

## The history of measles: from a 1912 genome to an antique origin

**Authors:** Ariane Düx<sup>1,2†</sup>, Sebastian Lequime<sup>3†</sup>, Livia Victoria Patrono<sup>1,2</sup>, Bram Vrancken<sup>3</sup>, Sengül Boral<sup>4</sup>, Jan F. Gogarten<sup>1,2</sup>, Antonia Hilbig<sup>1</sup>, David Horst<sup>4</sup>, Kevin Merkel<sup>1,2</sup>, Baptiste Prepoint<sup>2</sup>, Sabine Santibanez<sup>5</sup>, Jasmin Schlotterbeck<sup>2</sup>, Marc A. Suchard<sup>6</sup>, Markus Ulrich<sup>1</sup>, Navena Widulin<sup>7</sup>, Annette Mankertz<sup>5</sup>, Fabian H. Leendertz<sup>1</sup>, Kyle Harper<sup>8</sup>, Thomas Schnalke<sup>7</sup>, Philippe Lemey<sup>3‡</sup>, Sébastien Calvignac-Spencer<sup>1,2‡\*</sup>

### Affiliations:

<sup>1</sup> Epidemiology of Highly Pathogenic Microorganisms, Robert Koch Institute, Berlin, Germany

<sup>2</sup> Viral Evolution, Robert Koch Institute, Berlin, Germany

<sup>3</sup> KU Leuven Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory of Clinical and Evolutionary Virology, Leuven, Belgium

<sup>4</sup> Institute for Pathology, Charité, Berlin, Germany

<sup>5</sup> National Reference Centre Measles, Mumps, Rubella, Robert Koch Institute, Berlin, Germany

<sup>6</sup> Department of Biostatistics, Fielding School of Public Health and Departments of Biomathematics and of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

<sup>7</sup> Berlin Museum of Medical History at the Charité, Berlin, Germany

<sup>8</sup> Department of Classics and Letters, The University of Oklahoma, Norman, OK, USA

\*Correspondence to: CalvignacS@rki.de

† Equal contributions.

‡ Equal contributions.

## **Abstract:**

Many infectious diseases are thought to have emerged in humans after the Neolithic revolution. While it is broadly accepted that this also applies to measles, the exact date of emergence for this disease is controversial. Here, we sequenced the genome of a 1912 measles virus and used selection-aware molecular clock modeling to determine the divergence date of measles virus and rinderpest virus. This divergence date represents the earliest possible date for the establishment of measles in human populations. Our analyses show that the measles virus potentially arose as early as the 4<sup>th</sup> century BCE, rekindling the recently challenged hypothesis of an antique origin of this disease.

## **One Sentence Summary:**

Measles virus diverged from rinderpest virus in the 4<sup>th</sup> century BCE, which is compatible with an emergence of measles during Antiquity.

## **Main Text:**

Measles is a highly contagious viral disease that presents with rash, fever and respiratory symptoms. Before a live-attenuated vaccine was developed in the 1960s, the disease affected the vast majority of children (1, 2). Global vaccination campaigns resulted in a marked reduction of measles transmission and fatal cases and WHO has proclaimed an elimination goal. However, the disease still caused an estimated 110,000 deaths in 2017 (3) and incidence has recently been on the rise (4). Measles is caused by *Measles morbillivirus* (MeV), a negative sense single-stranded RNA virus from the family *Paramyxoviridae* (Order: *Mononegavirales*). MeV is an exclusively human pathogen whose closest relative was the now eradicated *Rinderpest morbillivirus* (RPV), a devastating cattle pathogen (5). It is generally accepted that measles emergence resulted from a

spill-over from cattle to humans, although the directionality of this cross-species transmission event has never been formally established (supplementary text S1, 6).

It is unclear when measles first became endemic in human populations, but assuming an origin in cattle, the earliest possible date of MeV emergence is defined by the MeV-RPV divergence time. Several studies have provided estimates for this date using molecular clock analyses (7-10), with the most reliable (and oldest) estimate falling at the end of the 9<sup>th</sup> century CE (mean: 899 CE; 95% highest posterior density (HPD) interval: 597 – 1144 CE) (8). Here, we reassess the MeV-RPV divergence time using advanced, selection-aware Bayesian molecular clock modelling (11) on a dataset of heterochronous MeV genomes including the oldest human RNA virus genome sequenced to date. We show that a considerably earlier emergence during Antiquity can no longer be excluded.

Our re-examination was prompted by the broadly accepted view that molecular dating based on tip date calibration, i.e. the method used in previous efforts to estimate the timing of MeV-RPV divergence, underestimates deep divergence times (8). Rapid short-term substitution rates captured by tip calibration can often not be applied over long evolutionary timescales, because of the effects of long-term purifying selection and substitution saturation. This causes a discrepancy between short- and long-term substitution rates, which is referred to as the time-dependent rate phenomenon (TDRP; 12, 13). Since measurement timescales matter, a first step to arrive at accurate estimates is to maximize the time depth of tip calibration, for example through the use of ancient viral sequences (14, 15).

RNA tends to be much less stable in the environment than DNA, making the recovery of MeV genetic material from archeological remains unlikely (16, 17). Pathology collections represent a more realistic source of MeV sequences that predate the oldest MeV genome – the genome of the Edmonston strain that was isolated in 1954 and attenuated to become the first measles vaccine.

We examined a collection of lung specimens gathered by Rudolf Virchow and his successors between the 1870s and 1930s and preserved by the Berlin Museum of Medical History at the Charité (Berlin, Germany), and identified a 1912 case diagnosed with fatal measles-related bronchopneumonia (Fig. 1, fig. S1, supplementary texts S2 and S3). To retrieve MeV genetic material from this specimen, we first heat treated 200mg of the formalin-fixed lung tissue to reverse macromolecule cross-links induced by formalin and subsequently performed nucleic acid extraction (18). Following DNase treatment and ribosomal RNA depletion, we built high-throughput sequencing libraries and shotgun sequenced them on Illumina® platforms. We generated 27,328,219 reads, of which 0.46% were mapped to MeV genome, representing 10,960 unique MeV reads. These allowed us to reconstruct an almost complete 1912 MeV genome: 15,257 of the 15,894 nucleotides in the MeV strain Edmonston (AF266288) were covered by at least 3 unique reads (11,988 nucleotides by at least 20 reads; mean coverage 54x).

In addition to the 1912 genome and the 1954 Edmonston genome, only 2 genomes have been determined from MeV isolated prior to 1990 (Mvi/Lyon.FRA/77: HM562899; T11wild: AB481087). We therefore searched the strain collection of the German National Reference Laboratory (Robert Koch Institute, Berlin, Germany) for pre-1990 isolates. We found two strains from the pre-vaccine era isolated in 1960 by the National Reference Laboratory of former Czechoslovakia in Prague (MVi/Prague.CZE/60/1 and MVi/Prague.CZE/60/2; 19). We performed serial passages of these strains and determined their genome sequences at a mean coverage of 109x and 70x, respectively. The two genomes were nearly identical, differing at only four sites.

We performed Bayesian and maximum likelihood (ML) phylogenetic analyses to investigate the phylogenetic placement of the 1912 and 1960 genomes with respect to 127 available MeV genomes. Tip-dated Bayesian phylogenetic trees placed the 1912 genome basal to all modern

genomes while the two genomes from 1960 clustered together with the Edmonston strain (genotype A; fig. S2). The basal placement of the 1912 genome was further confirmed by exploring the rooting of non-clock ML trees based on the relationship between sampling time and root-to-tip-divergence (both including and excluding this genome; supplementary text S4 and fig. S3 A and B). The relatedness of the 1912 and 1960 genomes to now extinct MeV lineages is in line with a marked reduction of MeV genetic diversity during the 20<sup>th</sup> century as a product of massive vaccination efforts.

Having extended the time depth of MeV tip calibration, we subsequently focused our attention on estimating the timing of MeV-RPV divergence. We assembled a dataset of 51 genomes comprising MeV (including one of the 1960 genomes and the 1912 genome), RPV and Peste des petits ruminants virus (PPRV, the closest relative to MeV-RPV) sequences, ensuring they represented the known genetic diversity of these viruses. This dataset was used to infer a time-scaled evolutionary history for these three species in a Bayesian phylogenetic framework. We constructed a series of evolutionary models with increasing degree of complexity to accommodate various sources of rate heterogeneity. Models ranged from a standard codon substitution model with a strict molecular clock assumption to a codon substitution model with time-varying selection pressure combined with a clade-specific rate for PPRV and additional branch-specific random effects on the substitution rate. Adequately accommodating different sources of rate heterogeneity is known to provide a better correction for multiple hits in genetic distance estimation and the potential of codon substitution modelling in recovering deep viral divergence has specifically been demonstrated (8). This was reflected in the increasingly older estimates of MeV-RPV and PPRV-MeV-RPV divergence times and wider 95% highest posterior density intervals in increasingly complex models (Fig. 2 and table S2). Parameter estimates of the substitution and clock models also provided evidence for a significant contribution of these

different sources of rate heterogeneity to model fit improvement (table S2, supplementary text S4).

We found a significantly negative coefficient for the time-dependent nonsynonymous/synonymous substitution rate ratio ( $\omega$ ) (11), indicating strong long-term purifying selection, a significantly positive coefficient for the fixed effect on the PPRV rate, indicating a faster evolutionary rate in this clade (as suggested by temporal signal analyses, fig. S4), and significant additional unexplained variation as modelled by the random effects. Our most complex model therefore provided the best description of the evolutionary process and pushed back the divergence date of MeV and RPV to Antiquity, with a mean estimate at 345 BCE [95% highest posterior density (HPD) interval: 1066 BCE - 319 CE] (Fig. 3A). These estimates were robust to including or estimating the age of the 1912 genome in the analyses; estimating its sampling date resulted in a posterior distribution in agreement with the known date (1923 CE [95% HPD interval: 1883 - 1962 CE]; supplementary text S5 and table S2).

While the divergence of MeV and RPV *predates* the emergence of measles in humans, it does not necessarily coincide with it. This new estimate should therefore be considered a lower bound for measles emergence, which is now compatible with the emergence of this disease during Antiquity. This raises the question of whether an earlier timing of measles emergence agrees with other sources of information.

The earliest clear clinical description of measles is often attributed to the Persian physician Rhazes, writing in the 10<sup>th</sup> century CE (20). But Rhazes was extremely familiar with all available medical literature at his time, and made use of earlier sources. Indian medical texts probably describe measles several centuries prior to Rhazes (21). While clear descriptions of measles are missing in the Hippocratic corpus and the Greek medical tradition (at least through the prolific second-century writer Galen), such absence alone cannot be decisive. Retrospective diagnosis from pre-modern medical texts is notoriously fraught, especially for diseases like measles whose

symptoms were easily confused with a variety of other conditions. Measles differential diagnosis remained a challenge well into more recent times (22). Therefore, any number of the large-scale “pestilences” described in first-millennium sources from Europe or China could reflect MeV outbreaks.

An antique origin of measles seems all the more plausible in the light of demographic changes that are compatible with our understanding of MeV (contemporary) epidemiology. Indeed, Antiquity witnessed an important upsurge in population sizes in Eurasia and South and East Asia. Populations large enough to support continuous MeV transmission, i.e. larger than the MeV critical community size (CCS) of 250,000-500,000 individuals (23-25), have existed in Eurasia since the first millennium BCE. Although population sizes derived from ancient documents or archaeological proxies are only estimates (supplementary text S6), there is broad agreement that a number of settlements in North Africa, India, China, Europe, and the Near East began to surpass the CCS for MeV by around 300 BCE, presumably for the first time in human history (Fig. 3B; 26).

Based on these considerations, our substantially older divergence estimate provides grounds for sketching a new model of MeV’s evolutionary history. Under this scenario, a bovine virus, the common ancestor to modern strains of RPV and MeV, circulated in large populations of cattle (and possibly wild ungulates) since its divergence from PPRV around the 3rd millennium BCE (mean estimate at 2821 BCE [95% HPD interval: 4177 - 1665 BCE]; Fig. 3A). As a fast-evolving RNA virus, it may have produced variants that were able to cross the species barrier on several occasions, but small human populations could only serve as dead-end hosts. Then, almost as soon as contiguous settlements reached sufficient sizes to maintain the virus’ continuous transmission (Fig. 3B), it emerged as a human pathogen, the progenitor of modern-day MeV. Concurrent human-bovine epidemics that are well documented in e.g. Roman sources from the 5<sup>th</sup> century

BCE on (27) may mark the early stages of MeV emergence when the ancestor of MeV was pathogenic to both cattle and humans. During the following centuries, introduction of MeV into naive populations and/or flare-ups of the disease might have caused some ancient epidemics whose etiology remains uncertain.

While our findings shed new light on the origin of measles, formally proving that the virus emerged during Antiquity would require archeological genomic evidence. Recently, a 1000-year-old plant RNA virus genome (*Zea mays chrysovirus 1*) was recovered from maize remains (28) and genetic material of parvovirus B19 was detected in early Neolithic skeletal remains, despite the relatively unstable nature of its single-stranded DNA genome (15). Such advances highlight that it may not be completely impossible for antique remains to still contain MeV RNA, especially if preserved under favorable circumstances, including natural mummification or preservation in cold environments (16, 17). While awaiting such direct evidence, we believe that the proposed model of MeV evolution constitutes a compelling working hypothesis.

## References and Notes:

1. A. D. Langmuir, Medical importance of measles. *American journal of diseases of children (1960)* **103**, 224-226 (1962).
2. W. J. Moss, Measles. *Lancet* **390**, 2490-2502 (2017).
3. A. Dabbagh, R. L. Laws, C. Steulet, L. Dumolard, M. N. Mulders, K. Kretsinger, J. P. Alexander, P. A. Rota, J. L. Goodson, Progress toward regional measles elimination—worldwide, 2000–2017. *Morbidity and Mortality Weekly Report* **67**, 1323 (2018).
4. WHO, Provisional data based on monthly data reported to WHO (Geneva) as of September 2019. [https://www.who.int/immunization/monitoring\\_surveillance/burden/vpd/surveillance\\_type/active/measles\\_monthlydata/en/](https://www.who.int/immunization/monitoring_surveillance/burden/vpd/surveillance_type/active/measles_monthlydata/en/).
5. P. Roeder, J. Mariner, R. Kock, Rinderpest: the veterinary perspective on eradication. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **368**, 20120139 (2013).



6. N. D. Wolfe, C. P. Dunavan, J. Diamond, Origins of major human infectious diseases. *Nature* **447**, 279-283 (2007).
7. Y. Furuse, A. Suzuki, H. Oshitani, Origin of measles virus: divergence from rinderpest virus between the 11th and 12th centuries. *Virology journal* **7**, 52 (2010).
8. J. O. Wertheim, S. L. Kosakovsky Pond, Purifying selection can obscure the ancient age of viral lineages. *Mol Biol Evol* **28**, 3355-3365 (2011).
9. M. Muniraju, M. Munir, A. R. Parthiban, A. C. Banyard, J. Bao, Z. Wang, C. Ayebazibwe, G. Ayelet, M. El Harrak, M. Mahapatra, G. Libeau, C. Batten, S. Parida, Molecular evolution of peste des petits ruminants virus. *Emerging infectious diseases* **20**, 2023-2033 (2014).
10. H. Kimura, M. Saitoh, M. Kobayashi, H. Ishii, T. Saraya, D. Kurai, H. Tsukagoshi, K. Shirabe, A. Nishina, K. Kozawa, M. Kuroda, F. Takeuchi, T. Sekizuka, H. Minakami, A. Ryo, M. Takeda, Molecular evolution of haemagglutinin (H) gene in measles virus. *Scientific reports* **5**, 11648 (2015).
11. J. V. Membrebe, M. A. Suchard, A. Rambaut, G. Baele, P. Lemey, Bayesian Inference of Evolutionary Histories under Time-Dependent Substitution Rates. *Mol Biol Evol* **36**, 1793-1803 (2019).
12. S. Y. Ho, S. Duchene, M. Molak, B. Shapiro, Time-dependent estimates of molecular evolutionary rates: evidence and causes. *Molecular ecology* **24**, 6007-6012 (2015).
13. P. Aiewsakun, A. Katzourakis, Time-Dependent Rate Phenomenon in Viruses. *Journal of virology* **90**, 7184-7195 (2016).
14. B. Muhlemann, T. C. Jones, P. B. Damgaard, M. E. Allentoft, I. Shevnina, A. Logvin, E. Usmanova, I. P. Panyushkina, B. Boldgiv, T. Bazartseren, K. Tashbaeva, V. Merz, N. Lau, V. Smrcka, D. Voyakin, E. Kitov, A. Epimakhov, D. Pokutta, M. Vicze, T. D. Price, V. Moiseyev, A. J. Hansen, L. Orlando, S. Rasmussen, M. Sikora, L. Vinner, A. Osterhaus, D. J. Smith, D. Glebe, R. A. M. Fouchier, C. Drosten, K. G. Sjogren, K. Kristiansen, E. Willerslev, Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature* **557**, 418-423 (2018).
15. B. Muhlemann, A. Margaryan, P. B. Damgaard, M. E. Allentoft, L. Vinner, A. J. Hansen, A. Weber, V. I. Bazaliiskii, M. Molak, J. Arneborg, W. Bogdanowicz, C. Falys, M. Sablin, V. Smrcka, S. Sten, K. Tashbaeva, N. Lynnerup, M. Sikora, D. J. Smith, R. A. M. Fouchier, C. Drosten, K. G. Sjogren, K. Kristiansen, E. Willerslev, T. C. Jones, Ancient human parvovirus B19 in Eurasia reveals its long-term

- association with humans. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 7557-7562 (2018).
16. O. Smith, M.T. Gilbert, “Ancient RNA” in *Paleogenomics: Genome-Scale Analysis of Ancient DNA*, C. Lindqvist, O. P. Rajora, Eds. (Springer, 2019).
  17. O. Smith, G. Dunshea, M. S. Sinding, S. Fedorov, M. Germonpre, H. Bocherens, M. T. P. Gilbert, Ancient RNA from Late Pleistocene permafrost and historical canids shows tissue-specific transcriptome survival. *PLoS biology* **17**, e3000166 (2019).
  18. M. T. Gilbert, T. Haselkorn, M. Bunce, J. J. Sanchez, S. B. Lucas, L. D. Jewell, E. Van Marck, M. Worobey, The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when? *PloS one* **2**, e537 (2007).
  19. S. Santibanez, A. Heider, E. Gerike, A. Agafonov, E. Schreier, Genotyping of measles virus isolates from central Europe and Russia. *Journal of medical virology* **58**, 313-320 (1999).
  20. R. Kim-Farley, “Measles,” in *The Cambridge World History of Human Disease*, K. Kiple, Ed. (Cambridge, 1993) 871-75.
  21. K.R.L. Gupta, translator, *Madhava Nidana: Ayurvedic System of Pathology* (Delhi, 1987).
  22. B. A. Cunha, Smallpox and measles: historical aspects and clinical differentiation. *Infectious disease clinics of North America* **18**, 79-100 (2004).
  23. M. S. Bartlett, Measles periodicity and community size. *Journal of the Royal Statistical Society. Series A (General)* **120**, 48-70 (1957).
  24. F. L. Black, Measles endemicity in insular populations: critical community size and its evolutionary implication. *Journal of theoretical biology* **11**, 207-211 (1966).
  25. M. J. Keeling, B. T. Grenfell, Disease extinction and community size: modeling the persistence of measles. *Science* **275**, 65-67 (1997).
  26. H. Inoue, A. Álvarez, E. N. Anderson, A. Owen, R. Álvarez, K. Lawrence, C. Chase-Dunn, Urban scale shifts since the bronze age: upsweeps, collapses, and semiperipheral development. *Social Science History* **39**, 175-200 (2015).
  27. C. A. Spinage, *Cattle Plague: A History* (New York, 2003).
  28. M. Peyambari, S. Warner, N. Stoler, D. Rainer, M. J. Roossinck, A 1,000-Year-Old RNA Virus. *Journal of virology* **93**, (2019).

29. I. Morris, *The Measure of Civilization* (Princeton, 2013).
30. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
31. S. Nurk, D. Meleshko, A. Korobeynikov, P. A. Pevzner, metaSPAdes: a new versatile metagenomic assembler. *Genome research* **27**, 824-834 (2017).
32. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
33. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410 (1990).
34. G. Jäger, ClipAndMerge. <https://github.com/apeltzer/ClipAndMerge>.
35. B. Institute, Picard. <http://broadinstitute.github.io/picard>.
36. A. Peltzer, DeDup. <https://github.com/apeltzer/DeDup>.
37. Geneious 11.1.5. <https://www.geneious.com>.
38. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013).
39. D. P. Martin, B. Murrell, M. Golden, A. Khoosal, B. Muhire, RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* **1**, vev003 (2015).
40. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268-274 (2015).
41. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007 (2016).
42. 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).
43. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic biology* **67**, 901-904 (2018).
44. O. Chernomor, B. Q. Minh, F. Forest, S. Klaere, T. Ingram, M. Henzinger, A. von Haeseler, Split diversity in constrained conservation prioritization using integer linear programming. *Methods in ecology and evolution* **6**, 83-91 (2015).

45. D. L. Ayres, M. P. Cummings, G. Baele, A. E. Darling, P. O. Lewis, D. L. Swofford, J. P. Huelsenbeck, P. Lemey, A. Rambaut, M. A. Suchard, BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics. *Systematic biology*, (2019).
46. Z. Yang, Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol Evol* **11**, 367-372 (1996).
47. A. D. Yoder, Z. Yang, Estimation of primate speciation dates using local molecular clocks. *Mol Biol Evol* **17**, 1081-1090 (2000).
48. B. Vrancken, A. Rambaut, M. A. Suchard, A. Drummond, G. Baele, I. Derdelinckx, E. Van Wijngaerden, A. M. Vandamme, K. Van Laethem, P. Lemey, The genealogical population dynamics of HIV-1 in a large transmission chain: bridging within and among host evolutionary rates. *PLoS computational biology* **10**, e1003505 (2014).
49. D. Pitt, N. Sevane, E. L. Nicolazzi, D. E. MacHugh, S. D. E. Park, L. Colli, R. Martinez, M. W. Bruford, P. Orozco-terWengel, Domestication of cattle: Two or three events? *Evolutionary applications* **12**, 123-136 (2019).
50. S. Morand, K. M. McIntyre, M. Baylis, Domesticated animals and human infectious diseases of zoonotic origins: domestication time matters. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **24**, 76-81 (2014).
51. J. C. Mariner, J. McDermott, J. A. Heesterbeek, A. Catley, P. Roeder, A model of lineage-1 and lineage-2 rinderpest virus transmission in pastoral areas of East Africa. *Preventive veterinary medicine* **69**, 245-263 (2005).
52. T. Barrett, P. B. Rossiter, Rinderpest: the disease and its impact on humans and animals. *Advances in virus research* **53**, 89-110 (1999).
53. W.H. McNeill, *Plagues and Peoples* (Anchor Press, 1976)
54. P.-P. Pastoret, K. Yamanouchi, U. Mueller-Doblies, M. M. Rweyemamu, M. Horzinek, T. Barrett, "Rinderpest – an old and worldwide story: history to c.1902" in *Rinderpest and Peste des Petits Ruminants: Virus Plagues of Large and Small Ruminants*, T. Barrett, P.-P. Pastoret, W. Taylor, Eds. (Elsevier, 2005).
55. J. Blancou, "Old prophylactic methods" in *Rinderpest and Peste des Petits Ruminants: Virus Plagues of Large and Small Ruminants*, T. Barrett, P.-P. Pastoret, W. Taylor, Eds. (Elsevier, 2005).

56. OIE, Peste des Petits Ruminants  
[https://www.oie.int/fileadmin/Home/eng/Animal\\_Health\\_in\\_the\\_World/docs/pdf/Disease\\_cards/PESTE\\_DES\\_PETITS\\_RUMINANTS.pdf](https://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/Disease_cards/PESTE_DES_PETITS_RUMINANTS.pdf)
57. D. G. Streicker, A. S. Turmelle, M. J. Vonhof, I. V. Kuzmin, G. F. McCracken, C. E. Rupprecht, Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science* **329**, 676-679 (2010).
58. T. Lembo, C. Oura, S. Parida, R. Hoare, L. Frost, R. Fyumagwa, F. Kivaria, C. Chubwa, R. Kock, S. Cleaveland, C. Batten, Peste des petits ruminants infection among cattle and wildlife in northern Tanzania. *Emerging infectious diseases* **19**, 2037-2040 (2013).
59. E. Couacy-Hymann, M. Y. Koffi, V. K. Kouadio, A. Mossoum, L. Kouadio, A. Kouassi, K. Asseman, P. H. Godji, P. Nana, Experimental infection of cattle with wild type peste-des-petits-ruminants virus - Their role in its maintenance and spread. *Research in veterinary science* **124**, 118-122 (2019).
60. J. F. Drexler, V. M. Corman, M. A. Muller, G. D. Maganga, P. Vallo, T. Binger, F. Gloza-Rausch, V. M. Cottontail, A. Rasche, S. Yordanov, A. Seebens, M. Knornschild, S. Oponng, Y. Adu Sarkodie, C. Pongombo, A. N. Lukashev, J. Schmidt-Chanasit, A. Stocker, A. J. Carneiro, S. Erbar, A. Maisner, F. Fronhoffs, R. Buettner, E. K. Kalko, T. Kruppa, C. R. Franke, R. Kallies, E. R. Yandoko, G. Herrler, C. Reusken, A. Hassanin, D. H. Kruger, S. Mathee, R. G. Ulrich, E. M. Leroy, C. Drosten, Bats host major mammalian paramyxoviruses. *Nat Commun* **3**, 796 (2012).
61. W. Ghawar, H. Pascalis, J. Bettaieb, J. Melade, A. Gharbi, M. A. Snoussi, D. Laouini, S. M. Goodman, A. Ben Salah, K. Dellagi, Insight into the global evolution of Rodentia associated Morbilli-related paramyxoviruses. *Scientific reports* **7**, 1974 (2017).
62. J. R. Pulliam, J. H. Epstein, J. Dushoff, S. A. Rahman, M. Bunning, A. A. Jamaluddin, A. D. Hyatt, H. E. Field, A. P. Dobson, P. Daszak, Agricultural intensification, priming for persistence and the emergence of Nipah virus: a lethal bat-borne zoonosis. *Journal of the Royal Society, Interface* **9**, 89-101 (2012).
63. K. Halpin, P. L. Young, H. E. Field, J. S. Mackenzie, Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus. *The Journal of general virology* **81**, 1927-1932 (2000).
64. T. Chandler, *Four thousand years of urban growth: an historical census* (Lewiston, 1987).
65. G. Modelski, *World Cities: -3000 to 2000* (Washington, D.C., 2003).

66. P. Bairoch, *Cities and economic development from the dawn of history to the present* (Chicago, 1988).
67. D. Pasciuti, C. Chase-Dunn, "Estimating The Population Sizes of Cities," 2002.  
<https://irows.ucr.edu/research/citemp/estcit/estcit.htm> Accessed 10/25/2019.
68. A. Cliff, P. Haggett, and M. Smallman-Raynor, *Measles: an historical geography of a major human viral disease: from global expansion to local retreat, 1840-1990* (Oxford, 1993).
69. M. D. Baron, T. Barrett, Rescue of rinderpest virus from cloned cDNA. *Journal of virology* 71, 1265-1271 (1997).

### **Acknowledgments:**

The Titan V GPU used for this research was donated by the NVIDIA Corporation.

### **Funding:**

The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no.~725422-ReservoirDOCS). PL acknowledges support by the Research Foundation -- Flanders ('Fonds voor Wetenschappelijk Onderzoek -- Vlaanderen', FWO: G066215N, G0D5117N and G0B9317N). S.L. and B.V. are postdoctoral research fellows funded by the FWO. M.A.S. was partially supported through National Institutes of Health grant U19 AI135995.

### **Author contributions:**

Conceptualization: PL, SCS; data curation: AD, SL, BV, MAS, PL, SCS; formal analysis: AD, SL, BV, JFG, MU, PL, SCS; funding acquisition: FHL, SCS; investigation: AD, LVP, SB, AH, DH, KM, BP and JS; methodology: SL, BV, MAS, PL; project administration: FHL, PL, SCS; resources: SS, NW, AM, FHL, TS, PL, SCS; software: SL, BV, MAS, PL; supervision: PL, SCS;

validation: AD, MU, JFG, SCS; visualization: SL; writing – original draft: AD, SL, KH, PL, SCS; writing – review and editing: all authors.

**Competing interests:** All authors declare no competing interests.

**Data and materials availability:** MeV sequences from 1912 and 1960 are deposited in GenBank under **XXX**. Alignments, trees, and BEAST xml files are available at **XXX**.

**Supplementary Materials:**

Materials and Methods

Supplementary text S1-S6

Figures S1-S5

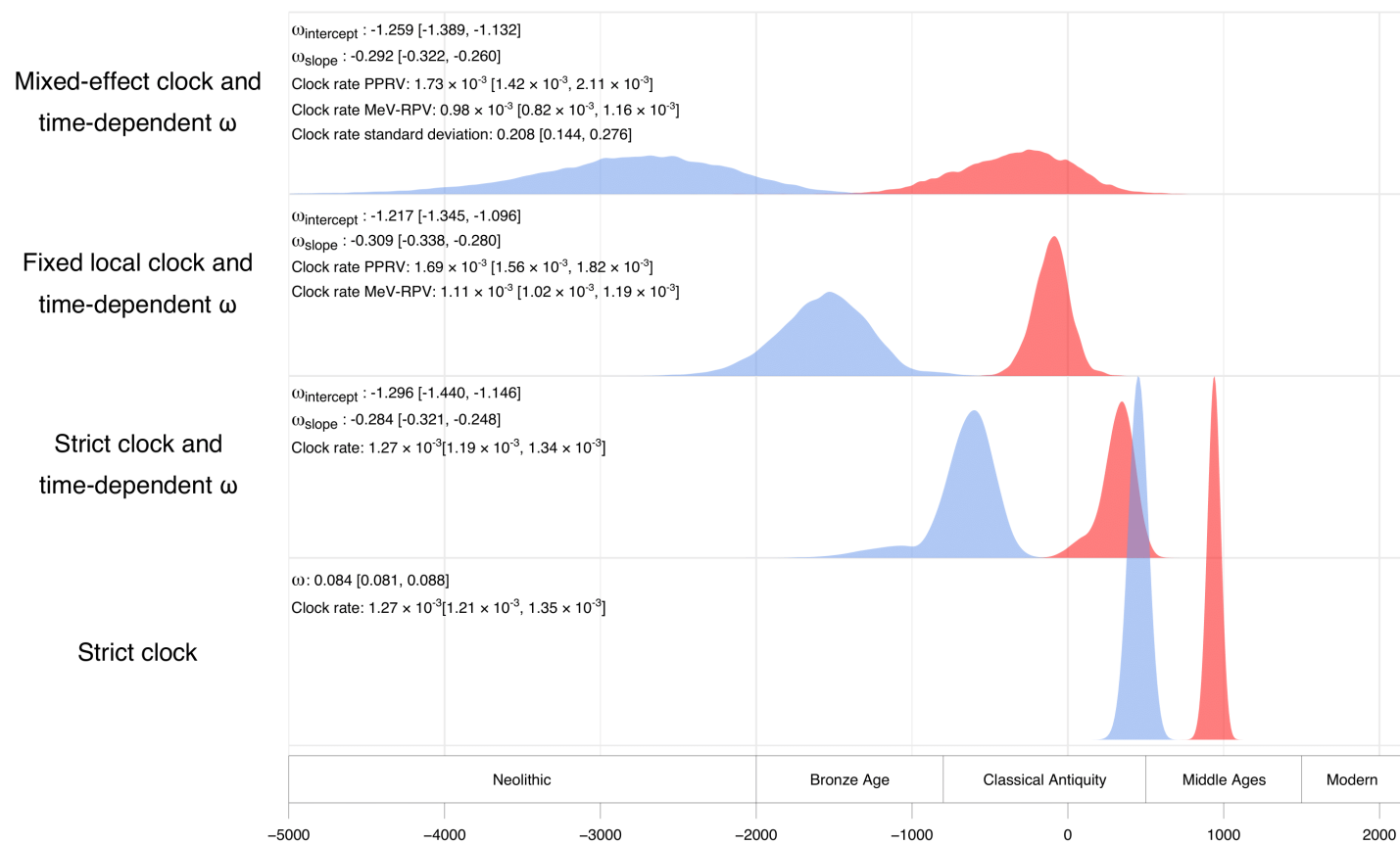
Tables S1-S2

References 30-69

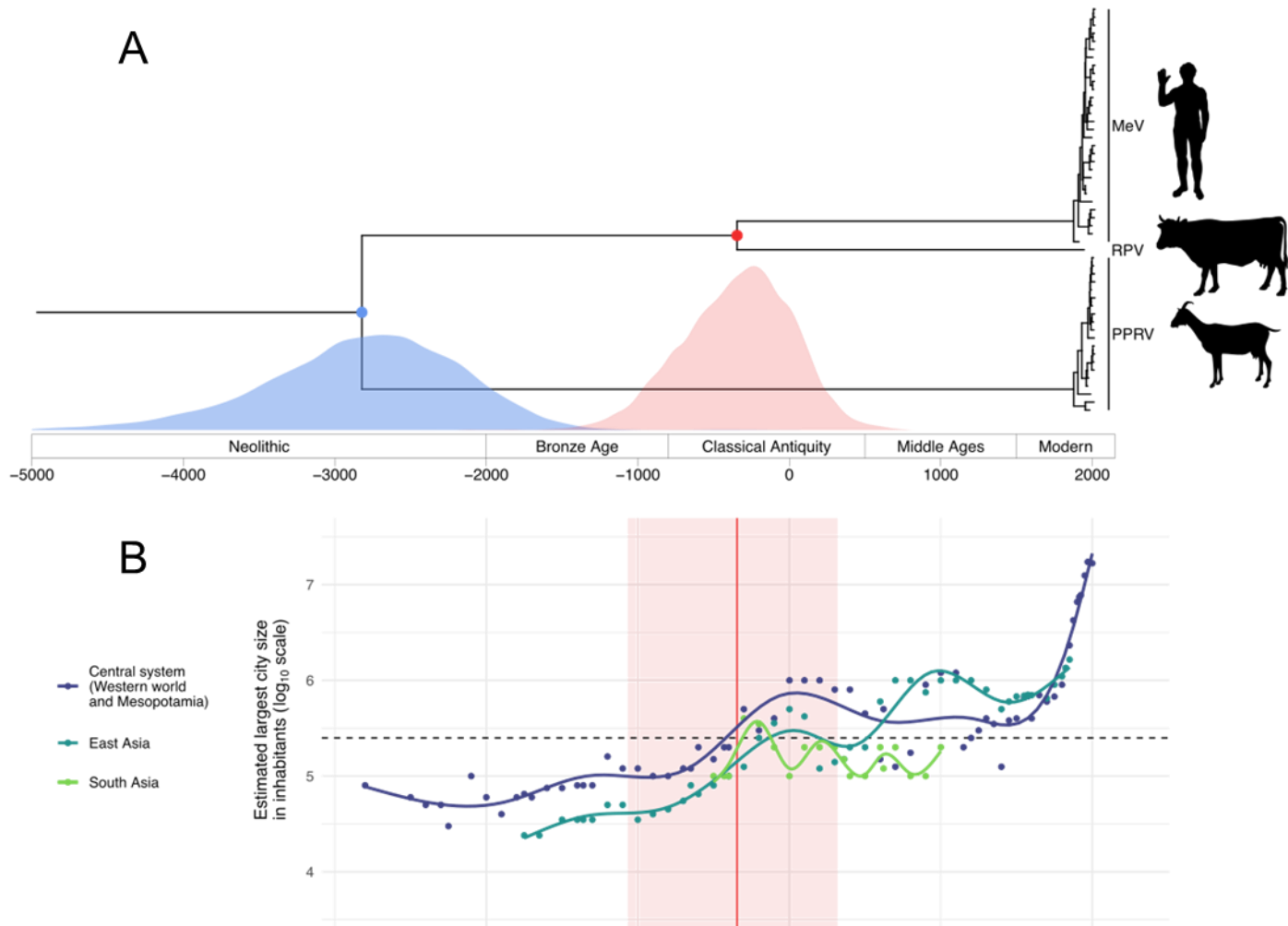


**Fig. 1. Formalin-fixed lung specimen collected in 1912 in Berlin from a 2-year old girl diagnosed with measles-related bronchopneumonia (museum object ID: BMM 655/1912).**





**Fig. 2. Influence of model complexity on the divergence time estimates between MeV and RPV (red), and between MeV/RPV and PPRV (blue).** Estimated parameter values (posterior mean and 95% highest posterior density interval) for each model are provided on the left. Historical divisions are based on classical "world history" periods.



**Fig. 3. Estimation of the divergence time between MeV and RPV, and between MeV/RPV and PPRV (A) and estimated size (log<sub>10</sub> scale) of the largest city in the Central system (dark blue), East Asia (petrol) and South Asia (green) over time (B).** (A) Maximum clade credibility (MCC) tree summarized from a Bayesian time-measured inference using tip-dating and accounting for long-term purifying selection. The red and blue points represent the divergence events of MeV and RPV and MeV/RPV and PPRV, respectively; the according divergence date estimates are reported below as marginal posterior distributions. Historical divisions are based on classical "world history" periods. (B) The red vertical line represents the divergence time estimate between MeV and RPV and the red area its 95% highest posterior density interval. The dashed horizontal line represents the classical threshold for MeV maintenance in a population (i.e. 250,000 individuals). Dots show data points according to Morris (29) and Inoue et al. (26). Each line represents the fit of a generalized additive model with a cubic spline smoothing function.