



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/206703/>

Version: Accepted Version

---

**Article:**

Fiddaman, Steven, Dimopoulos, Evangelos, Lebrasseur, Ophélie et al. (2023) Ancient chicken remains reveal the origins of virulence in Marek's 2 disease virus. *Science*. pp. 1276-1281. ISSN: 0036-8075

<https://doi.org/10.1126/science.adg2238>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

1  
2 **Title: Ancient chicken remains reveal the origins of virulence in Marek's**  
3 **disease virus**

4  
5 **Authors:**

6 Steven R Fiddaman<sup>1†\*</sup>, Evangelos A Dimopoulos<sup>2,3†</sup>, Ophélie Lebrasseur<sup>4,5</sup>, Louis du Plessis<sup>6,7</sup>, Bram  
7 Vrancken<sup>8,9</sup>, Sophy Charlton<sup>2,10</sup>, Ashleigh F Haruda<sup>2</sup>, Kristina Tabbada<sup>2</sup>, Patrik G Flammer<sup>1</sup>, Stefan  
8 Dascalu<sup>1</sup>, Nemanja Marković<sup>11</sup>, Hannah Li<sup>12</sup>, Gabrielle Franklin<sup>13</sup>, Robert Symmons<sup>14</sup>, Henriette Baron<sup>15</sup>,  
9 László Daróczi-Szabó<sup>16</sup>, Dilyara N Shaymuratova<sup>17</sup>, Igor V Askeyev<sup>17</sup>, Olivier Putelat<sup>18</sup>, Maria Sana<sup>19</sup>,  
10 Hossein Davoudi<sup>20</sup>, Homa Fathi<sup>20</sup>, Amir Saed Mucheshi<sup>21</sup>, Ali Akbar Vahdati<sup>22</sup>, Liangren Zhang<sup>23</sup>, Alison  
11 Foster<sup>24</sup>, Naomi Sykes<sup>25</sup>, Gabrielle Cass Baumberg<sup>2</sup>, Jelena Bulatović<sup>26</sup>, Arthur O Askeyev<sup>17</sup>, Oleg V  
12 Askeyev<sup>17</sup>, Marjan Mashkour<sup>20,27</sup>, Oliver G Pybus<sup>1,28</sup>, Venugopal Nair<sup>1,29</sup>, Greger Larson<sup>2‡</sup>, Adrian L  
13 Smith<sup>1\*‡</sup>, Laurent AF Frantz<sup>30,31\*‡</sup>

14 **Affiliations:**

15 <sup>1</sup>Department of Biology, University of Oxford, Oxford, UK

16 <sup>2</sup>The Palaeogenomics & Bio-Archaeology Research Network, Research Laboratory for Archaeology and  
17 History of Art, University of Oxford, Oxford, UK

18 <sup>3</sup>Department of Veterinary Medicine, University of Cambridge, Cambridge, UK

19 <sup>4</sup>Centre d'Anthropobiologie et de Génomique de Toulouse, Toulouse, France

20 <sup>5</sup>Instituto Nacional de Antropología y Pensamiento Latinoamericano, Ciudad Autónoma de Buenos Aires,  
21 Buenos Aires, Argentina

22 <sup>6</sup>Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

23 <sup>7</sup>Swiss Institute of Bioinformatics, Lausanne, Switzerland

24 <sup>8</sup>Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven,  
25 Belgium

26 <sup>9</sup>Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium

27 <sup>10</sup>BioArCh, Department of Archaeology, University of York, York, UK

28 <sup>11</sup>Institute of Archaeology, Belgrade, Serbia

- 29 <sup>12</sup>Institute of Immunity and Transplantation, University College London, London, UK
- 30 <sup>13</sup>Silkie Club of Great Britain, Charing, UK
- 31 <sup>14</sup>Fishbourne Roman Palace, Fishbourne, UK
- 32 <sup>15</sup>Leibniz-Zentrum für Archäologie, Mainz, Germany
- 33 <sup>16</sup>Medieval Department, Budapest History Museum, Budapest, Hungary
- 34 <sup>17</sup>Laboratory of Biomonitoring, The Institute of Problems in Ecology and Mineral Wealth, Tatarstan  
35 Academy of Sciences, Kazan, Russia
- 36 <sup>18</sup>Archéologie Alsace - PAIR, Bas-Rhin, France
- 37 <sup>19</sup>Departament de Prehistòria, Universitat Autònoma de Barcelona, Barcelona, Spain
- 38 <sup>20</sup>Bioarchaeology Laboratory, Central Laboratory, University of Tehran, Tehran, Iran
- 39 <sup>21</sup>Department of Art and Architecture, Payame Noor University (PNU), Tehran, Iran
- 40 <sup>22</sup>Provincial Office of the Iranian Center for Cultural Heritage, Handicrafts and Tourism Organisation,  
41 Bojnord, Iran
- 42 <sup>23</sup>Department of Archaeology, School of History, Nanjing University, China
- 43 <sup>24</sup>Headland Archaeology, Edinburgh, UK
- 44 <sup>25</sup>Department of Archaeology, University of Exeter, Exeter, UK
- 45 <sup>26</sup>Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden
- 46 <sup>27</sup>CNRS, National Museum Natural History Paris, Paris, France
- 47 <sup>28</sup>Department of Pathobiology and Population Sciences, Royal Veterinary College, London, UK
- 48 <sup>29</sup>Viral Oncogenesis Group, Pirbright Institute, Woking, UK
- 49 <sup>30</sup>Department of Veterinary Sciences, Ludwig Maximilian University of Munich, Munich, Germany
- 50 <sup>31</sup>School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
- 51 †joint-first author
- 52 ‡ co-senior authors

53 \*corresponding authors. Emails: [steven.fiddaman@biology.ox.ac.uk](mailto:steven.fiddaman@biology.ox.ac.uk); [adrian.smith@biology.ox.ac.uk](mailto:adrian.smith@biology.ox.ac.uk);  
54 [laurent.frantz@lmu.de](mailto:laurent.frantz@lmu.de)

55

56

57

58 **Abstract:**

59 The dramatic growth in livestock populations since the 1950s has altered the epidemiological and  
60 evolutionary trajectory of their associated pathogens. For example, Marek's disease virus (MDV), which  
61 causes lymphoid tumors in chickens, has experienced a marked increase in virulence over the last century.  
62 Today, MDV infections kill >90% of unvaccinated birds and controlling it costs >US\$1bn annually. By  
63 sequencing MDV genomes derived from archeological chickens, we demonstrate that it has been  
64 circulating for at least 1000 years. We functionally tested the *Meq* oncogene, one of 49 viral genes  
65 positively selected in modern strains, demonstrating that ancient MDV was likely incapable of driving  
66 tumor formation. Our results demonstrate the power of ancient DNA approaches to trace the molecular  
67 basis of virulence in economically relevant pathogens.

68

69 **One sentence summary:**

70 Functional paleogenomics reveals the molecular basis for increased virulence in Marek's Disease Virus.

71

72 **Main Text:**

73 Marek's Disease Virus (MDV) is a highly contagious alphaherpesvirus that causes a tumor-associated  
74 disease in poultry. At the time of its initial description in 1907, Marek's Disease (MD) was a relatively  
75 mild disease with low mortality, characterized by nerve pathology mainly affecting older individuals (1).  
76 However, over the course of the 20<sup>th</sup> century, MDV-related mortality has risen to >90% in unvaccinated  
77 chickens. To prevent this high mortality rate, the poultry industry spends more than US\$1 billion per year  
78 on health intervention measures, including vaccination (2).

79

80 The increase in virulence and clinical pathology of MDV infection has likely been driven by a  
81 combination of factors. Firstly, the growth in the global chicken population since the 1950s led to more  
82 viral replication, which increased the supply of novel mutations in the population. In addition, the use of  
83 imperfect (also known as 'leaky') vaccines that prevent symptomatic disease but do not prevent  
84 transmission of the virus likely shifted selective pressures and led to an accelerated rate of MDV  
85 virulence evolution (3). Combined, these factors have altered the evolutionary trajectory, resulting in  
86 modern hyper-pathogenic strains. To date, the earliest sequenced MDV genomes were sampled in the  
87 1960s (4), several decades after the first reports of MDV causing tumors (5). As a result, the genetic  
88 changes that contributed to the increase in virulence of MDV infection prior to the 1960s remain  
89 unknown.

90

91 **Marek's disease virus has been circulating in Europe for at least 1000 years**

92 To empirically track the evolutionary change in MDV virulence through time, we generated MDV  
93 genome sequences (serotype 1) isolated from the skeletal remains of archeological chickens. We first  
94 shotgun sequenced 995 archeological chicken samples excavated from >140 Western Eurasian  
95 archeological sites and screened for MDV reads using HAYSTAC (6) with a herpesvirus-specific  
96 database. Samples with any evidence of MDV reads were then enriched for viral DNA using a

97 hybridisation-based capture approach based on RNA baits designed to tile the entire MDV genome  
98 (excluding one copy of each of the terminal repeats and regions of low complexity). To validate the  
99 approach, we also captured and sequenced DNA from the feather of a modern Silkie chicken that  
100 presented MDV symptoms. As a negative control, we also included an ancient sample that displayed no  
101 evidence of MDV reads following screening (OL1214; Serbia, C14<sup>th</sup>-15<sup>th</sup>).

102  
103 Using the capture protocol we identified 15 ancient chickens with MDV-specific reads of  $\geq 25$ bp in  
104 length. This approach also yielded a  $\sim 4\times$  genome from a modern positive control. We found that the  
105 majority of uniquely mapped reads (i.e. 88-99%) generated from ancient samples classified as MDV-  
106 positive were  $\geq 25$ bp, while the majority (i.e. 53-100%) of uniquely mapped reads generated from samples  
107 considered MDV-negative were shorter than 25bp. In addition, samples considered MDV-positive yielded  
108 between 308 and 133,885 uniquely mapped reads ( $\geq 25$ bp) while samples considered MDV-negative  
109 (including a negative control; Table S2) yielded between 0 and 211 uniquely mapped reads of  $\geq 25$ bp.  
110 MDV-positive ancient samples ranged in depth of coverage from 0.13 $\times$  to 41.92 $\times$  (OL1385; Fig. 1a,  
111 Table S2), with seven genomes at  $\geq 2\times$  coverage.

112  
113 In all positive samples, the proportion of duplicated reads approached 100%, indicating that virtually all  
114 of the unique molecules in each library were sequenced at least once (Fig. S1). Reads obtained from  
115 MDV-positive ancient samples were characterized by chemical signatures of DNA damage typically  
116 associated with ancient DNA (Fig. S2). In contrast, reads obtained from our modern positive control did  
117 not show any evidence of DNA damage (Fig. S2). The earliest unequivocally MDV-positive sample (with  
118 4,760 post-capture reads  $\geq 25$ bp) was derived from a 10<sup>th</sup>-12<sup>th</sup> century chicken from Eastern France  
119 (Andlau in Fig. 1a; Table S2). Together, these results demonstrate that MDV strains have been circulating  
120 in Western Eurasian poultry for at least 1,000 years.

### 121 122 **Ancient MDV strains are basal to modern lineages**

123 To investigate the relationship between ancient and modern MDV strains, we built phylogenetic trees  
124 based on both neighbor-joining (NJ) and maximum likelihood (ML) methods. We first built trees using  
125 10 ancient genomes with at least 1% coverage at a depth of  $\geq 5\times$ , a modern positive control derived from  
126 the present study (OL1099), and 42 modern genomes from public sources (Table S3). Both NJ (Fig. 1b,  
127 Fig. S3) and ML trees (Fig. S4) match the previously described general topology (7), in which Eurasian  
128 and North American lineages were evident, along with a well-supported (bootstrap: 94) ancient clade (Fig  
129 1b). The same topology was also obtained when restricting our ML analysis to include only transversion  
130 sites (Fig. S5). Lastly, we built a tree using an outgroup (Meleagrid herpesvirus 1, accession:  
131 NC\_002641.1) to root our topology (Fig. S6). We obtained a well-supported topology showing that the  
132 ancient MDV sequences form a highly supported clade lying basal to all modern MDV strains (including  
133 the modern positive control OL1099).

134  
135 Next, we built a time-calibrated phylogeny using BEAST (v. 1.10; (8)) that included 31 modern genomes  
136 collected since 1968 (Table S3), and four ancient samples with an average depth of coverage >5×  
137 (OL1986, Castillo de Montsoriu, Spain, 1593 cal. CE; OL1385, Buda Castle, Hungary, 1802 cal. CE;  
138 OL1389, an additional Buda Castle sample from the same archeological context as OL1385; OL2272,  
139 Naderi Tepe, Iran, 1820 cal. CE; Table S1-S2, Fig. 1a). All of the ancient samples were phylogenetically  
140 basal to all modern MDV strains. The time of the most recent common ancestor (TMRCA) of the  
141 phylogeny was 1483 CE (95% HPD interval 1349 - 1576; Fig. 1c, Table S4).

142  
143 As previously reported (7) we found that, aside from a few exceptions, most Eurasian and North  
144 American MDV strains formed distinct clades (Fig. 1b), suggesting that there has been little recent  
145 transatlantic exchange of the virus. The inclusion of time-stamped ancient MDV sequences improved the  
146 accuracy of the molecular clock analysis, and pushed back the TMRCA of all modern MDV sequences,  
147 from 1922-1952 (7) to 1825 (95% HPD interval 1751 - 1895; Table S4). Our mean TMRCA also pre-  
148 dates a recent estimate that incorporated 26 modern MDV genomes from East Asian chickens (1880, 95%  
149 HPD 1772-1968; (9)), although confidence intervals considerably overlap. This phylogenetic analysis  
150 implies that the two major modern clades of MDV were established long before the earliest documented  
151 increases in MDV virulence in the 1920s. Furthermore, since birds infected with highly virulent MDV  
152 would not have survived a transatlantic crossing, a TMRCA of 1915 (95% HPD 1879 - 1947) for the  
153 North American samples is consistent with the virus having been transmitted at an early point in the  
154 trajectory to increased virulence. These results are also consistent with the hypothesis that Eurasian and  
155 North American MDV lineages independently evolved towards increased virulence (7).

### 156 157 **Virulence factors are among positively selected genes in the modern MDV lineage**

158 The rapid increase in MDV virulence could potentially have been driven by gene loss or gain which  
159 would have substantially altered the biology of the virus (10, 11). Analysis of a Hungarian, high coverage,  
160 MDV genome (OL1385; >41x) from the 18<sup>th</sup> - 19<sup>th</sup> century indicated that it possessed the full complement  
161 of genes present in modern sequences. This indicates that there was no gene gain or loss in either ancient  
162 or modern lineage (Fig. 2). We also found that all MDV miRNAs, some of which are implicated in  
163 pathogenesis and oncogenesis in modern strains (12), were intact and highly conserved in ancient strains  
164 (Table S5). Together, these results indicate that the acquisition of virulence most likely resulted not from  
165 changes in MDV genome content or organization, but from point mutations.

166  
167 In fact, considering sites at which we had coverage for at least two ancient genomes, we identified 158  
168 fixed single nucleotide polymorphism (SNPs) between the ancient and modern samples, of which 31 were  
169 found in intergenic regions and may be candidates for future study of MDV regulatory regions (Table  
170 S6). To assess the impact of positive selection on point mutations we performed a branch-site analysis in  
171 PAML (13) (ancient sequences as background lineage, modern sequences as foreground lineage) on open

172 reading frames (ORFs) using four ancient MDV genomes (OL1385, OL1389, OL1986 and OL2272).  
173 After controlling the false discovery rate using the Benjamini-Hochberg procedure (14), this analysis  
174 identified 49 ORFs with significant evidence for positive selection (Fig. 2; Table S7).

175  
176 Several positively selected loci identified in this analysis have previously been associated with MDV  
177 virulence in modern strains. Some of these are known immune modulators or potential targets of a  
178 protective response. This includes ICP4, a large transcriptional regulatory protein involved in innate  
179 immune interference. Interestingly, ICP4 appears to be an important target of T cell-mediated immunity  
180 against MDV in chickens possessing the B21 Major Histocompatibility Complex (MHC) haplotype (15),  
181 and it is plausible that sequence variation in important ICP4 epitopes could confer differential  
182 susceptibility to infection.

183  
184 We also identified signatures of positive selection in several genes encoding viral glycoproteins (gC, gE,  
185 gI, gK and gL). Glycoproteins are important targets for the immune response to MDV (16). In fact, the  
186 majority of MDV peptides presented on chicken MHC class II are derived from just four proteins (17), of  
187 which two were glycoproteins found to be under selection in our analysis (gE and gI). This result  
188 indicates that glycoproteins are likely under selection in MDV because they are immune targets. The  
189 limited scope of immunologically important MDV peptides presented by MHC class II may have  
190 important implications for vaccine development.

191  
192 Positive selection was also detected in the viral chemokine termed viral interleukin-8 (considered a  
193 functional ortholog of chicken CXCL13; (18)). Viral IL-8 is an important virulence factor that  
194 recruits B cells for lytic replication and CD4+ CD25+ T cells that are transformed to generate lymphoid  
195 tumors. Viruses that lack vIL-8 are severely impaired in the establishment of infection and generation of  
196 tumors through bird-to-bird transmission (19), so sequence variation in this gene could plausibly impact  
197 transmission.

198  
199 **The key oncogene of MDV has experienced positive selection and an ordered loss of tetraproline**  
200 **motifs**

201 Our selection scan also identified *Meq*, a transcription factor considered to be the master regulator of  
202 tumor formation in MDV (20). In fact, the *Meq* coding sequence had the greatest average pairwise  
203 divergence between ancient and modern strains across the entirety of the MDV genome (Fig. 2), implying  
204 there were numerous sequence changes along the branch leading to modern samples. Animal experiments  
205 have demonstrated that *Meq* is essential for tumor formation (20) and polymorphisms in this gene, even in  
206 the absence of variants elsewhere in the genome, are known to confer significant differences in strain  
207 virulence or vaccine breakthrough ability (21).

208

209 *Meq* exerts transcriptional control on downstream gene targets (both in the host and viral genome) via its  
210 C-terminal transactivation domain. This domain is characterized by PPPP (tetraproline) repeats spaced  
211 throughout the second half of the protein, and the number of tetraproline repeats is inversely proportional  
212 to the virulence of the MDV strain (22). The difference in the number of tetraproline repeats in most  
213 strains is the result of point mutations rather than deletion or duplication; these strains are considered  
214 ‘standard length’-*Meq* (339 amino acids). In some strains, however, tetraproline repeats have been  
215 duplicated (‘long’-*Meq* strains, 399 amino acids) or deleted (‘short’-*Meq* strains, 298 amino acids, or  
216 ‘very short’-*Meq*, 247 amino acids). These mutations have led to varying numbers of tetraproline repeats  
217 between strains.

218  
219 We did not find any evidence of duplication or deletion in ancient *Meq* sequences, indicating that there  
220 are ‘standard length’-*Meq*. We then identified point mutations in a database containing four ancient *Meq*  
221 sequences (OL1385, OL1389, OL1986 and OL2272) along with 408 modern ‘standard length’-*Meq*  
222 sequences (Table S8). This analysis demonstrated that ancient *Meq* possessed six intact tetraproline  
223 motifs while all modern ‘standard length’-*Meq* sequences had between two and five. All ancient *Meq*  
224 sequences had a unique additional intact tetraproline motif at amino acids 290-293. This tetraproline  
225 motif was disrupted by a point mutation – causing a Proline to Histidine change – in the recent  
226 evolutionary history of ‘standard length’-*Meq* MDV strains.

227  
228 To further explore the virulence-related disruption of tetraprolines in modern *Meq* sequences, we  
229 constructed a phylogeny of *Meq* sequences (Fig. 3a). Mapping the tetraproline content of each sequence  
230 on the phylogeny indicated that tetraprolines have been lost in a specific order. Following the universal  
231 disruption of the 6<sup>th</sup> tetraproline through a point mutation (at amino acids 290-293) at the base of the  
232 modern MDV lineage, the 4<sup>th</sup> tetraproline was disrupted at the base of two major lineages (amino acids  
233 216-219). Disruption of the 4<sup>th</sup> tetraproline was followed in seven independent lineages by the disruption  
234 of the 2<sup>nd</sup> tetraproline (amino acids 175-178), and then by the loss of either the 1<sup>st</sup> (amino acids 152-155)  
235 or the 5<sup>th</sup> tetraproline (amino acids 232-235) in six lineages (Fig. 3a-b).

236  
237 Interestingly, our analysis indicated that the 2<sup>nd</sup> and 4<sup>th</sup> tetraprolines (codons 176 and 217) were under  
238 positive selection (Table S7). Although there were some observations of virus lineages exhibiting an  
239 alternative loss order (e.g. the occasional loss of the 3<sup>rd</sup> tetraproline (amino acids 191-194) following the  
240 loss of the 4<sup>th</sup>), such lineages are not widespread, suggesting that they may become stuck in local fitness  
241 peaks and are outcompeted by lineages following the order described above. The independent  
242 recapitulation of this pattern in different lineages suggests loss of tetraproline motifs acts as a ratchet,  
243 whereby each subsequent loss results in an increase in virulence, and once lost, motifs are unlikely to be  
244 regained.

245  
246 **Ancient *Meq* is a weak transactivator that likely did not drive tumor formation**

247 The initial description of MD in 1907 did not mention tumors (1). Given the degree of sequence  
248 differentiation observed between ancient and modern *Meq* genes, it is possible that ancient MDV  
249 genotypes were incapable of driving lymphoid cell transformation. To test this hypothesis experimentally,  
250 we assessed whether ancient *Meq* possessed lower transactivation capabilities, compared to modern  
251 strains, in a cultured cell-based assay.

252  
253 To do so, we synthesized an ancient *Meq* gene based on our highest coverage ancient sample (OL1385;  
254 Buda Castle, Hungary; 1802 cal. CE) and experimentally tested its transactivation function. We also  
255 cloned ‘very virulent’ modern pathotype strains (RB1B and Md5), which each differ from ancient *Meq* at  
256 13-14 amino acid positions (Fig. 3c; Table S9). All the *Meq* proteins were expressed in cells alongside a  
257 chicken protein (c-Jun), with which *Meq* forms a heterodimer, and a luciferase reporter containing the  
258 *Meq* binding (AP-1) sequence.

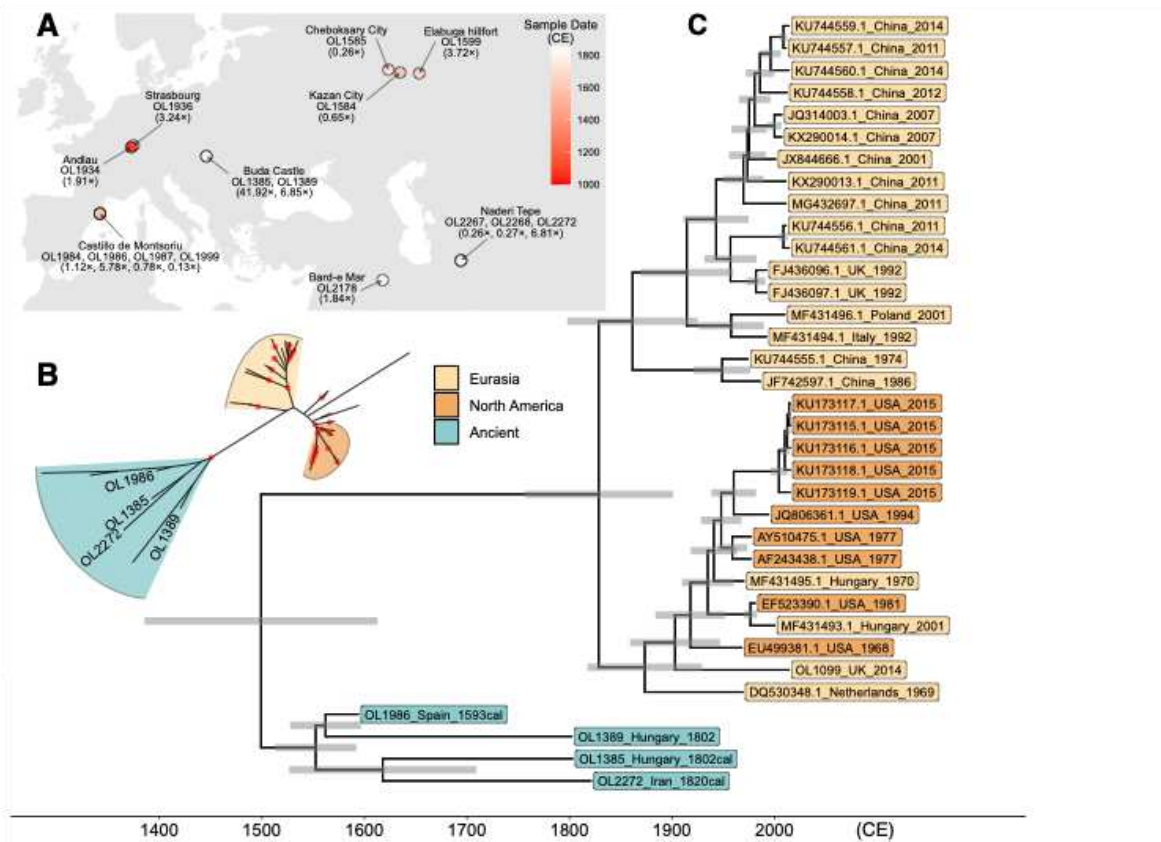
259  
260 Relative to the baseline signal, the transactivation of the ‘very virulent’ *Meq* strains RB1B and Md5 were  
261 7.5 and 10 times greater, respectively (Fig. 3d). Consistent with previous reports (23), removal of the  
262 partner protein, c-Jun, from RB1B resulted in severe abrogation of the transactivation capability (Fig. 3d).  
263 Ancient *Meq* exhibited a ~2.5-fold increase in transactivation relative to the baseline, but was  
264 substantially lower (3-4-fold) than *Meq* from the two ‘very virulent’ pathotypes (Fig. 3d). The ancient  
265 *Meq* was thus a demonstrably weaker transactivator than *Meq* from modern strains of MDV.

266  
267 Given that the transcriptional regulation of target genes (both host and virus) by *Meq* is directly related to  
268 oncogenicity (20, 23), it is likely that the weaker transactivation we demonstrate is associated with  
269 reduced or absent tumor formation. These data indicate that ancient MDV strains were unlikely to cause  
270 tumors, and were less pathogenic than modern strains. Ancient MDV likely established a chronic  
271 infection characterized by slower viral replication, low levels of viral shedding and low clinical  
272 pathology, which acted to facilitate maximal lifetime viral transmission in pre-industrialized, low-density  
273 settings.

## 274 275 **Conclusion**

276 Overall, our results demonstrate that Marek’s Disease Virus has been circulating in Western Eurasia for at  
277 least the last millennium. By reconstructing and functionally assessing ancient and modern genomes, we  
278 showed that ancient MDV strains were likely substantially less virulent than modern strains, and that the  
279 increase in virulence took place over the last century. Along with changes in several known virulence  
280 factors, we identified sequence changes in the *Meq* gene – the master regulator of oncogenesis – that  
281 drove its enhanced ability to transactivate its target genes and drive tumor formation. The historical  
282 perspective that our results provide can form the basis on which to rationally improve modern vaccines,  
283 and track or even predict future virulence changes. Lastly, our results highlight the utility of functional

284 paleogenomics to generate insights into the evolution and fundamental biological workings of pathogen  
285 virulence.



287

288 **Fig. 1. Locations of MDV-positive samples and time-scaled phylogeny.** (A) Map showing the

289 locations of screened archeological chicken samples that were positive for MDV sequence. Colored

290 circles indicate sample dates (either from calibrated radiocarbon dating or estimated from archeological

291 context; Table S1). Average sequencing depth following capture is given in parentheses under sample

292 names. If more than one sample was derived from the same site, this is indicated by a list of sample

293 identifiers (beginning 'OL') and sequencing depths in parentheses. (B) Unrooted neighbor-joining tree of

294 42 modern and 10 ancient genomes. Only the four high-coverage ancient samples used in our BEAST

295 analysis were labeled in this tree (Table S2). Nodes with bootstrap support of >90 are indicated by red

296 dots. (C) Time-scale maximum clade credibility tree of ancient and modern MDV sequences using the

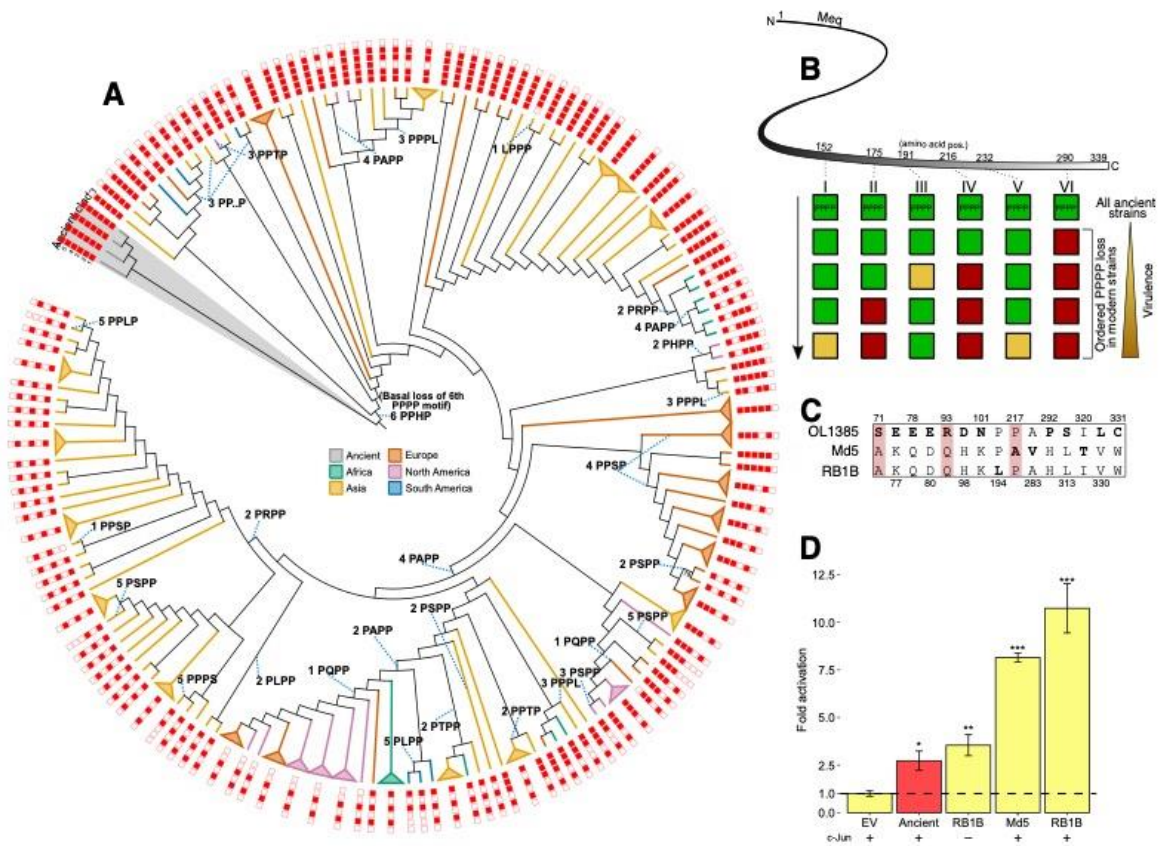
297 uncorrelated lognormal relaxed clock model (UCLD) and the general time-reversible (GTR) substitution

298 model. Gray bars indicate the 95% highest posterior density (HPD) for the age of each node. The 'cal'

299 suffix for ancient samples indicates that samples were radiocarbon dated and these dates used as priors for

300 the molecular clock analyses (24).





312  
 313 **Fig. 3. *Meq* has undergone ordered loss of tetraproline repeats and increased transactivation**  
 314 **ability.** (A) Phylogenetic analysis of 412 *Meq* sequences of standard length (1017 bp). The outermost  
 315 track shows the integrity of each tetraproline motif (filled squares = intact; open squares = disrupted). The  
 316 mutations that disrupt the tetraproline motif are linked by dotted blue lines (e.g. ‘4 PAPP’ indicates that  
 317 the 4<sup>th</sup> tetraproline motif is disrupted by a proline-to-alanine substitution in the second proline position. ‘3  
 318 PP..P’ denotes a deletion of the 3<sup>rd</sup> proline in the 3<sup>rd</sup> tetraproline motif). For a complete version of this  
 319 figure, see Fig. S7. (B) Proposed model for the most common ordered loss of tetraproline motifs in *Meq*.  
 320 Green and red boxes indicate presence and absence of an intact tetraproline, respectively. The yellow box  
 321 on the third row indicates that the 3<sup>rd</sup> tetraproline is occasionally lost after the 6<sup>th</sup>, but typically only in  
 322 terminal branches. The two yellow boxes in the bottom row indicate that it is either the 1<sup>st</sup> or 5<sup>th</sup>  
 323 tetraproline that is lost at this point. (C) Positions of amino acid differences between the ancient  
 324 Hungarian MDV strain (OL1385) and the two modern strains (RB1B and Md5). Positions that were also  
 325 found to be under positive selection are highlighted in red. (D) The transactivation ability of *Meq*  
 326 reconstructed from an ancient Hungarian MDV strain (OL1385) was compared to the transactivation  
 327 abilities of modern strains: RB1B and Md5 (‘very virulent’ pathotype). To show the effect of the partner  
 328 protein c-Jun on transactivation ability, the strongest transactivator RB1B was tested with (+) and without  
 329 (–) c-Jun. Transactivation ability is expressed as fold activation relative to baseline signal from an empty  
 330 vector (EV). Error bars are standard deviation, and statistical significance was determined using

331 Dunnett's test for comparing several treatment groups with a control. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P <$   
332 0.001.

333

334 **References**

- 335 1. J. Marek, Multiple Nervenentzündung (Polyneuritis) bei Hühnern. *Dtsch. Tierarztl. Wochenschr.*  
336 **15**, 417–421 (1907).
- 337 2. C. Morrow, F. Fehler, "5 - Marek's disease: A worldwide problem" in *Marek's Disease*, F.  
338 Davison, V. Nair, Eds. (Academic Press, Oxford, 2004), pp. 49–61.
- 339 3. A. F. Read, S. J. Baigent, C. Powers, L. B. Kgosana, L. Blackwell, L. P. Smith, D. A. Kennedy,  
340 S. W. Walkden-Brown, V. K. Nair, Imperfect Vaccination Can Enhance the Transmission of Highly  
341 Virulent Pathogens. *PLoS Biol.* **13**, e1002198 (2015).
- 342 4. C. S. Eidson, K. W. Washburn, S. C. Schmittle, Studies on acute Marek's disease. 9. Resistance  
343 to MD by inoculation with the GA isolate. *Poult. Sci.* **47**, 1646–1648 (1968).
- 344 5. N. Osterrieder, J. P. Kamil, D. Schumacher, B. K. Tischer, S. Trapp, Marek's disease virus: from  
345 miasma to model. *Nat. Rev. Microbiol.* **4**, 283–294 (2006).
- 346 6. E. A. Dimopoulos, A. Carmagnini, I. M. Velsko, C. Warinner, G. Larson, L. A. F. Frantz, E. K.  
347 Irving-Pease, HAYSTAC: A Bayesian framework for robust and rapid species identification in high-  
348 throughput sequencing data. *PLoS Comput. Biol.* **18**, e1010493 (2022).
- 349 7. J. Trimpert, N. Groenke, M. Jenckel, S. He, D. Kunec, M. L. Szpara, S. J. Spatz, N. Osterrieder,  
350 D. P. McMahon, A phylogenomic analysis of Marek's disease virus reveals independent paths to  
351 virulence in Eurasia and North America. *Evol. Appl.* **10**, 1091–1101 (2017).
- 352 8. A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti and  
353 the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
- 354 9. K. Li, Z. Yu, X. Lan, Y. Wang, X. Qi, H. Cui, L. Gao, X. Wang, Y. Zhang, Y. Gao, C. Liu,  
355 Complete genome analysis reveals evolutionary history and temporal dynamics of Marek's disease  
356 virus. *Front. Microbiol.* **13**, 1046832 (2022).
- 357 10. K. Majander, S. Pfrengle, A. Kocher, J. Neukamm, L. du Plessis, M. Pla-Díaz, N. Arora,  
358 G. Akgül, K. Salo, R. Schats, S. Inskip, M. Oinonen, H. Valk, M. Malve, A. Kriiska, P. Onkamo, F.  
359 González-Candelas, D. Kühnert, J. Krause, V. J. Schuenemann, Ancient Bacterial Genomes Reveal a  
360 High Diversity of *Treponema pallidum* Strains in Early Modern Europe. *Curr. Biol.* **30**, 3788–  
361 3803.e10 (2020).
- 362 11. B. Mühlemann, L. Vinner, A. Margaryan, H. Wilhelmson, C. de la Fuente Castro, M. E.  
363 Allentoft, P. de Barros Damgaard, A. J. Hansen, S. Holtsmark Nielsen, L. M. Strand, J. Bill, A.  
364 Buzhilova, T. Pushkina, C. Falys, V. Khartanovich, V. Moiseyev, M. L. S. Jørkov, P. Østergaard  
365 Sørensen, Y. Magnusson, I. Gustin, H. Schroeder, G. Sutter, G. L. Smith, C. Drosten, R. A. M.  
366 Fouchier, D. J. Smith, E. Willerslev, T. C. Jones, M. Sikora, Diverse variola virus (smallpox) strains

- 367 were widespread in northern Europe in the Viking Age. *Science*. **369** (2020),  
368 doi:10.1126/science.aaw8977.
- 369 12. M. Teng, Z.-H. Yu, A.-J. Sun, Y.-J. Min, J.-Q. Chi, P. Zhao, J.-W. Su, Z.-Z. Cui, G.-P.  
370 Zhang, J. Luo, The significance of the individual Meq-clustered miRNAs of Marek's disease virus in  
371 oncogenesis. *J. Gen. Virol.* **96**, 637–649 (2015).
- 372 13. Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**,  
373 1586–1591 (2007).
- 374 14. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and  
375 powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
- 376 15. A. R. Omar, K. A. Schat, Syngeneic Marek's disease virus (MDV)-specific cell-mediated  
377 immune responses against immediate early, late, and unique MDV proteins. *Virology*. **222**, 87–99  
378 (1996).
- 379 16. C. J. Markowski-Grimsrud, K. A. Schat, Cytotoxic T lymphocyte responses to Marek's  
380 disease herpesvirus-encoded glycoproteins. *Vet. Immunol. Immunopathol.* **90**, 133–144 (2002).
- 381 17. S. Halabi, M. Ghosh, S. Stevanović, H.-G. Rammensee, L. D. Bertzbach, B. B. Kaufer,  
382 M. C. Moncrieffe, B. Kaspers, S. Härtle, J. Kaufman, The dominantly expressed class II molecule  
383 from a resistant MHC haplotype presents only a few Marek's disease virus peptides by using an  
384 unprecedented binding motif. *PLoS Biol.* **19**, e3001057 (2021).
- 385 18. S. Haertle, I. Alzuheir, F. Busalt, V. Waters, P. Kaiser, B. B. Kaufer, Identification of the  
386 Receptor and Cellular Ortholog of the Marek's Disease Virus (MDV) CXC Chemokine. *Front.*  
387 *Microbiol.* **8**, 2543 (2017).
- 388 19. A. T. Engel, R. K. Selvaraj, J. P. Kamil, N. Osterrieder, B. B. Kaufer, Marek's disease  
389 viral interleukin-8 promotes lymphoma formation through targeted recruitment of B cells and CD4+  
390 CD25+ T cells. *J. Virol.* **86**, 8536–8545 (2012).
- 391 20. B. Lupiani, L. F. Lee, X. Cui, I. Gimeno, A. Anderson, R. W. Morgan, R. F. Silva, R. L.  
392 Witter, H.-J. Kung, S. M. Reddy, Marek's disease virus-encoded Meq gene is involved in  
393 transformation of lymphocytes but is dispensable for replication. *Proc. Natl. Acad. Sci. U. S. A.* **101**,  
394 11815–11820 (2004).
- 395 21. A. M. Conradie, L. D. Bertzbach, J. Trimpert, J. N. Patria, S. Murata, M. S. Parcells, B.  
396 B. Kaufer, Distinct polymorphisms in a single herpesvirus gene are capable of enhancing virulence  
397 and mediating vaccinal resistance. *PLoS Pathog.* **16**, e1009104 (2020).
- 398 22. K. G. Renz, J. Cooke, N. Clarke, B. F. Cheetham, Z. Hussain, A. F. M. Fakhru Islam, G.  
399 A. Tannock, S. W. Walkden-Brown, Pathotyping of Australian isolates of Marek's disease virus and  
400 association of pathogenicity with meq gene polymorphism. *Avian Pathol.* **41**, 161–176 (2012).

- 401 23. Z. Qian, P. Brunovskis, F. Rauscher 3rd, L. Lee, H. J. Kung, Transactivation activity of  
402 Meq, a Marek's disease herpesvirus bZIP protein persistently expressed in latently infected  
403 transformed T cells. *J. Virol.* **69**, 4037–4044 (1995).
- 404 24. See Supplementary Materials.
- 405 25. O. Putelat, "Archéozoologie" in *Strasbourg, Bas-Rhin. Rue de Lucerne – Rue du Jeu-de-*  
406 *Paume. Rapport de fouille préventive. Volume 1. Le système défensif primitif et le processus*  
407 *d'urbanisation d'un secteur du faubourg de la Krutenau du Moyen Âge à nos jours. Rapport de*  
408 *fouille préventive, Sélestat : Pôle d'Archéologie Interdépartemental Rhénan*, M. Werlé, Ed. (2015),  
409 pp. 98–174.
- 410 26. A. Cicović, D. Radičević, "Arheološka istraživanja srednjovekovnih nalazišta na Rudniku  
411 2009–2013. godine" in *Rudnik I, istraživanja srednjovekovnih nalazišta (2009-2013. godina)*,  
412 *Gornji Milanovac*, D. Radičević, A. Cicović, Eds. (2013), pp. 19–57.
- 413 27. N. Marković, J. Bulatović, "Rudnik 2009–2013: rezultati arheozoološke analize" in  
414 *Rudnik I, istraživanja srednjovekovnih nalazišta (2009-2013. godina)*, *Gornji Milanovac*, A.  
415 Cicović, D. Radičević, Eds. (2019), pp. 119–129.
- 416 28. H. Baron, Quasi Liber Et Pictura. Die Tierknochenfunde aus dem Gräberfeld an der  
417 Wiener Csokorgasse – eine anthrozoologische Studie zu den awarischen Bestattungssitten.  
418 *Monographien des RGZM.* **143** (2018).
- 419 29. O. Putelat, thesis, Université de Paris 1 Panthéon-Sorbonne (2015).
- 420 30. M. Popović, *Manastir Studenica – arheološka otkrića* (Republički zavod za zaštitu  
421 spomenika kulture, Arheološki institut, Beograd, 2015).
- 422 31. N. Marković, "Ishrana u manastiru Studenica: arheozoološka svedočanstva" in *Manastir*  
423 *Studenica – arheološka otkrića*, M. Popović, Ed. (Beograd: Republički zavod za zaštitu spomenika  
424 kulture i Arheološki institut, 2015), pp. 395–406.
- 425 32. I. Živaljević, N. Marković, M. Maksimović, Food worthy of kings and saints: fish  
426 consumption in the medieval monastery Studenica (Serbia). *anth.* **54**, 179–201 (2019).
- 427 33. Marković, N., Radišić, T. & Bikić, "Uloga živine u srednjovekovnoj ekonomiji manastira  
428 Studenice" in *Bioarheologija na Balkanu. Metodološke, komparativne i rekonstruktivne studije*  
429 *života u prošlosti*, M.-R. N. Vitezović S., Ed. (2016), pp. 99–116.
- 430 34. A. Saed Mucheshi, M. Nikzad, M. Zamani-Dadaneh, Rescue excavations at Bardeh Mar,  
431 Darian Dam area, Hawraman, Kurdistan, western Iran. *Proceedings of the 15th* (2017).
- 432 35. M. Mashkour, A. Mohaseb, S. Amiri, S. Beyzaiedoust, R. Khazaeli, H. Davoudi, H.  
433 Fathi, S. Komijani, A. Aliyari, H. Laleh, Archaeozoological Report of the Bioarchaeology  
434 Laboratory of the University of Tehran and the Osteology Department of the National Museum of

- 435 Iran, 2015-2016. *Proceedings of the 15th Annual Symposium on the Iranian Archaeology, 5-7 march*  
436 *2017, Tehran, Iranian Center for Archaeological Research*, 803–807 (2017).
- 437 36. P. J. Reimer, W. E. N. Austin, E. Bard, A. Bayliss, P. G. Blackwell, C. B. Ramsey, M.  
438 Butzin, H. Cheng, R. Lawrence Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, I. Hajdas,  
439 T. J. Heaton, A. G. Hogg, K. A. Hughen, B. Kromer, S. W. Manning, R. Muscheler, J. G. Palmer, C.  
440 Pearson, J. van der Plicht, R. W. Reimer, D. A. Richards, E. Marian Scott, J. R. Southon, C. S. M.  
441 Turney, L. Wacker, F. Adolphi, U. Büntgen, M. Capano, S. M. Fahrni, A. Fogtmann-Schulz, R.  
442 Friedrich, P. Köhler, S. Kudsk, F. Miyake, J. Olsen, F. Reinig, M. Sakamoto, A. Sookdeo, S.  
443 Talamo, The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP).  
444 *Radiocarbon*. **62**, 725–757 (2020).
- 445 37. J. Dabney, M. Knapp, I. Glocke, M.-T. Gansauge, A. Weihmann, B. Nickel, C.  
446 Valdiosera, N. García, S. Pääbo, J.-L. Arsuaga, M. Meyer, Complete mitochondrial genome  
447 sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc.*  
448 *Natl. Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).
- 449 38. M.-T. Gansauge, M. Meyer, Selective enrichment of damaged DNA molecules for  
450 ancient genome sequencing. *Genome Res.* **24**, 1543–1549 (2014).
- 451 39. C. Carøe, S. Gopalakrishnan, L. Vinner, S. S. T. Mak, M. H. S. Sinding, J. A. Samaniego,  
452 N. Wales, T. Sicheritz-Pontén, M. T. P. Gilbert, Single-tube library preparation for degraded DNA.  
453 *Methods Ecol. Evol.* **9**, 410–419 (2018).
- 454 40. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: fast  
455 approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. **29**, 1682–1684  
456 (2013).
- 457 41. M. Schubert, A. Ginolhac, S. Lindgreen, J. F. Thompson, K. A. S. Al-Rasheid, E.  
458 Willerslev, A. Krogh, L. Orlando, Improving ancient DNA read mapping against modern reference  
459 genomes. *BMC Genomics*. **13**, 178 (2012).
- 460 42. G. Jun, M. K. Wing, G. R. Abecasis, H. M. Kang, An efficient and scalable analysis  
461 framework for variant extraction and refinement from population-scale DNA sequence data. *Genome*  
462 *Res.* **25**, 918–925 (2015).
- 463 43. Broad Institute, *Picard toolkit* (Broad Institute, 2019;  
464 <http://broadinstitute.github.io/picard/>).
- 465 44. G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-  
466 Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S.  
467 Gabriel, M. A. DePristo, From FastQ data to high confidence variant calls: the Genome Analysis  
468 Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics*. **43**, 11.10.1–11.10.33 (2013).

- 469 45. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*.  
470 **9**, 357–359 (2012).
- 471 46. G. Tonkin-Hill, J. A. Lees, S. D. Bentley, S. D. W. Frost, J. Corander, Fast hierarchical  
472 Bayesian analysis of population structure. *Nucleic Acids Res.* **47**, 5539–5549 (2019).
- 473 47. J. Corander, P. Marttinen, Bayesian identification of admixture events using multilocus  
474 molecular markers. *Mol. Ecol.* **15**, 2833–2843 (2006).
- 475 48. M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, A. Rambaut,  
476 Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016  
477 (2018).
- 478 49. G. Yu, Using ggtree to Visualize Data on Tree-Like Structures. *Curr. Protoc.*  
479 *Bioinformatics.* **69**, e96 (2020).
- 480 50. M. Krzywinski, J. Schein, Í. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M.  
481 A. Marra, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645  
482 (2009).
- 483 51. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of  
484 large phylogenies. *Bioinformatics.* **30**, 1312–1313 (2014).
- 485 52. A. J. Page, B. Taylor, A. J. Delaney, J. Soares, T. Seemann, J. A. Keane, S. R. Harris,  
486 SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom.* **2**,  
487 e000056 (2016).
- 488 53. P. O. Lewis, A likelihood approach to estimating phylogeny from discrete morphological  
489 character data. *Syst. Biol.* **50**, 913–925 (2001).
- 490 54. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7:  
491 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- 492 55. P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg,  
493 T. Hamelryck, F. Kauff, B. Wilczynski, M. J. L. de Hoon, Biopython: freely available Python tools  
494 for computational molecular biology and bioinformatics. *Bioinformatics.* **25**, 1422–1423 (2009).
- 495 56. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for  
496 FASTA/Q File Manipulation. *PLoS One.* **11**, e0163962 (2016).
- 497 57. H. Li, seqtk Toolkit for processing sequences in FASTA/Q formats. *GitHub.* **767**, 69  
498 (2012).
- 499 58. S. Duchêne, D. Duchêne, E. C. Holmes, S. Y. W. Ho, The Performance of the Date-  
500 Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data. *Mol. Biol. Evol.* **32**,  
501 1895–1906 (2015).

- 502 59. A. Rieux, F. Balloux, Inferences from tip-calibrated phylogenies: a review and a practical  
503 guide. *Mol. Ecol.* **25**, 1911–1924 (2016).
- 504 60. M. Navascués, F. Depaulis, B. C. Emerson, Combining contemporary and ancient DNA  
505 in population genetic and phylogeographical studies. *Mol. Ecol. Resour.* **10**, 760–772 (2010).
- 506 61. M. Molak, M. A. Suchard, S. Y. W. Ho, D. W. Beilman, B. Shapiro, Empirical calibrated  
507 radiocarbon sampler: a tool for incorporating radiocarbon-date and calibration error into Bayesian  
508 phylogenetic analyses of ancient DNA. *Mol. Ecol. Resour.* **15**, 81–86 (2015).
- 509 62. F. Rodríguez, J. L. Oliver, A. Marín, J. R. Medina, The general stochastic model of  
510 nucleotide substitution. *J. Theor. Biol.* **142**, 485–501 (1990).
- 511 63. Z. Yang, Maximum likelihood phylogenetic estimation from DNA sequences with  
512 variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306–314 (1994).
- 513 64. A. J. Drummond, S. Y. W. Ho, M. J. Phillips, A. Rambaut, Relaxed phylogenetics and  
514 dating with confidence. *PLoS Biol.* **4**, e88 (2006).
- 515 65. B. Pfeifer, U. Wittelsbürger, S. E. Ramos-Onsins, M. J. Lercher, PopGenome: an  
516 efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936  
517 (2014).
- 518 66. D. K. Ajithdoss, S. M. Reddy, P. F. Suchodolski, L. F. Lee, H.-J. Kung, B. Lupiani, In  
519 vitro characterization of the Meq proteins of Marek’s disease virus vaccine strain CVI988. *Virus Res.*  
520 **142**, 57–67 (2009).
- 521 67. T. Huszár, I. Mucsi, T. Terebessy, A. Masszi, S. Adamkó, C. Jeney, L. Rosivall, The use  
522 of a second reporter plasmid as an internal standard to normalize luciferase activity in transient  
523 transfection experiments may lead to a systematic error. *J. Biotechnol.* **88**, 251–258 (2001).
- 524 68. K.-S. Chang, K. Ohashi, M. Onuma, Diversity (polymorphism) of the meq gene in the  
525 attenuated Marek’s disease virus (MDV) serotype 1 and MDV-transformed cell lines. *J. Vet. Med.*  
526 *Sci.* **64**, 1097–1101 (2002).
- 527 69. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and  
528 annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–5 (2016).
- 529 70. J. Sato, S. Murata, Z. Yang, B. B. Kaufer, S. Fujisawa, H. Seo, N. Maekawa, T.  
530 Okagawa, S. Konnai, N. Osterrieder, M. S. Parcells, K. Ohashi, Effect of Insertion and Deletion in  
531 the Meq Protein Encoded by Highly Oncogenic Marek’s Disease Virus on Transactivation Activity  
532 and Virulence. *Viruses.* **14** (2022), doi:10.3390/v14020382.
- 533 71. C. Firth, A. Kitchen, B. Shapiro, M. A. Suchard, E. C. Holmes, A. Rambaut, Using time-  
534 structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol. Biol. Evol.* **27**,  
535 2038–2051 (2010).

536 **Acknowledgments:**

537 This research used the University of Oxford's Advanced Research Computing, Queen Mary's Apocrita,  
538 and the Leibniz-Rechenzentrum (LRZ) High Performance Computing facility.

539

540 **Funding:**

541 European Research Council grant ERC-2019-StG-853272-PALAEOFARM or ERC-2013-StG-337574-  
542 UNDEAD or both (SRF, LAF, GL, ALS)

543 Wellcome Trust grant 210119/Z/18/Z (SRF, LAF)

544 Oxford Martin School grant ATR02370 (SRF, ALS, LdP, OGP)

545 AHRC grant AH/L006979/1 (GL, OL, NS)

546 European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie  
547 grant agreement no. 895107 (OL)

548 BBSRC grant number BB/M011224/1 (SD)

549 Ppostdoctoral grant (12U7121N) of the Research Foundation -- Flanders (Fonds voor Wetenschappelijk  
550 Onderzoek) (BV)

551

552

553 **Author contributions:**

554 Conceptualization: SRF, ALS, LAFF, GL

555 Methodology: SRF, EAD, ALS, LAFF, GL, BV, LdP, VN, OL, OGP

556 Sample provision: OL, NM, GF, RS, HB, LDS, DNS, IVA, OP, MS, HD, HF, ASM, AAV, AF,  
557 NS, JB, AOA, OVA, MM, VN

558 Investigation: SRF, EAD, OL, LdP, BV, SC, AFH, KT, PGF, SD, HL, GCB, OGP, VN, GL, ALS,  
559 LAFF

560 Visualization: SRF

561 Funding acquisition: LAFF, ALS, GL

562 Project administration: SRF, LAFF, ALS, GL

563 Supervision: LAFF, ALS, GL

564 Writing – original draft: SRF, LAFF, ALS, GL

565 Writing – review & editing: SRF, EAD, OL, LdP, BV, SC, AFH, KT, PGF, SD, NM, HL, GF, RS,  
566 HB, LDS, DNS, IVA, OP, MS, HD, HF, ASM, AAV, AF, NS, GCB, JB, AOA, OVA, MM, OGP,  
567 VN, GL, ALS, LAFF

568

569 **Competing interests:**

570 The authors declare that they have no competing interests.

571

572 **Data and materials availability:**

573 All MDV sequence data generated have been deposited in GenBank under accession PRJEB64489. Code  
574 is available from the following GitHub repository: <https://github.com/antonisdim/MDV>.  
575

576 **Supplementary Materials:**

577 Materials and Methods  
578 Supplementary Text  
579 Figs. S1 to S12  
580 Tables S4, S9 and S10  
581 Captions for Data S1  
582 References (25-71)

583

584 **Other Supplementary Materials for this manuscript include the following:**

585

586 Data S1, which comprises:

587

- Table S1: Sample metadata
- Table S2: Screening and capture sequencing results
- Table S3: Modern genome metadata
- Table S5: Integrity of miRNA sequences in ancient MDV
- Table S6: Fixed differences between ancient and modern MDV strains
- Table S7: PAML results
- Table S8: *Meq* sequence metadata
- Table S11: Metagenomic screening summary data
- Table S12: SNP summary table
- Table S13: Tip dates for BEAST analysis

588

589

590

591

592

593

594

595

596

Supplementary Materials for

**Ancient chicken remains reveal the origins of virulence in Marek's  
disease virus**

**Authors:**

Steven R Fiddaman<sup>1†\*</sup>, Evangelos A Dimopoulos<sup>2,3†</sup>, Ophélie Lebrasseur<sup>4,5</sup>, Louis du Plessis<sup>6,7</sup>, Bram Vrancken<sup>8,9</sup>, Sophy Charlton<sup>2,10</sup>, Ashleigh F Haruda<sup>2</sup>, Kristina Tabbada<sup>2</sup>, Patrik G Flammer<sup>1</sup>, Stefan Dascalu<sup>1</sup>, Nemanja Marković<sup>11</sup>, Hannah Li<sup>12</sup>, Gabrielle Franklin<sup>13</sup>, Robert Symmons<sup>14</sup>, Henriette Baron<sup>15</sup>, László Daróczi-Szabó<sup>16</sup>, Dilyara N Shaymuratova<sup>17</sup>, Igor V Askeyev<sup>17</sup>, Olivier Putelat<sup>18</sup>, Maria Sana<sup>19</sup>, Hossein Davoudi<sup>20</sup>, Homa Fathi<sup>20</sup>, Amir Saed Mucheshi<sup>21</sup>, Ali Akbar Vahdati<sup>22</sup>, Liangren Zhang<sup>23</sup>, Alison Foster<sup>24</sup>, Naomi Sykes<sup>25</sup>, Gabrielle Cass Baumberg<sup>2</sup>, Jelena Bulatović<sup>26</sup>, Arthur O Askeyev<sup>17</sup>, Oleg V Askeyev<sup>17</sup>, Marjan Mashkour<sup>20,27</sup>, Oliver G Pybus<sup>1,28</sup>, Venugopal Nair<sup>1,29</sup>, Greger Larson<sup>2‡</sup>, Adrian L Smith<sup>1\*‡</sup>, Laurent AF Frantz<sup>30,31\*‡</sup>

**Affiliations:**

<sup>1</sup>Department of Biology, University of Oxford, Oxford, UK

<sup>2</sup>The Palaeogenomics & Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford, UK

<sup>3</sup>Department of Veterinary Medicine, University of Cambridge, Cambridge, UK

<sup>4</sup>Centre d' Anthropobiologie et de Génomique de Toulouse, Toulouse, France

<sup>5</sup>Instituto Nacional de Antropología y Pensamiento Latinoamericano, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

<sup>6</sup>Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

<sup>7</sup>Swiss Institute of Bioinformatics, Lausanne, Switzerland

<sup>8</sup>Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium

<sup>9</sup>Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium

<sup>10</sup>BioArCh, Department of Archaeology, University of York, York, UK

- 29 <sup>11</sup>Institute of Archaeology, Belgrade, Serbia
- 30 <sup>12</sup>Institute of Immunity and Transplantation, University College London, London, UK
- 31 <sup>13</sup>Silkie Club of Great Britain, Charing, UK
- 32 <sup>14</sup>Fishbourne Roman Palace, Fishbourne, UK
- 33 <sup>15</sup>Leibniz-Zentrum für Archäologie, Mainz, Germany
- 34 <sup>16</sup>Medieval Department, Budapest History Museum, Budapest, Hungary
- 35 <sup>17</sup>Laboratory of Biomonitoring, The Institute of Problems in Ecology and Mineral Wealth, Tatarstan  
36 Academy of Sciences, Kazan, Russia
- 37 <sup>18</sup>Archéologie Alsace - PAIR, Bas-Rhin, France
- 38 <sup>19</sup>Departament de Prehistòria, Universitat Autònoma de Barcelona, Barcelona, Spain
- 39 <sup>20</sup>Bioarchaeology Laboratory, Central Laboratory, University of Tehran, Tehran, Iran
- 40 <sup>21</sup>Department of Art and Architecture, Payame Noor University (PNU), Tehran, Iran
- 41 <sup>22</sup>Provincial Office of the Iranian Center for Cultural Heritage, Handicrafts and Tourism Organisation,  
42 Bojnord, Iran
- 43 <sup>23</sup>Department of Archaeology, School of History, Nanjing University, China
- 44 <sup>24</sup>Headland Archaeology, Edinburgh, UK
- 45 <sup>25</sup>Department of Archaeology, University of Exeter, Exeter, UK
- 46 <sup>26</sup>Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden
- 47 <sup>27</sup>CNRS, National Museum Natural History Paris, Paris, France
- 48 <sup>28</sup>Department of Pathobiology and Population Sciences, Royal Veterinary College, London, UK
- 49 <sup>29</sup>Viral Oncogenesis Group, Pirbright Institute, Woking, UK
- 50 <sup>30</sup>Department of Veterinary Sciences, Ludwig Maximilian University of Munich, Munich, Germany
- 51 <sup>31</sup>School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
- 52 †joint-first author
- 53 ‡co-senior authors

54 \*corresponding authors. Emails: [steven.fiddaman@biology.ox.ac.uk](mailto:steven.fiddaman@biology.ox.ac.uk); [adrian.smith@biology.ox.ac.uk](mailto:adrian.smith@biology.ox.ac.uk);  
55 [laurent.frantz@lmu.de](mailto:laurent.frantz@lmu.de)

56 **This PDF file includes:**

57 Materials and Methods  
58 Supplementary Text  
59 Figs. S1 to S12  
60 Tables S4, S9 and S10  
61 Captions for Data S1  
62 References (25-71)

63  
64  
65 **Other Supplementary Materials for this manuscript include the following:**

66  
67 Data S1, which comprises:

- 68 ● Table S1: Sample metadata
- 69 ● Table S2: Screening and capture sequencing results
- 70 ● Table S3: Modern genome metadata
- 71 ● Table S5: Integrity of miRNA sequences in ancient MDV
- 72 ● Table S6: Fixed differences between ancient and modern MDV strains
- 73 ● Table S7: PAML results
- 74 ● Table S8: Meq sequence metadata
- 75 ● Table S11: Metagenomic screening summary data
- 76 ● Table S12: SNP summary table
- 77 ● Table S13: Tip dates for BEAST analysis

78  
79

80 **Materials and Methods:**

81 Archeological site descriptions

- 82 1. **Buda Castle**, Teleki Palace, Budapest, Hungary (13<sup>th</sup>-18<sup>th</sup> century; OL1385, OL1389; contact:  
83 László Daróczi-Szabó).

84 The former Teleki Palace is located in Buda Castle (Castle Hill, Budapest, Hungary), the medieval royal  
85 capital of Hungary, in close vicinity of the Royal Palace, in a general height of 158 meters above Baltic  
86 Sea level. Excavations of the remains of the palace were directed in 1999-2000 by Dorottya B. Nyékhelyi.  
87 Despite the relatively small area, a large number of archeological finds came to the surface, with a large  
88 majority dating between the 13<sup>th</sup> and 18<sup>th</sup> Century. Among other types of finds at least 100,000 animal  
89 remains were excavated (the analysis of the assemblage is ongoing). In general the assemblage shows a  
90 typical picture of animal use of the area: cattle dominating with around 45% of the finds, followed by  
91 sheep and/or goat and pig (20%), and others. The domination of domestic animals is almost total,  
92 although written sources prove that game was an important part of everyday diet: the most likely  
93 explanation is that large wild animals were deboned at the hunting site, and only the meat got to the  
94 markets. The relatively small number of bird and fish remains is because of the excavation methods: most  
95 of the finds were hand collected, and only a small part went through water sieving – here the number of  
96 these small finds drastically increased. Discrepancy from this generalities could be observed when finds  
97 connected to non-christian population were examined: pig is almost totally non-existent in layers  
98 supposedly connected to the local, 14<sup>th</sup> Century Jewish population of “Well 8”, and objects from the  
99 Ottoman Period (mid 16<sup>th</sup>-late 17<sup>th</sup> Century) (e.g. “Well 7”). In the latter next to the lack of pig remains a  
100 drastic increase of domestic small ruminants was noticed, as well as a decrease in cattle remains, the  
101 former becoming dominant.

- 102  
103 2. **Castillo de Montsoriu**, Spain (16<sup>th</sup> century; OL1984, OL1986, OL1999; contact: Maria Sana).

104  
105 Montsoriu Castle is located in the north-east of the Iberian Peninsula, on a hilltop at 650 m.a.s.l.  
106 (4114605800N, 213203000E). Archaeological excavations on-going since 1993 have been able to document  
107 the successive changes and re-structuring it underwent from the tenth century onwards, in consonance with  
108 the social and political changes in that long span of time. During the 2007 season, an abandoned cistern was  
109 excavated. This corresponds to the last stable occupation phase at the castle. It yielded an extremely well-  
110 preserved assemblage (UE 10955) consisting of different categories of organic and inorganic remains from  
111 the castle's larder. This sample results from a specific action carried out in a very short time. The bone  
112 sample recovered from the castle's cistern totals 10,922 remains, being representative of the food stored in  
113 the castle's larder. The deposition of this assemblage in the cistern would have been linked with the final  
114 abandonment of the castle in the last third of the sixteenth century. The inhabitants of the castle based animal  
115 resources exploitation strategies on domestic species (97% of NISP), including sheep and goats (29%), pigs  
116 (39%), cattle (32%) and poultry. Hunting was of lesser economic importance, and mainly involved wild  
117 rabbits, hares, red deer, boar, roe deer and fox. A large number of birds and fish are also consumed. A total

118 of 863 chicken bones were retrieved from UE10955, representing a total of 74 specimens. Most of the  
119 chicken remains correspond to females (61%). A total of 65 chicken remains show cut marks resulting from  
120 food processing and cooking for consumption.

121  
122 3. **Strasbourg Rue de Lucerne** - Rue du Jeu de Paume, France (16<sup>th</sup>- early 17<sup>th</sup> century; OL1936;  
123 contact: Olivier Putelat).

124 An archeological excavation was carried out « Rue de Lucerne – Rue du Jeu de Paume », in 2012-2013,  
125 in a suburb of the city of Strasbourg (France, Bas-Rhin). This excavation was carried out by the  
126 PAIR/Archéologie Alsace, on an area of 1119m<sup>2</sup>, under the direction of M. Werlé. An agro-pastoral  
127 building was uncovered, as well as latrines. The site yielded 3,124 terrestrial animal bones, and about  
128 1,300 fish remains. The filling of latrines 1067 (US 1059), from which the OL1936 sample originated, is  
129 dated to the end of the 16<sup>th</sup> Century. US 1059 yielded 499 remains of terrestrial animals (NISP: 397),  
130 including 45 chicken bones (11% of the NISP) and 32 remains of other birds, ducks and geese mainly  
131 (25).

132  
133 4. **Rudnik**, Serbia (14<sup>th</sup>-15<sup>th</sup> century; OL1008; contacts: Nemanja Marković and Jelena Bulatović).

134 The medieval settlement of Rudnik comprises four archaeological sites: Stacionar, Imanje Nikić, Imanje  
135 Marković, and Kojovača. The settlement is situated on the Rudnik Mountain, near the present-day town  
136 of Rudnik in central Serbia (approx. 95 km south of Belgrade (44°08'06.0"N 20°29'42.0"E)). Based on  
137 the current archaeological and written records, this area became particularly important in the second half  
138 of the 13<sup>th</sup> century, when German Saxon miners inhabited the country on the invitation of the Serbian  
139 king Stefan Uroš I Nemanjić (1241/1242–1276). New technologies in the exploitation and processing of  
140 ores provided a significant increase in the economy of medieval Serbia. The established square with an  
141 urban settlement developed later into an important mining and trade center, reaching its peak during the  
142 14<sup>th</sup> and the first part of the 15<sup>th</sup> century. The main mining activity was the silver extraction, while large  
143 amounts of lead and copper were also exploited. Most likely from the very beginning of the mining  
144 activities, the mint under the direct management of the Serbian ruler started working at medieval Rudnik.  
145 Historical records indicate that a colony of Dubrovnik traders was also established here. So far,  
146 archaeological excavations revealed three sacred buildings (two Orthodox and one Catholic churches),  
147 numerous profane buildings (some of which are monumental in scale), and numerous and varied movable  
148 archaeological finds. All these data indicate that during the last decades of the 13<sup>th</sup> to the middle of the  
149 15<sup>th</sup> century, Rudnik was an important developing settlement composed of native and newcomer  
150 populations of different social and religious characters (26).

151 Animal remains analyzed so far come from two Rudnik sites of Stacionar and Imanje Nikić. They were  
152 collected from cultural layers dated to the period between the 14<sup>th</sup> and the first half of the 15<sup>th</sup> century  
153 based on the findings of coins and pottery fragments. The Rudnik faunal assemblage comprises the

154 remains of mammals, birds, fish, and molluscs, of which 449 specimens were identified to a species or at  
155 least a genus level. Most of the animal remains belonged to the domestic animals. Caprines (sheep and  
156 goats taken together) are the most abundant taxa, followed by domestic cattle and domestic pigs. Other  
157 mammal remains were of horse, dog, cat, roe deer, red deer, hare, and squirrel. A small number of fish  
158 remains (common carp and sturgeon) and Jacob's scallop shell fragment were also present. Domestic  
159 chicken was the only bird species identified and is represented with 20 (4% of the total assemblage) bone  
160 fragments (27). The specimen OL1008 (a coracoid bone) was found within the cultural layer inside of an  
161 economic/residential building at the site of Imanje Nikić.

162       5. **Fishbourne** (Roman BC50-AD280; OL1128; contact: Robert Symmons).  
163 Discovered in 1960, Fishbourne Roman Palace is situated approximately 2.5km west of the city of  
164 Chichester, near the south coast of Britain. The palace itself is the largest domestic Roman building yet  
165 found north of the Alps, and today is a visitor center and museum displaying artifacts from the site and 16  
166 *in situ* mosaics. The site seemingly was occupied from the late Iron Age. Following the Roman invasion  
167 of 43 AD, the area was used as a military supply base. This later evolved into a “proto palace”, which  
168 included a large and luxurious bath house. Around 75 AD the proto palace was enlarged into the “Flavian  
169 palace”: a massive high-status complex comprising 4 wings arranged in a square around a formal garden  
170 with an informal garden to the south, leading down to nearby Chichester Harbour. The footprint of the  
171 building is approximately 21,000 square meters, although little is known about any ancillary structures  
172 that would undoubtedly have serviced the main complex. From the second century the building  
173 experienced a gradual period of decline and contraction before it was destroyed by fire in 280 AD.

174  
175 Sample OL1128 was recovered in 2002, during excavations approximately 30 m east of the northeast  
176 corner of the Flavian palace (site code FBE02, context 1040, catalog number CHCFB : FBE02/CB88).  
177 The context was interpreted as a midden and dated to the mid-late first century.

178  
179  
180       6. **Vienna Avar Cemetery**, Austria (7<sup>th</sup> / 8<sup>th</sup> century; OL1231; contact: Henriette Baron).

181 The Avar Cemetery at Vienna Csokorgasse was excavated under the direction of Ludwig Streinz and on  
182 behalf of the Historical Museum Vienna (today Wien Museum) in the years 1976 and 1976 within the  
183 scope of a building project. In the course of the excavation, 705 burials were unearthed and documented  
184 and the cemetery was excavated completely. According to Falko Daim and Ludwig Streinz, the cemetery  
185 was continuously used: the first burials in the northeast date to the Early Avar Period II (2<sup>nd</sup> quarter 7<sup>th</sup>  
186 century AD) and the last graves in the west and south of the burial area stem from the Late Avar Period II  
187 to III (2<sup>nd</sup> and 3<sup>rd</sup> third 8<sup>th</sup> century AD). The burial type – inhumation burials orientated west-east,  
188 displaying a variety of partially gender-specific burial goods – conforms to the usual customs of the

189 Middle and Late Avar Periods. For the most recent discussion of the site and its zooarchaeological  
190 findings, see (28).

191 In 491 (70 %) of the 705 burials animal bones were found. In most cases these were remains of chickens  
192 (319 graves, 45 %) and of sheep or goats (313, 44 %). Bones of cattle (240 burials, 34 %) and of pigs (84,  
193 12 %) also occurred in many burials. The chicken was interred in different degrees of skeletal  
194 completeness. Of the mentioned domestic mammals, primarily the thigh portion containing the femur was  
195 selected as a burial good. The faunal material also comprises some birds (domestic or greylag geese,  
196 Western jackdaw, Northern Goshawk, Eurasian skylark, a pigeon, white-tailed eagle, smew, gray  
197 partridge, Eurasian woodcock, as well as some unidentified bird remains), and some fish (pike, a wels  
198 catfish, and different cyprinids). Four outstanding rich Late Avar burials in the south of the cemetery area  
199 (burials 650, 690, 692, and 693) contained harnessed horses in an age fit for riding usage. In addition,  
200 three of these comprised complete skeletons of fully grown large male dogs. From the fourth equestrian  
201 burial (692) only a single dog tibia was recovered. Furthermore, a partial skeleton of a young puppy was  
202 found in the digging shaft of burial 650. Another one was recovered from a Middle Avar period child  
203 burial (burial 462). The equestrian graves contained grown up men with belt fittings and in one case  
204 (burial 690) two children and an adolescent, the latter presumably also with belt fittings.

205

206 7. **Kazan City**, Russia (16<sup>th</sup>-17<sup>th</sup> century; OL1584; contact: Dilyara Shaymuratova).

207 Kazan City is located in the Republic of Tatarstan in the east of the European part of Russia (55°47'26" N  
208 49°07'19" E), excavations were carried out on the territory of the modern backyard of the Kazan Federal  
209 University in 2002. Sample OL1584 was located in layers from the "Early Russian period" (16<sup>th</sup>-17<sup>th</sup>  
210 centuries). The study of the remains of bird bones was carried out by Igor Askeyev and Dilyara  
211 Shaymuratova.

212

213 8. **Cheboksary City**, Russia (16<sup>th</sup>-18<sup>th</sup> century; OL1585; contact: Dilyara Shaymuratova).

214 Cheboksary City is situated in the Republic of Chuvashia in the east of the European part of Russia  
215 (56°09'07" N 47°14'48" E), research and excavations of this archaeological site were carried out in 2004-  
216 2005 near the territory of the Vvedensky Cathedral. Sample OL1585 was located in layers of the 16<sup>th</sup>-18<sup>th</sup>  
217 centuries. Archaeozoological material was studied by Igor Askeyev and Dilyara Shaymuratova.

218

219 9. **Elabuga hillfort**, Russia (17<sup>th</sup>-18<sup>th</sup> century; OL1599; contact: Dilyara Shaymuratova).

220 Elabuga hillfort is located in the Republic of Tatarstan in the east of the European part of Russia  
221 (55°44'48" N 52°01'57" E), excavations were carried out on the territory of the location called "Chertovo

222 gorodishche" in 2003. Sample OL1599 was located in layers from the "Russian period" (17<sup>th</sup>-18<sup>th</sup>  
223 centuries). Research of bone remains of birds was carried out by Igor Askeyev and Dilyara  
224 Shaymuratova.

225  
226 10. **Andlau**, France (10<sup>th</sup>-12<sup>th</sup> century; OL1934; contact: Olivier Putelat).  
227 The village of Andlau (France, Bas-Rhin) is located in Alsace, about forty kilometers southwest of  
228 Strasbourg. An archaeological excavation was carried out in 2008 by the PAIR/Archéologie Alsace, at the  
229 "12 Cour de l'Abbaye", on an area of 900 m<sup>2</sup>, under the direction of A. Koziol. It concerned the gardens,  
230 located about thirty meters south of the Roman abbey church, in the center of the current village. This  
231 operation constitutes the first archaeological approach of the medieval abbey, which until then was known  
232 only from written sources. The latter indicates that this Benedictine monastery, reserved for women from  
233 the aristocracy, was founded in 879/880, by Richarde, the wife of the emperor *Charles le Gros*. The  
234 excavation yielded more than 7300 bone remains, 4304 of which are dated to phase A2 of the site (10<sup>th</sup>-  
235 12<sup>th</sup> century AD), from which the OL 1936 sample originated. Of these 4304, 2108 were determined, the  
236 85 chicken bones accounting for 4% of the NISP. It should be noted that the examination of some of the  
237 sieve refusals yielded several thousand fragments of eggshells, mainly attributable to the hen, as well as  
238 the discovery of a dwarf hen (29).

239  
240  
241 11. **Studenica Monastery**, Serbia (14<sup>th</sup>-15<sup>th</sup> century; OL1214; contacts: Nemanja Marković and  
242 Jelena Bulatović).

243 The Studenica Monastery is located on a flat plateau in the valley of the Studenica River, 49 km  
244 southwest of Kraljevo in southwestern Serbia (approx. 220 km from Belgrade; 43°29'11.4"N  
245 20°31'54.9"E). It represents one of the largest and richest medieval Orthodox monasteries in the country  
246 which has been on the UNESCO list of the world cultural heritage since 1986. It was founded at the end  
247 of the 12<sup>th</sup> century as an endowment and burial place of the progenitor of the Serbian medieval ruling  
248 dynasty of Nemanjić, the Grand Prince (*Veliki Župan*) Stefan Nemanja (1166–1196, †1199). Within its  
249 unique circular walls, the monastery complex contains: the Church of the Virgin Mary (the central sacred  
250 building), the King's Church, as well as the churches of St. John and St. Nicholas, several smaller sacred  
251 buildings, residential and economic facilities. Archaeological excavations (with minor and major  
252 interruptions) were conducted between 1949–2014. The study of everyday life in the medieval Studenica  
253 was part of the systematic archaeological research carried out in two phases between 1989–1998 and  
254 2010–2014. Several waste areas were discovered inside the monastery complex and along the outer side  
255 of the southeastern rampart in the immediate vicinity of the eastern gate (30). These waste areas contained  
256 a large amount of animal remains whose analysis provided the first insights into the strategies of animal  
257 exploitation in one Orthodox medieval monastery (31).

258 Faunal remains originated from midden areas both outside and inside the monastery walls. Based on the  
259 stratigraphic data, coins and ceramic fragments findings, these middens with faunal remains were dated to  
260 the period from the beginning of the 14<sup>th</sup> till the middle of the 15<sup>th</sup> century. Out of the total number of  
261 identified specimens (NISP=1949), 1527 belonged to mammals, 282 to birds, and 140 to fish. The  
262 majority of mammal remains (92%) were from economically important domesticates: sheep, goat, pig,  
263 and in a smaller number cattle. Hare is the best represented game species, whereas only one red deer tine  
264 fragment was found (31). The diversity of fish species and written sources (f.e., Studenica Typikon) point  
265 to the fact that the exploitation of fresh fish had an important place in the medieval economy of the  
266 monastery. Besides the remains of the fish available more or less locally (such as Wels catfish, carp, and  
267 pike), remains of migratory sturgeons (such as beluga, Russian sturgeon, and stellate sturgeon) which  
268 were probably transported from the Danube area about 200 km far from the monastery, were found too  
269 (32).

270 Bird remains were found in three of the four analyzed midden areas. Out of 282 bird remains, 243 were  
271 determined to a species level. Remains of the five species were identified: chicken, duck, goose, pigeon,  
272 and eagle. Remains of chicken are the most frequent and comprise 227 specimens (i.e., 93.5% of the total  
273 bird remains identified to the species level). Chicken age and sex profiles shows that older hens were the  
274 most numerous in all midden areas implying that they were exploited primarily for the egg production  
275 (33). The sample OL1214 (a coracoid) was recovered in the midden area formed on the Buildings V and  
276 VII ruins inside the monastery walls and dated to the last decade of the 14<sup>th</sup> century and the first decades  
277 of the 15<sup>th</sup> century.

278 12. **Naderi Tepe**, Iran (19<sup>th</sup> century; OL2267, OL2268, OL2272; contact: Marjan Mashkour).  
279 Three samples:

- 280 a. Tepe Naderi 2018 - Tr1, S2, F5, A12. DNA ID: OL2267
- 281 b. Tepe Naderi 2016 - Tr1, S2, F5, 34. Depth 290cm. DNA ID: OL2268
- 282 c. Tepe Naderi 2016 - Tr1, S2, F5, A41. Depth 330cm. DNA ID: OL2272

283  
284 Tepe Naderi is a 20m high artificial mound with a long occupational sequence that lasted from the late  
285 Chalcolithic/early Bronze Age through the early Iron Age (5128-158BP to 2752-27BP; Achaemenid to  
286 Sassanid period), a short medieval age settlement from 10<sup>th</sup>-11<sup>th</sup> Century CE, and eventually an  
287 occupational phase from Qajar period (18<sup>th</sup>-19<sup>th</sup> Century), when a fortress was built at the top of the  
288 mound overlooking a lower town which was a fortified enclosure.

289  
290 F5 is a bag-shaped pit dug into the medieval layers at the foot of the Tepe in the lower town. The pit is  
291 1.16m at the opening and 2.96m at the bottom (2.48m deep) in the south of Trench 1, from which dozens  
292 of fragments of blue-white stonepaste wares and many bone fragments were discovered. The pit was  
293 found and partly emptied during the first campaign in 2016 and partly in 2018. Therefore, it is plausible  
294 that samples from 2016 and 2018 are from the same individual. All samples are from the same context.

295  
296 At first, since the pit cuts through the medieval layers in the lower town at the foot of the tepe, it was  
297 attributed to the medieval period. However, this was later revised based on a finer study on the material  
298 and some radiometric dating. A calibrated radiocarbon date of a piece of animal bone places the pit within  
299 the range of 146-145 BP, which falls in the Qajar period. Two thermoluminescence dates of two pottery  
300 samples are 180-250 and 150-210 BP, or between 1700 and 1800 CE, which are consistent with the  
301 radiocarbon-dated animal bone.

302  
303 Further radiocarbon dating conducted in the present study corroborates the previous dating. The calibrated  
304 age range and associated probabilities for sample OL2272 is: 1698-1723 (23.6%); 1814-1835 (20.8%);  
305 1885-1910 (23.8%). See section below for full description of radiocarbon dates obtained in the present  
306 study.

307  
308 13. **Bard-e Mar** (Bem), Kurdistan (18<sup>th</sup>-early 21<sup>st</sup> century; OL2178; contact: Amir Saed Mucheshi  
309 and Marjan Mashkour).

310  
311 Bard-e Mar is situated in the Zargos region (35°9'39.63" N, 46°22'14.35" E, 790 m above sea level), on  
312 an old terrace of the Sirwan River in Sarvabad County, (a part of the Hawraman region), in the Kurdistan  
313 Province, Western Iran. Bard-e Mar is a small site of approximately two hectares with a steep slope  
314 towards the river. Bard-e Mar was recorded during the Darian Dam Archaeological Salvage Project  
315 (DDASP) directed by Fereidoun Biglari in 2014 and excavated by Amir Saed Mucheshi in 2015.  
316 Excavations at Bard-e Mar were conducted in two trenches, named Trench I and Trench II, both reaching  
317 the virgin soil. The cultural finds included potsherds, ground stones, clay pipes, rich faunal remains and  
318 rectangular and circular stone architectural structures. The Hawraman region has had a traditional and  
319 special masonry architectural style due to steep slopes of the mountains, and its continuity to the modern  
320 day is observable in evidence from Bard-e Mar. It is worthy to note that the diagnostic glazed pottery  
321 vessels were not discovered at site. As a result of relative and absolute chronological comparisons, two  
322 distinct periods were identified in Bard-e Mar, including the Middle Islamic (13<sup>th</sup> to 14<sup>th</sup> century A.D.)  
323 and the Late Islamic, the reign of Qajar Dynasty (18<sup>th</sup> to early-20<sup>th</sup> century A.D.). The finds from the  
324 Middle Islamic layers are less than its upper part, due to the reduction of the dimension of the trenches.  
325 The data from the Middle Islamic period are comparable to the western regions of Iran. The Qajar  
326 Dynasty/Late Islamic finds, especially clay pipes, are comparable with the Late Ottoman period in the  
327 Iraqi Kurdistan and Kurdish potteries reported from the Zagros region in the Late Islamic period (34).  
328 Faunal remains were studied by H. Davoudi, R. Khazaeli and M. Mashkour in the National Museum of  
329 Iran and Bioarchaeology Laboratory, Central Laboratory of the University of Tehran. A total of 2807  
330 pieces of animal bones have been discovered from excavations in Bard-e Mar, of which 1718 samples  
331 belong to Trench I and 1089 pieces belong to Trench II. The bones are generally placed in the group  
332 mammals and birds, and only one piece of fish bone was identified in the collection. Cattle, sheep and

333 goats are dominant, comprising 50% of the collection. A small amount of the collection derives from wild  
334 sheep and goats, deer, equines, foxes, forest otters and rabbits. A significant part of animal remains from  
335 the Late Islamic period belong to birds, of which some are domestic chickens and their relatives.  
336 Examining the animal remains of Bard-e Mar shows that its subsistence economy is based on animal  
337 husbandry and the people living in this area used goats, sheep and cattle as the main source of livelihood,  
338 a method that is still common in this region (35).  
339 In this paper, Sample OL2178 was analyzed, which belonged to the right femur of *Gallus gallus*, derived  
340 from Trench I, Locus 14, Code a#56, the Qajar Dynasty/Late Islamic period (18<sup>th</sup> to early 20<sup>th</sup> century  
341 A.D.).

#### 342 Description of bird used as positive control

343 To validate the shotgun sequencing and baiting procedure, we also processed DNA from a modern bird  
344 displaying symptoms of Marek's Disease from a naturally acquired infection. A feather from the bird, a  
345 Silkie chicken, was provided by Gabrielle Franklin of the Silkie Club of Great Britain. At the time of  
346 infection, the bird was unvaccinated and 3-4 years of age. Symptoms included a sudden loss of weight  
347 and dropped wing/paralysis on the left side. Following the development of these more severe symptoms,  
348 the bird was euthanized to prevent suffering.

#### 349 Radiocarbon dating

350 In total, four samples were radiocarbon dated ([Table S10](#)) using either the Oxford Research Laboratory  
351 for Archaeology and the History of Art or Beta Analytic. Uncalibrated dates were calibrated using  
352 IntCal20: Northern Hemisphere (36). In all cases, there was a degree of ambiguity in the measured sample  
353 ages, so raw plots are given below ([Fig. S8](#)). For the BEAST analysis, the mean date was used for all  
354 samples.

#### 355 Ancient DNA extraction

356 DNA extractions were performed in dedicated ancient DNA (aDNA) facilities at the PalaeoBARN  
357 (University of Oxford). DNA was extracted from chicken bone samples ([Table S1](#)) in a dedicated ancient  
358 DNA laboratory using the appropriate sterile techniques and equipment. Prior to grinding, ~0.5mm of the  
359 exterior surface of the bone was removed using a Dremel 3000 electric hand-drill. Extraction was carried  
360 out using the Dabney extraction protocol (37).

#### 361 Library preparation and preliminary sequencing

362 Illumina libraries were built following either the protocol in (38) or (39), but with the addition of a six  
363 base-pair barcode added to the IS1\_adapter.P5 and IS3\_adapter. P5+P7 adapter pair. The libraries were

364 then amplified on an Applied Biosystems StepOnePlus Real-Time PCR system to check that library  
365 building was successful, and to determine the minimum number of cycles to use during the indexing  
366 amplification PCR reaction. A six base-pair barcode was used during the indexing amplification reaction  
367 resulting in each library being double-barcoded with an “internal adapter” directly adjacent to the ancient  
368 DNA strand and which would be the first bases sequenced, and a traditional external barcode that would  
369 be sequenced during Illumina barcode sequencing. An additional sequencing library was built for sample  
370 OL1385, which had two external 6 bp indices appended to the P5 and P7 Illumina sequencing primers,  
371 and had no internal index tag. Up to 200 libraries with unique barcode combinations were pooled at  
372 equimolar levels (as determined by an Agilent Technologies 2200 TapeStation) and an 80 - 150 bp run  
373 was carried out on an Illumina HiSeq 2500/4000/X sequencer at Novogene or Macrogen.

#### 374 Preliminary analyses of sequencing data using HAYSTAC

375 We screened 995 ancient chicken (*Gallus gallus domesticus*) DNA libraries, with HAYSTAC (6), with  
376 default settings, for the Gallid Alphaherpesvirus 2 (MDV) genome, using a custom database of complete  
377 herpesvirus genomes obtained from in the NCBI RefSeq database (Table S11). This analysis identified 18  
378 samples with at least one MDV read (range: 1–785; [Table S2](#)). To ensure the validity of the positive  
379 identifications in samples with a higher number of confidently assigned MDV reads (>50), we ensured  
380 that the respective evenness of coverage ratio (genome coverage / fraction of genome covered) was less  
381 than 10, and that appropriate chemical damage patterns characteristic of ancient DNA were identified  
382 ((40); [Fig. S2](#)). These 18 samples, as well as a modern sample with MD symptoms (OL1099; positive  
383 control) and a sample for which we did not identify any MDV reads (OL1214; negative control) were  
384 subsequently genome captured (see below).

#### 385 Design of probes for in-solution capture

386 Custom DNA probes for a tiled baiting approach were designed and synthesized by Arbor BioScience,  
387 based on the *Gallid alphaherpesvirus 2* genome (strain RB-1B; accession EF523390.1, NCBI). Baits  
388 were only constructed for one copy of each of the two terminal repeats (coordinates 1-14004 and 165217-  
389 178246 in EF523390.1 were excluded), and regions of low complexity in the genome were masked .  
390 Oligonucleotide baits were designed approximately every 28 nucleotides, yielding 8403 50-mer probe  
391 sequences for the whole MDV target genome. The probes allowed for the hybridisation of both modern  
392 and ancient strains of MDV, while limiting carryover of endogenous DNA.

#### 393 In-solution capture and sequencing of captured libraries

394 Additional libraries were either amplified or re-built from extracts of the samples that were selected for  
395 targeted genome enrichment ([Table S2](#)), following the metagenomic screening for MDV DNA reads.  
396 These selected libraries were amplified again following the same indexing and PCR amplification

397 strategies as previously described, and were condensed in a final volume of 25  $\mu$ l. Libraries from samples  
398 OL1008, OL1128, OL1214, OL1231, OL1385, OL1585, OL1987, OL1999 were built once and were also  
399 amplified and captured once, whereas for samples OL1389, OL1584, OL1599, OL1934, OL1936,  
400 OL1984, OL1986, OL2178, OL2267, OL2268, OL2272 three different libraries were built from the same  
401 extract and each library was subsequently amplified and captured ([Table S2](#)). Any additional libraries  
402 built for the samples that were captured more than once, were indexed and amplified for sequencing,  
403 using conventional full-length P7 and P5 Illumina primers that were both indexed with external 6 bp  
404 indices, instead of having one internal and one external index (39).

405 MDV genome capture and paired-end sequencing on Illumina NovaSeq lanes were performed by Arbor  
406 BioScience. Specifically ~200 million 150 bp paired-end reads were generated per capture sequencing  
407 pool. As a positive control, one modern chicken feather sample (OL1099) which displayed symptoms  
408 characteristic of MD was also processed using the above procedure and yielded a 4 $\times$  genome. Sample  
409 OL1214 yielded no MDV-specific reads following screening and was included as a negative control in  
410 the genome capture protocol.

#### 411 MDV-positive samples

412 In total, we performed genome captures for 20 samples, comprising a modern positive control (OL1099),  
413 an ancient negative control (OL1214) and 18 ancient samples with varying degrees of MDV positivity  
414 ([Table S2](#)). Of these, 16 samples (including the positive control) were unambiguously positive for MDV  
415 DNA. The negative control was enriched to 451 post-capture reads, but a majority of these were short  
416 (<25bp) and therefore likely nonspecific ([Table S2](#)). Three further samples (OL1008, OL1231 and  
417 OL1128) yielded very few (<60) post-capture reads, and these were overwhelmingly short (<25bp), so  
418 were considered negative for MDV ([Table S2](#)).

419

#### 420 Capture efficiency

421

422 For MDV-positive samples, enrichment for MDV sequence ranged from 1.65 to 273.12-fold increase in  
423 unique DNA templates, allowing us to increase the depth of coverage of the ancient MDV isolates (range:  
424 0.13 - 41.92), reduce the percentage of missing sequence data and reconstruct consensus genomes ([Table](#)  
425 [S2](#)). The genome enrichment captured virtually all the unique DNA templates that were available in the  
426 analyzed libraries, as shown by the high percentage of duplicated reads post capture (range: 56.6 -  
427 99.98%; [Fig. S1](#), [Table S2](#)).

428

#### 429 Short read alignment of ancient data

430 The captured data were aligned with the bwa aln algorithm against the RB-1B (EF523390.1) MDV  
431 reference genome with default parameters apart from disabling the seed option (“-l 1024”) (41). Because

432 the MDV genome has terminal long (TRL) and short (TRS) repeat regions (14kbp and 13kbp,  
433 respectively), one repeat from each of these duplicated regions was masked (coordinates 1-14004 and  
434 165217-178246 in EF523390.1) in order to avoid low mapping quality stemming from a DNA read  
435 aligning to multiple positions, and to retain the maximum number of reads.

#### 436 Testing read sharing between chicken and MDV

437 To detect sequence homology between chicken and MDV genomes that could lead to false positive  
438 identification of ancient MDV-positive samples, we simulated short read sequencing data based on the  
439 RB-1B (EF523390.1) genome and mapped these regions against the chicken genome (GRCg7b). To  
440 simulate reads, ART was used (art\_illumina -ss HS10 -i [input\_fasta] -l 100 -f 50 -p -m 350 -s 50 -o  
441 [output\_fastq]). Reads were then mapped against the chicken genome using the bwa aln algorithm.

#### 442 Genotyping of ancient data

443 Aligned reads had 5 base pairs trimmed from both their 5' and 3' ends with bamUtil v 1.0.14 (42), and  
444 read group identifiers were added to the resulting bam files by using picard tools v 2.16.0 (43).

445 The GATK package (44) was used for variant calling and sample genotyping. HaplotypeCaller v 4.1.2.0  
446 was employed for SNP (single nucleotide polymorphism) calling with the following parameters: the  
447 EF523390.1 MDV genome (masked for TRL and TRS; see above) was used as reference, a value of 30  
448 for the minimum base quality (Q) was required, all sites were outputted in GVCF files, ploidy was set to  
449 1, and physical phasing was not allowed (gatk HaplotypeCaller --reference EF523390.1\_mask.fasta --  
450 min-base-quality-score 30 --output-mode EMIT\_ALL\_SITES --sample-ploidy 1 -ERC GVCF --do-not-  
451 run-physical-phasing true).

452 The GVCF files, from each individual sample, were aggregated into one with the CombineGVCFs v  
453 4.1.2.0 tool. The aggregated data were input into GenotypeGVCFs for joint genotyping of the different  
454 samples. The raw variants were further filtered using VariantFiltration, filtering in variants with a  
455 minimum base quality (Q) value of 30 and a minimum sequence depth of 5 reads (gatk VariantFiltration -  
456 -filter-expression "QUAL >= 30.0 && DP >= 5"). VariantsToTable was subsequently employed to  
457 convert the resulting VCF files containing the filtered SNPs to tables (Table S12).

#### 458 Generating fasta files

459 Consensus fasta files were also obtained from aligned reads in bam format using the htsbox pileup tool  
460 (<https://github.com/lh3/htsbox>). The following parameters were used: a minimum read length of 25 base  
461 pairs, 5 base pairs were trimmed from each end of an aligned read, a value of 30 for the minimum base  
462 quality (Q) and for the minimum read alignment quality (q) was set, for polymorphic sites the majority

463 frequency allele was called. Two sets of consensus fasta sequences were built for the newly captured  
464 samples, one set with  $\geq 1x$  and one with  $\geq 5x$  coverage depth per site (htsbox pileup -l 25 -T 5 -q 30 -Q  
465 30 -M-s 1/5).

466

#### 467 Publicly available (modern) data

468

469 Raw Illumina sequencing reads for samples KU173115.1, KU173116.1, KU173117.1, KU173118.1,  
470 KU173119.1 were downloaded from the SRA using fasterq-dump v 2.11.0 ([https://github.com/ncbi/sra-](https://github.com/ncbi/sra-tools)  
471 [tools](https://github.com/ncbi/sra-tools)), and raw reads for samples MF431493.1, MF431494.1, MF431495.1, MF431496.1 were provided  
472 to us by Jakob Trimpert (7). Modern samples AF243438.1, AY510475.1, DQ530348.1, EF523390.1,  
473 EU499381.1, FJ436096.1, FJ436097.1, JF742597.1, JQ314003.1, JQ806361.1, JX844666.1,  
474 KU744555.1, KU744556.1, KU744557.1, KU744558.1, KU744559.1, KU744560.1, KU744561.1,  
475 KX290013.1, KX290014.1, MG432697.1, did not have any associated raw sequencing reads deposited on  
476 the SRA, so full genome sequences were downloaded from NCBI's Nucleotide database using a custom  
477 biopython (v 1.79) script. Metadata associated with modern MDV strains can be found in [Table S3](#).

478

479 For publicly available sequenced MDV strains that have undergone serial passage under experimental  
480 conditions, only the lowest passage number was retained for time-calibrated phylogenetic analyses.  
481 Therefore the following accessions were excluded from our BEAST analysis (but were included in  
482 maximum likelihood phylogenies): JQ806362.1 (passage 31), JQ809691.1 (passage 41), JQ809692.1  
483 (passage 61), JQ820250.1 (passage 81), JQ836662.1 (passage 101), KT833852.1 (passage 70),  
484 KX290015.1 (passage 75), KX290016.1 (passage 110). We also excluded strains that were duplicated in  
485 the dataset (KT833851.1 and NC\_002229.3, both Md5). We further excluded accession AF147806.2,  
486 since it was isolated in 1964 and subsequently experimentally passaged in cell culture at least 45 times  
487 before being sequenced (4). On a maximum-likelihood phylogeny AF147806.2 falls on a long terminal  
488 branch and a root-to-tip divergence analysis further showed that it is enriched for substitutions compared  
489 to contemporary strains (see Fig. S9A and C and "Temporal signal" section below). Repeated passaging  
490 likely introduced mutations in this strain that obfuscates the temporal signal and makes it unsuitable for  
491 time-calibrated phylogenetic analyses. Accession MG518371.1 was also excluded from our BEAST  
492 dataset as it lay in a clade with a vaccine strain and thus likely represented contamination.

493

#### 494 Removing BAC sequences

495 MDV reference genomes with accession numbers KT833851.1, KT833852.1, and FJ436097.1 were  
496 contaminated with sequences from BAC (bacterial artificial chromosome) cloning vectors. BAC  
497 sequences were removed from these genomes with a custom biopython script, and their non-contaminated  
498 respective genomes were outputted in fasta format.

499 Masking of the modern genomes

500 Since the consensus MDV sequences in the ancient viral dataset had one of each of the terminal repeats  
501 (TRL and TRS) masked, we also masked the same terminal repeats in our modern viral sequence dataset.

502  
503 The modern MDV genomes, which were downloaded from the NCBI nucleotide database, (cleaned from  
504 BAC contamination) were split into 500 bp fragments, with a 5 bp overlap between each k-mer, using the  
505 pyfasta split v 0.5.2 tool (<https://github.com/brentp/pyfasta>). Each fragmented genome was subsequently  
506 aligned against the masked EF523390.1 MDV genome using the bowtie2 algorithm v 2.4.4 (21) with  
507 default settings. The resulting bam files for each modern MDV genome were sorted and passed to htsbox  
508 pileup with default settings (htsbox pileup -l 25 -M), in order to build new consensus fasta sequences that  
509 had the TRL and TRS repeats masked, as for the ancient MDV sequence dataset.

510  
511 For the modern samples for which sequencing reads were available, Bowtie2 v 2.4.4 (21) with default  
512 settings was used to align these short reads against the masked EF523390.1 MDV reference genome, and  
513 consensus fasta sequences were built using htsbox (htsbox pileup -l 25 -T 3 -q 30 -Q 30 -M -s 5).

514

515 Sequencing depth for sample OL1385

516  
517 To calculate per-base sequencing depth for the highest coverage sample (OL1385), samtools depth (v.  
518 1.10) was used. To plot the sequencing depth (Fig. S10), a custom R script was used to plot the mean  
519 depth across non-overlapping bins of 300bp.

520

521 Phylogenetic clustering analysis

522 The R package fastbaps v 1.0.4 (46) was used to identify phylogenetic clusters within our MDV dataset,  
523 using the baps algorithm to perform Bayesian hierarchical clustering (47), unconstrained by any  
524 phylogenetic tree. The 35 ancient and modern genomes used in the BEAST analysis were used as input  
525 for the lineage clustering analysis. The resulting lineages were compared and plotted side-by-side with the  
526 Maximum Credibility Tree, produced by TreeAnnotator v 1.10.4 from the BEAST software suite (48),  
527 using the R library ggtree v 3.2.0 (49); [Fig. S11](#)).

528

529 Positive selection analysis

530

531 The same alignments used for the BEAST analysis were also analyzed for positive selection using codeml  
532 implemented in PAML (v4.9; (13)). Open reading frames were sliced out of the genomic alignment with  
533 reference to coordinates of the annotated RB-1B genome (accession EF523390.1). Exons from multi-  
534 exonic genes were concatenated and ORFs on the reverse strand were reverse complemented. For ORFs

535 in the repeated regions, only the internal repeat was included in the analysis, meaning ORFs MDV000.5–  
536 MDV006.6 and MDV097.3–MDV103 were excluded. The branch-site test in PAML was used (model=2;  
537 NSsites=2; fix\_omega=0/1; omega = 1), specifying the modern MDV sequences as the ‘foreground’  
538 lineage and the ancient MDV sequences as the ‘background’ lineage. Pairs of log-likelihood values for  
539 each locus were used to perform a likelihood ratio test ( $2\Delta\ln L$ ) with d.f. = 1, then the final p-value  
540 calculated by dividing the original p-value by 2 (as described in the PAML manual). P-values were  
541 transformed using the Benjamini-Hochberg procedure in R (v4.2.1; (14)). Of the 154 non-redundant loci  
542 included in the analysis, 49 genes showed significant evidence (corrected  $p < 0.05$ ) for positive selection in  
543 modern MDV with respect to ancient MDV. Full PAML results are presented in [Table S7](#). Positive  
544 selection results were plotted on a circular representation of the RB1B genome in Fig. 2 using Circos  
545 (50).

#### 546 Maximum Likelihood trees and Multisequence Alignments

548 Maximum likelihood (ML) phylogenetic trees (Fig. S4) were built using RAxML v 8.2.9 (51), and  
549 Neighbour Joining (NJ) trees (Fig. S3) were constructed using seaview v4. The percentage of unknown  
550 bases (Ns) was calculated for each genome with the seqtk comp tool v 1.2.95, and all samples with at  
551 least 1% of the genome covered (1782 sites) were kept, in an effort to retain and analyze the maximum  
552 number of ancient MDV genomes. In total 42 modern MDV genomes from previous studies ([Table S3](#))  
553 and 11 captured MDV genomes from the present study were included. We used consensus fasta  
554 sequences built with htsbox for this analysis (see above - for the captured samples one set was built with  
555 coverage depth of 1x and one set with 5x). Of the 11 captured genomes that were included, 1 genome was  
556 from a modern sample (OL1099; sampled 2014) while the remaining 10 were obtained from  
557 archeological samples from the 14<sup>th</sup>-20<sup>th</sup> century ([Table S1](#)).

558 RAxML was used to build an unrooted tree with the following parameters: a random seed value for rapid  
559 bootstrapping, a parsimony-based starting tree, a GTRGAMMA substitution model, and 100 bootstrap  
560 runs (raxmlHPC-PTHREADS -f a -T 10 -x \$RANDOM -k -# 100 -p \$RANDOM -m GTRGAMMA).  
561 Seaview was used to build an NJ tree using the Jukes-Cantor model, with 100 bootstrap runs.

#### 562 Maximum likelihood tree based on transversions

563 A midpoint-rooted maximum likelihood tree based only on transversions (Fig. S5), was built using 4  
564 ancient MDV genomes (each with at least 80% coverage at 5x), 1 modern positive control from the  
565 current study, and 30 modern MDV genomes from public sources. All transition polymorphisms were set  
566 to “N” (unknown base) using a custom biopython script. All monomorphic positions were also removed  
567 using the snp-sites tool v 2.5.1 (52). A phylogenetic tree was then constructed using the Lewis correction

568 method for ascertainment bias (53) along with the ASC\_GTRGAMMA substitution model (raxmlHPC-  
569 PTHREADS -f a -T 10 -x \$RANDOM -k -# 100 -p \$RANDOM -m ASC\_GTRGAMMA --asc-corr lewis)

#### 570 Maximum likelihood tree rooted with Herpesvirus of Turkeys

571 To confirm the placement of the root in the tree based only on transversions we used genomic data from a  
572 *Meleagrid herpesvirus 1* (HVT; accession: NC\_002641.1) as an outgroup to build a rooted maximum-  
573 likelihood tree (Fig. S6). Mafft v 7.123 (54) was used to align the sequences of the outgroup to the MDV  
574 genomes (mafft --maxiterate 1000 --thread 5 --nwildcard). For the phylogenetic analysis we filtered out  
575 positions that were missing in the HVT genome (either “N” or gaps) using a custom biopython (55)  
576 script. We removed all the positions that were either monomorphic or unknown throughout all the  
577 samples in the alignment by using the snp-sites tool v 2.5.1 (52). A phylogenetic tree was constructed  
578 using the Lewis correction method for ascertainment bias (53) along with the GTRGAMMA substitution  
579 model (raxmlHPC-PTHREADS -f a -T 10 -x \$RANDOM -k -# 100 -p \$RANDOM -m GTRGAMMA), .

#### 580 Filtering overlapping open reading frames for BEAST analyses

581 To obtain divergence time with BEAST v 1.10.4 we removed overlapping open reading frames (ORF)  
582 based on a bed file containing ORF coordinates for EF523390.1 using a custom biopython script. Non-  
583 coding regions of EF523390.1 were also included, which resulted in a bed file that contained non-  
584 overlapping ORF and non-coding regions. These regions were extracted from the htsbox consensus fasta  
585 files for each of the ancient and modern genomes using the seqkit subseq tool of the seqkit package v  
586 0.12.1 (56). ORFs that were found on the minus (-) strand were reverse complemented with the seqtk seq  
587 tool v. 1.2.95 (57).

#### 588 Filtering for coverage for BEAST analyses

589 In order to maximize the information content of the sequence alignments provided to BEAST and reduce  
590 the noise introduced by missing data in our ancient samples, we used a custom biopython script to  
591 calculate the percentage of missing data within each genomic region included in the BEAST analysis (see  
592 above). Only ORFs and intergenic regions with at least 10% of sites in which were covered in all  
593 individual sequences in the alignment (i.e. 100% coverage: no missing data) were retained. We then  
594 concatenated the ORF and intergenic regions separately using the seqkit concat tool v. 0.12.1.

#### 595 Temporal signal

596 To assess whether our data possess a temporal signal we first examined the correlation between root-to-  
597 tip divergence (in substitutions/site) and sampling date, using maximum likelihood trees constructed in  
598 RAxML v. 8.2.9 (51). Sampling dates of ancient sequences were fixed to the means of the C14 calibrated

599 distributions and the maximum-likelihood trees were rooted based on the outgroup analysis (Fig. S6). A  
600 total of 36 MDV genomes were included in this analysis: 30 modern, 1 modern extensively passed  
601 strain (AF147806.2), 1 modern positive control from the present study (OL1099), and 4 ancient captured  
602 genomes from the present study (OL1385, OL1389, OL1986, OL2272). Four variations were used: (i) all  
603 strains (Fig. S9A), (ii) all strains except AF147806.2 (Fig. S9B), (iii) all modern strains (Fig. S9C) and  
604 (iv) all modern strains except AF147806.2 (Fig. S9D). This analysis confirmed that AF147806.2 (cell  
605 passaged strain) did not undergo clock-like evolution and can be considered as an outlier.

606 We further used a Bayesian date randomization test (58–60) (DRT) to examine the strength of the  
607 temporal signal, by permuting sampling dates among genomes and performing 100 replicate analyses. For  
608 the analyses the same dataset ( $n = 35$ , excluding AF147806.2) and BEAST model as below were used,  
609 using an uncorrelated lognormally distributed (UCLD) relaxed clock and fixing the sampling dates of the  
610 ancient sequences to the means of the C14 calibrated distributions. Chains were run for 50 million steps  
611 and the parameter sampling frequency set at every 5,000 steps. This resulted in  $ESS > 200$  for all  
612 parameters in all replicate analyses. The HPD intervals of the mean clock rates compared to an analysis  
613 with unpermuted sampling dates is shown in Fig. S12.

614 Regressing the root-to-tip divergence against sampling date showed a strong positive correlation (Fig. S9)  
615 and the Bayesian DRT further showed strong evidence for a temporal signal in our dataset (Fig. S12).

#### 616 BEAST analysis

617 The BEAST v 1.10.4 (48) package was used for the divergence dating and molecular clock rate  
618 estimation of the MDV phylogeny. A total of 35 MDV genomes were included in this analysis: 30  
619 modern, 1 modern positive control from the present study (OL1099), and 4 ancient captured genomes  
620 from the present study (OL1385, OL1389, OL1986, OL2272). Publically available sequenced MDV  
621 strains that have undergone serial passage under experimental conditions or contaminated with vaccine  
622 strains were excluded from this analysis (see "Publicly available (modern) data" section for accessions).  
623 Tip dates were provided to BEAST for both the modern and archaeological samples (Table S13). We  
624 used the probability density function of the radiocarbon date as prior for the tip dates of the 4 ancient  
625 samples. To do so we used the empirical calibrated radiocarbon sampler (ECRS; (61) as implemented in  
626 BEAST. We used the same prior for OL1385 (which was directly dated) and OL1389 (not directly dated)  
627 as the bones from which these sequences were derived were from the same archaeological context.

628 The concatenated ORF alignment was partitioned further into codon positions with a custom python  
629 script. Each partitioned alignment along with the concatenated intergenic region alignment, was used as  
630 input for BEAST. The tree topology was fixed to the topology of the outgroup-rooted maximum  
631 likelihood tree (without the outgroup) constructed only from transversion polymorphisms. An

632 independent GTR+ $\Gamma_4$  (62, 63) substitution model was used for each alignment position and a constant  
633 size coalescent model was specified as a tree prior.

634 We used an uncorrelated lognormally distributed (UCLD) relaxed clock model (64). The MCMC chain  
635 length was set at 300 million steps, and the parameter and tree sampling frequencies set at every 10,000  
636 steps. The BEAST parameter log files were inspected with Tracer to ensure convergence and successful  
637 mixing of the MCMC run (i.e. ESS>100). The 95% CI of the coefficient of variation estimated under the  
638 UCLD model excluded 0 (mean=0.56, 95% CI 0.36-0.8) indicating that the MDV sequences in this tree  
639 did not evolve under a strict clock.

#### 640 Pairwise divergence

641 Average pairwise divergence between ancient (OL1385) and modern sequences was computed using the  
642 R (v. 4.2.1) Ape package (65) and a window size of 100 bp and step size 25 bp. This analysis used the  
643 same genomic alignment used for other analyses in the study. Pairwise divergence was plotted on Fig. 2  
644 using Circos as described elsewhere.

#### 645 Functional validation

646 In order to functionally test the transactivation ability of the Meq oncogene in ancient vs. modern MDV  
647 strains, an *in vitro* study system was constructed. The full length coding sequence of Meq from the  
648 highest coverage ancient sample (OL1385; Buda Castle, Hungary), along with Meq from modern strains  
649 (RB1B and Md5), were codon optimized and purchased from Integrated DNA Technologies (IDT; IA,  
650 USA) with a 3× FLAG sequence appended to the N-terminus (Table S9). Since Meq preferentially forms  
651 a heterodimer with chicken c-Jun, the chicken c-Jun coding sequence was also synthesized with the  
652 addition of a 3× FLAG sequence on the N-terminus (UniProtKB accession: P18870.2). All Meq and c-Jun  
653 sequences were cloned into the mammalian expression vector pTarget (accession: AY540613) using the  
654 NotI and XmaI sites. A reporter construct comprising a 773 bp section of the Meq promoter (which  
655 contains the putative AP-1 binding site AGTCATGCATGACGT bound by Meq itself) upstream of the  
656 firefly luciferase gene on the pGL3-Basic vector backbone was a kind gift from Venugopal Nair  
657 (Pirbright Institute, UK). A further reporter construct to normalize transfection efficiency – pRL-TK –  
658 comprises the *Renilla* firefly luciferase downstream of the relatively weak constitutive HSV-thymidine  
659 kinase promoter and was purchased from Promega (UK). All vectors underwent whole-plasmid  
660 sequencing (Plasmidsaurus) prior to their use experimentally to validate sequence integrity.

661  
662 The chicken embryonic fibroblast cell line DF-1 was maintained in high glucose Dulbecco's Modified  
663 Eagle Medium with GlutaMAX (ThermoFisher, UK) and 10% fetal bovine serum (ThermoFisher, UK) in  
664 a humidified incubator at 37 °C and 5 % CO<sub>2</sub>. The day before transfection, cells were passaged and  
665 seeded into 96-well tissue culture plates at a density of 2×10<sup>4</sup> cells/well. Wells were transfected with a

666 Meq construct (240 ng), c-Jun (240 ng), Meq-pGL3 Basic reporter (200 ng) and pRL-TK reporter (4 ng)  
667 using 1.92 µl TransIT-2020 transfection reagent (Mirus; WI, USA). Twenty-four hours post transfection,  
668 cell supernatants were discarded and cells were washed in cold phosphate-buffered saline. Additional cell  
669 lysates were also prepared for immunoblotting of FLAG-tagged protein. Luciferase measurements were  
670 conducted using the Dual-Luciferase Reporter Assay System (Promega). Briefly, cells were incubated  
671 with 20 µl passive lysis buffer at room temperature for 15 minutes. In opaque white 96-well plates, the  
672 Dual-Luciferase assay was completed in a GloMAX plate reader (Promega), programmed to inject 100 µl  
673 Luciferase Assay Reagent II, then quench and read *Renilla* luciferase with 100 µl Stop & Glo Reagent,  
674 both with a 2-second delay and 10-second integration time.

675  
676 Over several experiments, it became apparent that the *Renilla* reporter did not serve as an adequate  
677 background control because *Renilla* luciferase varied considerably depending on whether Meq was  
678 present. *Renilla* luciferase measurements were substantially higher in wells containing Meq, suggesting  
679 that the TK promoter was being driven by the transactivation ability of Meq. The same phenomenon has  
680 been reported with pRL-SV40 constructs previously (66, 67). As a consequence, firefly luciferase values  
681 normalized to *Renilla* luciferase were artificially diminished in Meq-containing wells. Instead of  
682 normalizing to *Renilla* luciferase, we present the raw firefly luciferase measurements. We confirmed via  
683 immunoblotting that the expression of ancient Meq was comparable to that of other Meq constructs. The  
684 experiment was repeated on three independent occasions.

## 685 **Supplementary Text**

### 686 Sequence analyses of Meq gene: extended description

687 Given that Meq is known to be a major determinant of MDV virulence, we elected to consider this gene  
688 in more detail. Meq exerts transcriptional control on downstream targets via its C-terminal transactivation  
689 domain. This domain is characterized by PPPP (tetraproline) repeats, and the number of tetraproline  
690 repeats is inversely proportional to the virulence of the MDV strain (22). Standard Meq is 339 aa in  
691 length, but length variants of Meq exist: long(l)-Meq is 398 aa due to a tandem duplication in the  
692 transactivation domain; short(s)-Meq is 298 aa; very short(vs)-Meq is 247 aa (68). We conducted an  
693 analysis of 413 standard-length Meq sequences (comprising four ancient sequences and one modern  
694 sequence derived from the present study, and 408 modern sequences derived from public databases), and  
695 found that all modern Meq sequences have between two and five tetraproline repeats. However, all  
696 ancient sequences have an extra tetraproline repeat (totaling six) that has been disrupted in the modern  
697 lineage ([Table S8](#)).

698 We reconstructed the phylogeny of all Meq sequences using the RAxML (51) tree builder in Geneious (v.  
699 2019.2.3) with 100 bootstrap replicates. For each of the sequences, we also determined which of the

700 tetraproline repeats had been disrupted and the causative mutation for any disruptions. This information  
701 was plotted onto the Meq phylogeny using iTOL (v. 6.5.8; (69)) along with the internal branches where  
702 each disruption is likely to have taken place ([Fig. S7](#)).

703 Close inspection of the tetraproline motifs revealed that each motif has been lost at multiple points  
704 throughout MDV phylogeny, confirmed by the presence of independent disruptive mutations. Moreover,  
705 there is evidence that the loss of tetraproline motifs is ordered. Following the loss of the 6<sup>th</sup> tetraproline  
706 motif (which occurred in the common ancestor to all modern strains), the 4<sup>th</sup> tetraproline motif is lost  
707 independently in two major lineages of European and Asian/N. American strains. Next, the 2<sup>nd</sup>  
708 tetraproline motif is most commonly lost, followed by either the 5<sup>th</sup> or the 1<sup>st</sup> tetraproline motif, usually in  
709 terminal branches of the tree. The selective pressure for the ordered loss of the 4<sup>th</sup> then 2<sup>nd</sup> tetraproline  
710 appears to be particularly strong – in the major Asian/N. American lineage, the 2<sup>nd</sup> tetraproline was lost 6  
711 independent times. Intense selection pressure within the 2<sup>nd</sup> and 4<sup>th</sup> tetraproline motifs was also confirmed  
712 using a small number of Meq sequences derived from whole genomes in the positive selection analysis  
713 described above (where codons 176 and 217 were under selection in the 2<sup>nd</sup> and 4<sup>th</sup> tetraproline,  
714 respectively). Occasionally, the 3<sup>rd</sup> tetraproline is lost after the 6<sup>th</sup> tetraproline, but this typically occurs at  
715 terminal branches ([Fig. S7](#)).

716 The loss of tetraproline motifs appears to act as a ratchet, whereby each subsequent loss results in an  
717 increase in virulence, and once lost, motifs cannot be regained. This results in a stepwise scaling of the  
718 fitness landscape wherein the order of losses is important. Although there are some observations of virus  
719 lineages exhibiting an alternative loss order, such lineages are not widespread, suggesting that they may  
720 become stuck in local fitness peaks and are outcompeted by lineages following the order described above.

#### 721 Functional validation results: extended description

722  
723 Having identified the crucial oncogene Meq as being positively selected between the ancient and modern  
724 strains, we sought to test whether the polymorphisms translate into a change in Meq function. Meq is a  
725 regulator of transcriptional activity, and the strength of the transcriptional activation of target genes is  
726 strongly linked to virulence (70). In order to compare transcriptional activation between ancient and  
727 modern MDV strains, the Meq gene was cloned based on the sequence from the highest coverage ancient  
728 sample (OL1385; Buda Castle, Hungary) alongside Meq from three modern MDV strains (RB1B and  
729 Md5, both very virulent pathotypes). Meq forms a functional heterodimer with a chicken protein – c-Jun –  
730 which is required for transcriptional activation, and so chicken c-Jun was cloned and expressed alongside  
731 each Meq construct in DF-1 cells. As a reporter of transcriptional activity, the Meq promoter (which  
732 contains an AP-1 binding site) in an expression vector upstream of the firefly luciferase gene was co-  
733 transfected with the Meq/c-Jun pair.

734

735 As anticipated, Meq from the modern very virulent strains of MDV (RB-1B and Md5) showed the  
736 greatest transactivation ability, with RB-1B Meq exceeding a ten-fold enhancement of luciferase signal  
737 compared to empty vector (**Fig. 3c**). As previously reported, we found that c-Jun is important for Meq  
738 function, exemplified by RB1B Meq transactivation being severely abrogated without c-Jun co-  
739 expression. There was, however, a small transactivation effect of RB1B Meq in isolation, possibly as a  
740 result of endogenous c-Jun in DF-1 cells forming a small amount of functional heterodimer. Most  
741 significantly, we found that the ancient Meq sequence derived from the ancient Hungarian strain of MDV  
742 (OL1385), was a very weak transactivator. Transactivation from this construct was still significantly  
743 elevated relative to baseline, but was considerably less than Meq from modern strains. These findings  
744 support the hypothesis that ancient Meq, and thus ancient MDV, was considerably less virulent than  
745 modern strains.

746  
747

#### 748 BEAST results: extended description

749

750 The mean root age of the tree was estimated around 1483 AD, with the mean time to the most recent  
751 common ancestor (tMRCA) for the Eurasian and North American modern lineages tracing back to 1859  
752 and 1870 AD respectively. The mean UCLD clock rate was estimated to be 6.53E-6, an estimate that is in  
753 the same order of magnitude as other dsDNA viruses (71). The ages estimated for the ancient MDV  
754 samples OL1385 (1803), OL1389 (1802), OL1986 (1594) and OL2272 (1821) closely matched the  
755 estimated mean C14 dates (+/-1 year).

#### 756 Phylogenetic clustering analysis: extended description

757 A total of 4 phylogenetic clusters were identified by fastbaps ([Fig. S11](#)). The first cluster included the 4  
758 ancient samples with the highest genome coverage, followed by a Eurasian and a North American cluster,  
759 and a cluster of modern samples of both Eurasian and North American origin. As the clustering is  
760 performed without being constrained by the BEAST MCC tree, the identified phylogenetic clusters  
761 further corroborate the presence of a monophyletic Eurasian and North American lineages (as it can be  
762 observed in the topology estimated by BEAST) as well as the presence of a “transitional” lineage between  
763 the two modern MDV lineages.

764

#### 765 Read sharing between chicken and MDV: extended description

766 To test whether the chicken and MDV genomes share significant homology that could result in read  
767 sharing between the host and virus, we simulated short-read data from the MDV genome and mapped  
768 these against the chicken genome. We identified one region which shares significant homology between  
769 the chicken and MDV and could be a source of read sharing: MDV001/MDV080 is a RNA telomerase  
770 subunit (vTR), 435bp in length, that is presumed to be a recent gene capture from the chicken genome.

771 There was no evidence of significant read sharing between the chicken and MDV genomes – depth of  
772 coverage in our highest coverage ancient sample (OL1385) over the homologous region (coordinates:  
773 141336-141771; depth:  $58.8 \pm 24.8$ ) was in line with the wider region (coordinates:140000-142000; depth  
774  $77.0 \pm 47.9$ ).

775

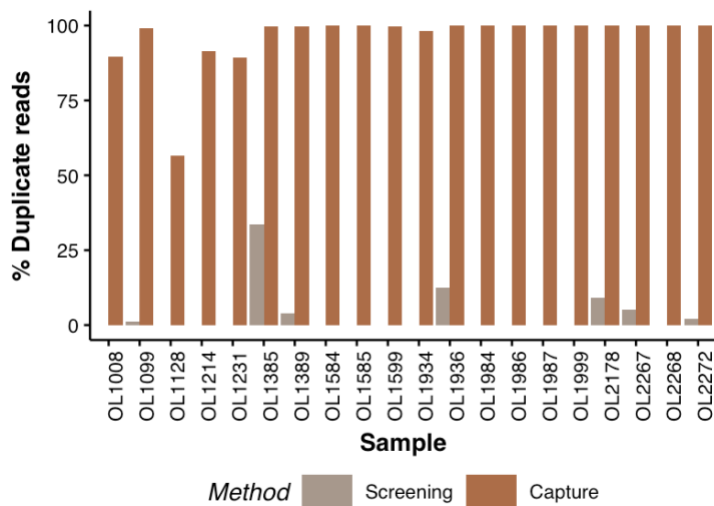
776

777

778

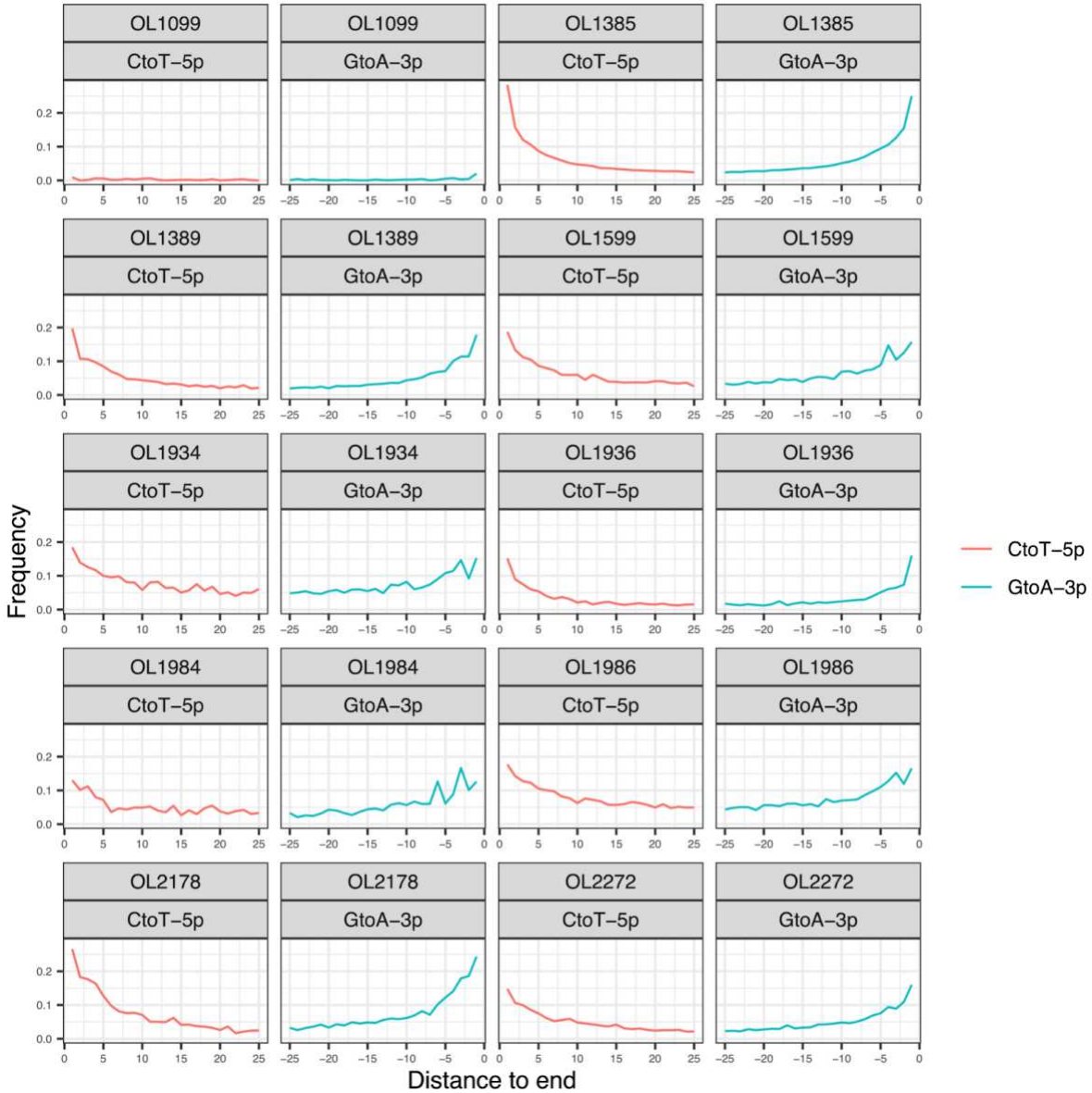
779 **Supplementary Figures**

780



781 **Fig. S1.**

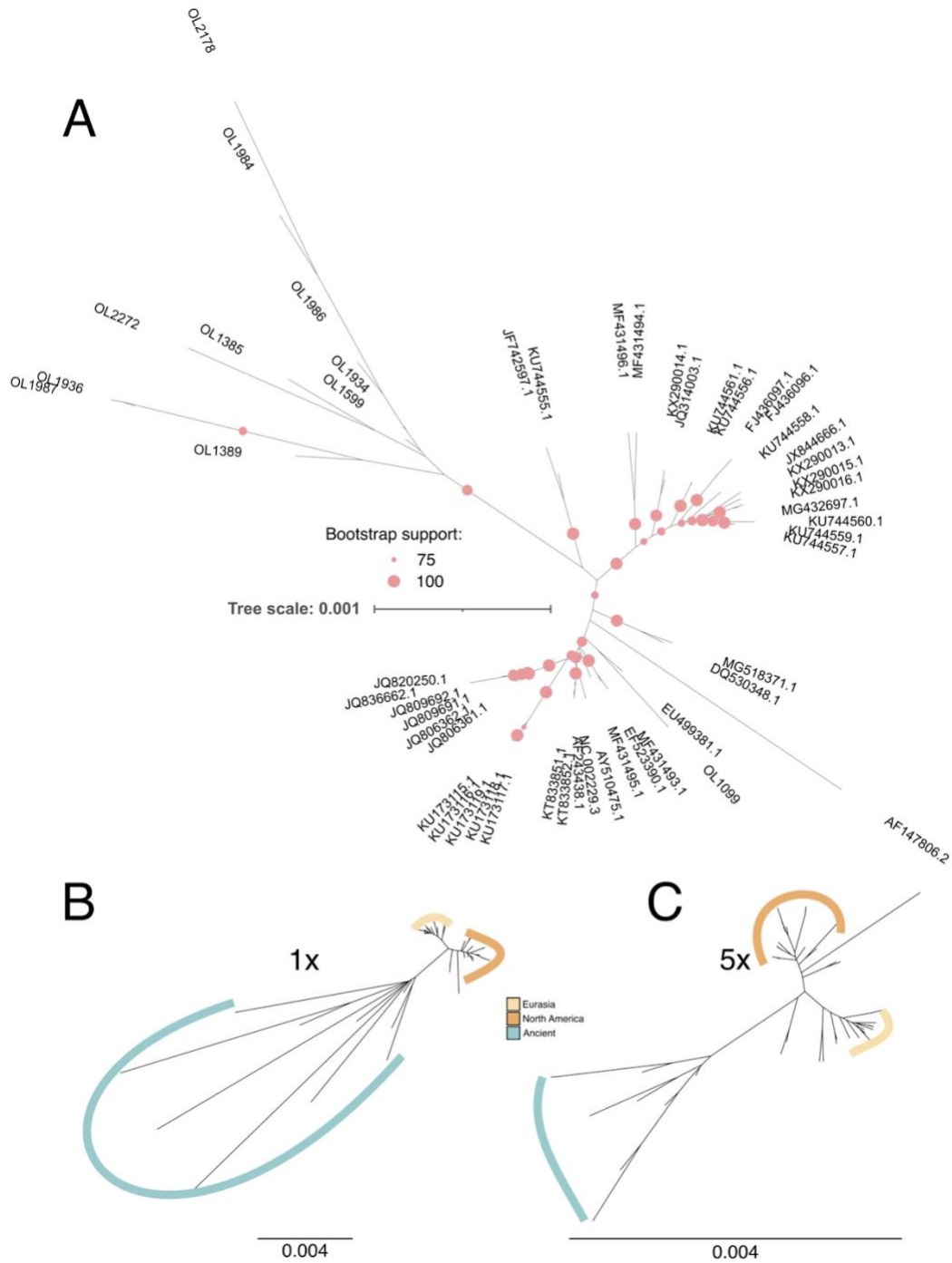
782 **Percentage of duplicated reads after sequencing of screening and bait-captured libraries.** In most  
783 cases, baiting of the sample resulted in near-saturation of duplicated reads, meaning almost all unique  
784 molecules in the sample were sequenced.



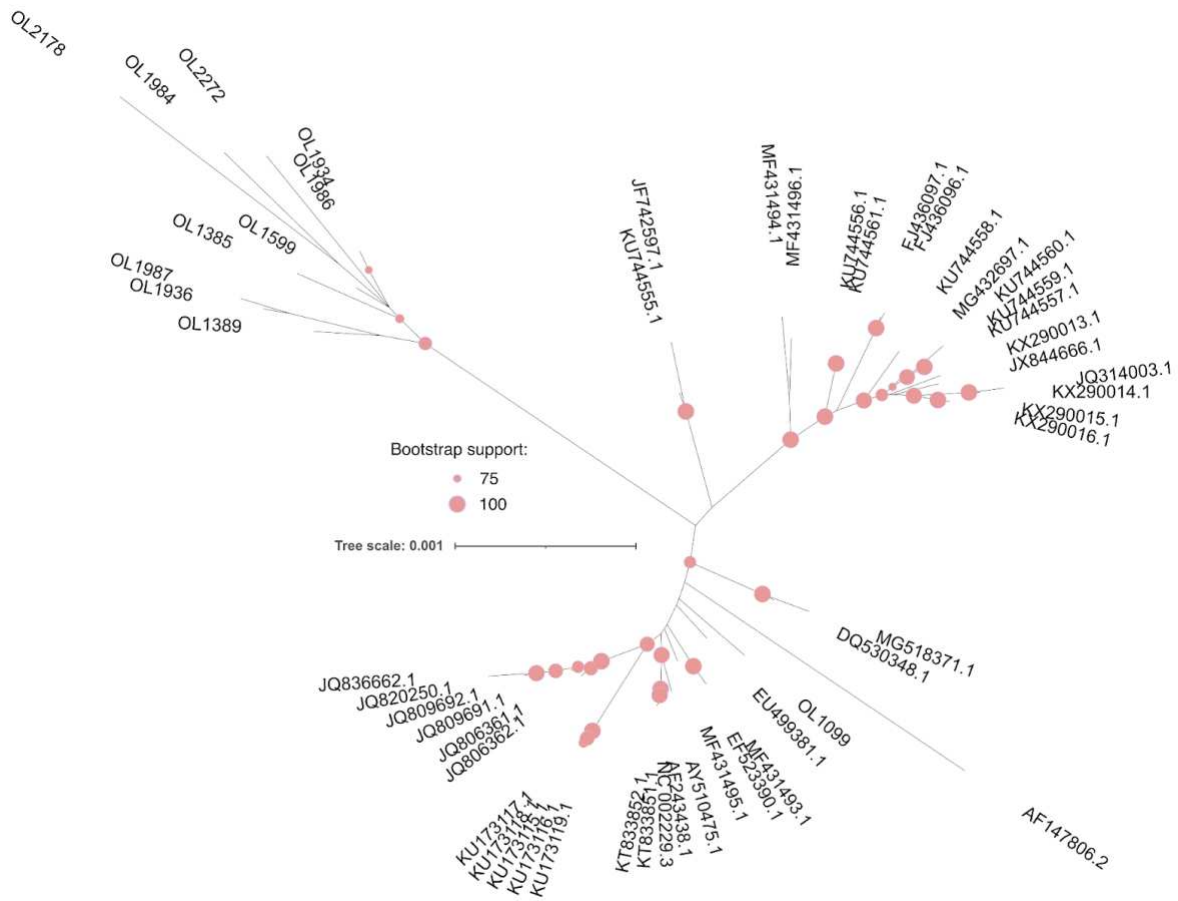
785

786 **Fig. S2.**

787 **DNA damage profiles of DNA derived from archaeological samples.** DNA reads from all samples  
 788 with genomic coverage >1x were analyzed using the MapDamage program (40) which assesses the rate of  
 789 C-to-T and G-to-A transitions as a result of the spontaneous deamination of cytosine. As a control, reads  
 790 from a modern sample (OL1099; 2014 CE) were also included.



791  
 792 **Fig. S3.**  
 793 **Neighbor-joining phylogenies of modern and ancient MDV genomes.** (A) Bootstraps of  $\geq 75$  are  
 794 displayed as filled circles. Captured samples (with ‘OL’ prefix) from the present study were included if  
 795 the genome had  $\geq 1\%$  coverage at 5x. Differences in branch lengths for ancient samples are highlighted  
 796 when using a depth of coverage threshold of 1x (B) and 5x (C).



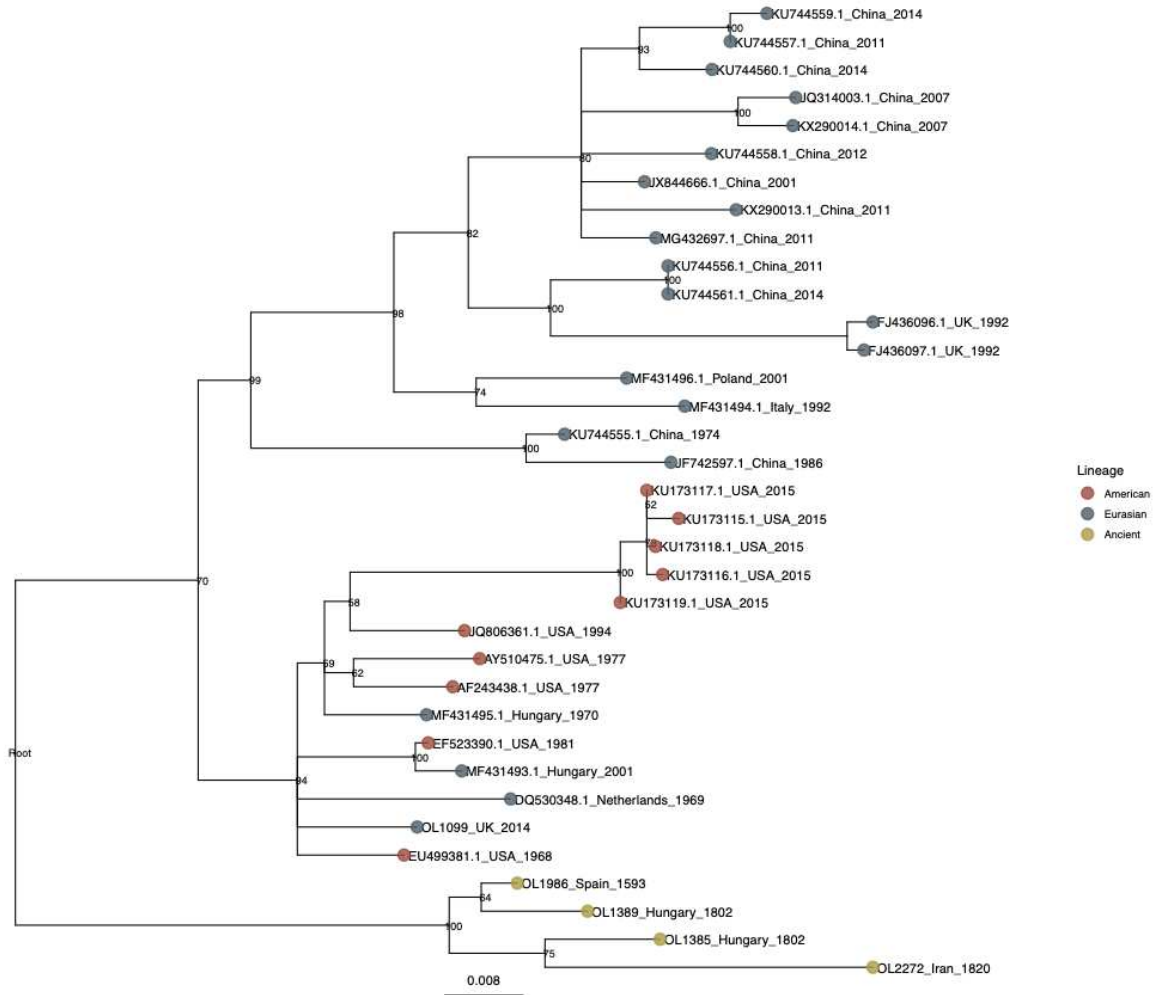
797

798 **Fig. S4.**

799 **Maximum likelihood phylogeny of modern and ancient MDV genomes.** Bootstraps of  $\geq 75$  are

800 displayed as filled circles. Captured samples (with 'OL' prefix) from the present study were included if

801 the genome had  $\geq 1\%$  coverage.



802

803 **Fig. S5.**

804 **Midpoint-rooted maximum likelihood tree using transversion SNPs only.** Tree was built using  
 805 RAxML (51), with ascertainment correction for SNPs, and included all ancient samples with a genomic  
 806 coverage of  $\geq 20\%$ . The placement of the root was confirmed using an outgroup (Fig. S6).

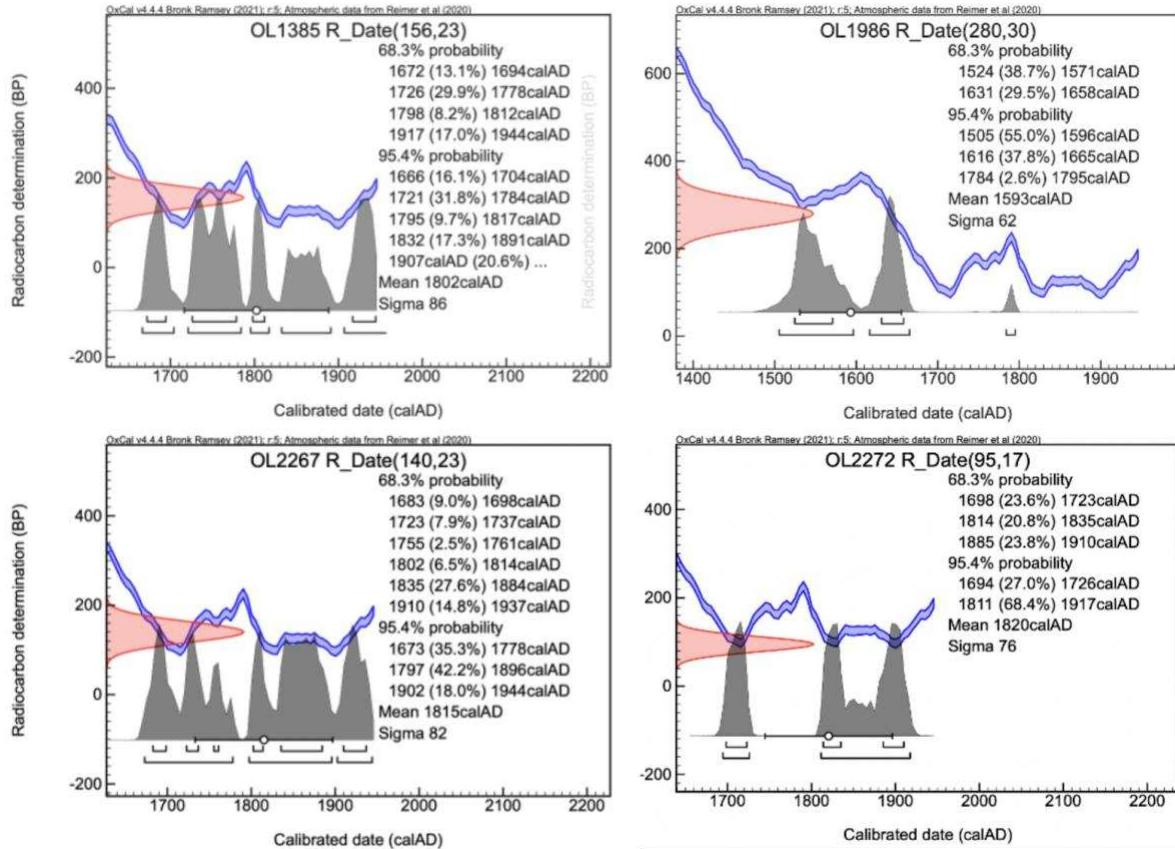




816

817 **Fig. S7.**

818 **Meq phylogeny showing ordered loss of tetraproline motifs.** The tree comprises 413 Meq sequences of  
 819 the standard length (339aa) built using RAxML with 100 bootstrap replicates and visualized in iTOL  
 820 (69). Around the edge, the integrity of each tetraproline motif is depicted (filled squares representing an  
 821 intact tetraproline and open squares for a disrupted tetraproline). Meq sequences from ancient samples  
 822 described herein are highlighted in gray, and the basal loss of the 6th tetraproline motif (common to all  
 823 modern strains) is shown in bold. The label connected by a blue dotted line indicates the polymorphism  
 824 that is found instead of tetraproline and the position of the tetraproline motif. For instance, ‘4 PAPP’  
 825 indicates that the 4<sup>th</sup> tetraproline motif is disrupted by a proline-to-alanine substitution in the second  
 826 proline position. ‘3 PP..P’ denotes a deletion of the 3<sup>rd</sup> proline in the 3<sup>rd</sup> tetraproline motif.



827

828 **Fig. S8.**

829 **Probability density plots for radiocarbon dating of ancient samples.** Calibration of dates was done in  
 830 OxCal (<https://c14.arch.ox.ac.uk/oxcal.html>) using the IntCal20: Northern Hemisphere (36) method.

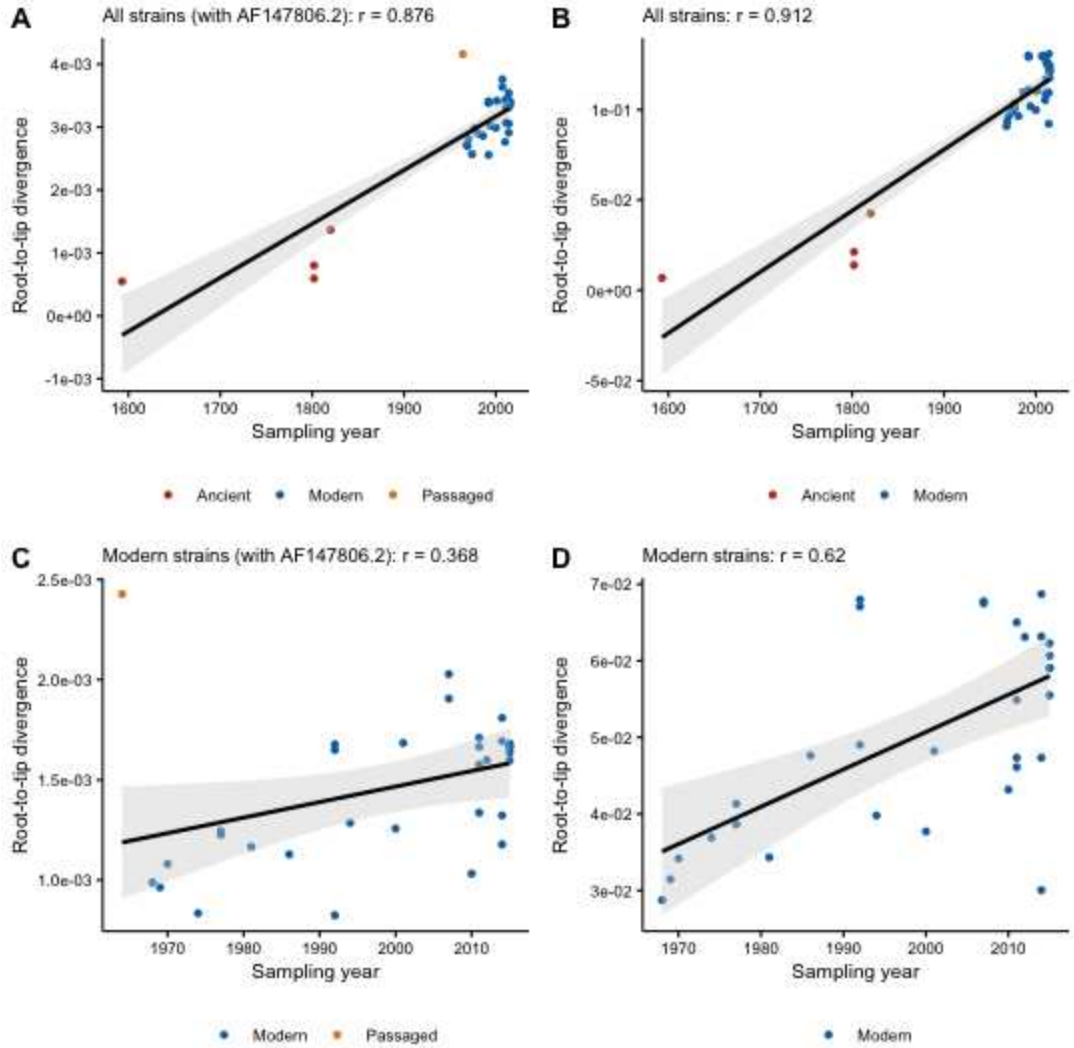
831

832

833

834

835

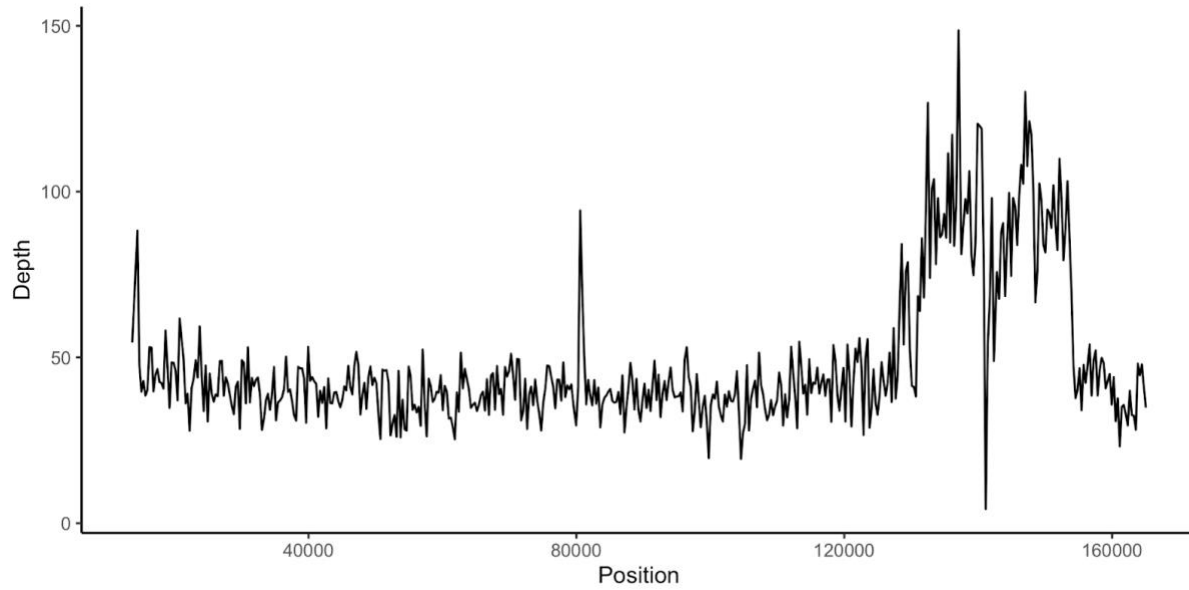


836

837 **Fig. S9.**

838 **Root-to-tip divergence against sampling date of Marek's Disease Virus strains and respective**  
 839 **Pearson correlation coefficients.** (A) Ancient and modern strains (including AF147806.2, which was  
 840 extensively passaged). (B) Ancient and modern strains (excluding AF147806.2). (C) Modern strains only  
 841 (including AF147806.2) (D) Modern strains only (excluding AF147806.2).

842



843

844 **Fig. S10.**

845 **Sequencing depth of the highest coverage ancient sample (OL1385).** Per-base sequencing depth

846 averaged over 300bp windows illustrating no significant drop-off in sequence depth across the genome.

847 The region of elevated coverage from ~130 kb is a duplicated region that encompasses the internal repeat

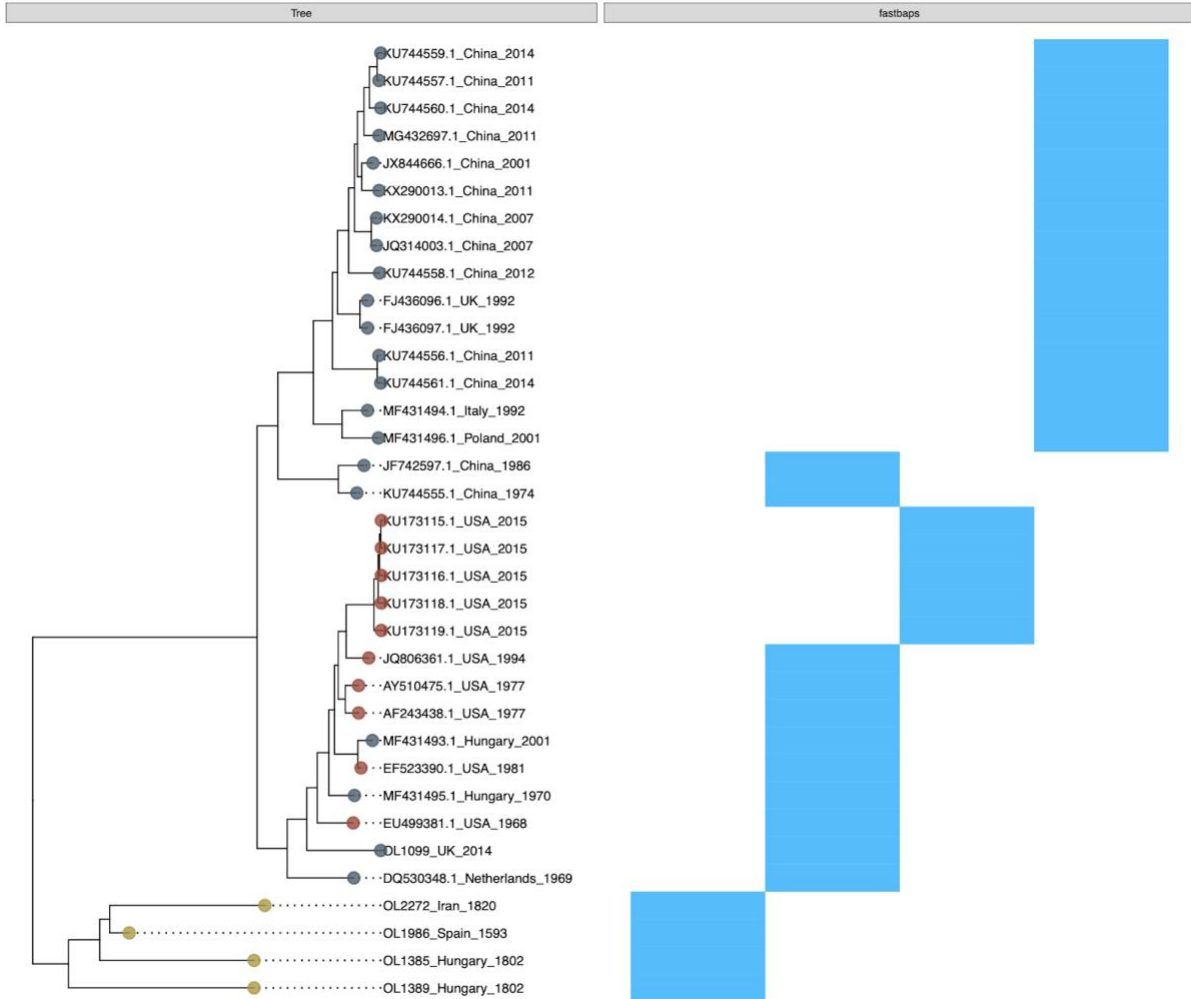
848 long (IRL) and the internal repeat short (IRS). The sharp decline in coverage at ~142 kb is a short

849 repetitive region that was not baited.

850

851

852



853

854 **Fig. S11.**

855 **Phylogenetic clustering analysis using fastbaps.** Phylogenetic clusters (blue bars) were identified in the  
 856 dataset using the fastbaps algorithm, under the baps prior model and unconstrained by a phylogenetic tree.

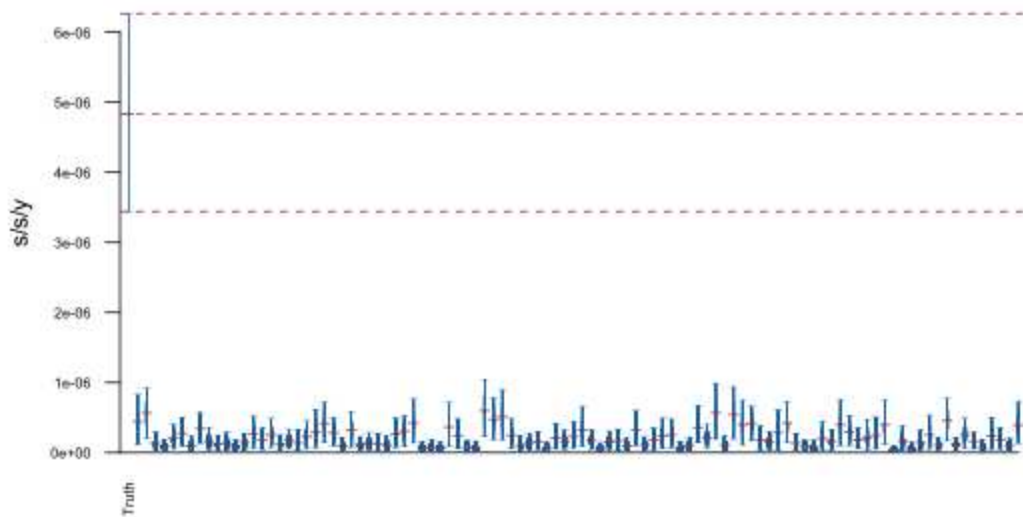
857

858

859

860

861



862

863 **Fig S12.**

864 **The Bayesian date randomization test (DRT) performed with 100 replicates, under a UCLD clock**

865 **model in BEAST v1.10.4.** The plot shows the posterior distributions for the mean clock rate, using the

866 true, unpermuted sampling dates (far left) and 100 replicates with sampling dates permuted among tips.

867 Distributions are truncated at the upper and lower limits of the 95% HPD interval and horizontal red lines

868 indicate the median estimates. The red dashed lines indicate the median and upper and lower limits of the

869 95% HPD interval of the clock rate inferred under the true sampling dates.

870

871

872

873 **Supplementary Tables**

874

875 **Table S4. Summary of estimated parameters from the ME and UCLD clock model BEAST runs.**

876

Mean ages tip dating runs	ME (95% HPD Interval)	UCLD (95% HPD Interval)
Root age	1453 (1301, 1569)	1483 (1349, 1576)
tMRCA(Modern)	1837 (1770 - 1902)	1825 (1751 - 1895)
tMRCA(American superclade)	1870 (1816, 1920)	1870 (1813, 1925)
tMRCA(Eurasian)	1866 (1807, 1921)	1859 (1791, 1919)
age OL1389	1803	1802
age OL1385	1802	1803
age OL1986	1593	1594
age OL2272	1820	1821
Mean clock rate		6.53E-6 (3.67E-6, 9.81E-6)
Clock rate Ancient/Background	1.33E-06 (8.43E-07, 2.03E-06)	
Clock rate American	1.63E-06 (1.03E-06, 2.70E-06)	
Clock rate Eurasian	2.33E-06 (1.37E-06, 3.77E-06)	
BF effect Ancient	inf	
BF effect North American	1.09	
BF effect Eurasian	5.25	

877

878

879

880

881

882  
883

884 **Table S9.**

<b>Strain</b>	<b>Year (CE)</b>	<b>Pathotype</b>	<b>Meq coding length</b>	<b>Sequence source</b>
OL1385 (ancient)	1802 ± 86	Unknown	1020 bp	This study
RB1B	1981	Very virulent	1020 bp	EF523390.1
Md5	1977	Very virulent	1020 bp	NC_002229.3

885  
886 **Meq sequences cloned for *in vitro* functional analysis.** Meq was cloned from the highest coverage  
887 ancient sample (OL1385) as well as from two modern strains.  
888

889 **Table S10.**

Sample	Country	Uncalibrated date (YBP)	Error +/-	Calibrated date	Error +/-	Laboratory	Lab Code
OL1385	Hungary	156	23	1802	86	Oxford RLAHA	OxA- 40466
OL2267	Iran	140	23	1815	82	Oxford RLAHA	OxA- 40467
OL2272	Iran	95	17	1820	76	Oxford RLAHA	OxA- 40491
OL1986	Spain	280	30	1593	62	Beta Analytic	Beta- 638008

890

891 **Radiocarbon dating of ancient chicken samples.** Raw dates were calibrated using IntCal20: Northern  
892 Hemisphere (36). See also [Figure S8](#) for probability density plots.

893

894

895

896 **Data S1.**

- 897 ● **Table S1: Sample metadata.** Metadata for all ancient samples sequenced in the present study.
- 898 ● **Table S2: Screening and capture sequencing results.** Sequencing statistics for all MDV-  
899 positive screened samples that were then submitted for in-solution capture.
- 900 ● **Table S3: Modern genome metadata.** Metadata for all modern MDV sequences included in  
901 phylogenetic analyses.
- 902 ● **Table S6: Fixed differences between ancient and modern MDV strains.** Table of single  
903 nucleotide polymorphisms found to be fixed between ancient strains (with at least 2  
904 representatives) and modern strains.
- 905 ● **Table S7: PAML results.** Results from the branch-site analysis of positive selection.
- 906 ● **Table S8: Meq metadata.** Metadata of all Meq sequences included in the Meq analysis, plus  
907 additional information about the tetraproline content for each sequence.
- 908 ● **Table S11: Pre-capture metagenomic IDs.** Summary of the pre-capture HAYSTAC results.
- 909 ● **Table S12: SNP summary table.** Summary table of all SNPs and depth of coverage in all  
910 ancient samples.
- 911 ● **Table S13: BEAST tip dates.** Tip dates used as priors in the BEAST analysis.

912

913 **References cited in Supplementary Materials**

- 914 25. O. Putelat, "Archéozoologie" in *Strasbourg, Bas-Rhin. Rue de Lucerne – Rue du Jeu-de-*  
915 *Paume. Rapport de fouille préventive. Volume 1. Le système défensif primitif et le processus*  
916 *d'urbanisation d'un secteur du faubourg de la Krutenau du Moyen Âge à nos jours. Rapport de*

- 917 *fouille préventive, Sélestat : Pôle d'Archéologie Interdépartemental Rhénan*, M. Werlé, Ed. (2015),  
918 pp. 98–174.
- 919 26. A. Cicović, D. Radičević, "Arheološka istraživanja srednjovekovnih nalazišta na Rudniku  
920 2009–2013. godine" in *Rudnik I, istraživanja srednjovekovnih nalazišta (2009-2013. godina)*,  
921 *Gornji Milanovac*, D. Radičević, A. Cicović, Eds. (2013), pp. 19–57.
- 922 27. N. Marković, J. Bulatović, "Rudnik 2009–2013: rezultati arheozoološke analize" in  
923 *Rudnik I, istraživanja srednjovekovnih nalazišta (2009-2013. godina)*, *Gornji Milanovac*, A.  
924 Cicović, D. Radičević, Eds. (2019), pp. 119–129.
- 925 28. H. Baron, *Quasi Liber Et Pictura. Die Tierknochenfunde aus dem Gräberfeld an der*  
926 *Wiener Csokorgasse – eine anthrozoologische Studie zu den awarischen Bestattungssitten.*  
927 *Monographien des RGZM. 143* (2018).
- 928 29. O. Putelat, thesis, Université de Paris 1 Panthéon-Sorbonne (2015).
- 929 30. M. Popović, *Manastir Studenica – arheološka otkrića* (Republički zavod za zaštitu  
930 spomenika kulture, Arheološki institut, Beograd, 2015).
- 931 31. N. Marković, "Ishrana u manastiru Studenica: arheozoološka svedočanstva" in *Manastir*  
932 *Studenica – arheološka otkrića*, M. Popović, Ed. (Beograd: Republički zavod za zaštitu spomenika  
933 kulture i Arheološki institut, 2015), pp. 395–406.
- 934 32. I. Živaljević, N. Marković, M. Maksimović, Food worthy of kings and saints: fish  
935 consumption in the medieval monastery Studenica (Serbia). *anth. 54*, 179–201 (2019).
- 936 33. Marković, N., Radišić, T. & Bikić, "Uloga živine u srednjovekovnoj ekonomiji manastira  
937 Studenice" in *Bioarheologija na Balkanu. Metodološke, komparativne i rekonstruktivne studije*  
938 *života u prošlosti*, M.-R. N. Vitezović S., Ed. (2016), pp. 99–116.
- 939 34. A. Saed Mucheshi, M. Nikzad, M. Zamani-Dadaneh, Rescue excavations at Bardeh Mar,  
940 Darian Dam area, Hawraman, Kurdistan, western Iran. *Proceedings of the 15th* (2017).
- 941 35. M. Mashkour, A. Mohaseb, S. Amiri, S. Beyzaiedoust, R. Khazaeli, H. Davoudi, H.  
942 Fathi, S. Komijani, A. Aliyari, H. Laleh, Archaeozoological Report of the Bioarchaeology  
943 Laboratory of the University of Tehran and the Osteology Department of the National Museum of  
944 Iran, 2015-2016. *Proceedings of the 15th Annual Symposium on the Iranian Archaeology, 5-7 march*  
945 *2017, Tehran, Iranian Center for Archaeological Research*, 803–807 (2017).
- 946 36. P. J. Reimer, W. E. N. Austin, E. Bard, A. Bayliss, P. G. Blackwell, C. B. Ramsey, M.  
947 Butzin, H. Cheng, R. Lawrence Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, I. Hajdas,  
948 T. J. Heaton, A. G. Hogg, K. A. Hughen, B. Kromer, S. W. Manning, R. Muscheler, J. G. Palmer, C.  
949 Pearson, J. van der Plicht, R. W. Reimer, D. A. Richards, E. Marian Scott, J. R. Southon, C. S. M.  
950 Turney, L. Wacker, F. Adolphi, U. Büntgen, M. Capano, S. M. Fahrni, A. Fogtmann-Schulz, R.

- 951 Friedrich, P. Köhler, S. Kudsk, F. Miyake, J. Olsen, F. Reinig, M. Sakamoto, A. Sookdeo, S.  
952 Talamo, The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP).  
953 *Radiocarbon*. **62**, 725–757 (2020).
- 954 37. J. Dabney, M. Knapp, I. Glocke, M.-T. Gansauge, A. Weihmann, B. Nickel, C.  
955 Valdiosera, N. García, S. Pääbo, J.-L. Arsuaga, M. Meyer, Complete mitochondrial genome  
956 sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc.*  
957 *Natl. Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).
- 958 38. M.-T. Gansauge, M. Meyer, Selective enrichment of damaged DNA molecules for  
959 ancient genome sequencing. *Genome Res.* **24**, 1543–1549 (2014).
- 960 39. C. Carøe, S. Gopalakrishnan, L. Vinner, S. S. T. Mak, M. H. S. Sinding, J. A. Samaniego,  
961 N. Wales, T. Sicheritz-Pontén, M. T. P. Gilbert, Single-tube library preparation for degraded DNA.  
962 *Methods Ecol. Evol.* **9**, 410–419 (2018).
- 963 40. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: fast  
964 approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. **29**, 1682–1684  
965 (2013).
- 966 41. M. Schubert, A. Ginolhac, S. Lindgreen, J. F. Thompson, K. A. S. Al-Rasheid, E.  
967 Willerslev, A. Krogh, L. Orlando, Improving ancient DNA read mapping against modern reference  
968 genomes. *BMC Genomics*. **13**, 178 (2012).
- 969 42. G. Jun, M. K. Wing, G. R. Abecasis, H. M. Kang, An efficient and scalable analysis  
970 framework for variant extraction and refinement from population-scale DNA sequence data. *Genome*  
971 *Res.* **25**, 918–925 (2015).
- 972 43. Broad Institute, *Picard toolkit* (Broad Institute, 2019);  
973 <http://broadinstitute.github.io/picard/>).
- 974 44. G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-  
975 Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S.  
976 Gabriel, M. A. DePristo, From FastQ data to high confidence variant calls: the Genome Analysis  
977 Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics*. **43**, 11.10.1–11.10.33 (2013).
- 978 45. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*.  
979 **9**, 357–359 (2012).
- 980 46. G. Tonkin-Hill, J. A. Lees, S. D. Bentley, S. D. W. Frost, J. Corander, Fast hierarchical  
981 Bayesian analysis of population structure. *Nucleic Acids Res.* **47**, 5539–5549 (2019).
- 982 47. J. Corander, P. Martinen, Bayesian identification of admixture events using multilocus  
983 molecular markers. *Mol. Ecol.* **15**, 2833–2843 (2006).

- 984 48. M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, A. Rambaut,  
985 Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016  
986 (2018).
- 987 49. G. Yu, Using ggtree to Visualize Data on Tree-Like Structures. *Curr. Protoc.*  
988 *Bioinformatics.* **69**, e96 (2020).
- 989 50. M. Krzywinski, J. Schein, Í. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M.  
990 A. Marra, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645  
991 (2009).
- 992 51. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of  
993 large phylogenies. *Bioinformatics.* **30**, 1312–1313 (2014).
- 994 52. A. J. Page, B. Taylor, A. J. Delaney, J. Soares, T. Seemann, J. A. Keane, S. R. Harris,  
995 SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom.* **2**,  
996 e000056 (2016).
- 997 53. P. O. Lewis, A likelihood approach to estimating phylogeny from discrete morphological  
998 character data. *Syst. Biol.* **50**, 913–925 (2001).
- 999 54. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7:  
1000 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- 1001 55. P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg,  
1002 T. Hamelryck, F. Kauff, B. Wilczynski, M. J. L. de Hoon, Biopython: freely available Python tools  
1003 for computational molecular biology and bioinformatics. *Bioinformatics.* **25**, 1422–1423 (2009).
- 1004 56. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for  
1005 FASTA/Q File Manipulation. *PLoS One.* **11**, e0163962 (2016).
- 1006 57. H. Li, seqtk Toolkit for processing sequences in FASTA/Q formats. *GitHub.* **767**, 69  
1007 (2012).
- 1008 58. S. Duchêne, D. Duchêne, E. C. Holmes, S. Y. W. Ho, The Performance of the Date-  
1009 Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data. *Mol. Biol. Evol.* **32**,  
1010 1895–1906 (2015).
- 1011 59. A. Rieux, F. Balloux, Inferences from tip-calibrated phylogenies: a review and a practical  
1012 guide. *Mol. Ecol.* **25**, 1911–1924 (2016).
- 1013 60. M. Navascués, F. Depaulis, B. C. Emerson, Combining contemporary and ancient DNA  
1014 in population genetic and phylogeographical studies. *Mol. Ecol. Resour.* **10**, 760–772 (2010).

- 1015 61. M. Molak, M. A. Suchard, S. Y. W. Ho, D. W. Beilman, B. Shapiro, Empirical calibrated  
1016 radiocarbon sampler: a tool for incorporating radiocarbon-date and calibration error into Bayesian  
1017 phylogenetic analyses of ancient DNA. *Mol. Ecol. Resour.* **15**, 81–86 (2015).
- 1018 62. F. Rodríguez, J. L. Oliver, A. Marín, J. R. Medina, The general stochastic model of  
1019 nucleotide substitution. *J. Theor. Biol.* **142**, 485–501 (1990).
- 1020 63. Z. Yang, Maximum likelihood phylogenetic estimation from DNA sequences with  
1021 variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306–314 (1994).
- 1022 64. A. J. Drummond, S. Y. W. Ho, M. J. Phillips, A. Rambaut, Relaxed phylogenetics and  
1023 dating with confidence. *PLoS Biol.* **4**, e88 (2006).
- 1024 65. B. Pfeifer, U. Wittelsbürger, S. E. Ramos-Onsins, M. J. Lercher, PopGenome: an  
1025 efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936  
1026 (2014).
- 1027 66. D. K. Ajithdoss, S. M. Reddy, P. F. Suchodolski, L. F. Lee, H.-J. Kung, B. Lupiani, In  
1028 vitro characterization of the Meq proteins of Marek’s disease virus vaccine strain CVI988. *Virus Res.*  
1029 **142**, 57–67 (2009).
- 1030 67. T. Huszár, I. Mucsi, T. Terebessy, A. Masszi, S. Adamkó, C. Jeney, L. Rosivall, The use  
1031 of a second reporter plasmid as an internal standard to normalize luciferase activity in transient  
1032 transfection experiments may lead to a systematic error. *J. Biotechnol.* **88**, 251–258 (2001).
- 1033 68. K.-S. Chang, K. Ohashi, M. Onuma, Diversity (polymorphism) of the meq gene in the  
1034 attenuated Marek’s disease virus (MDV) serotype 1 and MDV-transformed cell lines. *J. Vet. Med.*  
1035 *Sci.* **64**, 1097–1101 (2002).
- 1036 69. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and  
1037 annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–5 (2016).
- 1038 70. J. Sato, S. Murata, Z. Yang, B. B. Kaufer, S. Fujisawa, H. Seo, N. Maekawa, T.  
1039 Okagawa, S. Konnai, N. Osterrieder, M. S. Parcells, K. Ohashi, Effect of Insertion and Deletion in  
1040 the Meq Protein Encoded by Highly Oncogenic Marek’s Disease Virus on Transactivation Activity  
1041 and Virulence. *Viruses.* **14** (2022), doi:10.3390/v14020382.
- 1042 71. C. Firth, A. Kitchen, B. Shapiro, M. A. Suchard, E. C. Holmes, A. Rambaut, Using time-  
1043 structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol. Biol. Evol.* **27**,  
1044 2038–2051 (2010).
- 1045  
1046  
1047  
1048  
1049