



Scottish Environment, Food and Agriculture Research Institutes

Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley

Mascher, Martin; J. Schuenemann, Verena; Davidovich, Uri; Marom, Nimrod; Himmelbach, Axel; Hübner, Sarel; Korol, Abraham; David, Michal; Reiter, Ella; Riehl, Simone; Schreiber, Mona; Vohr, Samuel H.; Green, Richard E.; Dawson, Ian K.; Russell, Joanne; Kilian, Benjamin; Muehlbauer, Gary J.; Waugh, Robert; Fahima, Tzion; Krause, Johannes; Weiss, Ehud; Stein, Nils

Published in:
Nature Genetics

DOI:
[10.1038/ng.3611](https://doi.org/10.1038/ng.3611)

Publication date:
2016

Document Version
Peer reviewed version

Citation for published version (APA):

Mascher, M., J. Schuenemann, V., Davidovich, U., Marom, N., Himmelbach, A., Hübner, S., ... Stein, N. (2016). Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. *Nature Genetics*, 48(9), 1089-1093. DOI: 10.1038/ng.3611

General rights

Copyright and moral rights for the publications made accessible in SEFARI website are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from SEFARI website for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

1 **Genomic analysis of 6,000-year-old cultivated grain illuminates the**
2 **domestication history of barley**

3

4 Martin Mascher^{1,2,19}, Verena J. Schuenemann^{3,4,19}, Uri Davidovich⁵, Nimrod Marom⁶,
5 Axel Himmelbach¹, Sarel Hübner^{7,8}, Abraham Korol^{9,10}, Michal David¹¹, Ella
6 Reiter³, Simone Riehl^{3,4}, Mona Schreiber¹, Samuel H. Vohr¹², Richard E. Green¹², Ian
7 K. Dawson¹³, Joanne Russell¹³, Benjamin Kilian^{1,18}, Gary J. Muehlbauer^{14,15}, Robbie
8 Waugh^{13,16}, Tzion Fahima^{9,10}, Johannes Krause^{17,3,4}, Ehud Weiss¹¹, Nils Stein¹

9

10 1) Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben,
11 Seeland, Germany

12 2) German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig,
13 Leipzig, Germany

14 3) University of Tübingen, Institute for Archaeological Sciences, Tübingen, Germany

15 4) University of Tübingen, Senckenberg Center for Human Evolution and
16 Paleoenvironment, Tübingen, Germany

17 5) The Hebrew University, Institute of Archaeology, Jerusalem, Israel

18 6) University of Haifa, Laboratory of Archaeozoology, Zinman Institute of
19 Archaeology, Haifa, Israel

20 7) University of British Columbia, Department of Botany, Vancouver, Canada

21 8) Tel Hai College, Department of Biotechnology, Upper Galilee, Israel

22 9) University of Haifa, Institute of Evolution, Haifa, Israel

23 10) University of Haifa, Department of Evolutionary and Environmental Biology,
24 Haifa, Israel

25 11) Bar-Ilan University, The Martin (Szusz) Department of Land of Israel Studies and
26 Archaeology, Ramat-Gan, Israel

27 12) University of California, Santa Cruz, Department of Biomolecular Engineering,
28 Santa Cruz, USA

29 13) The James Hutton Institute, Cell and Molecular Sciences, Invergowrie, Dundee,
30 United Kingdom

31 14) University of Minnesota, Department of Plant Biology, St. Paul, USA

32 15) University of Minnesota, Department of Agronomy and Plant Genetics, St. Paul,
33 USA

34 16) University of Dundee, Division of Plant Sciences, Dundee United Kingdom

35 17) Max Planck Institute for the Science of Human History, Jena, Germany
36 18) present address: Bayer CropScience NV, BCS Breeding & Trait Development,
37 Zwijnaarde (Gent), Belgium

38 19) These authors contributed equally to this work.

39

40 Correspondence should be addressed to Tzion Fahima (fahima@research.haifa.ac.il),
41 Johannes Krause (krause@shh.mpg.de), Ehud Weiss (ehud.weiss@biu.ac.il) or Nils
42 Stein (stein@ipk-gatersleben.de).

43

44

45

46 **The cereal grass barley was domesticated about 10,000 years ago in the Fertile**
47 **Crescent and became a founder crop of Neolithic agriculture¹. Here, we report**
48 **genome sequences of five 6,000-year-old barley grains excavated at a cave in the**
49 **Judean Desert close to the Dead Sea. Comparison to whole exome sequence data**
50 **from a diversity panel of present-day barley accessions revealed the close affinity**
51 **of ancient samples to extant landraces from the Southern Levant and Egypt,**
52 **consistent with a proposed origin of domesticated barley in the Upper Jordan**
53 **Valley. Our findings suggest that barley landraces grown in present-day Israel in**
54 **the past six millennia have not experienced a major lineage turnover although**
55 **there is evidence for gene flow between cultivated and sympatric wild**
56 **populations. We show the utility of ancient genomes from desiccated**
57 **archaeobotanical remains in informing research into the origin, early**
58 **domestication and subsequent migration of crop species.**

59

60 Genetic analyses of ancient DNA can greatly inform research into the origin, initial
61 domestication and subsequent dispersal of crops and livestock as evidenced by studies
62 involving ancient DNA samples and genomic datasets of present-day populations of
63 cattle², swine³, dogs⁴ and maize⁵. Wheat and barley, founder crops of agriculture in
64 the ancient Near East and Europe, were domesticated in the Fertile Crescent, where
65 their wild relatives still thrive today^{1,6}. Our current knowledge of their domestication
66 is largely derived from morphological analysis of archaeobotanical remains¹ and the
67 population genetic analysis of present-day samples^{7,8}. Although domesticated wheat
68 and barley appear in the archaeological record by 10,000 calendar years before
69 present (cal BP)¹, the oldest verified DNA sequences to date were retrieved from
70 archaeobotanical specimens originating from Bronze Age China⁹ and Ancient
71 Egypt¹⁰. Claims about a small number of prehistoric wheat DNA molecules retrieved
72 from Mesolithic paleosol¹¹ have remained contentious^{12,13}. There have been no studies
73 where large quantities of ancient DNA sequences have been retrieved that could
74 underpin the comparison of modern and ancient samples of Old World cereals at a
75 genome-wide scale.

76

77 Here, we report genome sequences of five 6000-year old barley grains excavated at
78 Yoram Cave in the Judean Desert, Israel. Yoram Cave is part of a complex of three
79 difficult-to-access caves, located in the south-eastern cliff of the Masada Horst facing

80 the Dead Sea. High-resolution excavation (Online Methods, Fig. 1a, Supplementary
81 Figures 1, 2) revealed a single undisturbed anthropogenic layer of Chalcolithic origin
82 (ca. 6,200-5,800 cal BP). The rich plant assemblage of more than 100 taxa was well
83 preserved (Figure 1b), prompting us to attempt the retrieval of DNA sequences. We
84 selected ancient barley grains (Figure 1c, Supplementary Figure 3) for DNA
85 extraction because of barley's central role in ancient and modern agriculture and its
86 remarkable adaptive features that make it a model plant in domestication genomics.

87

88 DNA extractions were performed from ten bisected grains and spikelet remains,
89 whose other halves were subjected to direct radiocarbon dating, confirming the
90 Chalcolithic origin of the specimens (Table 1). Illumina sequencing of libraries
91 yielded between 7.3 and 21.5 million paired-end reads (Supplementary Table 1).
92 Based on the fraction of reads that could be aligned to the barley reference genome,
93 we estimated the content of endogenous DNA to range from 0.4 to 96.4 %. Sequence
94 reads of eight samples showed fragment sizes and damage patterns characteristic of
95 ancient DNA (Supplementary Table 1, Supplementary Figures 4, 5, 6), which
96 demonstrates the authenticity of the samples^{14,15}. Deamination-derived mismatches (C
97 -> T, G -> A) occurred towards the ends of reads with frequencies between 1.9 and
98 21.8 %. We only used five samples with on average more than 12 %
99 misincorporations¹⁵ at the first base of sequence fragments for further experiments.
100 Once the authenticity of these samples had been established, we treated five DNA
101 extracts containing a large fraction of endogenous DNA with uracil-DNA-glycosylase
102 (UDG) to reduce nucleotide misincorporations caused by ancient DNA damage by
103 removing deaminated cytosines¹⁶. Deep Illumina sequencing of the UDG-treated
104 libraries yielded between 82.5 million and 5.1 billion reads (Table 1).

105

106 We compared the ancient barley genome sequences to present-day accessions on the
107 basis of whole-exome capture¹⁷ sequence data from 267 entries from *ex situ*
108 collections representing extant populations of wild (*Hordeum vulgare* ssp.
109 *spontaneum*) and domesticated (*H. vulgare* ssp. *vulgare*) barley from across the range
110 of the species¹⁸ (Figure 2a). This dataset¹⁸ comprised 1,688,807 single nucleotide
111 polymorphisms (SNPs) (Table 1).

112

113 Principal component analysis¹⁹ has revealed fundamental patterns of population
114 structure across the present-day accessions¹⁸. The first principal component (PC1)
115 clearly differentiated wild and domesticated barleys, while PC2 represented the
116 variation in the wild barleys (Figure 2b). Least-square projection²⁰ of the ancient
117 samples onto the PCA axes defined by the extant samples revealed the close affinity
118 of ancient barley with present-day domesticated barley. The deep coverage of sample
119 JK3014 allowed us to ascertain the allelic status of the domestication genes *Non-*
120 *brittle rachis 1 (Btr1)* and *Non-brittle rachis 2 (Btr2)*. In domesticated barley, one of
121 *Btr1* and *Btr2* carry mutations that abolish the disarticulation of the spike at
122 maturity²¹. JK3014 had a wild-type *Btr2* haplotype, but carried the previously
123 described 1-bp deletion in the coding sequence of *Btr1* (Supplementary Figure 7),
124 consistent with a high frequency of this mutation in barleys of the Southern Levant²¹.
125 In agreement with the archaeobotanical classification of the ancient barley spike
126 remains as being of the two-rowed type (Figure 2c), the extant accessions closest to
127 the ancient samples were two-rowed domesticated barleys from the Southern Levant
128 and Egypt (Figure 2b, d). A putative two-rowed phenotype of sample JK3014 can also
129 be inferred from the allelic status of the *Six-rowed spike 1* gene²² (Supplementary
130 Figure 7).

131

132 Rare genetic variants can provide insights into the spatial structure of populations^{23,24}.
133 In inbreeding plants such as barley, isolation by distance²⁵ is common since gene flow
134 across larger geographic distances is limited. We identified rare variants with minor
135 allele counts of up to five across the ancient and extant barleys and determined the
136 number of rare alleles shared between pairs of sequenced samples. Transitions were
137 excluded from the analysis because deaminated cytosines cannot be repaired by UDG
138 treatment if they are methylated¹⁶ and thus can give rise to genotyping errors. The
139 extant landraces that shared at least 30 rare alleles with the ancient samples were two-
140 rowed accessions from Syria, Jordan, Egypt and Israel as well as six-rowed
141 accessions from North Africa (Supplementary Table 2). Eight wild barleys from Israel
142 also shared ≥ 30 rare alleles with the ancient samples (Supplementary Table 2).

143

144 We measured relatedness of ancient samples to each of the wild barley accessions
145 based on the level of identity by state (IBS) calculated across all SNPs. The
146 genetically closest wild accessions originated from a sampling site located in the

147 Upper Jordan Valley. We then calculated the geographic distance between Yoram
148 Cave and each wild accession, but no significant correlation between IBS and
149 geographic distance was found ($R = -0.17$, $P = 0.108$). However, when splitting the
150 data into geographically proximal (< 250 km) and distant (> 250 km) samples,
151 significant correlations between geographic distance and the relatedness score were
152 detected for both subsets of the tested wild barley samples (proximal: $R = 0.74$, $P = <$
153 0.001 ; distant: $R = -0.34$, $P = 0.006$) (Figure 3a). The same analysis was conducted
154 between wild barleys and extant landraces from the Fertile Crescent (Figure 3b, c)
155 pinpointing the Upper Jordan Valley as a peak for genetic similarity with
156 domesticated barley. Conducting this analysis using only data from the deeply
157 sequenced JK3014 yielded similar results (Supplementary Figure 8). The comparison
158 of modern cultivars and landraces from outside of the Fertile Crescent (Europe, North
159 Asia) to our wild barley panel accessions pinpointed accessions from the Upper
160 Jordan Valley as the most closely related (Supplementary Table 3). The Israel-Jordan
161 area was proposed earlier as one (though not the only) center of origin of
162 domesticated barley^{8,26}. This hypothesis is supported by two archeological sites, Tel
163 Aswad and Ohalo II, with the earliest traces of barley cultivation^{27,28}, which are
164 within 80 km of the extant wild barley accessions in our panel that are genetically
165 closest to the ancient samples.

166

167 Although self-fertilization is predominant in barley²⁹, wild barley is fully interfertile
168 with the domesticated crop and evidence for hybridization between the two has been
169 reported^{7,30}. To ascertain whether the genetic similarity between ancient and extant
170 landraces is the outcome of shared ancestry or the result of later hybridization
171 between local wild barley and domesticated forms, we performed a model-based
172 assignment of present-day and ancient samples to two ancestral groups corresponding
173 to wild and domesticated barleys using ADMIXTURE³¹ considering only transversion
174 variants. The analysis with two ancestral populations confirmed the strong
175 differentiation between wild and domesticated accessions observed in PCA, with a
176 perfect correlation between the domestication status and the assignment to ancestral
177 populations. Thus, all ancient samples were assigned to the domesticated group
178 (Figure 4a). However, wild ancestry coefficients of two ancient samples were 4.2 and
179 8.7 %. Present-day landraces from the Levant (Israel, Jordan, Lebanon and Syria) also
180 showed an elevated fraction of wild ancestry (6.8 %). By contrast, the average wild

181 ancestry for European landraces was only 0.14 %. We also performed ADMIXTURE
182 runs with the number of ancestral population (K) set to five because this K value had
183 the lowest cross validation error (Supplementary Figure 9). In this analysis,
184 domesticated barley was separated into three and wild barley into two clusters
185 (Supplementary Figure 10). The deeply sequenced ancient sample JK3014 had 7.6 %
186 ancestry from a cluster predominantly composed of wild accessions, while the
187 average ancestry fraction in this cluster was only 1.7 % for domesticated barley.
188 These observations suggest gene flow between wild and cultivated barley in regions
189 of sympatry.

190

191 We used D-statistics³² to corroborate the hypothesis of archaic admixture between
192 wild and domesticated barley populations in the Levant. We considered five
193 categories: ancient barley, extant wild barley from the Levant (Israel, Syria, Jordan,
194 Lebanon), extant landraces from the Levant, extant European landraces and outgroup
195 *Hordeum pubiflorum*³³. We calculated D for each ancient sample separately
196 (Supplementary Table 4) and focus here on the results for the single deeply sequenced
197 sample JK3014 (Figure 4b). D(extant Levantine landraces, extant European landraces,
198 JK3014, outgroup) was significantly positive, confirming the close affinity of the
199 ancient sample with the present-day Levantine landraces. The comparisons D(extant
200 Levantine landraces, extant European landraces, Levantine wild barley, outgroup) and
201 D(JK3014, extant European landraces, Levantine wild barley, outgroup) were also
202 significantly positive, indicating admixture between wild and domesticated barleys
203 from Israel after the lineage leading to Levantine landraces split from the progenitors
204 of European landraces. The ancient sample did not show a closer affinity to extant
205 Levantine wild barley than present-day landraces from this region, as D(JK3014,
206 extant Levantine landraces, Levantine wild barley, outgroup) was not significantly
207 different from zero. These findings indicate that the genomes of both ancient and
208 present-day cultivated barley from the Levant show traces of archaic gene flow from
209 sympatric wild accessions after the split between Levantine and European landraces,
210 supporting the notion of hypothetical hybridization events between domesticated
211 barley and sympatric wild stands^{7,34}. As a consequence of this demographic scenario,
212 the homogenization of the allele frequencies in sympatric wild and domesticated
213 barley through bi-directional gene flow may complicate inferences about the origin(s)
214 and domestication history³⁴ at the fine genomic scale, while key domestication genes

215 (such as the *btr* genes) are resistant to wild introgression due to a strong selection
216 against shattering spikes. Despite hybridization events between wild and domesticated
217 barleys in the last six millennia, the overall picture is that the genomes of extant
218 Levantine landraces have remained remarkably similar to how they were 6000 years
219 ago. This is despite climate change³⁵ and anthropogenic transformations of local flora
220 and fauna, including changes in agricultural practices³⁶, which might have favored the
221 introduction of landraces from other regions that were better adapted to the changing
222 agricultural environment. Whereas we found no indications of major lineage
223 turnovers in the barley crop in the Southern Levant (as has, for example, been
224 observed in Near Eastern pig populations^{37,38}), the eventful history of this region
225 makes it likely that the farmers that grew cereals there several millennia ago were not
226 the ancestors of those that tended the present-day landraces³⁹. One can speculate that
227 conquerors and immigrants did not bring their crop seeds from their old homelands,
228 but favored locally adapted landraces.

229

230 Expanding on previous studies that reported the PCR amplification and sequencing of
231 single genes from ancient wheat and barley samples^{10,40}, our results show that very
232 ancient desiccated plant remains preserved under hot and arid conditions contain
233 sufficient amounts of endogenous DNA to underpin genome-wide population genetic
234 analyses in the context of diversity panels of extant individuals. Our analysis
235 demonstrates the value of archaeogenomics in supporting contemporary genetic-based
236 phylogeographic studies in exploring crop origins, and shows that 6,000 years BP
237 domesticated barley appeared remarkably similar to proximate extant landraces,
238 indicating that the major domestication events had occurred by that time.

239

240 **URLs**

241 Novosort, <http://www.novocraft.com>

242 R package mapdata, <https://cran.r-project.org/web/packages/mapdata/index.html>

243

244 **Accession Codes**

245 Raw read files of the ancient samples can be retrieved from the European Nucleotide
246 Archive (ENA) under project ID PRJEB12197. The SNP genotype matrix is available
247 under DOI 10.5447/IPK/2016/6. Passport information for the extant barley panel can
248 be retrieved from DOI 10.5447/IPK/2016/3. DOIs were registered with e!DAL⁴¹.

249

250 **Acknowledgments**

251 We gratefully acknowledge Birgit Schäfer for providing photographs of barley spikes
252 and Anne Fiebig for sequence data submission. This work was supported by a grant of
253 the Israel Science Foundation (1179/13) to EW, funding from the Endowed Chair in
254 Molecular Genetics Applied to Crop Improvement at the University of Minnesota and
255 the Triticeae Coordinated Agricultural Project, USDA-NIFA #2011-68002-30029 to
256 GJM and core funding of IPK Gatersleben to NS and MM. RW, JR and IKD were
257 supported by Research Programme funding from the Scottish Government and the
258 University of Dundee (RW). NM and UD acknowledge a grant from the Irene Levi
259 Sela CARE Foundation.

260

261 **Authors' contributions**

262 EW, TF, NS, and JK conceived the study. EW, TF, NS, JK, VJS and MM designed
263 experiments. NM, UD, MD, SR and EW performed excavations and archaeobotanical
264 analyses. VJS, AH and ER performed the ancient DNA experiments. MM, SH, AK,
265 MS, SHV and REG analyzed data. JR, MM, IKD, BK, GJM, NS and RW provided
266 exome capture data. MM, VJS, AH, SR, TF, JK, EW and NS wrote the manuscript
267 with input from all co-authors. All authors read and approved the manuscript.

268

269 **Competing interests**

270 The authors declare no competing financial interests.

271

272 **References for the main text**

- 273 1. Zohary, D., Hopf, M. & Weiss, E. *Domestication of Plants in the Old World:*
274 *The origin and spread of domesticated plants in Southwest Asia, Europe, and*
275 *the Mediterranean Basin*, (Oxford University Press on Demand, 2012).
- 276 2. Troy, C.S. *et al.* Genetic evidence for Near-Eastern origins of European cattle.
277 *Nature* **410**, 1088-91 (2001).
- 278 3. Larson, G. *et al.* Ancient DNA, pig domestication, and the spread of the
279 Neolithic into Europe. *Proc Natl Acad Sci U S A* **104**, 15276-81 (2007).
- 280 4. Skoglund, P., Ersmark, E., Palkopoulou, E. & Dalen, L. Ancient wolf genome
281 reveals an early divergence of domestic dog ancestors and admixture into
282 high-latitude breeds. *Curr Biol* **25**, 1515-9 (2015).
- 283 5. da Fonseca, R.R. *et al.* The origin and evolution of maize in the Southwestern
284 United States. *Nature Plants* **1**(2015).
- 285 6. Harlan, J.R. & Zohary, D. Distribution of wild wheats and barley. *Science*
286 **153**, 1074-80 (1966).
- 287 7. Poets, A.M., Fang, Z., Clegg, M.T. & Morrell, P.L. Barley landraces are
288 characterized by geographically heterogeneous genomic origins. *Genome Biol*
289 **16**, 173 (2015).
- 290 8. Badr, A. *et al.* On the origin and domestication history of Barley (*Hordeum*
291 *vulgare*). *Mol Biol Evol* **17**, 499-510 (2000).
- 292 9. Li, C. *et al.* Ancient DNA analysis of desiccated wheat grains excavated from
293 a Bronze Age cemetery in Xinjiang. *Journal of Archaeological Science* **38**,
294 115-119 (2011).
- 295 10. Palmer, S.A., Moore, J.D., Clapham, A.J., Rose, P. & Allaby, R.G.
296 Archaeogenetic evidence of ancient nubian barley evolution from six to two-
297 row indicates local adaptation. *PLoS One* **4**, e6301 (2009).
- 298 11. Smith, O. *et al.* Archaeology. Sedimentary DNA from a submerged site
299 reveals wheat in the British Isles 8000 years ago. *Science* **347**, 998-1001
300 (2015).
- 301 12. Weiss, C.L., Dannemann, M., Prufer, K. & Burbano, H.A. Contesting the
302 presence of wheat in the British Isles 8,000 years ago by assessing ancient
303 DNA authenticity from low-coverage data. *Elife* **4**(2015).
- 304 13. Kistler, L. *et al.* Thermal age, cytosine deamination and the veracity of 8,000
305 year old wheat DNA from sediments. *bioRxiv* (2015).
- 306 14. Briggs, A.W. *et al.* Patterns of damage in genomic DNA sequences from a
307 Neandertal. *Proc Natl Acad Sci U S A* **104**, 14616-21 (2007).
- 308 15. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Paabo, S. Temporal
309 patterns of nucleotide misincorporations and DNA fragmentation in ancient
310 DNA. *PLoS One* **7**, e34131 (2012).
- 311 16. Briggs, A.W. *et al.* Removal of deaminated cytosines and detection of in vivo
312 methylation in ancient DNA. *Nucleic Acids Res* **38**, e87 (2010).
- 313 17. Mascher, M. *et al.* Barley whole exome capture: a tool for genomic research in
314 the genus *Hordeum* and beyond. *Plant J* **76**, 494-505 (2013).
- 315 18. Russell, J. *et al.* Adaptation of barley to different environments revealed in the
316 exomes of a range- wide collection of landraces and wild relatives. *Nature*
317 *Genetics in press*(2016).
- 318 19. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis.
319 *PLoS Genet* **2**, e190 (2006).
- 320 20. Reich, D., Thangaraj, K., Patterson, N., Price, A.L. & Singh, L.
321 Reconstructing Indian population history. *Nature* **461**, 489-94 (2009).

- 322 21. Pourkheirandish, M. *et al.* Evolution of the Grain Dispersal System in Barley.
323 *Cell* **162**, 527-39 (2015).
- 324 22. Komatsuda, T. *et al.* Six-rowed barley originated from a mutation in a
325 homeodomain-leucine zipper I-class homeobox gene. *Proc Natl Acad Sci U S*
326 *A* **104**, 1424-9 (2007).
- 327 23. Schiffels, S. *et al.* Iron Age and Anglo-Saxon genomes from East England
328 reveal British migration history. *bioRxiv*, 022723 (2015).
- 329 24. Mathieson, I. & McVean, G. Differential confounding of rare and common
330 variants in spatially structured populations. *Nat Genet* **44**, 243-6 (2012).
- 331 25. Nordborg, M. *et al.* The pattern of polymorphism in *Arabidopsis thaliana*.
332 *PLoS Biol* **3**, e196 (2005).
- 333 26. Morrell, P.L. & Clegg, M.T. Genetic evidence for a second domestication of
334 barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proc Natl Acad Sci U S*
335 *A* **104**, 3289-94 (2007).
- 336 27. W. van Zeist & Bakker-Heeres, J.A.H. Archaeological studies in the Levant 1.
337 Neolithic sites in the Damascus basin: Aswad, Ghoraifé, Ramad. .
338 *Palaeohistoria* **24**, 165-256 (1985).
- 339 28. Snir, A. *et al.* The Origin of Cultivation and Proto-Weeds, Long Before
340 Neolithic Farming. *PLoS One* **10**, e0131422 (2015).
- 341 29. Abdel-Ghani, A.H., Parzies, H.K., Omary, A. & Geiger, H.H. Estimating the
342 outcrossing rate of barley landraces and wild barley populations collected
343 from ecologically different regions of Jordan. *Theor Appl Genet* **109**, 588-95
344 (2004).
- 345 30. Russell, J. *et al.* Analysis of >1000 single nucleotide polymorphisms in
346 geographically matched samples of landrace and wild barley indicates
347 secondary contact and chromosome-level differences in diversity around
348 domestication genes. *New Phytol* **191**, 564-78 (2011).
- 349 31. Alexander, D.H., Novembre, J. & Lange, K. Fast model-based estimation of
350 ancestry in unrelated individuals. *Genome Res* **19**, 1655-64 (2009).
- 351 32. Durand, E.Y., Patterson, N., Reich, D. & Slatkin, M. Testing for ancient
352 admixture between closely related populations. *Mol Biol Evol* **28**, 2239-52
353 (2011).
- 354 33. Jakob, S.S. & Blattner, F.R. A chloroplast genealogy of hordeum (poaceae):
355 Long-term persisting haplotypes, incomplete lineage sorting, regional
356 extinction, and the consequences for phylogenetic inference. *Mol Biol Evol* **23**,
357 1602-12 (2006).
- 358 34. Harlan, J.R. On the origin of barley. *USDA Agriculture Handbook* **338**, 9-31
359 (1968).
- 360 35. Russell, J. *et al.* Genetic diversity and ecological niche modelling of wild
361 barley: refugia, large-scale post-LGM range expansion and limited mid-future
362 climate threats? *PLoS One* **9**, e86021 (2014).
- 363 36. Frumin, S., Maeir, A.M., Kolska Horwitz, L. & Weiss, E. Studying Ancient
364 Anthropogenic Impacts on Current Floral Biodiversity in the Southern Levant
365 as reflected by the Philistine Migration. *Sci Rep* **5**, 13308 (2015).
- 366 37. Meiri, M. *et al.* Ancient DNA and population turnover in southern levantine
367 pigs--signature of the sea peoples migration? *Sci Rep* **3**, 3035 (2013).
- 368 38. Giuffra, E. *et al.* The origin of the domestic pig: independent domestication
369 and subsequent introgression. *Genetics* **154**, 1785-91 (2000).
- 370 39. Haber, M. *et al.* Genome-wide diversity in the levant reveals recent structuring
371 by culture. *PLoS Genet* **9**, e1003316 (2013).

- 372 40. Bilgic, H., Hakki, E.E., Pandey, A., Khan, M.K. & Akkaya, M.S. Ancient
373 DNA from 8400 Year-Old Catalhoyuk Wheat: Implications for the Origin of
374 Neolithic Agriculture. *PLoS One* **11**, e0151974 (2016).
- 375 41. Arend, D. *et al.* e!DAL--a framework to store, share and publish research data.
376 *BMC Bioinformatics* **15**, 214 (2014).
377

378 **Figure legends for the main text**

379

380 **Figure 1: Ancient plant remains excavated at Yoram Cave.** Plan of the Southern
381 Chamber of Yoram Cave showing the excavation grid and sub-units **(a)**. Photograph
382 of Locus 3 in Square A2 during excavation – note the excellent dry preservation of
383 rope, reeds, seeds and pellets **(b)**. Photograph of a well-preserved, desiccated barley
384 grain found at Yoram Cave **(c)**.

385

386 **Figure 2: Ancient barley samples are closely related to present-day landraces**
387 **from the Levant.** Ancient barley sequences were compared to exome sequence data
388 of a present-day diversity panel. **(a)** The collection sites of landraces (black circles)
389 and wild barleys (blue circles) are shown. Masada is marked with a red circle. **(b)**
390 Principal component analysis showing ancient samples projected on the present-day
391 diversity panel. The inset magnifies the PCA space around the ancient samples. ISR,
392 SYR, LBN, EGY represent closely related landraces from Israel, Syria, Lebanon and
393 Egypt, respectively. The proportion of variance explained by each principal
394 component is indicated in parentheses. **(c)** Well-preserved rachis of two-rowed
395 domesticated barley from Yoram Cave **(d)**. Spike of an individual of a present day
396 two-rowed landrace barley (accession HOR8658) that is among the barleys most
397 closely related to the ancient DNA sample from Yoram Cave.

398

399 **Figure 3: Relationship between genetic similarity and geographic distance.**
400 Scatter plot of genetic similarity and geographic distance between 91 extant wild
401 barley accessions sampled range-wide including the Fertile Crescent and **(a)**
402 archaeological samples found at Yoram Cave, **(b)** two-row cultivated landrace from
403 Israel, and **(c)** two-row cultivated landrace from Egypt. The geographic position
404 attributed to each sample is: A (31.3141 N, 35.353 E), B (31.7156 N, 35.1871 E), C
405 (31.193 N 29.904 E). Correlation coefficients and P-values for the geographically
406 proximate and distant subsets are indicated in blue and red, respectively.

407

408 **Figure 4: Gene flow between wild and domesticated barleys in the Levant.**

409 **(a)** Wild ancestry coefficients of landraces from different geographic regions and
410 ancient barleys as determined by ADMIXTURE. The wild ancestry proportion is
411 shown in black for extant samples and in red for ancient samples. **(b)** D statistics

412 for different quadruples of barley populations (P_1 , P_2 , P_3 , O). Positive D values
413 indicate that P_1 shares more derived allele with P_3 than P_2 does. Black bars
414 indicate ± 3 standard errors (SE), gray bars ± 1 SE. Population names are
415 abbreviated as follows: LevDom (Levantine landraces), LevWild (Levantine wild
416 barley), Euro (European landraces), O (*H. pubiflorum*, outgroup). JK3014 is a
417 deeply sequenced ancient sample.

418 **Table 1: Summary of ancient barley samples used for genetic analyses.**

419

Sample name	Radiocarbon age ¹	Number of raw reads	Percentage of mapped reads	Percentage of unique reads	Average read depth ²	Number of called SNPs
JK2281	5290 ± 27	256.1 M	31.4%	26.0%	0.54	162,110
JK3009	5034 ± 36	89.2 M	61.7%	67.3%	0.46	133,365
JK3010	5032 ± 37	94.7 M	62.3%	57.9%	0.96	278,505
JK3013	5227 ± 37	82.5 M	49.4%	39.8%	0.19	18,949
JK3014	4988 ± 36	5,131.2 M	86.4%	28.3%	20	1,283,396

420

421 ¹Uncalibrated radiocarbon years before the present, see for Supplementary Table 1 for calibrated dates.

422 ²Average read depth in target regions of the exome capture assay for JK2281-JK3013. The mode of the coverage distribution is given for

423 JK3014.

424 **Online Methods**

425

426 **Archaeology**

427

428 Yoram Cave

429 Yoram Cave is archaeologically significant as one of the rare cave sites with a
430 single layer of human occupation according to current radiocarbon dating
431 findings. Unlike most other Judean Desert caves, there are no findings from the
432 later Roman and Byzantine Periods. In addition, it is one of the rare cave sites
433 that has not suffered from modern looting or hyena burrowing. It is the only
434 Chalcolithic cave site in the Judean Desert that has been excavated by high-
435 resolution sampling methods. The cave's plant assemblage has been preserved
436 by drying, supporting use in possible DNA-based studies.

437 Location and description of the cave:

438 Yoram Cave is part of a cave-complex ("Masada caves – South"), with
439 three caves located in the south-eastern cliff of Masada Horst (Supplementary
440 Figure 2), facing the Dead Sea. Access to the cave complex is relatively difficult as
441 it requires walking along narrow goat paths on a sharp incline rock-fall. The
442 Yoram Cave entrance is on an almost vertical cliff, some 4m above a goat trail at
443 its base (Fig. 1B).

444 Water sources are scarce. Some small rock depressions, holding flood
445 waters for a few months are found about 150 meters southwest of the
446 cave complex. The nearest permanent springs are in Tze'elim canyon, some 5 km
447 walking distance northward.

448 The cave entrance is 2.9 m wide with a fieldstone wall stretching along the
449 entrance. The interior has two rooms (Figure 1a and Supplementary Figure 1).
450 The northern room is approximately 7 m long and between 3.5 and 5 m wide,
451 and contains large boulders. The southern room consists of three areas: an
452 entrance (A), a short corridor (B), and a small inner cubicle (C). The latter room's
453 maximum length is 6.5 m, and its width is between 2 and 2.5 m. (Figure 1a,
454 Supplementary Figures 1, 2). The heights of these 3 areas range between 0.2 and
455 1.8 m. Most human activities and related plant remains were found in the
456 southern room.

457 Stratigraphy

458 Excavations in the southern room revealed three phases (from top to bottom):

459 1. A biogenic layer, mostly the result of nesting activities of large birds of prey
460 (possibly Bearded Vulture, *Gypaetus barbatus*, or members of the eagle family),
461 including large and small twigs, bones and droppings;

462 2. An anthropogenic layer, representing the Chalcolithic period of human
463 activity (Fig. S2); and

464 3. A pre-floor Chalcolithic layer.

465 Scant evidence of modern human activity was discovered on top of layer 1. An
466 initial round of radiocarbon dating validated the excavators' field observation of
467 a Chalcolithic origin of the anthropogenic layer (ca. 6,200-5,800 cal. BP). The
468 biogenic layer was dated to the LB/IAI period.

469 Reeds (possibly *Typha/Phragmites* spp. see Schick et al.⁴²) were found
470 abundantly among the anthropogenic layer plant remains (Figure 1b). The
471 appearance of reeds alongside various rope segments (plants still to be
472 identified) and a small mat section in one of the excavated samples hint that
473 simple mats probably covered the cave floor. Such mats indicate preparations for
474 prolonged stay in the cave, rather than a chance occupation. The human-built
475 wall in the entrance is further evidence for the prudent use of the cave.

476

477 Excavation and sampling

478 Excavation was conducted by a high-resolution excavation method, with the
479 excavated space being divided into sub-units (Figure 1a). These were
480 meticulously sampled, with each sample going through a sorting procedure,
481 using 1 mm and 100 µm mesh sieves. During sifting, various categories of finds
482 were separated (e.g. archaeological artifacts, macro- and microecofacts –
483 archaeozoological and archaeobotanical remains), which were packed
484 separately. Additional separation was undertaken on plant material, with 1 liter
485 of sediment out of each excavated bucket from the "anthropogenic" loci (#1, 3, 4,
486 5, 7, 9) was kept for archaeobotanical analysis in the laboratory. Larger samples
487 were also retained when plant remains were visible to the naked eye during
488 excavation.

489 Mapping of Yoram Cave (Fig 1a, Supplementary Fig. 1) was performed by the
490 excavation team headed by Uri Davidovich and Nimrod Marom in 2007 using
491 standard cave mapping equipment, including a Leica Disto D3 laser inclinometer
492 and a Silva Ranger 3 prismatic compass; the grade of mapping was 5C. Field
493 maps were later graphically edited using Limelight software.

494

495 **DNA extraction and library preparation**

496

497 A panel of 13 samples was initially selected for this study consisting of eight
498 barley grains, two barley ear fragments, two wheat emmer grains and one
499 emmer ear fragment. All subsequent sampling procedures, DNA extractions and
500 library preparations were carried out in clean room facilities dedicated to
501 ancient DNA research at Tuebingen University. During the sampling process, all
502 samples weighing more than 15 mg were divided into two parts, of which one
503 part was used for subsequent DNA extraction and the other one sent for
504 radiocarbon dating at Curt-Engelhorn-Zentrum Archaeometrie gGmbH
505 Mannheim (Germany). The DNA extractions were conducted on 5 to 30 mg of
506 plant material using the PTB extraction protocol detailed by Kistler⁴³ with the
507 following modifications: all samples were extracted twice (E1 and E2). After a
508 first incubation for two hours at 37°C the plant remains were pelleted, the
509 supernatants were taken off and stored at 4°C over night. Plant pellets were
510 resuspended in extraction buffer a second time and incubated over night at 37°C.
511 All extracts were purified simultaneously on the next day.

512 For library preparation, a well established protocol by Meyer and Kircher⁴⁴ was
513 used to convert a 20 µl aliquot of each DNA extract into double-stranded Illumina
514 libraries. Adaptor ligation to the fragments was quantified using a quantitative
515 PCR with the primers IS7 and IS8⁴⁴, the reagents of the DyNAmo Flash SYBR
516 Green qPCR Kit (Biozym) and the Lightcycler 96 (Roche). Then, double indexed
517 libraries were created by adding sample specific barcodes to both library
518 adapters via amplification⁴⁵ followed by another quantification assay using the
519 primers IS5 and IS6⁴⁴ to estimate the efficiency of the indexing PCR. All
520 extraction and library blanks were treated accordingly. These libraries were
521 used subsequently for initial shotgun sequencing.

522 For genome-wide shot-gun sequencing and enrichment additional libraries for
523 the extracts JK2281E1, JK2281E2, JK3009E1, JK3010E1, JK3013E1 and
524 JK3014E1 (Table 1, Supplementary Table 1) were prepared from 50 µl aliquots
525 of all DNA extracts following the methods described above^{44,45} with one
526 modification: all extracts and blanks were treated with uracil-DNA glycosylase
527 (UDG) and endonuclease VIII during the library preparation to avoid potential
528 sequencing artifacts that are caused by the characteristic ancient DNA damage
529 pattern due to the deamination of cytosine to uracil over time¹⁶.

530 For all indexed libraries, a second amplification was carried out in 100 µl
531 reactions using 5µl library template, 4 units AccuPrime Taq DNA Polymerase
532 High Fidelity (Invitrogen), 1 unit 10X AccuPrime buffer (containing dNTPs) and
533 0.3 µM IS5 and IS6 primers⁴⁴. The following thermal profile was performed: 2-
534 min initial denaturation at 94°C, followed by 4 to 17 cycles consisting of 30-sec
535 denaturation at 94°C, a 30-sec annealing at 60°C and a 2-min elongation at 68°C
536 with a final 5-min elongation at 68°C. After amplification the products were
537 purified using MinElute spin columns (Qiagen) according to the manufacturer's
538 protocol and quantified using an Agilent Bioanalyzer DNA 1000 Chip.

539 All libraries for initial and genome-wide shotgun sequencing were then diluted
540 to 10nM and pooled in equimolar amounts. Initial shotgun sequencing of
541 libraries was undertaken with an Illumina HiSeq 2500 platform, using a paired-
542 end dual index run with 2*100+7+7 cycles and the manufacturer's protocols for
543 multiplex sequencing (TruSeq PE Cluster Kit v3-cBot-HS). Genome-wide shotgun
544 sequencing of the UDG treated libraries was performed on an Illumina NextSeq
545 500 platform with 2 x 150+8+8 cycles using the NextSeq High Output reagent kit
546 v1 and the manufacturer's protocol for multiplex sequencing.

547 The UDG-treated libraries from the extracts JK2281E1, JK2281E2 (Table 1,
548 Supplementary Table 1) were treated separately: after the second amplification
549 the libraries were enriched using a sequence capture assay for the barley
550 exome¹⁷ as described by Himmelbach *et al.*⁴⁶ with one modification: the
551 concentration of the DNA fragments recovered from the capture was determined
552 by quantitative PCR using the primers IS5 and IS6⁴⁴, the SYBR Green PCR Master
553 Mix (Qiagen, Hilden) and the 7900 HT Fast Real-Time PCR system (Applied
554 Biosystems).

555 After a dilution to 10 nM the sequencing was carried out on the Illumina HiSeq
556 2500 platform as described by Mascher *et al.*¹⁷ with a paired-end single index
557 run using 101+6+100 cycles and the manufacturer's protocols for multiplex
558 sequencing (TruSeq PE Cluster Kit v3-cBot-HS).

559 Four additional UDG-treated libraries of the JK3014E1 were produced for deeper
560 sequencing as detailed previously, diluted to 10 nM and pooled in equimolar
561 amounts together with the already sequenced JK3014E1 UDG-library. The
562 sequencing of the pools was conducted on a HiSeq4000 platform with 2 x
563 75+8+8 cycles using the HiSeq 3000/4000 PE Cluster Kit, HiSeq 3000/4000 SBS
564 Kit and the manufacturer's protocol for multiplex sequencing.

565 Raw sequence reads have been uploaded to EMBL ENA short read archive
566 (accession: PRJEB12197).

567

568 **Processing and mapping of sequence reads**

569

570 Overlapping paired-end reads were merged using scripts provided by Kircher⁴⁷
571 for samples JK2279 – JK2284 or with leeHom⁴⁸ (using the parameter “—
572 ancientdna”) for the other samples. Length distribution of the merged reads was
573 calculated using AWK and the Unix tools “sort” and “uniq” as described in
574 Supplement S4 of Gallego Llorente *et al.*⁴⁹. Merged reads were aligned to (i) the
575 whole-genome shotgun assembly of barley cv. Morex⁵⁰ and (ii) the chloroplast
576 genome assembly of cv. Morex⁵¹ with BWA-MEM version 0.7.12⁵² using default
577 parameters. Conversion to BAM format and calculation of mapping statistics
578 were performed with SAMtools⁵³. Sorting of BAM files and duplicate removal
579 was done with novosort (Novocraft Technologies Sdn Bhd, Malaysia). Nucleotide
580 misincorporation profiles were generated with mapDamage version 2.0⁵⁴ for the
581 nuclear and the chloroplast genome. Genotypes of five UDG-treated ancient
582 samples at 1,688,807 known SNP positions¹⁸ were called using single-sample
583 variant calling with SAMtools (version 0.1.19, commands samtools mpileup and
584 bcftools view)⁵⁵ using only reads with mapping quality above Q30 and
585 considering bases with quality above Q20. SNP positions were retained if their
586 quality score was at least 30 and they were covered by at least two reads. In the
587 case of the deeply sequenced sample JK3014, we also required the read depth to

588 be not larger than 30 (= x 3 the mode of the coverage distribution in exome
589 capture target regions). Heterozygous calls were set to missing. In the deeply
590 sequenced ancient DNA sample JK3014, 0.7 % of variants were called
591 heterozygous (compared to 1.7 % in the extant samples). In an inbreeding crop,
592 the divergence between parental haplotypes of an individual is very low.
593 Contamination with DNA of extant barley would thus become evident in an
594 elevated fraction of heterozygous calls. The absence of such a pattern lends
595 further support to the authenticity of the ancient samples. Coverage statistics
596 were calculated with SAMtools⁵³ and BEDTools⁵⁶.

597

598 **Population genetic analysis**

599

600 Principal component analysis (PCA) was performed with EIGENSOFT 6.0.1¹⁹ for
601 five ancient barley samples and 228 extant barley *ex situ* accessions with clear
602 domestication status and well-described geographic origins¹⁸. Least-square
603 projection as implemented in the smartPCA program of EIGENSOFT was used to
604 project the ancient samples onto the PCA axes defined by the extant samples.

605 To investigate the relatedness between the Yoram Cave samples (and extant
606 landraces) and wild barley accessions representing the entire natural
607 distribution range, the corresponding data was extracted from the filtered SNP
608 table. Relatedness between the archeological samples and each wild accession
609 was measured by the level of IBS calculated across all SNPs using the SNPRelate
610 package⁵⁷ in R. For each wild accession, relatedness to the five archeological
611 samples was averaged using the geometrical mean to obtain one relatedness
612 score (RS). We then calculated the geographic distance (GD) between the Yoram
613 Cave location (latitude: 31.314 N, longitude: 35.353 E) and the sampling position
614 of each wild accession based on its coordinates, converting the distances
615 between coordinate positions to kilometres using the rough conversion metrics
616 of 1 degree = 111 km. To capture the change in the correlation coefficient sign
617 observed between geographic distance and genetic relatedness, the data were
618 split into geographically proximate and distant categories using the most related
619 wild accession coordinates rounded up to the nearest 50 km as a break point.

620 Model-based ancestry estimation was performed with ADMIXTURE (ref. 31). For
621 each K from 1 to 10, twenty replicate ADMIXTURE runs were performed on the
622 genotype matrix of 233 samples (228 geo-referenced extant accessions of known
623 domestication status¹⁸ plus five ancient samples) using only transversion
624 variants with a present genotype call for at least one ancient samples. Before
625 running ADMIXTURE, linkage disequilibrium pruning was done with the R
626 package SNPRelate⁵⁷ using the function snpgdsLDpruning() with the parameters
627 “ld.threshold=0.4, slide.max.bp=100000, slide.max.n=50”. Replicate ADMIXTURE
628 runs were combined with CLUMPP⁵⁸.

629 D statistics were calculated using ADMIXtools⁵⁹ after the SNP matrix had been
630 converted to the EIGENSOFT format with the SNPRelate⁵⁷ function
631 snpgdsGDS2Eigen(). The barley relative *Hordeum pubiflorum* was used as an
632 outgroup. We used exome sequencing reads of *H. pubiflorum* published by
633 Mascher et al. (ref. 17) to call genotypes at variant positions with SAMtools.
634 Read alignments at the *Btr1/2* and *Vrs1* loci were manually inspected with
635 SAMtools “tview”.

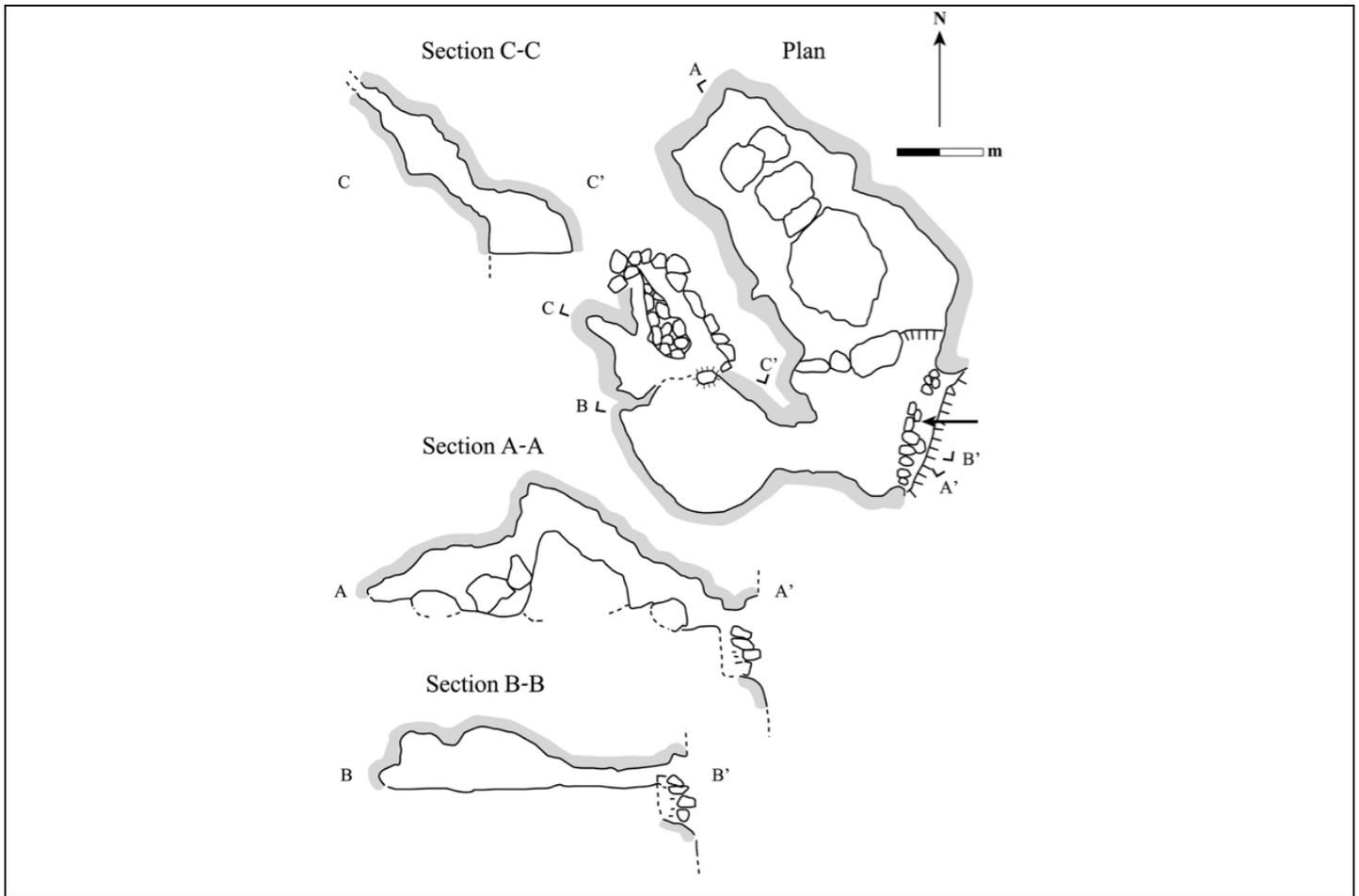
636 The map in Figure 2a was generated with the R package ‘mapdata.’
637

638 **Additional references**

639

- 640 42. Schick, T. The early basketry and textiles from caves in the northern Judean
641 desert. *Part II'Atiqot* **41**, 223-239 (2002).
- 642 43. Kistler, L. Ancient DNA extraction from plants. *Methods Mol Biol* **840**, 71-9
643 (2012).
- 644 44. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly
645 multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010**,
646 pdb prot5448 (2010).
- 647 45. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes
648 inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids*
649 *Res* **40**, e3 (2012).
- 650 46. Himmelbach, A., Knauff, M. & Stein, N. Plant sequence capture optimised for
651 Illumina sequencing. *Bio Protoc* **4**, e1166 (2014).
- 652 47. Kircher, M. Analysis of high-throughput ancient DNA sequencing data.
653 *Methods Mol Biol* **840**, 197-228 (2012).
- 654 48. Renaud, G., Stenzel, U. & Kelso, J. leeHom: adaptor trimming and merging
655 for Illumina sequencing reads. *Nucleic Acids Res* **42**, e141 (2014).
- 656 49. Gallego Llorente, M. *et al.* Ancient Ethiopian genome reveals extensive
657 Eurasian admixture throughout the African continent. *Science* **350**, 820-2
658 (2015).
- 659 50. International Barley Genome Sequencing Consortium. A physical, genetic and
660 functional sequence assembly of the barley genome. *Nature* **491**, 711-6
661 (2012).
- 662 51. Saski, C. *et al.* Complete chloroplast genome sequences of *Hordeum vulgare*,
663 *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with
664 other grass genomes. *Theor Appl Genet* **115**, 571-90 (2007).
- 665 52. Li, H. Aligning sequence reads, clone sequences and assembly contigs with
666 BWA-MEM. *arXiv preprint arXiv:1303.3997* (2013).
- 667 53. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools.
668 *Bioinformatics* **25**, 2078-9 (2009).
- 669 54. Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P.L. & Orlando, L.
670 mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage
671 parameters. *Bioinformatics* **29**, 1682-4 (2013).

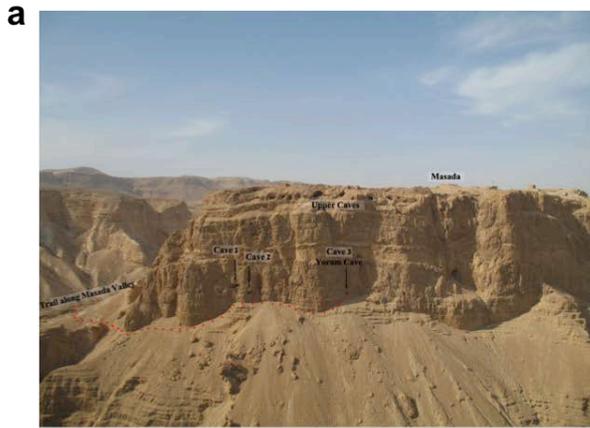
- 672 55. Li, H. A statistical framework for SNP calling, mutation discovery, association
673 mapping and population genetical parameter estimation from sequencing data.
674 *Bioinformatics* **27**, 2987-2993 (2011).
- 675 56. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for
676 comparing genomic features. *Bioinformatics* **26**, 841-2 (2010).
- 677 57. Zheng, X. *et al.* A high-performance computing toolset for relatedness and
678 principal component analysis of SNP data. *Bioinformatics* **28**, 3326-8 (2012).
- 679 58. Jakobsson, M. & Rosenberg, N.A. CLUMPP: a cluster matching and
680 permutation program for dealing with label switching and multimodality in
681 analysis of population structure. *Bioinformatics* **23**, 1801-6 (2007).
- 682 59. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065-
683 93 (2012).
- 684
- 685



Supplementary Figure 1

Plan of Yoram Cave – top plan and sections.

Arrows indicates the entrance to the cave and the human-made wall across it. Note the boulders in the northern room and the relatively horizontal surface in the southern room. Gray shaded borders indicate that the edge of the cave is cut into rock.



Supplementary Figure 2

Photographs from the excavation.

Masada Southern Cave complex, with three caves located in the south-eastern cliff of the Masada Horst **(a)**. The dotted red line marks the easiest trail giving access to the caves. The entrance of Yoram Cave, facing south-east, in an almost vertical cliff, some 4 meters above the trail leading to the cave **(b)**. The south room during excavation of locus 3 **(c)**.

JK3009



JK3010



JK3011



JK3012



JK3013

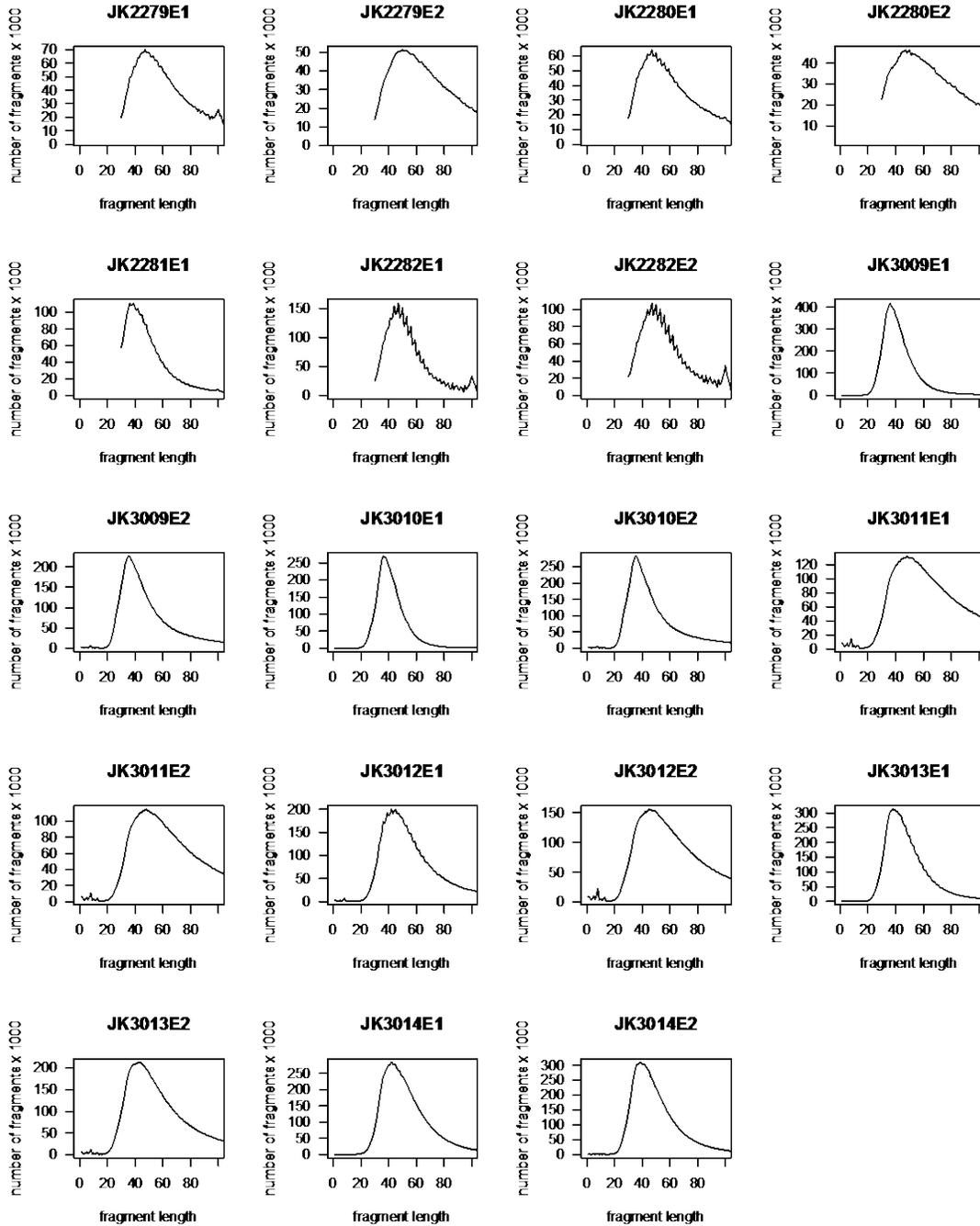


JK3014



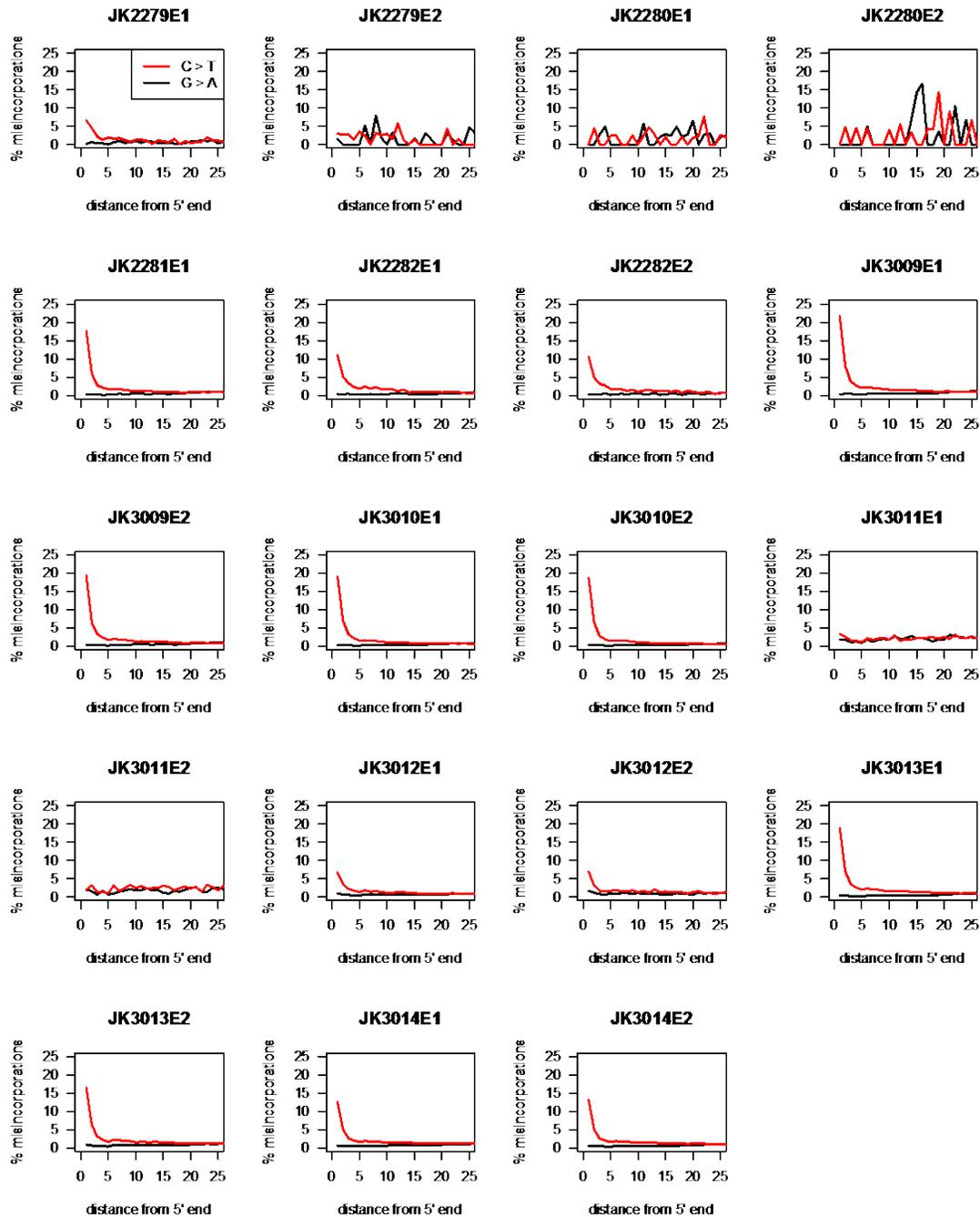
Supplementary Figure 3

Photographs of ancient barley grains used for DNA extraction.



Supplementary Figure 4

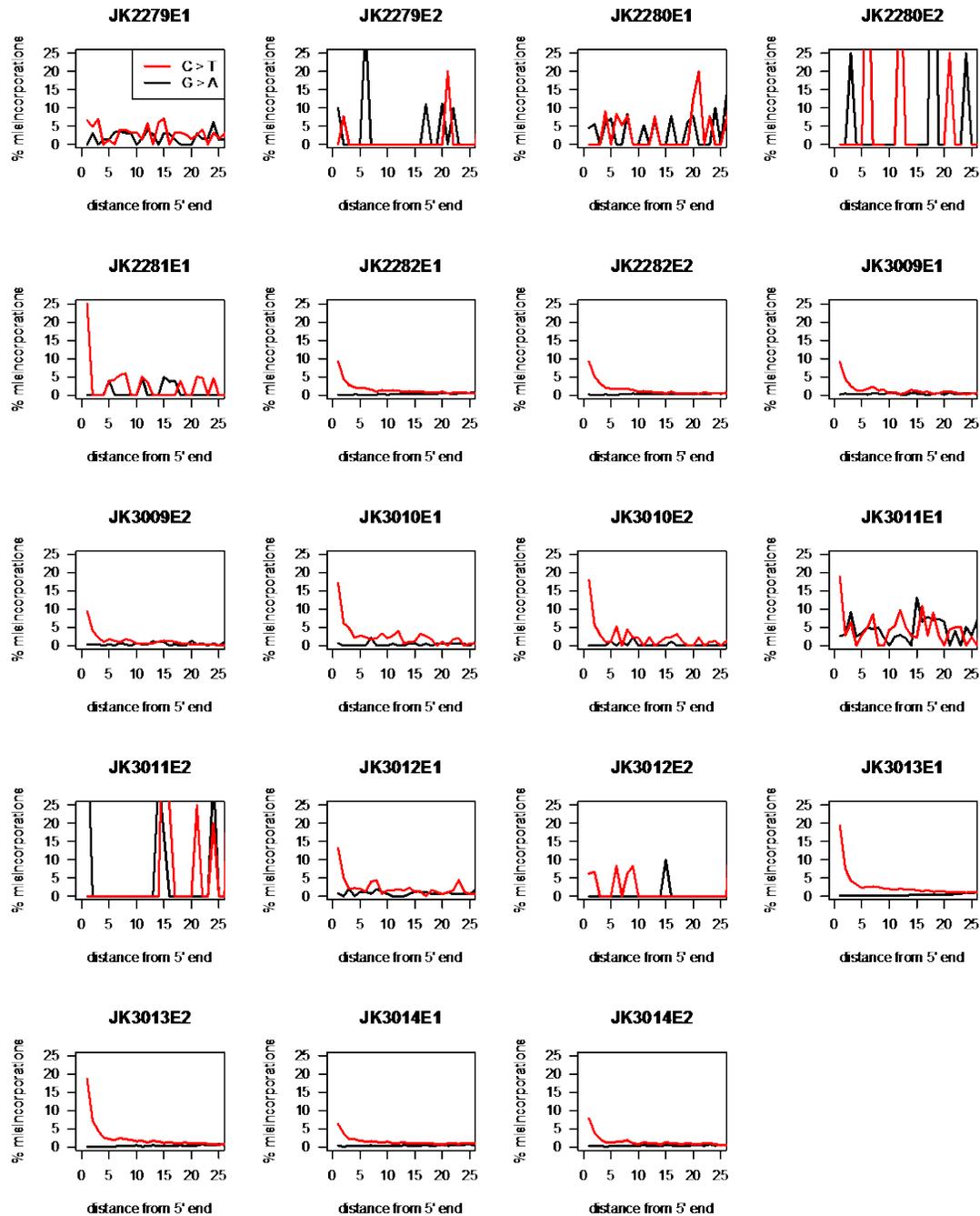
Length distribution of sequence fragments.



Supplementary Figure 5

Nucleotide misincorporation profiles in reads mapped to the whole-genome shotgun assembly of barley cv. Morex.

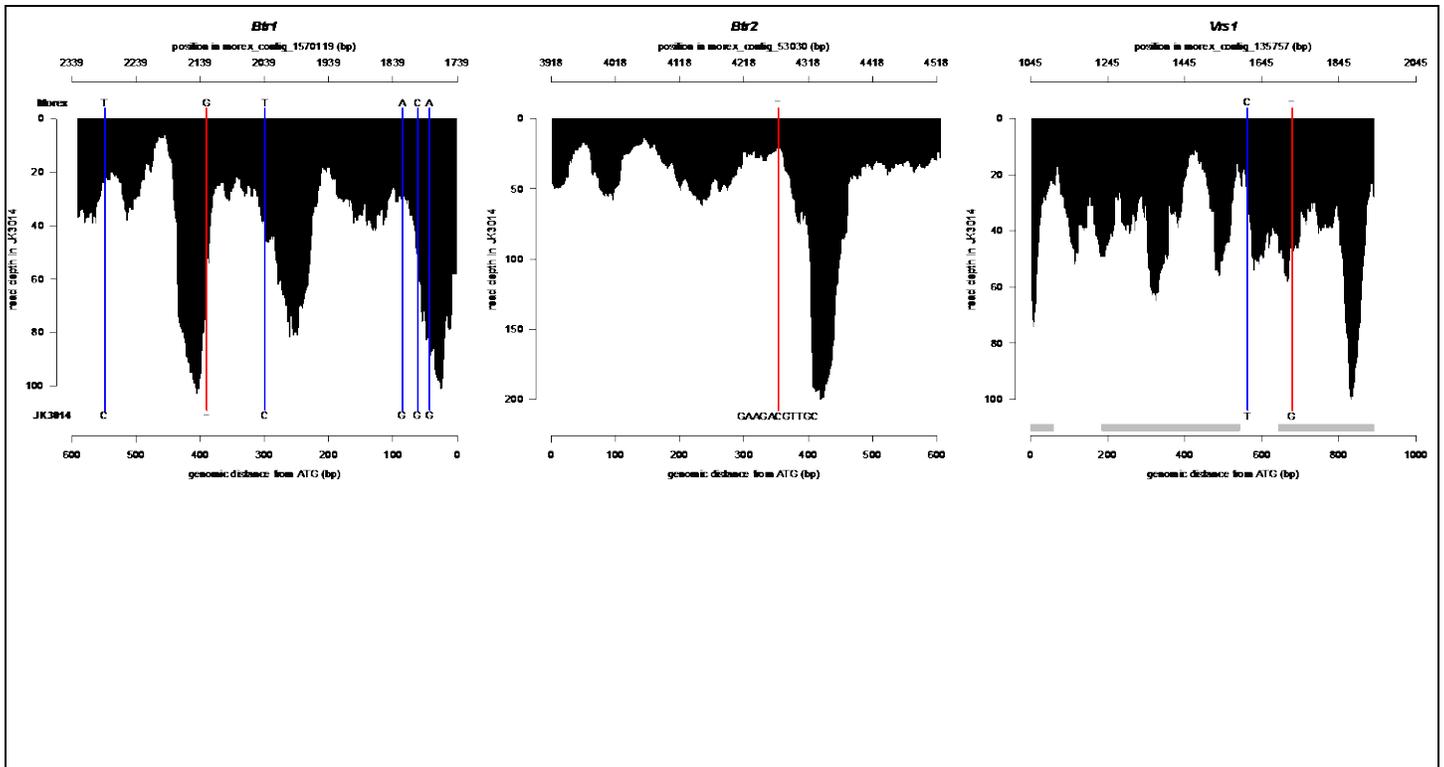
The proportion of C > T misincorporations (red) is compared to the G > A baseline (black).



Supplementary Figure 6

Nucleotide misincorporation profiles in reads mapped to the barley chloroplast genome assembly.

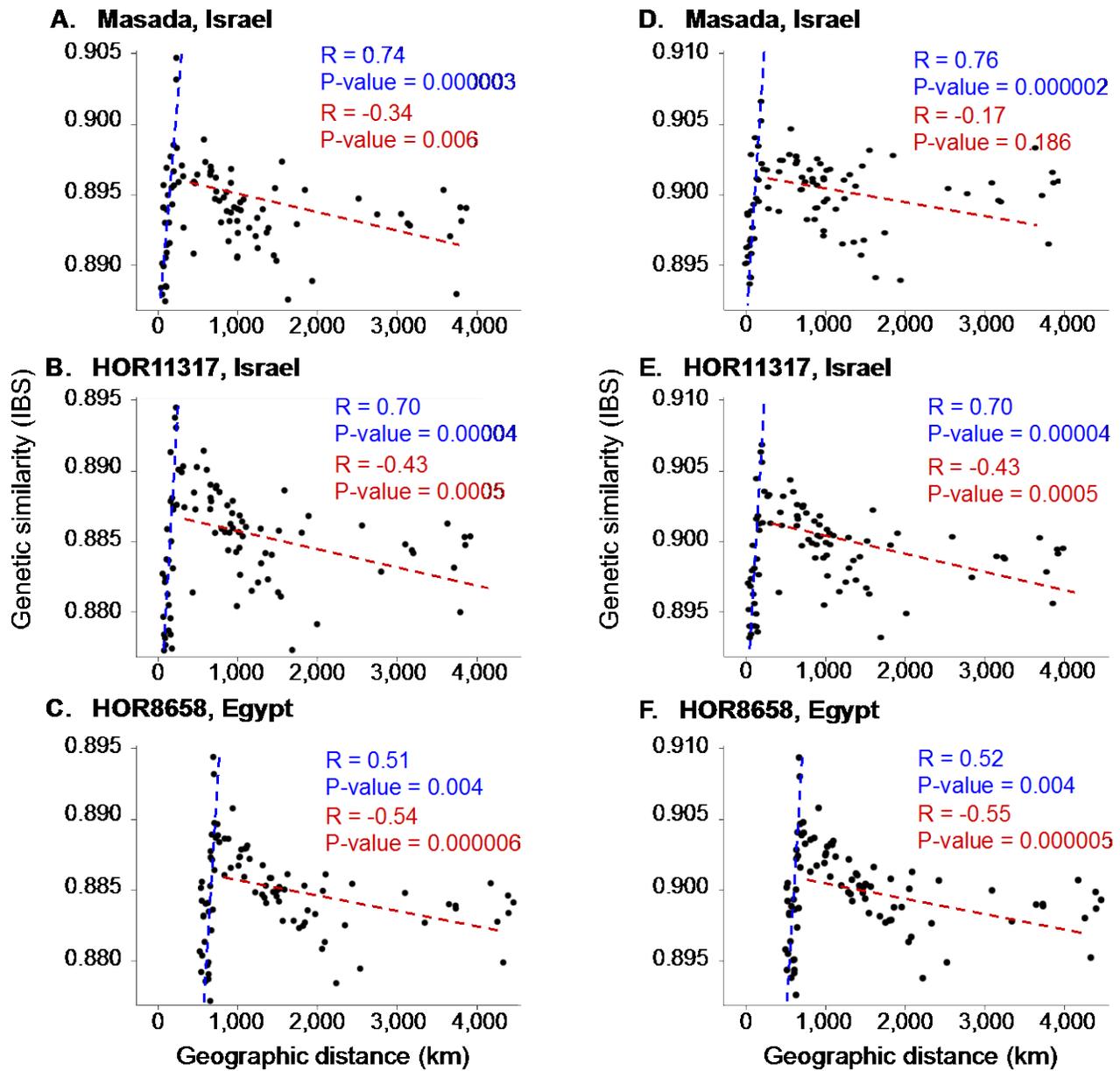
The proportion of C > T substitutions (red) is compared to the G > A baseline (black).



Supplementary Figure 7

Haplotype of *Btr1*, *Btr2* and *Vrs1*

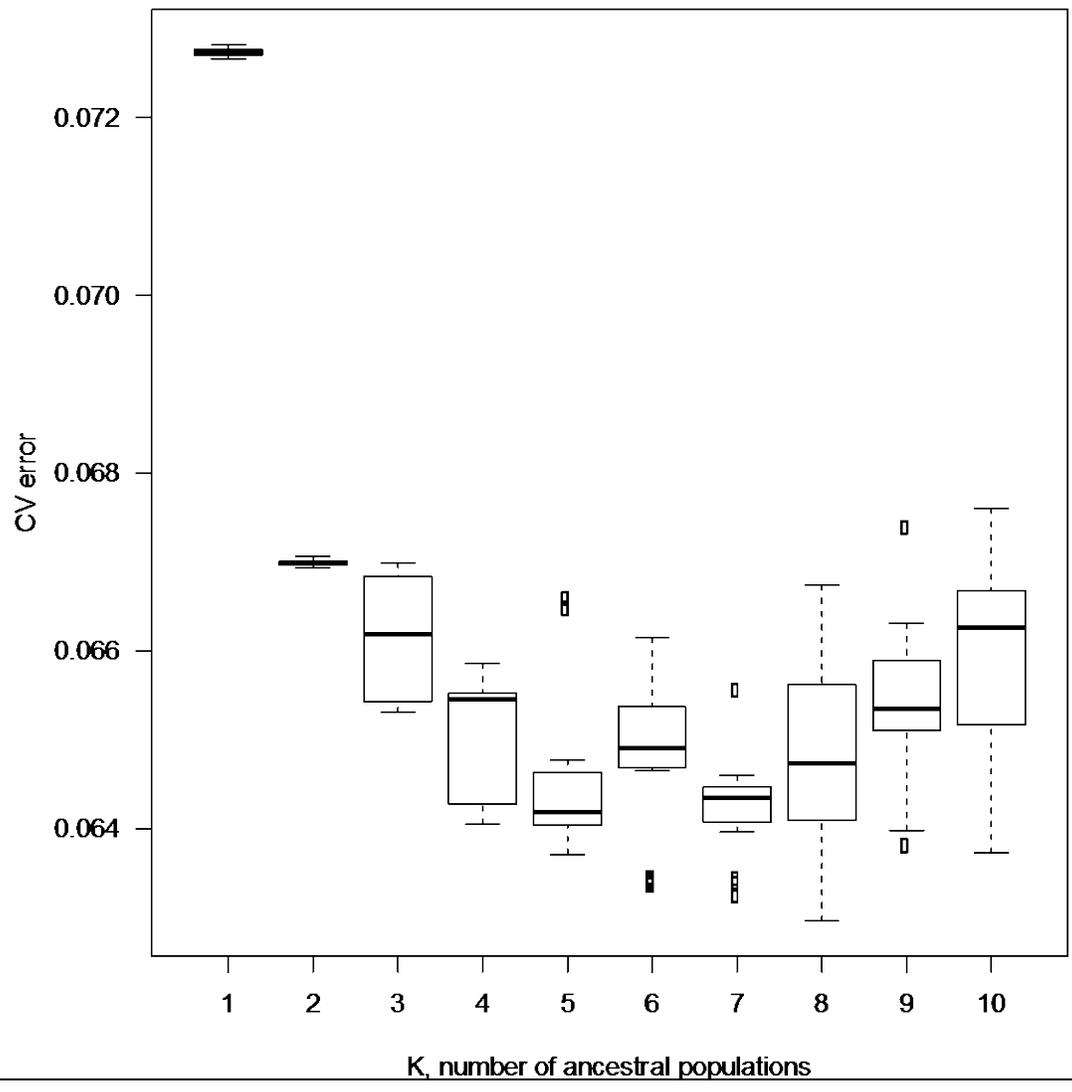
Contigs of the Morex WGS assembly harboring these genes were identified by BLAST searches. The read depth in the deeply sequenced sample JK3014 is shown in black. The position in the respective WGS contigs is indicated in the upper axis. The distance (in the genomic sequence) from the start codon of each gene is shown in the bottom axis. *Btr1* and *Btr2* are single-exon genes. Gray bars indicate the positions of the three exons of *Vrs1*. The positions of SNPs are highlighted by vertical lines. The Morex allele is shown above, and the JK3014 allele below the lines. SNPs with a previously reported functional effect are shown in red. JK3014 carries a loss-of-function allele of *Btr1*, while *Btr2* and *Vrs1* are wild-type alleles. The coding sequence of *Btr1* and *Btr2* is identical to the haplotype of cv. Haruna Nijo as reported by Komatsuda et al., 2015 (NCBI GenBank accession KR813337.1). The sequence of *Vrs1* matches the *Vrs1.b2* allele as designated by Komatsuda et al., 2007.



Supplementary Figure 8

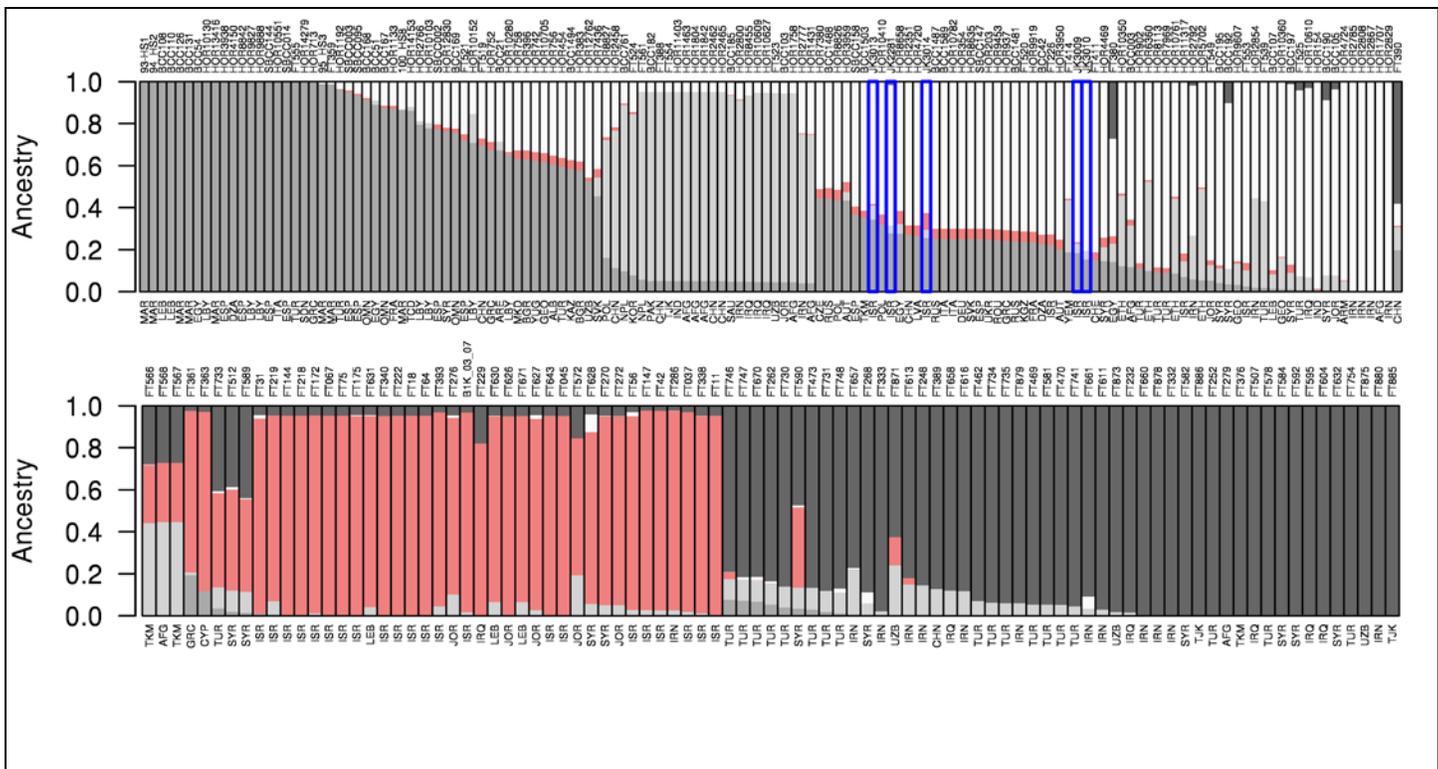
Relationship between genetic similarity and geographic distance

Scatter plot of genetic similarity and geographic distance between 91 extant wild barley accessions sampled across the Fertile Crescent and (A) archaeological sample JK3014 found at Yoram Cave and sequenced to higher depth using all SNPs, (B) two-row cultivated landrace from Israel, (C) a two-rowed cultivated landrace from Egypt, (D) the ancient sample JK3014 found at Yoram Cave and sequenced to higher depth excluding transition SNPs, (E) two-rowed cultivated landraces from Israel excluding transition SNPs, and (F) two-rowed cultivated landraces from Egypt excluding transition SNPs. The geographic position attributed to each sample is: A, D (31.3141 N, 35.353 E); B,E (31.7156 N, 35.1871 E); C,F (31.193 N 29.904 E). Correlation coefficients and P-values for the geographically proximate and distant subsets are indicated in blue and red, respectively.



Supplementary Figure 9

Cross-validation error of ADMIXTURE analysis. Box plots were calculated from 20 replicate runs for each K.



Supplementary Figure 10

ADMIXTURE analysis for K=5 for domesticated (top) and wild (bottom) samples.

Colors in both panels correspond to the same ancestral populations. Sample names and countries of origin are indicated above and below the plots, respectively. Ancient samples are highlighted by blue borders.

Supplementary Table 1. Sequencing statistics of ancient barley DNA without UDG treatment. Read statistics after UDG treatment are provided in Table 1.

Sample	C14 age ¹	Cal age ²	Extraction	Raw reads	Merged reads	% mapped	% unique	% damaged
JK2279	5065 ± 27	3942-3802	1	8,208,520	3,548,936	31.5%	18.0%	6.6%
JK2279	5065 ± 27	3942-3802	2	7,809,034	3,240,281	2.8%	27.2%	3.0%
JK2280	NA	NA	1	7,410,426	3,192,474	18.6%	8.4%	4.4%
JK2280	NA	NA	2	8,537,472	3,150,021	4.7%	15.0%	4.8%
JK2281	5290 ± 27	4226-4047	1	7,257,940	3,210,243	56.9%	55.8%	17.7%
JK2282	5275 ± 27	4225-4006	1	9,601,504	4,513,912	95.9%	9.8%	11.1%
JK2282	5275 ± 27	4225-4006	2	8,401,768	3,772,270	75.7%	10.4%	10.5%
JK3009	5034 ± 36	3939-3775	1	18,769,156	8,976,322	67.5%	77.5%	21.8%
JK3009	5034 ± 36	3939-3775	2	15,308,126	7,037,725	33.2%	68.6%	19.5%
JK3010	5032 ± 37	3940-3773	1	11,237,658	5,427,102	71.6%	78.1%	19.2%
JK3010	5032 ± 37	3940-3773	2	17,609,772	8,071,918	37.8%	58.3%	18.8%
JK3011	5115 ± 37	3968-3812	1	19,100,026	8,175,478	0.7%	54.4%	3.4%
JK3011	5115 ± 37	3968-3812	2	15,648,476	6,838,637	0.4%	58.9%	1.9%
JK3012	4987 ± 37	3886-3707	1	17,736,280	7,985,422	42.6%	12.3%	6.6%
JK3012	4987 ± 37	3886-3707	2	20,000,020	8,788,418	2.8%	24.4%	7.0%
JK3013	5227 ± 37	4048-3976	1	19,482,948	9,219,005	60.7%	50.2%	18.9%
JK3013	5227 ± 37	4048-3976	2	21,479,860	9,742,769	11.7%	48.9%	16.5%
JK3014	4988 ± 36	3886-3708	1	20,069,866	9,681,200	87.8%	81.9%	12.5%
JK3014	4988 ± 36	3886-3708	2	20,901,314	9,939,663	73.3%	74.4%	13.1%

¹Uncalibrated radiocarbon years before the present; ²Calibrated age, calendar years before Christ

Supplementary Table 2: Rare allele sharing between ancient and modern samples. The table lists extant samples that share at least 30 rare alleles (minor allele count ≤ 5) with the ancient samples. Only transversion variants were considered for this analysis.

sample	country of origin	domestication status	number of shared alleles
HOR4469	Syria	landrace 2-row	89
HOR2766	Libya	landrace 6-row	63
FT549	Jordan	landrace 2-row	56
HOR8658	Egypt	landrace 2-row	56
FT75	Israel	spontaneum	45
HOR10280	Libya	landrace 2-row	42
BCC107	Lebanon	landrace 2-row	40
FT037	Israel	spontaneum	38
FT31	Israel	spontaneum	38
FT340	Israel	spontaneum	38
100_HS8	Morocco	landrace 6-row	37
HOR10103	Libya	landrace 6-row	36
FT218	Israel	spontaneum	35
HOR11317	Israel	landrace 2-row	35
HOR14153	Chad	landrace 6-row	35
FT276	Jordan	spontaneum	34
FT631	Lebanon	spontaneum	32
FT393	Israel	spontaneum	30

Supplementary Table 4: D statistics for different comparisons between wild and domesticated barleys from the Levant, European landraces and ancient samples. Standard errors and Z scores (D/SE) are given in columns SE and Z.

P1	P2	P3	Outgroup	D	all SNPs			transversions only		
					SE	Z	D	SE	Z	
Levantine landraces	European landraces	Levant wild	H. pubiflorum	0.037	0.004328	8.546	0.0342	0.005323	6.418	
Levantine landraces	European landraces	JK2281	H. pubiflorum	0.1076	0.019226	5.594	0.1007	0.021775	4.623	
Levantine landraces	European landraces	JK3009	H. pubiflorum	0.0898	0.015975	5.62	0.0499	0.01801	2.768	
Levantine landraces	European landraces	JK3010	H. pubiflorum	0.1671	0.016772	9.963	0.1534	0.018811	8.155	
Levantine landraces	European landraces	JK3013	H. pubiflorum	0.0955	0.027177	3.516	0.0986	0.037957	2.598	
Levantine landraces	European landraces	JK3014	H. pubiflorum	0.1864	0.015474	12.045	0.1854	0.016067	11.538	
JK2281	European landraces	Levant wild	H. pubiflorum	0.061	0.01632	3.737	0.0116	0.021853	0.531	
JK2281	Levantine landraces	Levant wild	H. pubiflorum	0.0265	0.01447	1.83	0.0018	0.020225	0.088	
JK3009	European landraces	Levant wild	H. pubiflorum	0.0562	0.0166	3.386	0.0662	0.023755	2.789	
JK3009	Levantine landraces	Levant wild	H. pubiflorum	0.019	0.014787	1.288	0.0617	0.022587	2.731	
JK3010	European landraces	Levant wild	H. pubiflorum	0.0366	0.010696	3.424	0.0227	0.015691	1.445	
JK3010	Levantine landraces	Levant wild	H. pubiflorum	0.0031	0.009611	0.325	-0.0068	0.014167	-0.48	
JK3013	European landraces	Levant wild	H. pubiflorum	0.0883	0.028889	3.058	0.1282	0.056598	2.266	
JK3013	Levantine landraces	Levant wild	H. pubiflorum	0.0809	0.029933	2.704	0.1068	0.055261	1.933	
JK3014	European landraces	Levant wild	H. pubiflorum	0.0413	0.007807	5.288	0.0385	0.010813	3.558	
JK3014	Levantine landraces	Levant wild	H. pubiflorum	0.0077	0.007189	1.066	0.0068	0.009839	0.696	