN population.

Examination of runs of homozygosity (ROH) above 500 kb in length in WC1 demonstrated that he shared a similar ROH distribution with European and Aegean Neolithics, as well as modern-day Europeans (Fig. 3, A and B). However, of all ancient samples considered, WC1 displays the lowest total length of short ROH, suggesting that he was descended from a relatively large HG population. In contrast, the ROH distributions of the HG Kotias from Georgia, and Loschbour from Luxembourg, indicate prolonged periods of small ancestral population size (8).

We also developed a method to estimate heterozygosity ($\hat{\theta}$) in 1-Mb windows that takes into account postmortem damage and is unbiased even at low coverage (9) (Fig. 3, C and D). The mean $\hat{\theta}$ in WC1 was higher than in HG individuals (Bichon and Kotias), similar to that in Bronze Age individuals from Hungary and modern Europeans, and lower than in ancient (10) and modern Africans. Multidimensional scaling on a matrix of centered Spearman correlations of local $\hat{\theta}$ across the whole genome again puts WC1 closer to modern populations than to ancient foragers, indicating that both the mean and distribution of diversity over the genome are more similar to those of modern populations (Fig. 3E). However, WC1 does have an excess of long ROH segments (>1.6 Mb), relative to Aegean and European Neolithics (Fig. 3B). This includes several very long (7 to 16 Mb) ROH segments (Fig. 3A), confirmed by low $\hat{\theta}$ estimates in those regions (Fig. 3C). These regions do not show reduced coverage in WC1 nor a reduction in diversity in other samples, with the exception of the longest such segment where we find reduced diversity in modern and HG individuals, although less extended than in WC1 (7) (Fig. 3B). This observed excess of long segments of reduced heterozygosity could be the result of cultural practices such as consanguinity and endogamy, or demographic constraints such as a recent or ongoing bottleneck (11).

The extent of population genetic structure in Neolithic southwestern Asia has important implications for the origins of farming. High levels of structuring would be expected under a scenario of localized independent domestication processes by distinct populations, whereas low structure would be more consistent with a single population origin of farming or a diffuse homogeneous domestication process, perhaps

¹Palaeogenetics Group, Johannes Gutenberg University Mainz, 55099 Mainz, Germany. ²Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK. ³Department of Biology, University of Fribourg, 1700 Fribourg, Switzerland. ⁴Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. ⁵Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland. ⁶Department of Bioinformatics and Genetics, Swedish Museum of Natural History, SE-10405, Stockholm, Sweden. ⁷Unit of Human Evolutionary Genetics, Institut Pasteur, 75015 Paris, France. ⁸BioArCh, Department of Archaeology, University of York, York YO10 5YW, UK. ⁹Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794-5245, USA. ¹⁰Department of Archaeology, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran. ¹¹UMR 7041 ArScAn-VEPMO, Maison de l'Archéologie et de l'Ethnologie, 21 allée de l'Université, 92023 Nanterre, France. ¹²Department of Genetics and Evolution-Anthropology Unit, University of Geneva, 1211 Geneva, Switzerland. ¹³Samuel Jordan Center for Persian Studies and Culture, University of California-Irvine, Irvine, CA 92697-3370, USA. ¹⁴Paleolithic Department, National Museum of Iran, 113617111, Tehran, Iran. ¹⁵Institute of Archaeology, University College London, London WC1H 0PY, UK. ¹⁶CNRS/MNHN/SUS-UMR 7209, Archéozoologie et Archéobotanique, Sociétés, Pratiques et Environnements, Département Ecologie et Gestion de la Biodiversité, 55 rue Buffon, 75005 Paris, France.

*These authors contributed equally to this work. †Corresponding author. Email: jburger@uni-mainz.de (J.B.); g.hellenthal@ucl.ac.uk (G.H.); daniel.wegmann@unifr.ch (D.W.)



Fig. 1. Map of prehistoric Neolithic and Iron Age Zagros genome locations. Colors indicate isochrones, with numbers giving approximate arrival times of the Neolithic culture (in years BCE).

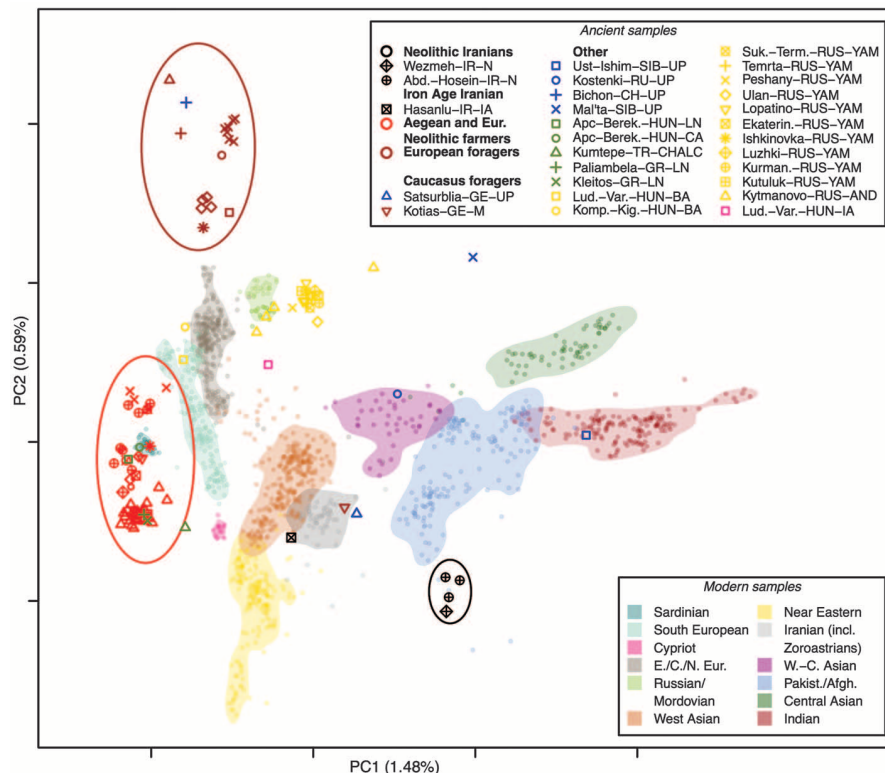
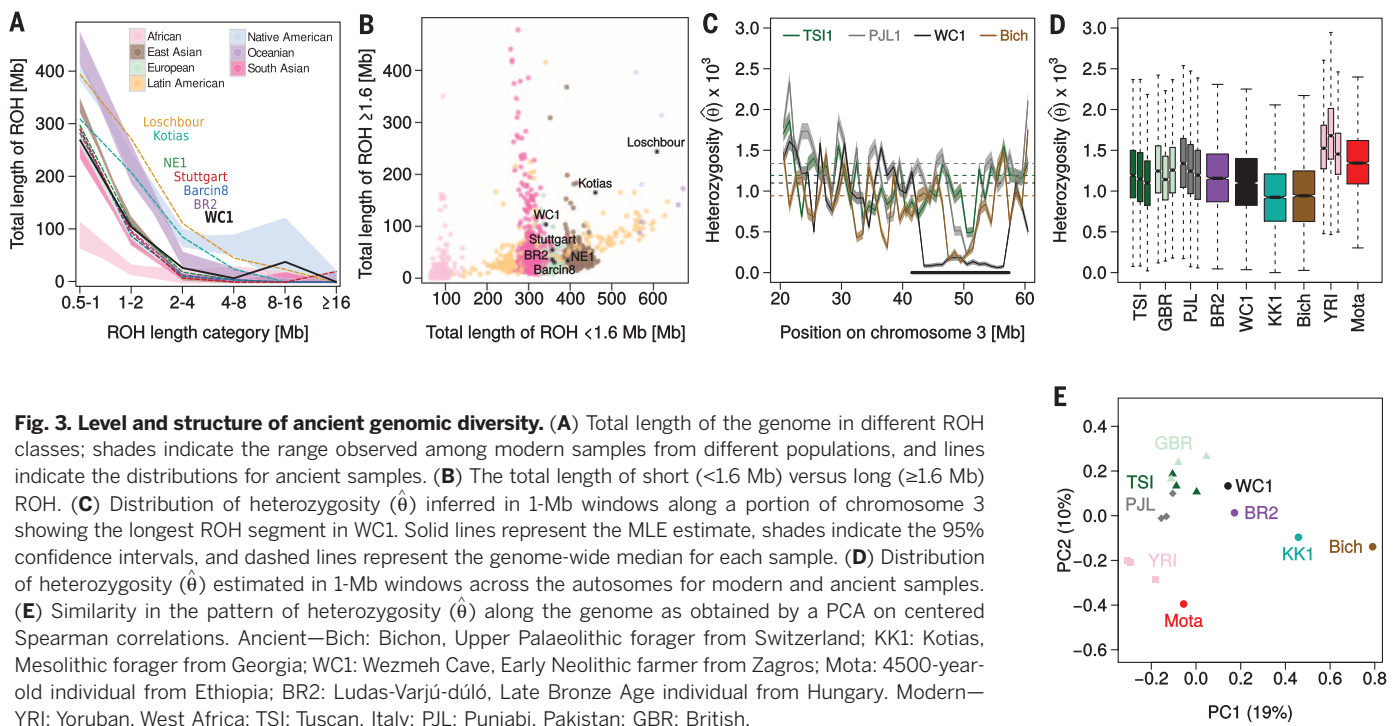


Fig. 2. PCA plot of Zagros, European, and Near and Middle Eastern ancient genomes. Comparison of ancient and modern genomes shows that Neolithic Zagros genomes form a discrete genetic cluster close to modern Pakistani and Afghan genomes but distinct from the genomes of other Neolithic farmers and European hunter-gatherers. See animation S1 for an interactive three-dimensional version of the PCA, including the third principal component.

involving high rates of gene flow across the entire Neolithic core zone. The ancient Zagros individuals show stronger affinities to Caucasus HGs (table S17.1), whereas Neolithic Aegeans showed closer affinities to other European HGs (tables S17.2 and S17.3). Formal tests of admixture of the form $f_3(\text{Neo_Iranian, HG; Anatolia_Neolithic})$ were all positive with Z-scores above 15.78 (table S17.6), indicating that Neolithic northwestern Anatolians did not descend from a population formed by the mixing of Zagros Neolithics and known HG groups. These results suggest that Neolithic populations from northwestern Anatolia and the Zagros descended from distinct ancestral populations. Furthermore, although the Caucasus HGs are genetically closest to EN Zagros individuals, they also share unique ancestry with eastern, western, and Scandinavian European HGs (table S16.1), indicating that they are not the direct ancestors of Zagros Neolithics.

The significant differences between ancient Iranians, Anatolian/European farmers, and European HGs suggest a pre-Neolithic separation. Assuming a mutation rate of 5×10^{-10} per site per year (12), the inferred mean split time for Anatolian/European farmers (as represented by Bar8, 4) and European HGs (Loschbour) ranged from 33,000 to 39,000 years ago [combined 95% confidence interval (CI) 15,000 to 61,000 years ago], whereas the preceding divergence of the ancestors of Neolithic Iranians (WCI) occurred 46,000 to 77,000 years ago (combined 95% CI 38,000 to 104,000 years ago) (13) (fig. S48 and tables S34 and S35). Furthermore, the European

Downloaded from https://www.science.org on September 25, 2022



HGs were inferred to have an effective population size (N_e) that was ~ 10 to 20% of either Neolithic farming group, consistent with the ROH and $\hat{\theta}$ analyses.

Levels of inferred Neanderthal ancestry in WC1 are low (fig. S22 and table S21), but fall within the general trend described recently in Fu *et al.* (14). Fu *et al.* (14) also inferred a basal Eurasian ancestry component in the Caucasus HG sample Satsurblia when examined within the context of a “base model” for various ancient Eurasian genomes dated from $\sim 45,000$ to 7,000 years ago. We examined this base model using ADMIXTUREGRAPH (6) and inferred almost twice as much basal Eurasian ancestry for WC1 as for Satsurblia (62 versus 32%) (fig. S52), with the remaining ancestry derived from a population most similar to ancient north Eurasians such as Mal’ta1 (15). Thus, Neolithic Iranians appear to derive predominantly from the earliest known Eurasian population branching event (7).

“Chromosome painting” and an analysis of recent haplotype sharing using a Bayesian mixture model (7) revealed that, when compared to 160 to 220 modern groups, WC1 shared a high proportion ($>95\%$) of recent ancestry with individuals from the Middle East, Caucasus, and India. We also compared WC1’s haplotype-sharing profile to that of three high-coverage Neolithic genomes from northwestern Anatolia (Bar8; Barcin, Fig. 4), Germany (LBK; Stuttgart), and Hungary (NE1; Polgár-Ferenci-hát). Unlike WC1, these Anatolian and European Neolithics shared ~ 60 to 100% of recent ancestry with modern groups sampled from southern Europe (figs. S24, S30, and S32 to S37; table S22).

We also examined recent haplotype sharing between each modern group and ancient Neo-

lithic genomes from Iran (WC1) and Europe (LBK, NE1), HG genomes sampled from Luxembourg (Loschbour) and the Caucasus (KK1; Kotias), a 4500-year-old genome from Ethiopia (Mota) and Ust’-Ishim, and a 45,000-year-old genome from Siberia. Modern groups from south, central, and northwestern Europe shared haplotypes predominantly with European Neolithic samples LBK and NE1, and European HGs, whereas modern Near and Middle Eastern, as well as southern Asian samples, had higher sharing with WC1 (figs. S28 and S29). Modern Pakistani, Iranian, Armenian, Tajikistani, Uzbekistani, and Yemeni samples were inferred to share $>10\%$ of haplotypes with WC1. This was true even when modern groups from neighboring geographic regions were added as potential ancestry surrogates (figs. S26 and S27 and table S23). Iranian Zoroastrians had the highest inferred sharing with WC1 out of all modern groups (table S23). Consistent with this, outgroup f_3 statistics indicate that Iranian Zoroastrians are the most genetically similar to all four Neolithic Iranians, followed by other modern Iranians (Fars), Balochi (southeastern Iran, Pakistan, and Afghanistan), Brahui (Pakistan and Afghanistan), Kalash (Pakistan), and Georgians (figs. S12 to S15). Interestingly, WC1 most likely had brown eyes, relatively dark skin, and black hair, although Neolithic Iranians carried reduced pigmentation-associated alleles in several genes and derived alleles at 7 of the 12 loci showing the strongest signatures of selection in ancient Eurasians (3) (tables S29 to S33). Although there is a strong Neolithic component in these modern south Asian populations, simulation of allele sharing rejected full population continuity under plausible ancestral population

sizes, indicating some population turnover in Iran since the Neolithic (7).

While Early Neolithic samples from eastern and western southwest Asia differ conspicuously, comparisons to genomes from Chalcolithic Anatolia and Iron Age Iran indicate a degree of subsequent homogenization. Kumtepe6, a ~ 6750 -year-old genome from northwestern Anatolia (16), was more similar to Neolithic Iranians than to any other non-Iranian ancient genome (figs. S17 to S20 and table S18.1). Furthermore, our male Iron Age genome (F38; 971 to 832 BCE; sequenced to 1.9 \times) from Tepe Hasanlu in northwestern Iran shares greatest similarity with Kumtepe6 (fig. S21) even when compared to Neolithic Iranians (table S20). We inferred additional non-Iranian or non-Anatolian ancestry in F38 from sources such as European Neolithics and even post-Neolithic Steppe populations (table S20). Consistent with this, F38 carried a N1a subclade mitochondrial DNA (mtDNA), which is common in early European and northwestern Anatolian farmers (3). In contrast, his Y chromosome belongs to subhaplogroup R1b1a2a2, also found in five Yamnaya individuals (17) and in two individuals from the Poltavka culture (3). These patterns indicate that post-Neolithic homogenization in southwestern Asia involved substantial bidirectional gene flow between the east and west of the region, as well as possible gene flow from the Steppe.

Migration of people associated with the Yamnaya culture has been implicated in the spread of Indo-European languages (17, 18), and some level of Near Eastern ancestry was previously inferred in southern Russian pre-Yamnaya populations (3). However, our analyses suggest that Neolithic Iranians were unlikely to be the main source of Near Eastern ancestry in the Steppe population

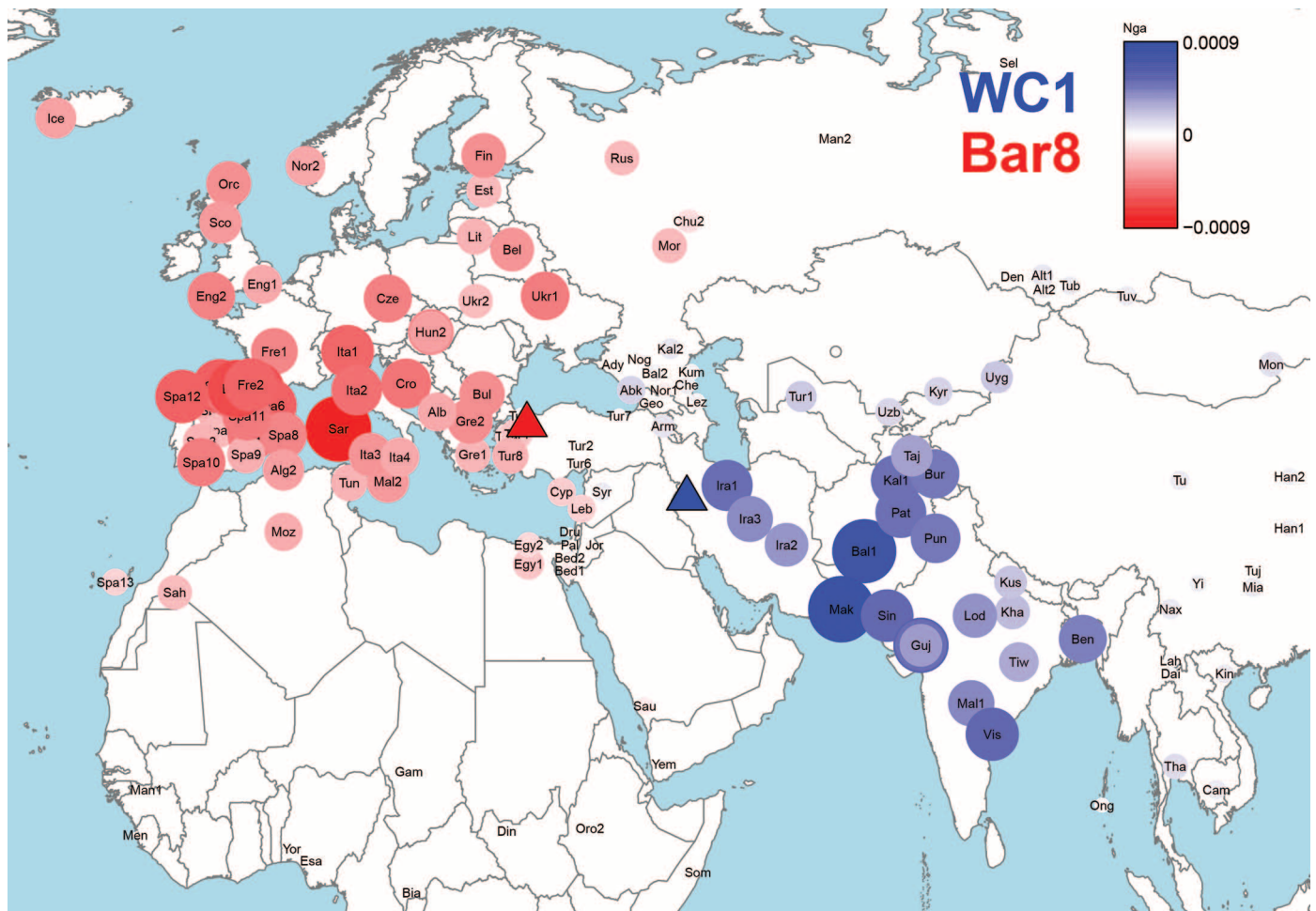


Fig. 4. Modern-day peoples with affinity to WC1. Modern groups with an increasingly higher (respectively lower) inferred proportion of haplotype sharing with the Iranian Neolithic Wezmeh Cave (WC1, 7455 to 7082 cal BCE, blue triangle) compared to the Anatolian Neolithic Barçın genome (Bar8; 6212 to 6030 cal BCE, red triangle) are depicted with an increasingly stronger blue or red color, respectively. Circle sizes illustrate the relative absolute proportion of this difference between WC1 versus Bar8. The key for the modern group labels is provided in table S24.

(table S20) and that this ancestry in pre-Yamnaya populations originated primarily in the west of southwest Asia.

We also inferred shared ancestry between Steppe and Hasanlu Iron Age genomes that was distinct from EN Iranians (table S20) (7). In addition, modern Middle Easterners and South Asians appear to possess mixed ancestry from ancient Iranian and Steppe populations (tables S19 and S20). However, Steppe-related ancestry may also have been acquired indirectly from other sources (7), and it is not clear if this is sufficient to explain the spread of Indo-European languages from a hypothesized Steppe homeland to the region where Indo-Iranian languages are spoken today. Yet the affinities of Zagros Neolithic individuals to modern populations of Pakistan, Afghanistan, Iran, and India is consistent with a spread of Indo-Iranian languages, or of Dravidian languages (which includes Brahui), from the Zagros into southern Asia, in association with farming (19).

The Neolithic transition in southwest Asia involved the appearance of different domestic

species, particularly crops, in different parts of the Neolithic core zone, with no single center (20). Early evidence of plant cultivation and goat management between the 10th and the 8th millennium BCE highlights the Zagros as a key region in the Neolithization process (1). Given the evidence of domestic species movement from east to west across southwest Asia (21), it is surprising that EN human genomes from the Zagros are not closely related to those from northwestern Anatolia and Europe. Instead they represent a previously undescribed Neolithic population. Our data show that the chain of Neolithic migration into Europe does not reach back to the eastern Fertile Crescent, also raising questions about whether intermediate populations in southeastern and Central Anatolia form part of this expansion. Nevertheless, it seems probable that the Zagros region was the source of an eastern expansion of the southwestern Asian domestic plant and animal economy. Our inferred persistence of ancient Zagros genetic components in modern day south Asians lends weight to a strong demic component to this expansion.

REFERENCES AND NOTES

1. S. Riehl, M. Zeidi, N. J. Conard, *Science* **341**, 65–67 (2013).
2. R. Matthews, W. Matthews, Y. Mohammadifard, Eds., *The Earliest Neolithic of Iran: 2008 Excavations at Sheikh-e Abad and Jani Central Zagros* (Oxbow Books, Oxford, 2013), vol. 1.
3. I. Mathieson et al., *Nature* **528**, 499–503 (2015).
4. Z. Hofmanová et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6886–6891 (2016).
5. G. Müldner, Isotope analysis of animal bone, in *The Earliest Neolithic of Iran: 2008 Excavations at Sheikh-e Abad and Jani*, R. Matthews, W. Matthews, Y. Mohammadifard, Eds. (Oxbow Books, 2013).
6. N. Patterson et al., *Genetics* **192**, 1065–1093 (2012).
7. Materials and methods are available as supplementary materials on Science Online.
8. E. R. Jones et al., *Nat. Commun.* **6**, 8912 (2015).
9. A. Kousathanas et al., Inferring heterozygosity from ancient and low coverage genomes. *bioRxiv* 10.1101/046748 (2016).
10. M. G. Llorente et al., *Science* **350**, 820–822 (2015).
11. T. J. Pemberton et al., *Am. J. Hum. Genet.* **91**, 275–292 (2012).
12. A. Kong et al., *Nature* **488**, 471–475 (2012).
13. M. W. Nachman, S. L. Crowell, *Genetics* **156**, 297–304 (2000).
14. Q. Fu et al., *Nature* **534**, 200–205 (2016).
15. M. Raghavan et al., *Nature* **505**, 87–91 (2014).

16. A. Omrak *et al.*, *Curr. Biol.* **26**, 270–275 (2016).
 17. W. Haak *et al.*, *Nature* **522**, 207–211 (2015).
 18. M. E. Allentoft *et al.*, *Nature* **522**, 167–172 (2015).
 19. J. P. Mallory, D. Q. Adams, Eds., *Encyclopedia of Indo-European Culture* (Taylor & Francis, London, 1997).
 20. D. Q. Fuller, G. Willcox, R. G. Allaby, *J. Exp. Bot.* **63**, 617–633 (2012).
 21. A. Scheu *et al.*, *BMC Genet.* **16**, 54 (2015).

ACKNOWLEDGMENTS

This paper is a product of the Palaeogenome Analysis Team (PAT). F.Br. was supported by funds of Johannes Gutenberg University Mainz given to J.Bu., Z.H., and R.M. are supported by a Marie Curie Initial Training Network (BEAN—Bridging the European and Anatolian Neolithic, GA no. 289966). C.S. was supported by the European Union (EU) SYNTHESIS—Synthesis of Systematic Resources, GA no. 226506-CP-CSA-INFRA, and Deutsche Forschungsgemeinschaft (DFG) (BO 4119/1). A.S. was supported by the EU CodeX Project no. 295729. M.C. was supported by Swiss NSF grant 31003A_156853. A.K. and D.W. were supported

by Swiss NSF grant 31003A_149920. S.L. is supported by Biotechnology and Biological Sciences Research Council (grant no. BB/L009382/1). L.v.D. is supported by CoMPLEX via the Engineering and Physical Sciences Research Council (grant no. EP/F500351/1). G.H. is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant no. 098386/Z/12/Z) and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. M.G.T. and Y.D. are supported by a Wellcome Trust Senior Research Fellowship awarded to M.G.T. J.Bu. is grateful for support by the high-performance computing cluster MOGON (funded by DFG; INST 247/602-1 FUGG). L.M.C. is funded by the Irish Research Council (GOIPG/2013/1219). M.M. was supported by the UMR 7209, CNRS National Museum of Natural History—Sorbonne University and ANR-14-CE03-0008-01-CNRS ANR Kharman. O.M. was supported by Institut Français de Recherche en Iran (October 2015). F.Bi., M.M., O.M., and H.D. thank the National Museum of Iran and especially J. Nokandeh, director of National Museum of Iran. We thank N. Patterson for early access to the latest version of

ppgraph. Accession numbers: Mitochondrial genome sequences are deposited in GenBank (KX353757–KX353761). Genomic data are available at ENA (European Nucleotide Archive) with the accession number PRJEB14180 in BAM format. Iranian Zoroastrian and Fars genotype data are available in plink format at figshare: <https://dx.doi.org/10.6084/m9.figshare.3470135>.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6298/499/suppl/DC1
 Materials and Methods
 Figs. S1 to S52
 Tables S1 to S37
 References (22–162)
 Animation S1

7 April 2016; accepted 5 July 2016
 Published online 14 July 2016
 10.1126/science.aaf7943

STRUCTURAL BIOLOGY

Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor

Jian Lei,^{1,2} Guido Hansen,¹ Christoph Nitsche,^{3*} Christian D. Klein,³
 Linlin Zhang,^{1,2} Rolf Hilgenfeld^{1,2†}

The ongoing Zika virus (ZIKV) outbreak is linked to severe neurological disorders. ZIKV relies on its NS2B/NS3 protease for polyprotein processing; hence, this enzyme is an attractive drug target. The 2.7 angstrom crystal structure of ZIKV protease in complex with a peptidomimetic boronic acid inhibitor reveals a cyclic diester between the boronic acid and glycerol. The P2 4-aminomethylphenylalanine moiety of the inhibitor forms a salt-bridge with the nonconserved Asp⁸³ of NS2B; ion-pairing between Asp⁸³ and the P2 residue of the substrate likely accounts for the enzyme's high catalytic efficiency. The unusual dimer of the ZIKV protease:inhibitor complex seen in the crystal may provide a model for assemblies formed at high local concentrations of protease at the endoplasmic reticulum membrane, the site of polyprotein processing.

Previously considered a rare and mild pathogen for humans (1), Zika virus (ZIKV) infection has recently been found to be responsible for neurological disorders in a substantial portion of patients. The infection can trigger Guillain-Barré syndrome (2), and prenatal ZIKV infection is responsible for a dramatically increased number of microcephaly cases in fetuses and newborn children (3). The World Health Organization (WHO) recently declared the association of ZIKV infec-

tion with these neurological disorders a Public Health Emergency of International Concern (4). There are no vaccines or antiviral drugs available for protection from or treatment of ZIKV infection.

ZIKV is a member of the genus *Flavivirus* in the Flaviviridae family of RNA viruses. Its ~10.7-kb single-stranded RNA genome of positive polarity encodes a single polyprotein, which, by analogy to other flaviviruses, is assumed to be cleaved by host-cell proteases (signalase and furin) and the viral NS2B/NS3 protease into three structural (C, prM/M, and E) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (fig. S1). Similar to other flavivirus proteases, such as those of dengue virus (DENV) and West Nile virus (WNV), the mature form of ZIKV protease consists of the N-terminal domain of NS3, which carries the catalytic triad Ser¹³⁵-His⁵¹-Asp⁷⁵, and the membrane-bound NS2B (a sequence alignment is available in fig. S2). Crystallization of this complex has not been

successful so far for any flavivirus protease, but it has been shown that a construct comprising ~40 hydrophilic residues of NS2B and ~185 residues of NS3, covalently linked via a Gly₄-Ser-Gly₄ sequence, displays strong peptidolytic activity (5). Crystal structures of the free form of this protease construct (“NS2B-NS3^{pro}”) usually reveal an “open conformation” featuring a well-ordered NS3^{pro} core and a flexible NS2B part that shows only limited interaction with NS3^{pro}, whereas inhibitor (and presumably substrate) binding induces a pronounced conformational change of NS2B yielding a more compact, “closed” form (6, 7).

We expressed in *Escherichia coli* a DNA construct corresponding to the NS2B-NS3^{pro}-coding region of the Brazilian ZIKV isolate BeH823339 (GenBank accession number KU729217.2) (8). This construct codes for residues 49 to 95 of ZIKV NS2B, the C terminus of which is covalently linked via Gly₄-Ser-Gly₄ to the N terminus of NS3 (residues 1 to 170). The recombinant enzyme obtained is a mixture of monomer, disulfide-linked dimer (here designated “SS-dimer”) and—to a lesser extent—higher oligomers (fig. S3). The double mutant Cys⁸⁰Ser/Cys¹⁴³Ser leads to loss of the disulfide bond, which occurs between Cys¹⁴³ residues of different polypeptide chains, as revealed by our x-ray structure. The SS-dimer and the monomer obtained by reduction with tris(2-carboxyethyl)phosphine (TCEP) (fig. S3) as well as the Cys⁸⁰Ser/Cys¹⁴³Ser mutant of ZIKV NS2B-NS3^{pro} are hyperactive against the standard flavivirus protease substrate benzoyl-norleucine-lysine-lysine-arginine 7-amino-4-methylcoumarin (Bz-Nle-Lys-Lys-Arg-AMC), with a very low Michaelis constant (K_m) and a specific catalytic efficiency (k_{cat}/K_m) more than 20 times higher than for the WNV enzyme (Table 1).

In order to elucidate the molecular basis of this hyperactivity, and to provide a starting point for structure-based drug design efforts, we have crystallized ZIKV NS2B-NS3^{pro} in the closed form and determined its x-ray structure at 2.7 Å resolution. Containing two molecules (“A” and “B”) per asymmetric unit of the crystal, the structure reveals the same chymotrypsin-like fold for the NS3^{pro} domain as seen previously for other flavivirus proteases, with the NS2B polypeptide

¹Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany. ²German Center for Infection Research (DZIF), Hamburg–Lübeck–Borstel–Riems Site, University of Lübeck, Germany. ³Medicinal Chemistry, Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany.

*Present address: Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia.

†Corresponding author. Email: hilgenfeld@biochem.uni-luebeck.de

Early Neolithic genomes from the eastern Fertile Crescent

Farnaz Broushaki, Mark G. Thomas, Vivian Link, Saioa López, Lucy van Dorp, Karola Kirsanow, Zuzana Hofmanová, Yoan Diekmann, Lara M. Cassidy, David Díez-del-Molino, Athanasios Kousathanas, Christian Sell, Harry K. Robson, Rui Martiniano, Jens Blöcher, Amelie Scheu, Susanne Kreutzer, Ruth Bollongino, Dean Bobo, Hossein Davoudi, Olivia Munoz, Mathias Currat, Kamyar Abdi, Fereidoun Biglari, Oliver E. Craig, Daniel G. Bradley, Stephen Shennan, Krishna R. Veeramah, Marjan Mashkour, Daniel Wegmann, Garrett Hellenthal, and Joachim Burger

Science, 353 (6298), • DOI: 10.1126/science.aaf7943

Near Eastern genomes from Iran

The genetic composition of populations in Europe changed during the Neolithic transition from hunting and gathering to farming. To better understand the origin of modern populations, Broushaki *et al.* sequenced ancient DNA from four individuals from the Zagros region of present-day Iran, representing the early Neolithic Fertile Crescent. These individuals unexpectedly were not ancestral to early European farmers, and their genetic structures did not contribute significantly to those of present-day Europeans. These data indicate that a parallel Neolithic transition probably resulted from structured farming populations across southwest Asia.

Science, this issue p. 499

View the article online

<https://www.science.org/doi/10.1126/science.aaf7943>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)