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

Authors:

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

Affiliations:

[1]Department of Biology, University of Oxford, Oxford, UK

[2]The Palaeogenomics & Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford, UK

[3]Department of Veterinary Medicine, University of Cambridge, Cambridge, UK

[4]Centre d'Anthropobiologie et de Génomique de Toulouse, Toulouse, France

[5]Instituto Nacional de Antropología y Pensamiento Latinoamericano, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

[6]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

[7]Swiss Institute of Bioinformatics, Lausanne, Switzerland

[8]Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium

[9]Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium

[10]BioArCh, Department of Archaeology, University of York, York, UK

[11]Institute of Archaeology, Belgrade, Serbia

[12]Institute of Immunity and Transplantation, University College London, London, UK

[13]Silkie Club of Great Britain, Charing, UK

[14]Fishbourne Roman Palace, Fishbourne, UK

[15]Leibniz-Zentrum für Archäologie, Mainz, Germany

[16]Medieval Department, Budapest History Museum, Budapest, Hungary

[17]Laboratory of Biomonitoring, The Institute of Problems in Ecology and Mineral Wealth, Tatarstan Academy of Sciences, Kazan, Russia

[18]Archéologie Alsace - PAIR, Bas-Rhin, France

[19]Departament de Prehistòria, Universitat Autònoma de Barcelona, Barcelona, Spain

[20]Bioarchaeology Laboratory, Central Laboratory, University of Tehran, Tehran, Iran

[21]Department of Art and Architecture, Payame Noor University (PNU), Tehran, Iran

[22]Provincial Office of the Iranian Center for Cultural Heritage, Handicrafts and Tourism Organisation, Bojnord, Iran

[23]Department of Archaeology, School of History, Nanjing University, China

[24]Headland Archaeology, Edinburgh, UK

[25]Department of Archaeology, University of Exeter, Exeter, UK

[26]Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden

[27]CNRS, National Museum Natural History Paris, Paris, France

[28]Department of Pathobiology and Population Sciences, Royal Veterinary College, London, UK

[29]Viral Oncogenesis Group, Pirbright Institute, Woking, UK

[30]Department of Veterinary Sciences, Ludwig Maximilian University of Munich, Munich, Germany

[31]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK

[†]joint-first author

[‡] co-senior authors

[*]corresponding authors. Emails: steven.fiddaman@biology.ox.ac.uk; adrian.smith@biology.ox.ac.uk; laurent.frantz@lmu.de

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

One sentence summary:
Functional paleogenomics reveals the molecular basis for increased virulence in Marek's Disease Virus.

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

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

Marek's disease virus has been circulating in Europe for at least 1000 years
To empirically track the evolutionary change in MDV virulence through time, we generated MDV genome sequences (serotype 1) isolated from the skeletal remains of archeological chickens. We first shotgun sequenced 995 archeological chicken samples excavated from >140 Western Eurasian archeological sites and screened for MDV reads using HAYSTAC(*6*) with a herpesvirus-specific database. Samples with any evidence of MDV reads were then enriched for viral DNA using a

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

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

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

Ancient MDV strains are basal to modern lineages

To investigate the relationship between ancient and modern MDV strains, we built phylogenetic trees based on both neighbor-joining (NJ) and maximum likelihood (ML) methods. We first built trees using 10 ancient genomes with at least 1% coverage at a depth of ≥5x, a modern positive control derived from the present study (OL1099), and 42 modern genomes from public sources (Table S3). Both NJ (Fig. 1b, Fig. S3) and ML trees (Fig. S4) match the previously described general topology(*7*), in which Eurasian and North American lineages were evident, along with a well-supported (bootstrap: 94) ancient clade (Fig 1b). The same topology was also obtained when restricting our ML analysis to include only transversion sites (Fig. S5). Lastly, we built a tree using an outgroup (Meleagrid herpesvirus 1, accession: NC_002641.1) to root our topology (Fig. S6). We obtained a well-supported topology showing that the ancient MDV sequences form a highly supported clade lying basal to all modern MDV strains (including the modern positive control OL1099).

Next, we built a time-calibrated phylogeny using BEAST (v. 1.10;(*8*)) that included 31 modern genomes collected since 1968 (Table S3), and four ancient samples with an average depth of coverage >5× (OL1986, Castillo de Montsoriu, Spain, 1593 cal. CE; OL1385, Buda Castle, Hungary, 1802 cal. CE; OL1389, an additional Buda Castle sample from the same archeological context as OL1385; OL2272, Naderi Tepe, Iran, 1820 cal. CE; Table S1-S2, Fig. 1a). All of the ancient samples were phylogenetically basal to all modern MDV strains. The time of the most recent common ancestor (TMRCA) of the phylogeny was 1602 CE (95% HPD interval 1486 - 1767; Fig. 1c, Table S4).

As previously reported(*7*) we found that, aside from a few exceptions, most Eurasian and North American MDV strains formed distinct clades (Fig. 1b), suggesting that there has been little recent transatlantic exchange of the virus. The inclusion of time-stamped ancient MDV sequences improved the accuracy of the molecular clock analysis, and pushed back the TMRCA of all modern MDV sequences, from 1922-1952(*7*) to 1881 (95% HPD interval 1822 - 1929; Table S4). Our mean TMRCA of modern MDV is concordant with a recent estimate that incorporated 26 modern MDV genomes from East Asian chickens (1880, 95% HPD 1772-1968;(*9*)). This phylogenetic analysis implies that the two major modern clades of MDV were likely established before the earliest documented increases in MDV virulence in the 1920s. Furthermore, since birds infected with highly virulent MDV would not have survived a transatlantic crossing, a TMRCA of 1938 (95% HPD 1914 - 1958) for the clade containing the earliest North American sample (CU2, 1968; accession: EU499381.1) could be consistent with the virus having been transmitted prior to the most significant virulence increases leading up to the 1960s. These results are also consistent with the hypothesis that Eurasian and North American MDV lineages independently evolved towards increased virulence(*7*).

Virulence factors are among positively selected genes in the modern MDV lineage

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

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

PAML(*13*) (ancient sequences as background lineage, modern sequences as foreground lineage) on open reading frames (ORFs) using four ancient MDV genomes (OL1385, OL1389, OL1986 and OL2272). After controlling the false discovery rate using the Benjamini-Hochberg procedure(*14*), this analysis identified 49 ORFs with significant evidence for positive selection (Fig. 2; Table S7).

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

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

Positive selection was also detected in the viral chemokine termed viral interleukin-8 (considered a functional ortholog of chicken CXC ligand 13;(*18*)). Viral IL-8 is an important virulence factor that recruits B cells for lytic replication and CD4+ CD25+ T cells that are transformed to generate lymphoid tumors. Viruses that lack vIL-8 are severely impaired in the establishment of infection and generation of tumors through bird-to-bird transmission(*19*), so sequence variation in this gene could plausibly impact transmission.

The key oncogene of MDV has experienced positive selection and an ordered loss of tetraproline motifs

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

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

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

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

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

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

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

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

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

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

Conclusion

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

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

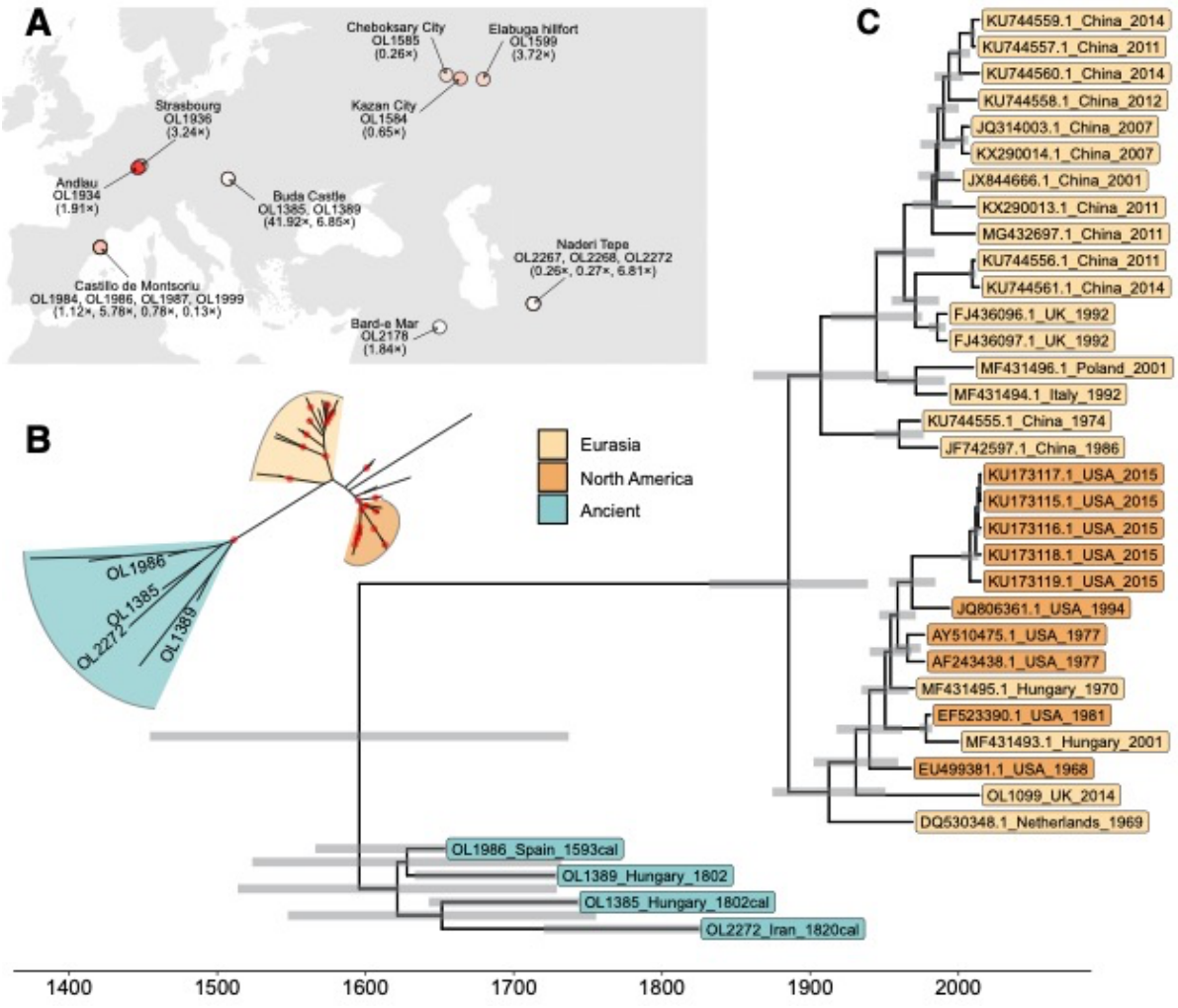

Fig. 1. Locations of MDV-positive samples and time-scaled phylogeny. (**A**) Map showing the locations of screened archeological chicken samples that were positive for MDV sequence. Colored circles indicate sample dates (either from calibrated radiocarbon dating or estimated from archeological context; Table S1). Average sequencing depth following capture is given in parentheses under sample names. If more than one sample was derived from the same site, this is indicated by a list of sample identifiers (beginning 'OL') and sequencing depths in parentheses. (B) Unrooted neighbor-joining tree of 42 modern and 10 ancient genomes. Only the four high-coverage ancient samples used in our BEAST analysis were labeled in this tree (Table S2). Nodes with bootstrap support of >90 are indicated by red dots. (C) Time-scaled maximum clade credibility tree of ancient and modern MDV sequences using the uncorrelated lognormal relaxed clock model (UCLD) and the general time-reversible (GTR) substitution model. Gray bars indicate the 95% highest posterior density (HPD) for the age of each node. The 'cal' suffix for ancient samples indicates that samples were radiocarbon dated and these date distributions were used as priors for the molecular clock analyses(*24*).

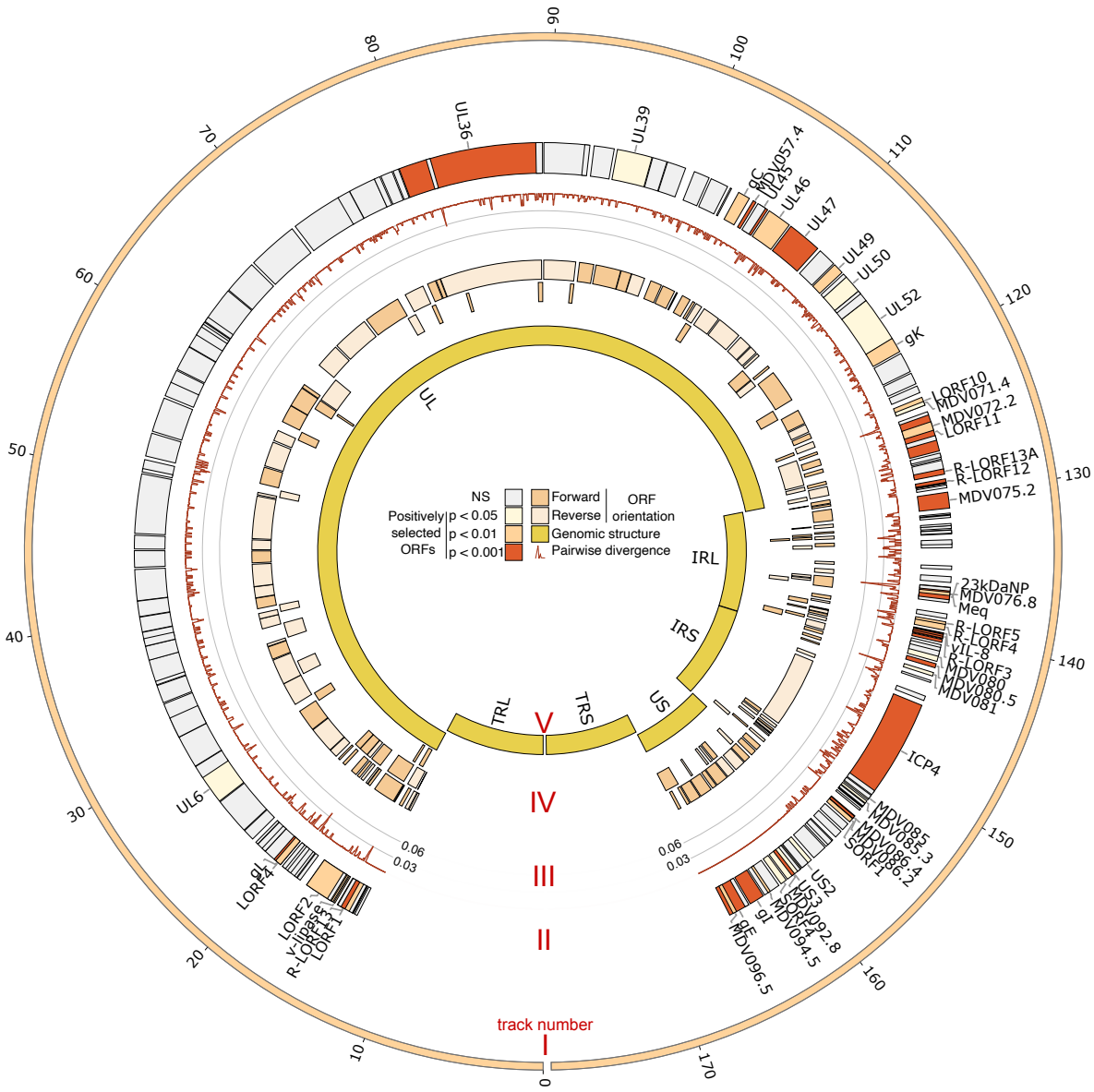

Fig. 2. Branch-site selection analysis of MDV genomes. The MDV genome is represented as a circular structure with gross genomic architecture displayed on the innermost track (track V) and genomic coordinates shown on the outermost track (units: $\times 10^3$ kb; track I). Since the long terminal repeat (TRL) and short terminal repeat (TRS) are copies of the long internal repeat (IRL) and the short internal repeat (IRS), respectively, selection analysis excluded the TRL and the TRS regions, leaving only the unique long (UL) and unique short (US) regions along with the two internal repeats. Results of the positive selection analysis are displayed on track II, where open reading frames (ORFs) are shaded according to the strength of statistical support (corrected P-values) for positive selection. Sliding window average pairwise divergence between ancient and modern samples is shown on track III, and ORF orientation is shown on track IV.

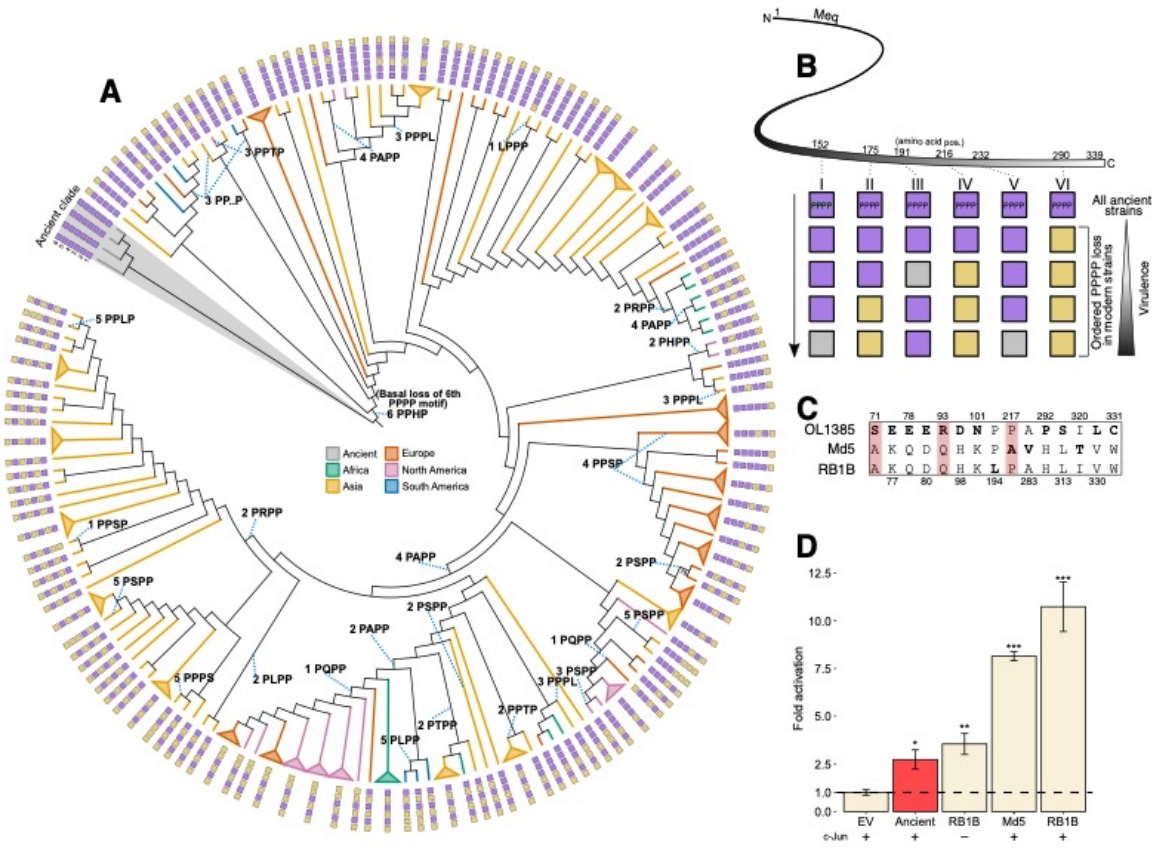

Fig. 3. *Meq* has undergone ordered loss of tetraproline repeats and increased transactivation ability. (A) Phylogenetic analysis of 412 *Meq* sequences of standard length (1017 bp). The outermost track shows the integrity of each tetraproline motif (purple squares = intact; yellow squares = disrupted).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

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

References

1. J. Marek, Multiple Nervenentzündung (Polyneuritis) bei Hühnern. *Dtsch. Tierarztl. Wochenschr.* **15**, 417–421 (1907).

2. C. Morrow, F. Fehler, "5 - Marek's disease: A worldwide problem" in *Marek's Disease*, F. Davison, V. Nair, Eds. (Academic Press, Oxford, 2004), pp. 49–61.

3. A. F. Read, S. J. Baigent, C. Powers, L. B. Kgosana, L. Blackwell, L. P. Smith, D. A. Kennedy, S. W. Walkden-Brown, V. K. Nair, Imperfect Vaccination Can Enhance the Transmission of Highly Virulent Pathogens. *PLoS Biol.* **13**, e1002198 (2015).

4. C. S. Eidson, K. W. Washburn, S. C. Schmittle, Studies on acute Marek's disease. 9. Resistance to MD by inoculation with the GA isolate. *Poult. Sci.* **47**, 1646–1648 (1968).

5. N. Osterrieder, J. P. Kamil, D. Schumacher, B. K. Tischer, S. Trapp, Marek's disease virus: from miasma to model. *Nat. Rev. Microbiol.* **4**, 283–294 (2006).

6. E. A. Dimopoulos, A. Carmagnini, I. M. Velsko, C. Warinner, G. Larson, L. A. F. Frantz, E. K. Irving-Pease, HAYSTAC: A Bayesian framework for robust and rapid species identification in high-throughput sequencing data. *PLoS Comput. Biol.* **18**, e1010493 (2022).

7. J. Trimpert, N. Groenke, M. Jenckel, S. He, D. Kunec, M. L. Szpara, S. J. Spatz, N. Osterrieder, D. P. McMahon, A phylogenomic analysis of Marek's disease virus reveals independent paths to virulence in Eurasia and North America. *Evol. Appl.* **10**, 1091–1101 (2017).

8. A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).

9. K. Li, Z. Yu, X. Lan, Y. Wang, X. Qi, H. Cui, L. Gao, X. Wang, Y. Zhang, Y. Gao, C. Liu, Complete genome analysis reveals evolutionary history and temporal dynamics of Marek's disease virus. *Front. Microbiol.* **13**, 1046832 (2022).

10. K. Majander, S. Pfrengle, A. Kocher, J. Neukamm, L. du Plessis, M. Pla-Díaz, N. Arora, G. Akgül, K. Salo, R. Schats, S. Inskip, M. Oinonen, H. Valk, M. Malve, A. Kriiska, P. Onkamo, F. González-Candelas, D. Kühnert, J. Krause, V. J. Schuenemann, Ancient Bacterial Genomes Reveal a High Diversity of Treponema pallidum Strains in Early Modern Europe. *Curr. Biol.* **30**, 3788–3803.e10 (2020).

11. B. Mühlemann, L. Vinner, A. Margaryan, H. Wilhelmson, C. de la Fuente Castro, M. E. Allentoft, P. de Barros Damgaard, A. J. Hansen, S. Holtsmark Nielsen, L. M. Strand, J. Bill, A. Buzhilova, T. Pushkina, C. Falys, V. Khartanovich, V. Moiseyev, M. L. S. Jørkov, P. Østergaard Sørensen, Y. Magnusson, I. Gustin, H. Schroeder, G. Sutter, G. L. Smith, C. Drosten, R. A. M. Fouchier, D. J. Smith, E. Willerslev, T. C. Jones, M. Sikora, Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age. *Science*. **369** (2020), doi:10.1126/science.aaw8977.

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. Zhang, J. Luo, The significance of the individual Meq-clustered miRNAs of Marek's disease virus in oncogenesis. *J. Gen. Virol.* **96**, 637–649 (2015).

13. Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

14. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).

15. A. R. Omar, K. A. Schat, Syngeneic Marek's disease virus (MDV)-specific cell-mediated immune responses against immediate early, late, and unique MDV proteins. *Virology*. **222**, 87–99 (1996).

16. C. J. Markowski-Grimsrud, K. A. Schat, Cytotoxic T lymphocyte responses to Marek's disease herpesvirus-encoded glycoproteins. *Vet. Immunol. Immunopathol.* **90**, 133–144 (2002).

17. S. Halabi, M. Ghosh, S. Stevanović, H.-G. Rammensee, L. D. Bertzbach, B. B. Kaufer, M. C. Moncrieffe, B. Kaspers, S. Härtle, J. Kaufman, The dominantly expressed class II molecule from a resistant MHC haplotype presents only a few Marek's disease virus peptides by using an unprecedented binding motif. *PLoS Biol.* **19**, e3001057 (2021).

18. S. Haertle, I. Alzuheir, F. Busalt, V. Waters, P. Kaiser, B. B. Kaufer, Identification of the Receptor and Cellular Ortholog of the Marek's Disease Virus (MDV) CXC Chemokine. *Front. Microbiol.* **8**, 2543 (2017).

19. A. T. Engel, R. K. Selvaraj, J. P. Kamil, N. Osterrieder, B. B. Kaufer, Marek's disease viral interleukin-8 promotes lymphoma formation through targeted recruitment of B cells and CD4+ CD25+ T cells. *J. Virol.* **86**, 8536–8545 (2012).

20. B. Lupiani, L. F. Lee, X. Cui, I. Gimeno, A. Anderson, R. W. Morgan, R. F. Silva, R. L. Witter, H.-J. Kung, S. M. Reddy, Marek's disease virus-encoded Meq gene is involved in transformation of lymphocytes but is dispensable for replication. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 11815–11820 (2004).

21. A. M. Conradie, L. D. Bertzbach, J. Trimpert, J. N. Patria, S. Murata, M. S. Parcells, B. B. Kaufer, Distinct polymorphisms in a single herpesvirus gene are capable of enhancing virulence and mediating vaccinal resistance. *PLoS Pathog.* **16**, e1009104 (2020).

22. K. G. Renz, J. Cooke, N. Clarke, B. F. Cheetham, Z. Hussain, A. F. M. Fakhrul Islam, G. A. Tannock, S. W. Walkden-Brown, Pathotyping of Australian isolates of Marek's disease virus and association of pathogenicity with meq gene polymorphism. *Avian Pathol.* **41**, 161–176 (2012).

23. Z. Qian, P. Brunovskis, F. Rauscher 3rd, L. Lee, H. J. Kung, Transactivation activity of Meq, a Marek's disease herpesvirus bZIP protein persistently expressed in latently infected transformed T cells. *J. Virol.* **69**, 4037–4044 (1995).

24. See supplementary information.

25. O. Putelat, "Archéozoologie" in *Strasbourg, Bas-Rhin. Rue de Lucerne – Rue du Jeu-de-Paume. Rapport de fouille préventive. Volume 1. Le système défensif primitif et le processus d'urbanisation d'un secteur du faubourg de la Krutenau du Moyen Âge à nos jours. Rapport de fouille préventive, Sélestat : Pôle d'Archéologie Interdépartemental Rhénan*, M. Werlé, Ed. (2015), pp. 98–174.

26. A. Cicović, D. Radičević, "Arheološka istraživanja srednjovekovnih nalazišta na Rudniku 2009–2013. godine" in *Rudnik 1, istraživanja srednjovekovnih nalazišta (2009-2013. godina), Gornji*

Milanovac, D. Radičević, A. Cicović, Eds. (2013), pp. 19–57.

27. N. Marković, J. Bulatović, "Rudnik 2009–2013: rezultati arheozoološke analize" in *Rudnik 1, istraživanja srednjovekovnih nalazišta (2009-2013. godina), Gornji Milanovac*, A. Cicović, D. Radičević, Eds. (2019), pp. 119–129.

28. H. Baron, Quasi Liber Et Pictura. Die Tierknochenfunde aus dem Gräberfeld an der Wiener Csokorgasse – eine anthrozoologische Studie zu den awarischen Bestattungssitten. *Monographien des RGZM*. **143** (2018).

29. O. Putelat, thesis, Université de Paris 1 Panthéon-Sorbonne (2015).

30. M. Popović, *Manastir Studenica – arheološka otkrića* (Republički zavod za zaštitu spomenika kulture, Arheološki institut, Beograd, 2015).

31. N. Marković, "Ishrana u manastiru Studenica: arheozoološka svedočanstva" in *Manastir Studenica – arheološka otkrića*, M. Popović, Ed. (Beograd: Republički zavod za zaštitu spomenika kulture i Arheološki institut, 2015), pp. 395–406.

32. I. Živaljević, N. Marković, M. Maksimović, Food worthy of kings and saints: fish consumption in the medieval monastery Studenica (Serbia). *anth*. **54**, 179–201 (2019).

33. Marković, N., Radišić, T. & Bikić, "Uloga živine u srednjovekovnoj ekonomiji manastira Studenice" in *Bioarheologija na Balkanu. Metodološke, komparativne i rekonstruktivne studije života u prošlosti*, M.-R. N. Vitezović S., Ed. (2016), pp. 99–116.

34. A. Saed Mucheshi, M. Nikzad, M. Zamani-Dadaneh, Rescue excavations at Bardeh Mar, Darian Dam area, Hawraman, Kurdistan, western Iran. *Proceedings of the 15th* (2017).

35. M. Mashkour, A. Mohaseb, S. Amiri, S. Beyzaiedoust, R. Khazaeli, H. Davoudi, H. Fathi, S. Komijani, A. Aliyari, H. Laleh, Archaeozoological Report of the Bioarchaeology Laboratory of the University of Tehran and the Osteology Department of the National Museum of Iran, 2015-2016. *Proceedings of the 15th Annual Symposium on the Iranian Archaeology, 5-7 march 2017, Tehran, Iranian Center for Archaeological Research*, 803–807 (2017).

36. P. J. Reimer, W. E. N. Austin, E. Bard, A. Bayliss, P. G. Blackwell, C. B. Ramsey, M. Butzin, H. Cheng, R. Lawrence Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, I. Hajdas, T. J. Heaton, A. G. Hogg, K. A. Hughen, B. Kromer, S. W. Manning, R. Muscheler, J. G. Palmer, C. Pearson, J. van der Plicht, R. W. Reimer, D. A. Richards, E. Marian Scott, J. R. Southon, C. S. M. Turney, L. Wacker, F. Adolphi, U. Büntgen, M. Capano, S. M. Fahrni, A. Fogtmann-Schulz, R. Friedrich, P. Köhler, S. Kudsk, F. Miyake, J. Olsen, F. Reinig, M. Sakamoto, A. Sookdeo, S. Talamo, The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon*. **62**, 725–757 (2020).

37. J. Dabney, M. Knapp, I. Glocke, M.-T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. García, S. Pääbo, J.-L. Arsuaga, M. Meyer, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).

38. M.-T. Gansauge, M. Meyer, Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome Res.* **24**, 1543–1549 (2014).

39. C. Carøe, S. Gopalakrishnan, L. Vinner, S. S. T. Mak, M. H. S. Sinding, J. A. Samaniego, N. Wales, T. Sicheritz-Pontén, M. T. P. Gilbert, Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* **9**, 410–419 (2018).

40. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. **29**, 1682–1684 (2013).

41. M. Schubert, A. Ginolhac, S. Lindgreen, J. F. Thompson, K. A. S. Al-Rasheid, E. Willerslev, A. Krogh, L. Orlando, Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics*. **13**, 178 (2012).

42. G. Jun, M. K. Wing, G. R. Abecasis, H. M. Kang, An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* **25**, 918–925 (2015).

43. Broad Institute, *Picard toolkit* (Broad Institute, 2019; http://broadinstitute.github.io/picard/).

44. G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, M. A. DePristo, From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics*. **43**, 11.10.1–11.10.33 (2013).

45. Code DOI: 10.5281/zenodo.10022436

46. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. **9**, 357–359 (2012).

47. G. Tonkin-Hill, J. A. Lees, S. D. Bentley, S. D. W. Frost, J. Corander, Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res.* **47**, 5539–5549 (2019).

48. J. Corander, P. Marttinen, Bayesian identification of admixture events using multilocus molecular markers. *Mol. Ecol.* **15**, 2833–2843 (2006).

49. M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, A. Rambaut, Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol*. **4**, vey016 (2018).

50. G. Yu, Using ggtree to Visualize Data on Tree-Like Structures. *Curr. Protoc. Bioinformatics*. **69**, e96 (2020).

51. M. Krzywinski, J. Schein, İ. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M. A. Marra, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

52. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. **30**, 1312–1313 (2014).

53. A. J. Page, B. Taylor, A. J. Delaney, J. Soares, T. Seemann, J. A. Keane, S. R. Harris, SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom*. **2**, e000056 (2016).

54. P. O. Lewis, A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* **50**, 913–925 (2001).

55. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

56. P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, M. J. L. de Hoon, Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*. **25**, 1422–1423 (2009).

57. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLoS One*. **11**, e0163962 (2016).

58. H. Li, seqtk Toolkit for processing sequences in FASTA/Q formats. *GitHub*. **767**, 69 (2012).

59. S. Duchêne, D. Duchêne, E. C. Holmes, S. Y. W. Ho, The Performance of the Date-Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data. *Mol. Biol. Evol.* **32**, 1895–1906 (2015).

60. A. Rieux, F. Balloux, Inferences from tip-calibrated phylogenies: a review and a practical guide. *Mol. Ecol.* **25**, 1911–1924 (2016).

61. M. Navascués, F. Depaulis, B. C. Emerson, Combining contemporary and ancient DNA in population genetic and phylogeographical studies. *Mol. Ecol. Resour.* **10**, 760–772 (2010).

62. S. Duchene, P. Lemey, T. Stadler, S. Y. W. Ho, D. A. Duchene, V. Dhanasekaran, G. Baele, Bayesian Evaluation of Temporal Signal in Measurably Evolving Populations. *Mol. Biol. Evol.* **37**, 3363–3379 (2020).

63. G. Baele, P. Lemey, M. A. Suchard, Genealogical Working Distributions for Bayesian Model Testing with Phylogenetic Uncertainty. *Syst. Biol.* **65**, 250–264 (2016).

64. M. Molak, M. A. Suchard, S. Y. W. Ho, D. W. Beilman, B. Shapiro, Empirical calibrated radiocarbon sampler: a tool for incorporating radiocarbon-date and calibration error into Bayesian phylogenetic analyses of ancient DNA. *Mol. Ecol. Resour.* **15**, 81–86 (2015).

65. F. Rodríguez, J. L. Oliver, A. Marín, J. R. Medina, The general stochastic model of nucleotide substitution. *J. Theor. Biol.* **142**, 485–501 (1990).

66. Z. Yang, Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306–314 (1994).

67. A. J. Drummond, S. Y. W. Ho, M. J. Phillips, A. Rambaut, Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).

68. B. Pfeifer, U. Wittelsbürger, S. E. Ramos-Onsins, M. J. Lercher, PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936 (2014).

69. D. K. Ajithdoss, S. M. Reddy, P. F. Suchodolski, L. F. Lee, H.-J. Kung, B. Lupiani, In vitro characterization of the Meq proteins of Marek's disease virus vaccine strain CVI988. *Virus Res.* **142**, 57–67 (2009).

70. T. Huszár, I. Mucsi, T. Terebessy, A. Masszi, S. Adamkó, C. Jeney, L. Rosivall, The use of a second reporter plasmid as an internal standard to normalize luciferase activity in transient transfection experiments may lead to a systematic error. *J. Biotechnol.* **88**, 251–258 (2001).

71. K.-S. Chang, K. Ohashi, M. Onuma, Diversity (polymorphism) of the meq gene in the attenuated Marek's disease virus (MDV) serotype 1 and MDV-transformed cell lines. *J. Vet. Med. Sci.* **64**, 1097–1101 (2002).

72. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–5 (2016).

73. J. Sato, S. Murata, Z. Yang, B. B. Kaufer, S. Fujisawa, H. Seo, N. Maekawa, T. Okagawa, S. Konnai, N. Osterrieder, M. S. Parcells, K. Ohashi, Effect of Insertion and Deletion in the Meq Protein Encoded by Highly Oncogenic Marek's Disease Virus on Transactivation Activity and Virulence. *Viruses*. **14** (2022), doi:10.3390/v14020382.

74. C. Firth, A. Kitchen, B. Shapiro, M. A. Suchard, E. C. Holmes, A. Rambaut, Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol. Biol. Evol.* **27**, 2038–2051 (2010).

Acknowledgments:

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

Funding:

European Research Council grant ERC-2019-StG-853272-PALAEOFARM or ERC-2013-StG-337574-UNDEAD or both (SRF, LAF, GL, ALS)
Wellcome Trust grant 210119/Z/18/Z (SRF, LAF)
Oxford Martin School grant ATR02370 (SRF, ALS, LdP, OGP)
AHRC grant AH/L006979/1 (GL, OL, NS)
European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 895107 (OL)
BBSRC grant number BB/M011224/1 (SD)
Postdoctoral grant (12U7121N) of the Research Foundation -- Flanders (Fonds voor Wetenschappelijk Onderzoek) (BV)

Author contributions:

Conceptualization: SRF, ALS, LAFF, GL

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

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

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

Visualization: SRF

Funding acquisition: LAFF, ALS, GL

Project administration: SRF, LAFF, ALS, GL

Supervision: LAFF, ALS, GL

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

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

Competing interests:

The authors declare that they have no competing interests.

Data and materials availability:

All MDV sequence data generated have been deposited in GenBank under accession PRJEB64489. Code is available from the following GitHub repository, DOI: 10.5281/zenodo.10022436

Supplementary Materials:

Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S4, S9 and S10
Captions for Data S1

Other Supplementary Materials for this manuscript include the following:

Data S1, which comprises:

- 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