1 **Title**

2 # Neolithic and Medieval virus genomes reveal complex evolution of Hepatitis B

3

4 **Authors**

5 Ben Krause-Kyora[1,2,†*], Julian Susat[1*], Felix M. Key[2], Denise Kühnert[3,2], Esther Bosse[1,4],
6 Alexander Immel[1,2], Christoph Rinne[5], Sabin-Christin Kornell[1], Diego Yepes[4], Sören
7 Franzenburg[1], Henrike O. Heyne[6], Thomas Meier[7], Sandra Lösch[8], Harald Meller[9], Susanne
8 Friederich[9], Nicole Nicklisch[9,10], Kurt Werner Alt[9,10,11,12], Stefan Schreiber[1,13], Andreas
9 Tholey[4], Alexander Herbig[2], Almut Nebel[1], Johannes Krause[2,†]

10

11 **Affiliations**

12 [1]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany.
13 [2]Max Planck Institute for the Science of Human History, Jena, Germany.
14 [3]Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, Zurich,
15 Switzerland
16 [4]Systematic Proteomics & Bioanalytics, Institute for Experimental Medicine, Kiel University, Kiel,
17 Germany.
18 [5]Institute of Pre- and Protohistoric Archaeology, Kiel University, Kiel, Germany.
19 [6]Broad Institute, Stanley Center for Psychiatric Research, Cambridge, MA, USA.
20 [7]Institute for Pre- and Protohistory and Near Eastern Archaeology, Heidelberg University, Heidelberg,
21 Germany.
22 [8]Department of Physical Anthropology, Institute of Forensic Medicine, University of Bern, Bern,
23 Switzerland.
24 [9]State Office for Heritage Management and Archaeology Saxony-Anhalt and State Museum of
25 Prehistory, Halle, Germany.
26 [10]Danube Private University, Krems, Austria.
27 [11]Department of Biomedical Engineering, University Hospital Basel, University of Basel, Basel,
28 Switzerland.
29 [12]Integrative Prehistory and Archaeological Science, University of Basel, Basel, Switzerland.
30 [13]Clinic for Internal Medicine I, University Hospital of Schleswig-Holstein, Kiel, Germany.

31

32 †Corresponding author E-mails: b.krause-kyora@ikmb.uni-kiel.de, krause@shh.mpg.de
33 *These authors contributed equally to this work
34

35

36 **Abstract**

37 The hepatitis B virus (HBV) is one of the most widespread human pathogens known today,
38 yet its origin and evolutionary history are still unclear and controversial. Here, we report the
39 analysis of three ancient HBV genomes recovered from human skeletons found at three
40 different archaeological sites in Germany. We reconstructed two Neolithic and one medieval
41 HBV genomes by *de novo* assembly from shotgun DNA sequencing data. Additionally, we
42 observed HBV-specific peptides using paleo-proteomics. Our results show that HBV
43 circulates in the European population for at least 7000 years. The Neolithic HBV genomes
44 show a high genomic similarity to each other. In a phylogenetic network, they do not group
45 with any human-associated HBV genome and are most closely related to those infecting
46 African non-human primates. These ancient virus forms appear to represent distinct lineages
47 that have no close relatives today and possibly went extinct. Our results reveal the great
48 potential of ancient DNA from human skeletons in order to study the long-time evolution of
49 blood borne viruses.

50

51

52 **Introduction**

53 The hepatitis B virus (HBV) is one of the most widespread human pathogens, with worldwide
54 over 250 million people being infected, and an annual death toll of about 1 million globally
55 (WHO, 2017). Infection of liver cells with HBV leads to acute hepatitis B, which is self-
56 limiting in about 90-95% of cases. In about 5-10% of infected individuals virus clearance fails
57 and patients develop chronic infection of hepatitis B, which puts them at lifelong elevated risk
58 for liver cirrhosis and liver cancer (hepatocellular carcinoma). HBV is usually transmitted by
59 contact with infectious blood, in highly endemic countries often during birth (WHO, 2017).
60 HBV has a circular, partially double-stranded DNA genome of about 3.2kbp that encodes four
61 overlapping open reading frames (P, pre-S/S, pre-C/C, and X). Based on the genomic
62 sequence diversity, HBVs are currently classified into 8 genotypes (A-H) and numerous
63 subgenotypes that show distinct geographic distributions (Castelhano et al., 2017). All
64 genotypes are hypothesised to be primarily the result of recombination events (Littlejohn et
65 al., 2016; Simmonds and Midgley, 2005). To a lesser extent, HBV evolution is also driven by
66 the accumulation of point mutations (Schaefer 2007, Araujo 2015).
67 Despite being widespread and well-studied, the origin and evolutionary history of HBV is still
68 unclear and controversial (Littlejohn et al., 2016, Souza et al., 2014). HBVs in non-human
69 primates (NHP), for instance in chimpanzees and gorillas, are phylogenetically closely related
70 to, and yet distinct from, human HBV isolates, supporting the notion of an Africa origin of the
71 virus (Souza et al., 2014). Molecular-clock based analyses dating the origin of HBV have
72 resulted in conflicting estimates with some as recent as about 400 years ago (Zhou and
73 Holmes, 2007, Souza et al., 2014). These observations have raised doubts about the suitability
74 of molecular dating approaches for reconstructing the evolution of HBV (Bouckaer et al.,
75 2103, Souza et al., 2014). Moreover, ancient DNA (aDNA) research on HBV-infected
76 mummies from the 16th century AD revealed a very close relationship between the ancient
77 and modern HBV genomes (Kahila Bar-Gal et al., 2012, Patterson Ross et al., 2018),
78 indicating a surprising lack of temporal genetic changes in the virus during the last 500 years
79 (Patterson Ross et al., 2018). Therefore, diachronic aDNA HBV studies, in which both the
80 changes in the viral genome over time as well as the provenance and age of the archaeological
81 samples, are needed to better understand the origin and evolutionary history of the virus.
82 Here, we report the analysis of three complete HBV genomes recovered from human skeletal
83 remains from the prehistoric Neolithic and Medieval Periods in Central Europe. Our results
84 show that HBV already circulated in the European population more than 7000 years ago.

85 Although the ancient forms show a relationship to modern isolates they appear to represent
86 distinct lineages that have no close modern relatives and are possibly extinct today.
87

88 **Results and Discussion**

89 We detected evidence for presence of ancient HBV in three human tooth samples as part of a
90 metagenomic screening for viral pathogens that was performed on shotgun sequencing data
91 from 45 skeletons using the metagenomic alignment software MALT (Vagene et al., 2018).
92 The remains of the individuals were excavated from the Neolithic sites of Karsdorf
93 (Linearbandkeramik [LBK], 5056–4959 cal BC) and Sorsum (Tiefstichkeramik group of the
94 Funnel Beaker culture, 3335-3107 cal BC), the medieval cemetery of Petersberg/Kleiner
95 Madron (1020-1116 cal AD), all located in Germany (Fig. A, Supplementary Figure S1-S3).
96 After the three aDNA extracts had appeared HBV-positive in the initial virus screening, they
97 were subjected to deep-sequencing without any prior enrichment resulting in 367 to 419
98 million reads per sample (table 1). Analysis of the human DNA recovered from Karsdorf (3-
99 fold genomic coverage) revealed that the sample clusters tightly with other contemporary
100 early Neolithic individuals from the LBK (Supplementary Figure S12). The genetic makeup
101 of the early LBK agriculturalists was previously found quite distinct from the preceding
102 western hunter gatherers of Europe. The genetic shift between both populations was
103 interpreted as a result of early farmers migrating from Western-Anatolia into Central Europe
104 introducing agriculture (Lazaridis et al. 2014, Haak et al. 2015). The almost 2000 years
105 younger Sorsum individual (1.2-fold genomic coverage) is most closely related with
106 individuals from the contemporary Funnel Beaker culture that inhabited Northern Germany at
107 the end of the 4th millenium BCE (Supplementary Figure S12). This population was
108 previously shown to be quite admixed, as a result of a spatial and temporal overlap of early
109 Neolithic farmers and remaining western hunter gatherers for almost 2000 years (Bollongino
110 et al. 2013, Haak et al. 2015). The Petersberg individual (2.9-fold genomic coverage) showed,
111 genetic affinities with modern day central European populations, typical for individuals from
112 that time. All three ancient human individuals are therefore in agreement with the
113 archaeological evidence and radiocarbon dates for their respective time of origin. Together
114 with typical aDNA damage patterns (Supplementary Figures S4-S5), the human population
115 genetic investigation supports the ancient origin of the obtained datasets.

116 For successful HBV genome reconstruction, we mapped all metagenomic sequences to 16
117 HBV reference genomes (8 human genotypes (A-H) and 8 NHP genotypes from Africa and
118 Asia) that are representative of the current HBV strain diversity (Supplementary Table S6).
119 The mapped reads were used for a *de novo* assembly, resulting in contigs from which one
120 ancient HBV consensus sequence per sample was constructed. The consensus genomes are
121 3161 (46-fold coverage), 3182 (47-fold coverage), and 3183 (105-fold coverage) nucleotides

122 in length, which falls in the length range of modern HBV genomes and suggests that we
123 successfully reconstructed the entire ancient HBV genomes (table 1, Supplementary Figure
124 S6-S8). Further, when we conducted liquid chromatography-mass spectrometry (LC-MS)
125 based bottom-up proteomics on tooth material from the three individuals, we identified in the
126 Karsdorf and Petersberg samples a peptide that is part of the very stable HBV core protein,
127 supporting the presence and active replication of HBV in the individuals' blood
128 (Supplementary Figure S18).
129 Phylogenetic network analysis was carried out with a dataset comprised of 493 modern HBV
130 strains representing the full genetic diversity. Strikingly, the Neolithic HBV genomes did not
131 group with any human strain in the phylogeny. Instead, they branched off in two lineages and
132 were most closely related to the African NHP genomes (Fig. B, 93% similarity). Although the
133 two Neolithic strains were recovered from humans who had lived about two thousand years
134 apart, they showed a higher genomic similarity to each other than to any other human or NHP
135 genotype. Still, their genomes differed by 6% from each other and may therefore be
136 considered representatives of two separate lineages. They did, however, differ less than 8%
137 from the African NHP strains and should therefore not be called a separate genotype. The
138 genome from the 1000-year-old Petersberg individual clustered with modern D4 genotypes.
139 Owing to continuous recombination over time, different gene segments or modules of the
140 ancestral genomes can show up in various subsequent virus generations. Such precursors have
141 been postulated (Simmonds and Midgley, 2005) and their existence is supported by the results
142 of our recombination analysis (Supplementary Figure S14-S17, source data 1). Some
143 fragments of the Karsdorf sequences appeared to be very similar to modern human (G, E) and
144 African NHP genotypes, and the Sorsum genome partially showed a high similarity to the
145 human genotypes G, E and B. (Supplementary Figure S14-S15, S17, source data 1). Given the
146 close relationship between the two Neolithic virus genomes, it is also conceivable that the
147 older HBV from Karsdorf could have been a distant source for the younger Sorsum virus
148 (Supplementary Figure S14-S15, S17, source data 1). The closer relationship between the
149 Neolithic and the NHP strains compared to other human strains is noteworthy and may have
150 involved reciprocal cross-species transmission at one or possibly several times in the past
151 (Simmonds and Midgley, 2005, Souza et al., 2014, Rasche et al. 2016).
152 Taken together, our results demonstrate that HBV already existed in Europeans 7000 years
153 ago and that its genomic structure closely resembled that of modern hepatitis B viruses. Both
154 Neolithic viruses fall between the present-day modern human and the known NHP diversity.
155 Therefore, it can be hypothesized that although the two Neolithic HBV strains are no longer

156 observed today and thus may reflect two distinct clades that went extinct, they could still be
157 closely related to the remote ancestors of the present-day genotypes, which is supported by
158 signs of ancient recombination events. More ancient precursors, intermediates and modern
159 strains of both humans and NHPs need to be sequenced to disentangle the complex evolution
160 of HBV. As this evolution is characterized by recombination and point mutations and may
161 further be complicated by human-ape host barrier crossing (Simmonds and Midgley, 2005,
162 Souza et al., 2014, Rasche et al. 2016), genetic dating is not expected to yield meaningful
163 results. This is additionally supported by a TempEst analysis (Rambaut et al., 2016) that
164 shows very little temporal signal (Supplementary Figure S10). It should, however, be noted
165 that the oldest genome (Karsdorf) was found in an individual that belonged to a population of
166 early farmers that had migrated in the previous few hundred years from the Near East into
167 central Europe. One might speculate that the close proximity to recently domesticated
168 animals, changes in subsistence strategy as well as the adopted sedentary lifestyle might have
169 contributed to the spread of HBV within Neolithic human populations.
170 Based on our analysis, HBV DNA can reliably be detected in tooth samples that are up to
171 7000 years old. Ancient HBV has so far only been identified in soft tissue from two 16th-
172 century mummies (Kahila Bar-Gal et al., 2012, Patterson Ross et al., 2018). The aDNA
173 analysis of HBV from prehistoric skeletons, which facilitates evolutionary studies on a
174 temporal scale as deep, has not been described up to now. One explanation for the difficulty
175 of a molecular HBV diagnosis in bones is that the virus infection does not leave lesions on
176 skeletal remains that would allow researchers to select affected individuals *a priori*, as it is the
177 case for instance for leprosy (Schuenemann et al., 2013). The diagnosis of an HBV infection
178 in skeletal populations is purely a chance finding and is thus more probable in a large-scale
179 screening.
180 Overall, HBV biomolecules seem to be well preserved in teeth: Avoiding biases from DNA
181 capture and reference-based mapping we could reconstruct three HBV genomes by *de novo*
182 assembly from shotgun data and even observed HBV-specific peptides. The ratio of HBV
183 genomes to the human genome in our samples was rather high and similar in all three samples
184 (Karsdorf 35:1, Sorsum 40.2:1 and Petersberg 16:1). As there is no evidence that HBV DNA
185 is more resistant to postmortem degradation than human DNA, the high rate of HBV
186 compared to human DNA may reflect the disease state in the infected individuals at the time
187 of death. High copy numbers of viral DNA in the blood of infected individuals are associated
188 with acute HBV infection, or reactivation of chronic HBV. Thus, it seems likely that the death
189 of the ancient individuals is related to the HBV infection, but might not be the direct cause of

190 death as fulminant liver failure is rather rare in modern day patients. The HBV infection
191 might have instead contributed to other forms of lethal liver failure such as cirrhosis or liver
192 cancer.
193
194 In view of the unexpected complexity of our findings, we envisage future diachronic HBV
195 studies that go beyond the temporal and geographic scope of our current work.
196

197 **Materials and Methods**

198

199 ***Human remains***

200 The LBK settlement of **Karsdorf**, Saxony-Anhalt, Germany, is located in the valley of the
201 river Unstrut. Between 1996-2010 systematic excavations were conducted at Karsdorf that led
202 to the discovery of settlements and graves from the Neolithic to the Iron Age (Behnke, 2007,
203 2011, 2012). The LBK is represented by 24 longhouses in north-west to south-east orientation
204 that were associated with settlement burials (Veit, 1996). The investigated individual 537 is a
205 male with an age at death of around 25-30 years (Supplementary Figure S1), dated to 5056–
206 4959 cal BC (KIA 40357 – 6116 ± 32 BP) (Brandt et al., 2014, Nicklisch, 2017).

207 The gallery grave of **Sorsum**, Lower-Saxony, Germany, is typologically dated to the
208 Tiefstichkeramik (group of the Funnelbeaker culture). Sorsum is exceptional as it was built
209 into the bedrock. During the excavations (1956-1960) of the grave chamber around 105
210 individuals were recovered (Claus, 1983, Czarnetzki, 1966). Individual XLVII 11 analyzed in
211 this study is a male (Supplementary Figure S2) and dates to 3335-3107 cal BC (MAMS
212 33641 – 4501 ± 19 BP).

213 The medieval cemetery on the **Petersberg/Kleiner Madron**, Bavaria, Germany, lies on a hill
214 top at 850 meters asl and 400 meters above the floor of the Inn Valley. On the eastern part of
215 the cemetery, which is under discussion here, members of a priory were buried that was most
216 likely established in the late 10th century. Written sources document its existence from 1132
217 onwards (Meier, 1998). During systematic excavations (1997-2004) in the southeastern part
218 of the churchyard 99 graves with a higher, but hardly determinable number of individuals
219 were uncovered. The examined individual in grave 820 is a male with an age at death of
220 around 65-70 years (Lösch, 2009 – Supplementary Figure S3) dating to 1020-1116 cal AD
221 (MAMS 33642 – 982 ± 17 BP).

222

223 ***DNA extraction and sequencing***

224 The DNA extractions and pre-PCR steps were carried out in clean room facilities dedicated to
225 aDNA research. Teeth were used for the analyses. The samples from Petersberg and Sorsum
226 were processed in the Ancient DNA Laboratory at Kiel University and the sample from
227 Karsdorf in the Ancient DNA Laboratory of the Max Planck Institute for the Science of
228 Human History (MPI SHH) in Jena. All procedures followed the guidelines on contamination
229 control in aDNA studies (Warinner et al., 2017, Key et al., 2017). The teeth were cleaned in
230 pure bleach solution to remove potential contaminations prior to powdering. Fifty milligrams

231 of powder were used for extraction following a silica-based protocol (Dabney et al., 2013).
232 Negative controls were included in all steps.
233 From each sample, double-stranded DNA sequencing libraries (UDGhalf) were prepared
234 according to an established protocol for multiplex high-throughput sequencing (Meyer and
235 Kircher, 2010). Sample-specific indices were added to both library adapters via amplification
236 with two index primers. Extraction and library blanks were treated in the same manner. For
237 the initial screening, the library of the individual from Karsdorf was sequenced on 1/50 of a
238 lane on the HiSeq 3000 (2x75 bp) at the MPI SHH in Jena and the libraries from Petersberg
239 and Sorsum were sequenced on the Illumina HiSeq 4000 (2x75 bp) platform at the Institute of
240 Clinical Molecular Biology, Kiel University, using the HiSeq v4 chemistry and the
241 manufacturer's protocol for multiplex sequencing. Deep-sequencing for each of the three
242 samples was carried out on 2 lanes on the Illumina HiSeq 4000 platform at the Institute of
243 Clinical Molecular Biology, Kiel University.
244
245 ***Metagenomics data processing, screening, and analyses***
246 The datasets for the three ancient samples comprised paired-end reads. The adapter sequences
247 were removed and overlapping paired-end reads were merged with ClipAndMerge which is a
248 module of the EAGER pipeline (Peltzer et al., 2016). The metagenomic viral screening was
249 carried out using MALT (Vagene et al., 2018) and the NCBI viral RefSeq database. All three
250 samples showed HBV-specific reads. In order to obtain all HBV related sequencing reads we
251 mapped against a multi-fasta reference containing one representative of each genotype (A-H)
252 and eight ape strains using BWA (Li and Durbin, 2010) (Supplementary Table S6). Mapped
253 reads were extracted from the BAM file, converted to FASTQ and a *de novo* assembly using
254 SPAdes (Bankevich et al., 2012) was carried out. Resulting contigs for each K-value where
255 checked and the k-value that spawned the longest contigs was selected as criteria for further
256 analysis. The contigs were re-mapped with BWA against the multi-fasta reference. The
257 resulting alignment was visually inspected in IGV v 2.3.92 (Thorvaldsdóttir et al., 2013) to
258 archive information about contig order and direction. Based on that information, a consensus
259 sequence was constructed from the contigs.
260 We assembled a comprehensive set of reference genomes using 5497 non-recombinant
261 genomes available at hpvdb (https://hbvdb.ibcp.fr/HBVdb/HBVdbDataset?seqtype=0) and a
262 previously defined set of 74 ape-infecting HBV genomes. In order to reduce the actual
263 number of genomes used for subsequent inferences but retain the full range of known HBV
264 diversity, we clustered all sequences using UClust v 1.1.579 (Edgar et al., 2010). We

265 extracted the centroid sequences based on a sequence identity of at least 97%, which resulted
266 in 493 representative genomes. Those genomes together with all available ancient genomes
267 were aligned using Geneious version 10.1.2 (Kearse et al., 2012) with a 65% similarity cost
268 matrix, a gap open penalty of 12 and a gap extension penalty of 3. The multiple sequence
269 alignment was stripped of any sites (columns) that had gaps in more than 95% of sequences.
270 The complete alignment including all modern and ancient genomes is available as multi-fasta
271 in source data 2. The alignment was used to construct a network with the software SplitsTree
272 v4 (Huson and Bryant, 2006), creating a NeighborNet (Bryant and Moulton, 2004) with
273 uncorrected P distances.
274

275 ***Recombination analysis***

276 We performed recombination analysis using all modern full reference genomes (n=493) and
277 five ancient genomes used for the network analysis (see above). The methods RDP,
278 GENECOV, Chimera, MaxChi, BootScan, SiScan, 3Seq within RDP v4 (Martin et al., 2015)
279 with a window size of 100 nt and the parameter set to circular genome with and without
280 outgroup reference (results are provided in source data 1) and SimPlot v 3.5.1 (Lole et al.
281 1999, Supplementary Figure S14-S17) were applied to the data set.
282

283 ***Human population genetic analyses***

284 Mapping of the adapter-clipped and merged FASTQ files to the human reference genome
285 hg19 was done using BWA (Li and Durbin, 2010) using a reduced mapping stringency of "-n
286 0.01" and the mapping quality parameter "q 30". The mapped sequencing data was
287 transformed into the *Eigenstrat* format (Price et al., 2006) and merged with a dataset of
288 1.233.013 SNPs (Haak et al., 2015, Mathieson et al., 2015). Using the software Smartpca
289 (Patterson et al., 2006) the three samples and previously published ancient populations were
290 projected onto a base map of genetic variation calculated from 32 West Eurasian populations
291 (Supplementaryt Figure S11-S13).
292

293 ***Sex determination***

294 Sex determination was assessed based on the ratio of sequences aligning to the X and Y
295 chromosomes compared to the autosomes (Skoglund et al., 2013).
296

297 ***LC-MS based bottom-up proteomics***

298 Proteins were extracted from powdered tooth samples (50 mg) using a modified filter-aided
299 sample preparation (FASP) protocol as previously described (Cappellini et al., 2013,
300 Warinner et al., 2014). Samples were digested using trypsin and analyzed by LC-MS/MS.
301 Protein identification was performed using the SequestHT (Thermo Scientific) search engine
302 in a combined database comprising the full Swiss protein database (468,716 entries), a
303 hepatitis B data base (7 entries) and a common contaminant list. Further details regarding the
304 LC-MS/MS analysis and database search parameters are given in the supplementary
305 information and Supplementary Figure S19.
306

307 **Acknowledgements**

308 We are grateful to the following people and institutions for providing samples, support, and
309 advice: Bodo Krause-Kyora, Hildegard Nelson (Referat A1 Archäologische Dokumentation,
310 Niedersächsisches Landesamt für Denkmalpflege), Ulrike Weller (Sammlungsverwaltung
311 Archäologie Landesmuseum Hannover) and Britta Steer for technical assistance with
312 proteomics sample preparation.

313 This work was supported by the Collaborative Research Centre 1266 *Scales of*
314 *Transformation*, the Excellence Cluster 306 *Inflammation at Interfaces*, the Medical Faculty
315 of Kiel University, the Max Planck Society and the European Research Council (ERC)
316 starting grant APGREID (to J.K.). Excavations and analysis of the archaeological site of
317 Karsdorf were supported by the German Research Foundation (DFG) Grant of Kurt W. Alt
318 (Al 287-7-1) and Harald Meller (Me 3245/1-1).

319

320 **Declaration of interests**

321 All other authors declare that they have no conflicts of interest.

322

323 **Accession numbers**

324 Raw sequence read files have been deposited at the European Nucleotide Archive under
325 accession no. PRJEB24921

326

327 **References**
328

329 Araujo NM. 2015. Hepatitis B virus intergenotypic recombinants worldwide: An overview. Infect
330 Genet Evol. 36:500-510. DOI: 10.1016/j.meegid.2015.08.024. PMID: 26299884
331
332 Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI,
333 Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.
334 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J
335 Comput Biol.19:455-477. doi: 10.1089/cmb.2012.0021. PMID: 22506599
336
337 Behnke HJ. 2007. Süßes aus der Bronzezeit – Siedlungsgunst im Unstruttal bei Karsdorf,
338 Burgenlandkreis. Die Ausgrabungen 2004. Archäologie in Sachsen-Anhalt N.F. 4. 63-86.
339
340 Behnke HJ. 2011. Erste Siedler der Linienbandkeramik in der Karsdorfer Feldflur. Ergebnisse der
341 Ausgrabungen im Jahr 2005. Archäologie in Sachsen-Anhalt N.F. 5. 184-199.
342
343 Behnke HJ. 2012. Siedlungsgunst im Unstruttal bei Karsdorf, Burgenlandkreis: Ergebnisse der
344 Grabungen 2006 und 2007. Archäologie in Sachsen-Anhalt N.F.6. 35-70.
345
346 Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkosová Z, Powell A,
347 Burger J. 2013. 2000 years of parallel societies in Stone Age Central Europe. Science 342:479-81.
348 DOI: 10.1126/science.1245049. PMID 24114781
349
350 Bouckaert R, Alvarado-Mora MV, Pinho JR. 2013. Evolutionary rates and HBV: issues of rate
351 estimation with Bayesian molecular methods. Antiviral Therapy 18:497-503 DOI: 10.3851/IMP2656.
352 PMID: 23792904
353
354 Brandt G, Knipper C, Nicklisch N, Ganslmeier R, Klamm M, Alt KW. 2014. Settlement burials at the
355 Karsdorf LBK site, Saxony-Anhalt, Germany - biological ties and residential mobility. In: A. Whittle,
356 Penny Bickle (eds) Early farmers. The view from archaeology und Science. Oxford University Press,
357 British Academy, 95-114.
358
359 Bryant D, Moulton V. 2004. Neighbor-net: an agglomerative method for the construction of
360 phylogenetic networks. Mol Biol Evol. 2004 Feb;21(2):255-65. DOI:
361 https://doi.org/10.1093/molbev/msh018. PMID: 14660700
362
363 Cappellini E, Gentry A, Palkopoulou E, Ishida Y, Cram D, Roos AM, Watson M, Johansson US,
364 Fernholm B, Agnelli P, Barbagli F, Littlewood DTJ, Kelstrup ChD, Olsen JV, Lister AM, Roca AL,
365 Dalén L, Gilbert MT. 2013. Resolution of the type material of the Asian elephant, Elephas maximus
366 Linnaeus, 1758 (Proboscidea, Elephantidae). Zoological Journal of the Linnean Society. 170:222–232.
367 DOI: 10.1111/zoj.12084
368
369 Castelhano N, Araujo NM, Arenas M. 2017. Heterogeneous recombination among Hepatitis B virus
370 genotypes. Infection, Genetics and Evolution 54:486-490. DOI: 10.1016/j.meegid.2017.08.015,
371 PMID: 28827173
372
373 Czarnetzki A. 1966. Die menschlichen Skelettreste aus vier neolithischen Steinkisten Hessens und
374 Niedersachsens.
375
376 Claus M. 1983. Das neolithische Felsenkammergrab auf dem Halsberg bei Sorsum, Stadt Hildesheim.
377 Die Kunde N. F. 34/35, 1983/84, 91-122.
378
379 Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, Valdiosera C, García N, Pääbo
380 S, Arsuaga JL, Meyer M. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene

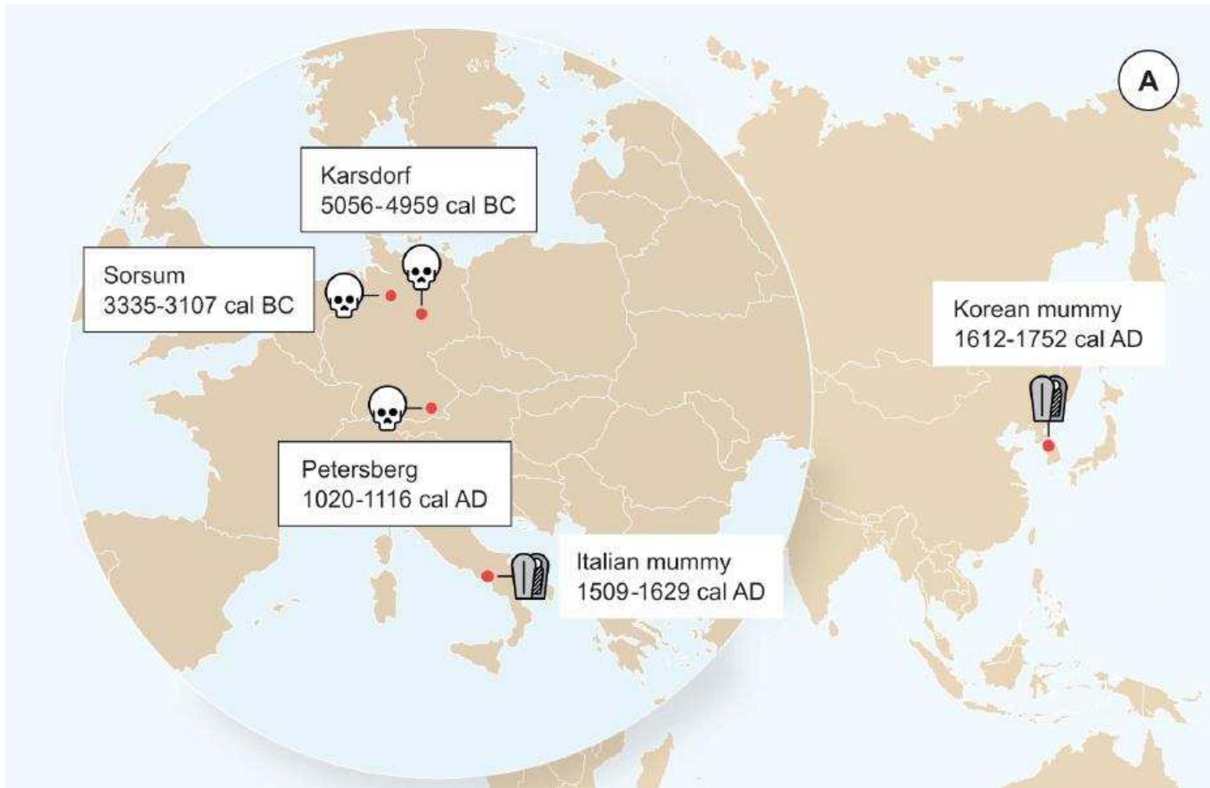
381 cave bear reconstructedfrom ultrashort DNA fragments. Proc Natl Acad Sci USA.110:15758-63. DOI:
382 10.1073/pnas.1314445110. PMID: 24019490
383
384 Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics.
385 26:2460-1. DOI: 10.1093/bioinformatics/btq461. PMID: 20709691
386
387 Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney
388 E, Stewardson K, Fu Q, Mittnik A, Bánffy E, Economou C, Francken M, Friederich S, Pena RG,
389 Hallgren F, Khartanovich V, Khokhlov A, Kunst M, Kuznetsov P, Meller H, Mochalov O, Moiseyev
390 V, Nicklisch N, Pichler SL, Risch R, Rojo Guerra MA, Roth C, Szécsényi-Nagy A, Wahl J, Meyer M,
391 Krause J, Brown D, Anthony D, Cooper A, Alt KW, Reich D. 2015. Massive migration from the
392 steppe was a source for Indo-European languages in Europe. Nature. 522:207-211. DOI:
393 10.1038/nature14317. PMID: 25731166
394
395 Huson DH, Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies.
396 Molecular Biology and Evolution. 23:254-267. PMID: 16221896
397
398 Kahila Bar-Gal G, Kim MJ, Klein A, Shin DH, Oh CS, Kim JW, Kim TH, Kim SB, Grant PR, Pappo
399 O, Spigelman M, Shouval D. 2012. Tracing hepatitis B virus to the 16th century in a Korean mummy.
400 Hepatology 56:1671-80. DOI: 10.1002/hep.25852. PMID: 22610996
401
402 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,
403 Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A. 2012. Geneious Basic: an
404 integrated and extendable desktop software platform for the organization and analysis of sequence
405 data. Bioinformatics. 28,1647-1649. DOI: 10.1093/bioinformatics/bts199. PMID: 22543367
406
407 Key FM, Posth C, Krause J, Herbig A, Bos KI. 2017. Mining Metagenomic Data Sets for Ancient
408 DNA: Recommended Protocols for Authentication. Trends Genet. 33:508-520. DOI:
409 10.1016/j.tig.2017.05.005. Epub 2017 Jul 5. PMID: 28688671
410
411 Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG,
412 Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI, Nordenfelt S, Li H, de
413 Filippo C, Prüfer K, Sawyer S, Posth C, Haak W, Hallgren F, Fornander E, Rohland N, Delsate D,
414 Francken M, Guinet JM, Wahl J, Ayodo G, Babiker HA, Bailliet G, Balanovska E, Balanovsky O,
415 Barrantes R, Bedoya G, Ben-Ami H, Bene J, Berrada F, Bravi CM, Brisighelli F, Busby GB, Cali F,
416 Churnosov M, Cole DE, Corach D, Damba L, van Driem G, Dryomov S, Dugoujon JM, Fedorova SA,
417 Gallego Romero I, Gubina M, Hammer M, Henn BM, Hervig T, Hodoglugil U, Jha AR, Karachanak-
418 Yankova S, Khusainova R, Khusnutdinova E, Kittles R, Kivisild T, Klitz W, Kučinskas V,
419 Kushniarevich A, Laredj L, Litvinov S, Loukidis T, Mahley RW, Melegh B, Metspalu E, Molina J,
420 Mountain J, Näkkäläjärvi K, Nesheva D, Nyambo T, Osipova L, Parik J, Platonov F, Posukh O,
421 Romano V, Rothhammer F, Rudan I, Ruizbakiev R, Sahakyan H, Sajantila A, Salas A, Starikovskaya
422 EB, Tarekegn A, Toncheva D, Turdikulova S, Uktveryte I, Utevska O, Vasquez R, Villena M,
423 Voevoda M, Winkler CA, Yepiskoposyan L, Zalloua P, Zemunik T, Cooper A, Capelli C, Thomas
424 MG, Ruiz-Linares A, Tishkoff SA, Singh L, Thangaraj K, Villems R, Comas D, Sukernik R, Metspalu
425 M, Meyer M, Eichler EE, Burger J, Slatkin M, Pääbo S, Kelso J, Reich D, Krause J. 2014. Ancient
426 human genomes suggest three ancestral populations for present-day Europeans. Nature 513:409-13.
427 DOI: 10.1038/nature13673. PMID 25230663
428
429 Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform.
430 Bioinformatics. 26:589-95. DOI: 10.1093/bioinformatics/btp698. PMID: 20080505
431
432 Littlejohn M, Locarnini S, Yuen L. 2016. Origins and Evolution of Hepatitis B Virus and Hepatitis D
433 Virus. Cold Spring Harb Perspect Med. 6:a021360. DOI: 10.1101/cshperspect.a021360, PMID:
434 26729756
435

436 Lösch S. 2009. Paläopathologisch-anthropologische und molekulare Untersuchungen an
437 mittelalterlichen und frühneuzeitlichen Bevölkerungsgruppen. Dissertation zum Erwerb des
438 Doktorgrades der Humanbiologie an der Medizinischen Fakultät der Ludwig-Maximilians-Universität
439 zu München.
440
441 Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard
442 HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-
443 infected seroconverters in India, with evidence of intersubtype recombination. J Virol. 73:152-60.
444 PMID: 9847317
445
446 Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: Detection and analysis of
447 recombination patterns in virus genomes. Virus Evolution 1:vev003. DOI: 10.1093/ve/vev00. PMID:
448 27774277
449
450 Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson
451 K, Fernandes D, Novak M, Sirak K, Gamba C, Jones ER, Llamas B, Dryomov S, Pickrell J, Arsuaga
452 JL, de Castro JM, Carbonell E, Gerritsen F, Khokhlov A, Kuznetsov P, Lozano M, Meller H,
453 Mochalov O, Moiseyev V, Guerra MA, Roodenberg J, Vergès JM, Krause J, Cooper A, Alt KW,
454 Brown D, Anthony D, Lalueza-Fox C, Haak W, Pinhasi R, Reich D. 2015 Genome-wide patterns of
455 selection in 230 ancient Eurasians. Nature. 528:499-503. DOI: 10.1038/nature16152. PMID:
456 26595274
457
458 Meier T. 1998. Ein mittelalterliches Kloster mit Friedhof auf dem Kleinen Madron/Petersberg bei
459 Flintsbach a. Inn. Landkreis Rosenheim, Oberbayern. Das Archäologische Jahr in Bayern 1998: 127-
460 129.
461
462 Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target
463 capture and sequencing. Cold Spring Harb. Protoc. 2010:pdb.prot5448. DOI: 10.1101/pdb.prot5448.
464 PMID: 20516186
465
466 Nicklisch N. 2017. Spurensuche an Skeletten. Paläodemografische und epidemiologische
467 Untersuchungen an neolithischen und frühbronzezeitlichen Bestattungen aus dem Mittelelbe-Saale-
468 Gebiet im Kontext populationsdynamischer Prozesse. Forschungsberichte des Landesmuseums für
469 Vorgeschichte Halle, Band 11.
470
471 Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis.
472 PLoS Genet. 2:e190. PMID: 17194218
473
474 Patterson Ross Z, Klunk J, Fornaciari G, Giuffra V, Duchêne S, Duggan AT, Poinar D, Douglas MW,
475 Eden JS, Holmes EC, Poinar HN. 2018. The paradox of HBV evolution as revealed from a 16th
476 century mummy. PLoS Pathog. 4;14:e1006750. DOI: 10.1371/journal.ppat.1006750. PMID:
477 29300782
478
479 Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, Nieselt K. 2016. EAGER: efficient ancient
480 genome reconstruction. Genome Biol. 17:60. DOI: 10.1186/s13059-016-0918-z. PMID: 27036623
481
482 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal
483 components analysis corrects for stratification in genome-wide association studies. Nat Genet. 38:904-
484 909. PMID: 16862161
485
486 Rambaut A, T. T. Lam, L. Max Carvalho, and O. G. Pybus, "Exploring the temporal structure of
487 heterochronous sequences using TempEst (formerly Path-O-Gen)," Virus Evol, vol. 2, no. 1, p.
488 vew007, Jan. 2016.
489
490

491 Rasche A, Souza BFCD, Drexler JF. 2016. Bat hepadnaviruses and the origins of primate hepatitis B
492 viruses. Curr Opin Virol. 16:86-94. DOI: 10.1016/j.coviro.2016.01.015. PMID: 26897577
493
494 Schaefer S. 2007. Hepatitis B virus taxonomy and hepatitis B virus genotypes. World J Gastroenterol.
495 13:14-21. PMID: 17206751
496
497 Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jäger G, Bos KI, Herbig A, Economou C,
498 Benjak A, Busso P, Nebel A, Boldsen JL, Kjellström A, Wu H, Stewart GR, Taylor GM, Bauer P, Lee
499 OY, Wu HH, Minnikin DE, Besra GS, Tucker K, Roffey S, Sow SO, Cole ST, Nieselt K, Krause J.
500 2013 Genome-wide comparison of medieval and modern Mycobacterium leprae. Science. 341:179-
501 183. DOI: 10.1126/science.1238286. PMID: 23765279
502
503 Simmonds P, Midgley S. 2005. Recombination in the genesis and evolution of hepatitis B virus
504 genotypes. J Virol. 79:15467-76. DOI: 10.1128/JVI.79.24.15467-15476.2005, PMID: 16306618
505
506 Skoglunda P, Storå J, Götherström A, Jakobsson M. 2013. Accurate sex identification of ancient
507 human remains using DNA shotgun sequencing. Journal of Archaeological Science. 12:4477-4482.
508 DOI: https://doi.org/10.1016/j.jas.2013.07.004
509
510 Souza BF, Drexler JF, Lima RS, Rosário Mde O, Netto EM. 2014. Theories about evolutionary origins
511 of human hepatitis B virus in primates and humans. Braz J Infect Dis. 18:535-43. DOI:
512 10.1016/j.bjid.2013.12.006, PMID: 24726560
513
514 Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-
515 performance genomics data visualization and exploration. Brief Bioinform. 14:178-192. DOI:
516 10.1093/bib/bbs017. PMID: 22517427
517
518 Vågene ÅJ, Herbig A, Campana MG, Robles García NM, Warinner C, Sabin S, Spyrou MA,
519 Andrades Valtueña A, Huson D, Tuross N, Bos KI, Krause J. 2018. Salmonella enterica genomes from
520 victims of a major sixteenth-century epidemic in Mexico. Nat Ecol Evol.2:520-528. DOI:
521 10.1038/s41559-017-0446-6. PMID: 29335577
522
523 Veit U. 1996. Studien zum Problem der Siedlungsbestattungen im europäischen Neolithikum,
524 Münster, New York: Waxmann
525
526 Warinner C, Rodrigues JF, Vyas R, Trachsel C, Shved N, Grossmann J, Radini A, Hancock Y, Tito
527 RY, Fiddyment S, Speller C, Hendy J, Charlton S, Luder HU, Salazar-García DC, Eppler E, Seiler R,
528 Hansen LH, Castruita JA, Barkow-Oesterreicher S, Teoh KY, Kelstrup CD, Olsen JV, Nanni P, Kawai
529 T, Willerslev E, von Mering C, Lewis CM Jr, Collins MJ, Gilbert MT, Rühli F, Cappellini E. 2014.
530 Pathogens and host immunity in the ancient human oral cavity. Nat Genet. 46:336-44. DOI:
531 10.1038/ng.2906. PMID: 24562188
532
533 Warinner C, Herbig A, Mann A, Fellows Yates JA, Weiß CL, Burbano HA, Orlando L, Krause J.
534 2017. A Robust Framework for Microbial Archaeology. Annu Rev Genomics Hum Genet. 18:321-
535 356. DOI: 10.1146/annurev-genom-091416-035526. PMID: 28460196
536
537 WHO, Global hepatitis report, 2017
538
539 Zhou Y, Holmes EC. 2007. Bayesian estimations of the evolutionary rate and age of hepatitis B virus.
540 J Mol Evol 65:197-205. DOI:10.1007/s00239-007-0054-1, PMID: 17684696
541
542

543 **Figure Legends**

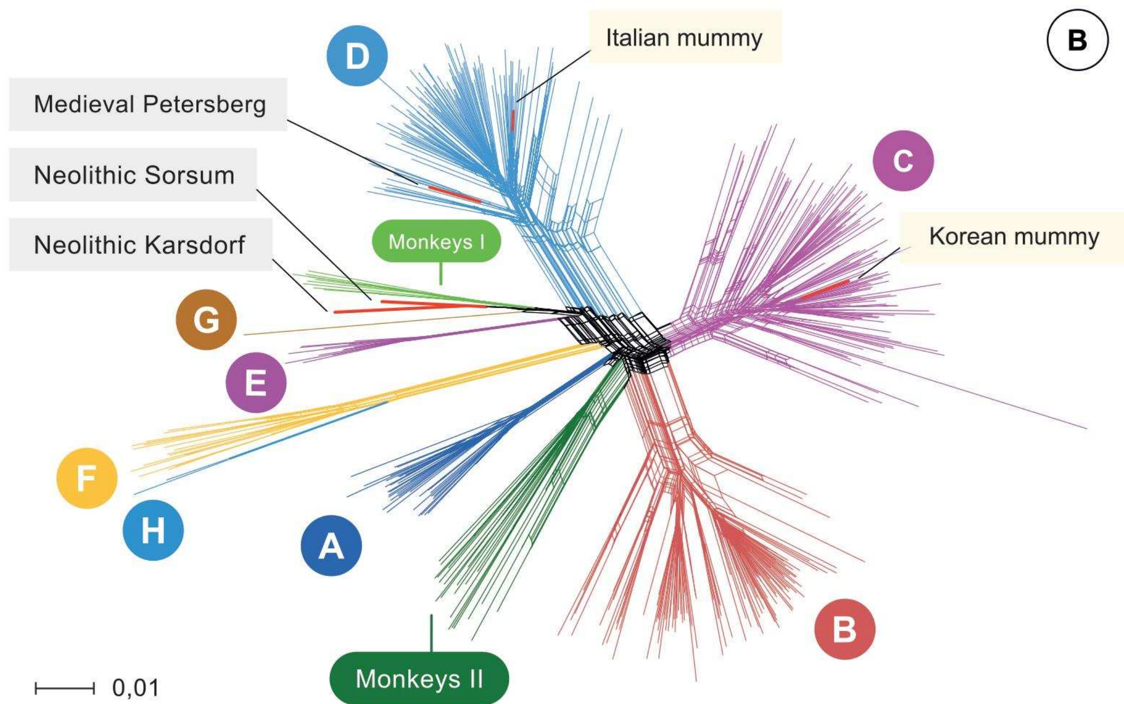
544



545

546 **Fig A. Origin of samples**

547 Geographic location of the samples from which ancient HBV genomes where isolated.
548 Radiocarbon dates of the specimens is given in 2 sigma range. Icon indicate the sample
549 material (tooth or mummy). HBV genomes obtained in this study indicated by black frame.



550

551 **Fig B Network**

552 Network of 495 modern, two published ancient genomes (light yellow box), and three ancient
553 hepatitis B virus (HBV) obtained in this study (grey box). Colors indicating the 8 human
554 HBV Genotypes (A-H), two monkey genotypes (Monkeys I, African apes and Monkeys II,
555 Asian monkeys) and ancient genomes (red).

556

557

558 **Tables**

559

560 **Table 1.** Results of the genome reconstruction.

561

	# Merged reads	Length of HBV consensus sequence	mean HBV coverage	Gaps in the consensus sequence at nt position	# mapped reads HBV	# mapped reads human	mean human coverage	human genomes/ HBVgenomes
Karsdorf	386,780,892	3183	104X	2157-2175; 3107-3128; 3133-3183	10,718	122,568,310	2.96X	1 : 35.1
Sorsum	367,574,767	3182	47X	**-**	3,249	9,856,001	1.17X	1 : 40.2
Petersberg	419,413,082	3161	46X	880-1000; 1232-1329; 1331-1415; 1420-1581; 1585-1598	2,125	105,476,677	2.88X	1 : 16

562 # - number

563 nt - nucleotide