Resolution of the type material of the Asian elephant, *Elephas maximus* Linnaeus, 1758 (Proboscidea, Elephantidae)

ENRICO CAPPELLINI[1]*, ANTHEA GENTRY[2], ELEFTHERIA PALKOPOULOU[3,4], YASUKO ISHIDA[5], DAVID CRAM[6], ANNA-MARIE ROOS[7], MICK WATSON[8], ULF S. JOHANSSON[9], BO FERNHOLM[9], PAOLO AGNELLI[10], FAUSTO BARBAGLI[10], D. TIM J. LITTLEWOOD[2], CHRISTIAN D. KELSTRUP[11], JESPER V. OLSEN[11], ADRIAN M. LISTER[2], ALFRED L. ROCA[5], LOVE DALÉN[3] and M. THOMAS P. GILBERT[1,12]

[1]*Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark*
[2]*Natural History Museum, Cromwell Road, London SW7 5BD, UK*
[3]*Department of Bioinformatics and Genetics, Swedish Museum of Natural History, SE-10405 Stockholm, Sweden*
[4]*Department of Zoology, Stockholm University, SE-10691 Stockholm, Sweden*
[5]*Department of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[6]*Jesus College, Turl Street, Oxford OX1 3DW, UK*
[7]*Lincoln School of Humanities, University of Lincoln, Brayford Pool, Lincoln LN6 7TS, UK*
[8]*The Roslin Institute, University of Edinburgh, Midlothian EH25 9RG, UK*
[9]*Department of Zoology, Swedish Museum of Natural History, SE-10405 Stockholm, Sweden*
[10]*Natural History Museum of Florence, via Romana 17, 50125 Florence, Italy*
[11]*Novo Nordisk Foundation Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3b, 2200 Copenhagen, Denmark*
[12]*Ancient DNA Laboratory, Murdoch University, South St, Perth, Western Australia 6150, Australia*

The understanding of Earth's biodiversity depends critically on the accurate identification and nomenclature of species. Many species were described centuries ago, and in a surprising number of cases their nomenclature or type material remain unclear or inconsistent. A prime example is provided by *Elephas maximus*, one of the most iconic and well-known mammalian species, described and named by Linnaeus (1758) and today designating the Asian elephant. We used morphological, ancient DNA (aDNA), and high-throughput ancient proteomic analyses to demonstrate that a widely discussed syntype specimen of *E. maximus*, a complete foetus preserved in ethanol, is actually an African elephant, genus *Loxodonta*. We further discovered that an additional *E. maximus* syntype, mentioned in a description by John Ray (1693) cited by Linnaeus, has been preserved as an almost complete skeleton at the Natural History Museum of the University of Florence. Having confirmed its identity as an Asian elephant through both morphological and ancient DNA analyses, we designate this specimen as the lectotype of *E. maximus*.

*Corresponding author. E-mail: ecappellini@gmail.com

INTRODUCTION

The International Code of Zoological Nomenclature (ICZN; 1999) establishes the starting date for zoological nomenclature as 1 January 1758, the year when Edition 10 of the *Systema Naturae* of Carl Linnaeus (1758) was published. References cited by Linnaeus form an integral part of the description of his species, and material that the descriptions and/or illustrations of the cited authors were based on is considered syntypic (forming part of the name-bearing type series), whether or not it was examined by Linnaeus and whether or not it still exists (ICZN Article 72.4.1). Thus, both Linnaeus' own specimens and descriptions, as well as those of the earlier authors he cited, are of equal status for zoological nomenclature. Like his predecessors, Linnaeus (1758) did not distinguish between Asian and African elephants. Amongst the authors he cited in his description of *Elephas maximus*, the figure in Gesner (1551), reproduced in Aldrovandi (1616), very likely represents an African elephant, whereas figures in Jonston (1650), pls 7–9 show an Asian elephant (Supporting Information S1), thus suggesting that *E. maximus* might have a composite type series. The African elephant was established as two separate species later: *Elephas africanus* Blumenbach, 1797 (currently *Loxodonta africana*) and *Elephas cyclotis* Matschie, 1900 (currently *Loxodonta cyclotis*).

A most remarkable specimen referred to by Linnaeus (1764) in his description of the Swedish King Adolf Fredrik's collection, is the well-preserved, near-complete, body of an elephant foetus in a spirit jar, held today at the Swedish Museum of Natural History (NRM) in Stockholm. This specimen was initially owned by the Dutch West India Company, which operated in West Africa, the Americas, and the Pacific. The foetus later became part of Albertus Seba's natural history collection, one of the richest of its time (Supporting Information S2). Between 1734 and 1765, Seba (1734, 1759, 1765) published descriptions with plates of his specimens in his 'Thesaurus'. The elephant foetus is clearly illustrated in vol. 1 (plate 111, fig. 1; 1734, here reproduced in Fig. 1A), and it is recorded (p. 175) as originating in Africa.

The Swedish sovereign and his wife bought the specimen either at the April 1752 auction of Seba's collection, or shortly thereafter from one of the other purchasers (Lovén, 1887; Engel, 1961; Boeseman, 1970). The elephant foetus was kept at Drottningholm Palace from 1773 to 1801 and was finally placed in what became the NRM (Fig. 1B). The provenance of the sample is without doubt: NRM inventories record the specimen as coming from Drottningholm, originally from Seba's cabinet, and it was mentioned by Linnaeus (1764) in his published account of the King's collection.

Despite its importance as a syntype of *E. maximus*, the species identity of the foetus has been questioned. Sundevall (1857) and Lönnberg (1904) both believed that, although it was labelled *Elephas maximus*, the specimen was identifiable as an African elephant, genus *Loxodonta*. Together with the African origin for the specimen, noted by Seba (1734), this again would suggest that *E. maximus* has a composite type series.

Here, we sought to further resolve the identity of the foetal syntype specimen of *E. maximus*. We used ancient protein and DNA sequencing (see Methods) to conclusively establish the taxonomic identity of Seba's elephant foetus as an African rather than an Asian elephant. Second, in examining John Ray's (1693) text cited by Linnaeus, reporting the description of an elephant skeleton, we found archival evidence indicating that this specimen still exists today in the collection of the Natural History Museum of Florence. Using morphological and molecular evidence, we confirm that it is the skeleton of an Asian elephant eligible for designation as the lectotype of *E. maximus* Linnaeus.

METHODS

SAMPLING

The elephant foetus, kept in the collection of the Swedish Museum of Natural History of Stockholm (catalogue number NRM 532062), was gently removed from its jar and laid on a tray. The specimen presented an extended thoracic–abdominal longitudinal opening.
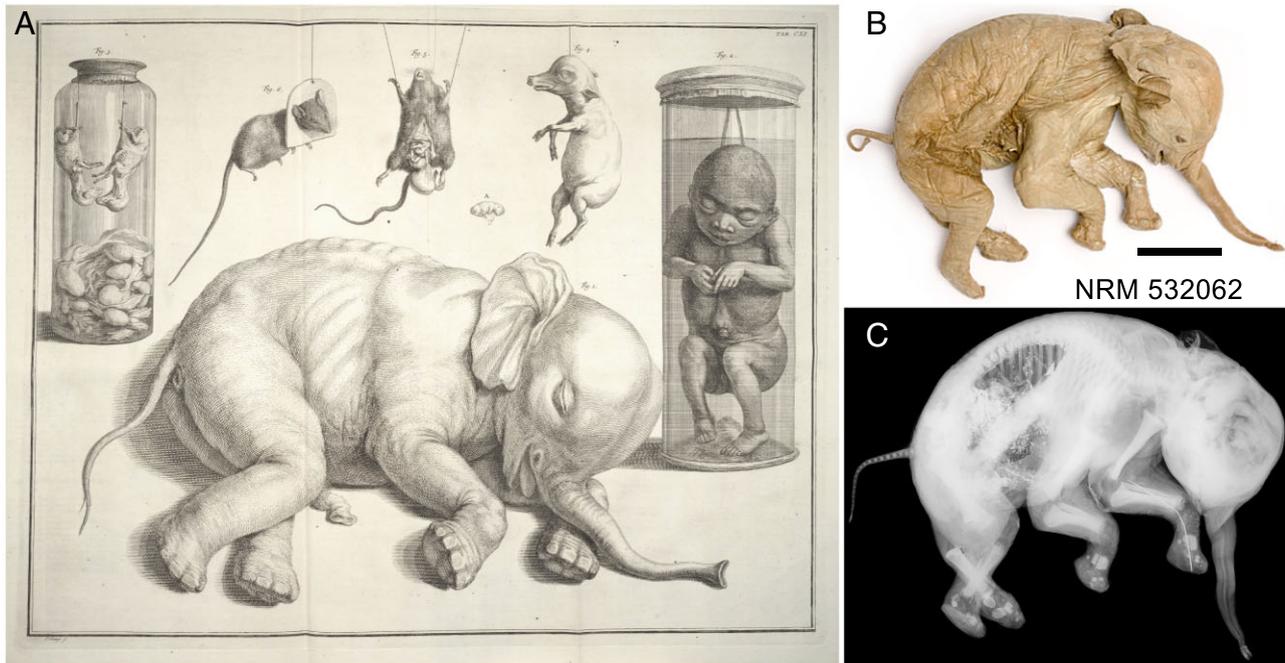
Figure 1. The elephant foetus, NRM 532062 from the Swedish Museum of Natural History, referred to by Linnaeus. A, reproduction of the foetus illustration in Figure 1, Plate 111, vol. 1 of Albertus Seba's '*Locupletissimi rerum naturalium thesauri accurata descriptio, et iconibus artificiosissimis expressio, per universam physices historiam . . .*' (1734). B, recent photograph, and C, radiograph of the same specimen. Scale bar in panel B equals approximately 10 cm.

Operating through the opening, a small, approximately 5 × 5 × 3 mm, oesophagus fragment and the terminal portion of a cartilaginous rib of similar size were removed and transferred to Copenhagen for ancient protein analysis. A second rib fragment was collected and stored using the same procedure for aDNA analysis in Stockholm.

SAMPLE PREPARATION FOR PROTEIN SEQUENCING

For ancient protein analysis, oesophagus and rib subsamples, of 64 and 39 mg (wet weight) respectively, were cut in a chemical hood, on a sterile Petri dish, using a single-use, disposable, sterile scalpel. A negative control containing no sample was prepared and analysed following the same procedure used for the ancient samples. Samples were dried in a centrifugal evaporator to remove residual ethanol. Tryptic peptides were generated using a filter-aided sample preparation protocol (Wisniewski *et al.*, 2009). Protein digestion was started by adding 4 μL 0.5 μg/μL sequencing-grade trypsin solution (Promega) and incubating at 37 °C overnight at pH 7.50–8.00. Tryptic peptides were then acidified with 10% trifluoroacetic acid to a final concentration of 0.2–0.8% to reach a pH < 2.00 and immobilized by C-18 solid phase extraction on Stage Tips (Rappsilber, Ishihama & Mann, 2003) as previously described (Cappellini *et al.*, 2012).

NANOLC-ESI-HIGH RESOLUTION TANDEM MASS SPECTROMETRY ANALYSIS

The liquid chromatography–mass spectrometry (LC-MS) system consisted of an EASY-nLC system (Thermo Scientific, Odense, Denmark) connected to a Q Exactive mass spectrometer (Thermo Scientific, Bremen, Germany) through a nano electrospray ion (ESI) source. Five μL of each peptide sample were auto-sampled onto and directly separated in a 15 cm analytical column (75 μm inner diameter) in-house packed with 3 μm C18 beads (Reprosil-AQ Pur, Dr. Maisch), using a 130-min linear gradient from 5 to 26% acetonitrile followed by a steeper linear 20-min gradient from 26 to 48% acetonitrile. Throughout the gradients a fixed concentration of 0.5% acetic acid and a flow rate of 250 nL min^{-1} were set. The effluent from the high-performance LC was directly electrosprayed into the mass spectrometer by applying 2.0 kV through a platinum-based liquid-junction. The settings used were as described for the 'sensitive' acquisition by Kelstrup *et al.* (2012). The mass spectrometry proteomics data have been deposited in the

ProteomeXchange Consortium (http://proteomecentral .proteomexchange.org) via the PRIDE partner repository (Vizcaíno *et al*., 2013) with the data set identifier PXD000423.

Spectra matching – MaxQuant search

Raw files of the two processed samples, generated during spectra acquisition, were merged and searched on a workstation [64523 tandem mass spectrometry (MS/MS) spectra, 31561 and 32962 from oesophagus and the rib samples, respectively, were submitted for analysis] using the MaxQuant algorithm v. 1.2.2.5 (Cox & Mann, 2008) and the Andromeda peptide search engine (Cox *et al*., 2011), initially against the target/reverse protein list included in the *L. africana* reference proteome (downloaded from UniProtKB, www.uniprot.org, on 29 February 2012), and then, in a separate search, against the complete list of proteins available on the same date in UniProtKB after taxonomic restriction to *E. maximus*. In every search, spectra were also matched against the common contaminants such as wool keratins and porcine trypsin, downloaded from UniProt.

To identify more genus-diagnostic tryptic peptides, we adopted a proteogenomics approach (Sanders *et al*., 2011) that: (1) extracted, from the *L. africana* genome (loxAfr3.67 downloaded from Ensembl Genome Browser, www.ensembl.org), the DNA sequences coding for the tryptic peptides confidently identified by matching the spectra from the Seba elephant against the African elephant reference proteome; (2) extended for 120 bases each of the ends of the identified coding regions; (3) assembled, if possible, the DNA sequences previously generated into short contigs, to obtain the coding DNA sequence for the ancient proteins identified on the basis of several contiguous peptides; (4) mapped a set of shotgun Illumina DNA reads, generated at the Roslin Institute from two modern Asian elephants, against the *L. africana* DNA extended reads and short contigs coding the peptides identified in the Seba foetus; and (5) filtered to retain alternate, homozygous, nonsynonymous single nucleotide polymorphisms (SNPs) in the positions coding for the ancient foetus tryptic peptides, allowing us to identify which of the peptides identified in the ancient sample had genus-diagnostic potential. Detailed description of this procedure is reported as Supporting Information S3.

Identification of modern DNA sequences coding for the *Loxodonta*/*Elephas* diagnostic peptides identified by mass spectrometry-based protein sequencing

African savannah elephant sequences of *HBB/D*, *IARS*, *SPR, GSTP* and *COL6A3* (full names are listed in Table 1) were identified in the National Center for Biotechnology Information (NCBI) *Loxodonta africana* genome Trace Archives (http://www.ncbi .nlm.nih.gov/Traces/home) using MegaBlast (Zhang *et al*., 2000) and also NCBI GenBank entries (*COL6A3*: XM_003417935, *GSTP*: XM_003419394, *IARS*: XM_003421084, *SPR*: XM_003413564, *HBB/D*: XM_003421554). African savannah elephant sequences were aligned with their Asian elephant (*E. maximus*) homologues, obtained as described in Supporting Information S3.5–6, and primers were designed at conserved regions to amplify genus-diagnostic nonsynonymous codon mutations, using the software PRIMER3 (http://fokker.wi.mit.edu/ primer3/input.htm; Rozen & Skaletsky, 2000). One or two sets of primers were designed to amplify each region (Supporting Information Table S3.4).

PCR and sequencing reactions were performed using M13 tailed primers (M13 forward tail sequence at the 5′ end of the forward primers and M13 reverse tail sequence at the 5′ end of the reverse primers), following the protocol by Ishida *et al*. (2011). Six African savannah elephants, five African forest elephants, and five Asian elephants were sequenced for each of the five genes (Supporting Information Table S3.5). Sequences were aligned to *L. africana* predicted mRNA (GenBank numbers as above) and the coding regions were translated using the ExPASy-translate tool (http://web.expasy.org/translate/; Artimo *et al*., 2012) to identify and confirm genus-diagnostic amino acid differences.

Amplification and sequencing of mtDNA, including diagnostic sites distinguishing *Loxodonta* from *Elephas*

Sequences of 4258 bp of mtDNA (from *MT-ND5* to control region) were aligned for 653 *Loxodonta* (GenBank accession numbers: JQ438119–JQ438771; Ishida *et al*., 2013) and 73 *E. maximus* (unpubl. data). Three short regions that contained genus-diagnostic nucleotide sites were identified, and primers were designed with conserved sequences at 3′ ends, using the software PRIMER3 (Supporting Information Table S3.6; Rozen & Skaletsky, 2000).

In anticipation of their application to ancient DNA, primers were tested using highly diluted modern elephant DNA. Human genomic DNA was also tested as a negative control to establish that primers would not amplify human DNA, which sometimes may contaminate ancient samples. All work involving modern DNA was performed at the University of Illinois at Urbana-Champaign, with primers ordered separately for use by the ancient DNA laboratory (Ishida *et al*., 2011).

Table 1. Genus-diagnostic tryptic peptides supporting the attribution of the ancient elephant foetus to *Loxodonta*. Genus-distinctive amino acids are in bold. Their position refers to the protein sequence identified by the accession number reported.

N	Sequence	Accession number	Protein name	Length (aa)	Mass (Da)	Charges	PEP	MaxQuant score	Matched spectra	Genus-specific aa substitution position*
1	**FFEHFGDLSTAEAVLHNAK**	P02085/Q45XJ0	Haemoglobin sub. Beta/Delta	19	2132.033	2,3,4	7.35E-15	161.24	7	52 (D)
2	EAALVDVVNDGVEDLR	G3TPD1	Glutathione S-transferase P-like	16	1712.858	2	7.92E-12	141.46	6	93 (A)
3	LYLINSPVVR	G3TAH7	Isoleucyl-tRNA synthetase, cytoplasmic	10	1172.692	2	0.003211	102.61	1	630 (M)
4	**NMMFQVLAAEEPTVR**	G5E6U0	Sepiapterin reductase-like	15	1734.843	2	3.34E-05	119.01	2	185 (V)
5	VAPLQGVLPSLLAPLR	G3TBD6/G3U3T4	Collagen alpha-3(VI)	16	1643.013	2,3	3.44E-06	123.38	9	208 (M)

PEP, posterior error probability; aa, amino acid; *the corresponding *Elephas*-distinctive amino acid substitutions are reported in parentheses.

DNA AMPLIFICATION AND SEQUENCING OF mtDNA FROM THE HISTORICAL SEBA AND FLORENCE ELEPHANT SAMPLES

Given the degraded nature of DNA in ancient specimens, pre-PCR laboratory work was performed in a dedicated ancient DNA laboratory at the Swedish Museum of Natural History (Stockholm). Negative controls and repeated (once or more) amplification were used to minimize the effects of potential contamination and to monitor damaged sites. DNA extraction was performed on a small, approximately $5 \times 5 \times 3$ mm, piece of soft-tissue from the elephant foetus, using a QIAamp DNA micro kit (Qiagen) and following the manufacturer's protocol for isolation of genomic DNA from tissues. To minimize alteration of the ancient specimen, a bone microcore, 5 mm diameter, was collected from the left humerus of the Florence elephant skeleton using a dentist's trephine (Asa Dental, Italy). Approximately 100–200 mg of bone powder obtained from the microcore were used for DNA extraction, following a modified version of protocol C in Yang *et al*. (1998). The extraction buffer included 1 M urea instead of sodium dodecyl sulphate and the post-lysis extraction buffer was concentrated using a 30 K MWCO Vivaspin filter (Sartorius). Silica spin columns were used for final DNA purification.

For the Seba elephant foetus, three pairs of mitochondrial DNA primers mentioned above (Supporting Information Table S3.6) were used to amplify DNA from the specimen. Additionally, two short fragments of the nuclear genes Biglycan (*BGN*) and Phosphorylase kinase alpha subunit 2 (*PHKA2*) that contain diagnostic sites that distinguish African forest from African savannah elephants were also targeted, using primer pairs BGN-s1F2/BGN-s1R2 and PHK-s1F/PHK-s1R, respectively (Supporting Information Table S3.7; Ishida *et al*., 2011). For PCR amplification of these DNA fragments, the M13 tailed primers were used to increase amplicon length and the quality of the sequences (Supporting Information Table S3.7; Ishida *et al*., 2011). We also included M13 tails in some of the mitochondrial DNA primer pairs.

For the Florence skeleton, 741 bp of mitochondrial DNA were amplified in seven overlapping fragments, using primer pairs from Barnes *et al*. (2007). This region includes the 3′ end of the cytochrome *b* gene (*MT-CYTB*), two tRNA genes: mitochondrially encoded tRNA threonine (*MT-TT*) and mitochondrially encoded tRNA proline (*MT-TP*), and the first hypervariable part of the control region (CR1).

PCR reactions had a final concentration of $1 \times$ PCR buffer, 0.2 µM of each primer, 200 µM deoxyribonucleotide triphosphates, 2.5 mM $MgCl_2$, 0.1 mg mL^{-1} bovine serum albumin, 2 units HotStarTaq DNA polymerase (Qiagen), purified water, and 2 µL DNA extract

in a total volume of 25 µL mix. PCR conditions included an initial denaturation step of 10 min at 95 °C, followed by 55 cycles of denaturation of 30 s at 94 °C, annealing of 30 s at 50 or 52 °C (depending on the primer pair; Supporting Information Tables S3.6, S3.7), extension of 30 s at 72 °C, and a final extension step of 7 min at 72 °C. Amplicons were purified with Exonuclease I and FastAP Thermosensitive Alkaline Phosphatase. Sequencing was performed in both directions on an ABI3130xl (Applied Biosystems) automated sequencer at the Molecular Systematics Laboratory, at the Swedish Museum of Natural History (Stockholm). Sequences were assembled in GENEIOUS 5.0.1 (http://www.geneious.com).

RESULTS

MORPHOLOGY OF THE SEBA ELEPHANT FOETUS

Morphological comparison of foetal organisms can be challenging. Nevertheless, preliminary inspection suggested that the foetus could be assigned to *Loxodonta*, based on comparatively large ears and two finger-like processes on the tip of the trunk (Fig. 1B; Shoshani, 2000). X-ray images of the specimen (Fig. 1C) were also taken to enable counting of the ribs and the thoracic, lumbar, and caudal vertebrae, which differ in number between African and Asian elephants (Shoshani, 2000). However, not all of these elements could be individually distinguished on the X-ray images owing to the incomplete development of the animal.

ANALYSIS OF ANCIENT PROTEINS

To resolve the identity of the Seba elephant foetus, we analysed DNA and sequenced proteins (Cappellini *et al.*, 2012) using a shotgun proteomics approach from a small section of the oesophagus and the terminal portion of a cartilaginous rib. Nano-liquid chromatography coupled with high-resolution MS/MS was used to generate a spectra data set from tryptic peptides and identify, as genus-diagnostic markers, the non-identical homologous ones present in publicly available protein lists for both elephant genera. The total mass of proteins recovered before digestion was 126 and 90 µg, whereas after digestion it was 72 and 60 µg, for oesophagus and rib extracts, respectively. However, as the nuclear genome of *E. maximus* was not publicly available at the time of analysis, and the list of publicly available protein sequences from this species is quite limited, at only *c.* 200, only seven proteins and 48 peptides identified in the Seba elephant sample were available for both *L. africana* and *E. maximus*. Despite such a limited set of homologous proteins and peptides, we did nevertheless identify one variable tryptic peptide from haemoglobin subunit Beta/Delta (Opazo *et al.*, 2009): FFEHFGDLSTAE$_{52}$AVLHNAK,

with start-end amino acid positions 41–59 in both UniProt accessions P02085 and P02084, for *L. africana* and *E. maximus*, respectively. This protein had a putative genus-diagnostic amino acidic substitution at position 52, namely Glu (E) in the African and Asp (D) in the Asian elephant. As the spectrum reported in Figure 2 demonstrates, the assigned peptide sequence clearly indicated that the Seba foetus had the African elephant version of the peptide. Parameters supporting the identification of the genus-diagnostic peptide and haemoglobin subunit Beta/ Delta are reported in Table 1 and Supporting Information Table S3.8, respectively.

The proteogenomics approach returned four nonsynonymous SNPs falling within regions coding for the peptides identified in the Seba elephant sample. At each of the four additional genus-diagnostic peptides, the Seba elephant matched *Loxodonta* and not *Elephas*. Proteomics parameters supporting the identification of the additional four genus-diagnostic peptides, and of the corresponding *L. africana* proteins that they are part of, are reported in Tables 1 and S3.8, respectively, whereas examples of their MS/MS spectra are visible in Supporting Information Figures S3.1–4.

PCR-amplification and Sanger-sequencing of the DNA regions coding for the five genus-diagnostic tryptic peptides identified in the elephant foetus, applied to 16 modern elephant individuals of diverse geographical provenance, confirmed the *Loxodonta*/ *Elephas* diagnostic capacity of these peptides (Supporting Information Table S3.5).

ANCIENT DNA ANALYSIS

Three regions of mtDNA were successfully amplified and sequenced for the Seba foetal elephant specimen, using primer pairs Ele-ND5-F1/R1, Ele-ND5-F3/R3, and Ele-cytB-1/R1. The three sequences were compared to *Loxodonta* and *Elephas* sequences. All diagnostic sites matched *Loxodonta* and not *Elephas* sequences. Two of the Seba elephant mtDNA fragment sequences (Ele-ND5-F1/R1 and Ele-ND5-F3/R3) included a site diagnostic for the west-central subclade, and matched the character state present in the west-central subclade at all informative sites (Supporting Information Table S3.9; Ishida *et al.*, 2013). By contrast, three to six character state mismatches separated the Seba foetus sequence from the seven other major mtDNA subclades identified in African elephants (Ishida *et al.*, 2013). The geographical distribution of this subclade notably corresponds to the west-central region of Africa in which the Dutch West India Company had conducted its African operations since the first half of the 17th century. Identification of the species of the elephant foetus as African savannah or African forest elephant was not
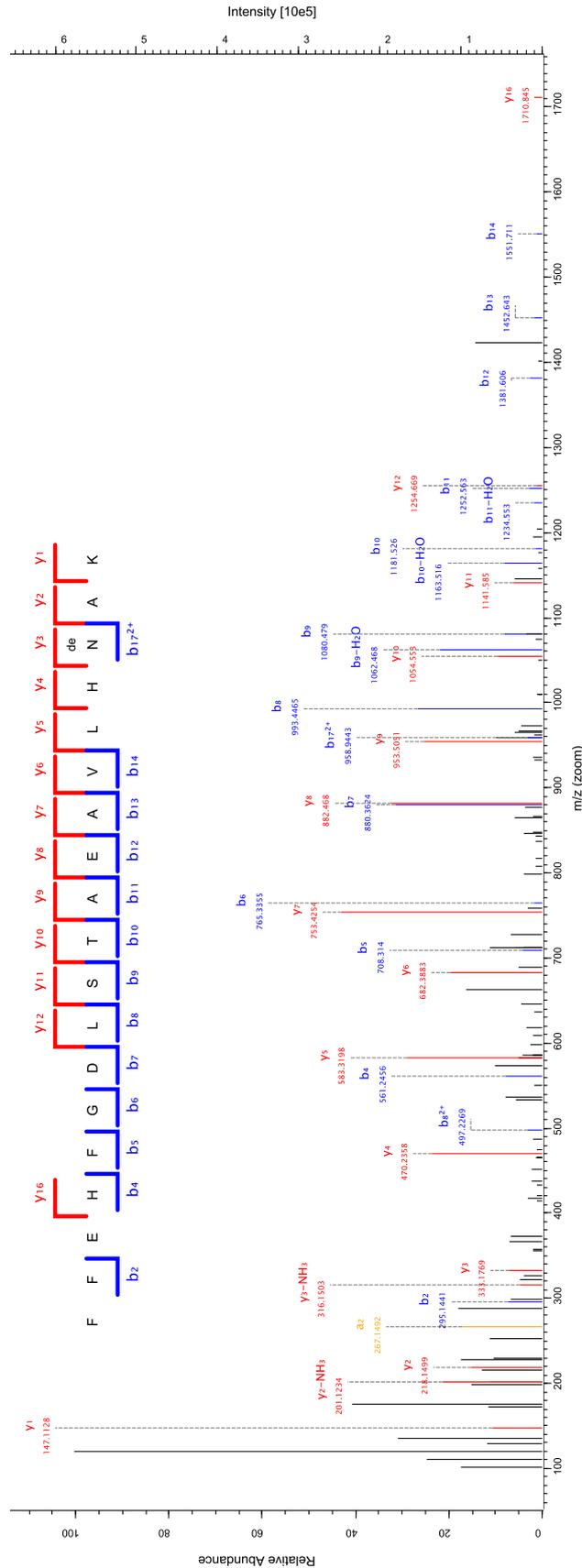
Figure 2. Tandem mass spectrum, generated from the analysis of the Seba elephant foetus sample, supporting identification of the genus-diagnostic tryptic peptide $F_{41}FEHFGDLSTAE_{52}AVLHNAK_{59}$, from haemoglobin subunit Beta/Delta, (UniProt accession number: P02085). $\Delta m/z_{b_{12}-b_{11}}$ and $\Delta m/z_{y_8-y_7} = 129.043$ support the presence of Glu in position 52. Peaks in the spectrum indicate the detected mass-to-charge ratios of peptide fragmentation products. Peaks that support assignment of the spectrum to the amino acidic sequence in the figure are colored (blue and red indicate b- and y-series fragments, respectively). Resulting fragmentation pattern is overlaid on the sequence.
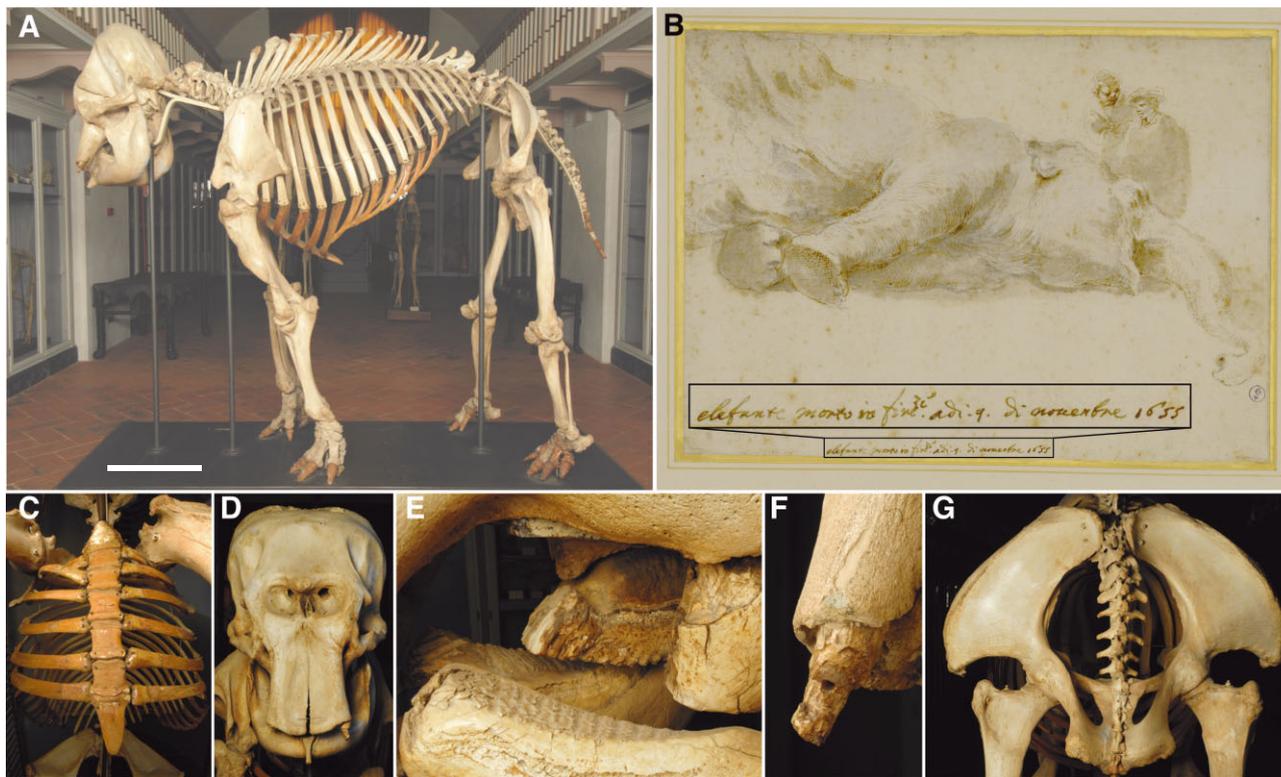
Figure 3. A, elephant skeleton MZUF-734 at the Natural History Museum of the University of Florence. B, drawing by Stefano Della Bella: 'Elephant died in Florence on November 9ᵗʰ 1655', inset (translation from Italian by E.C.). Anatomical and mounting evidence in support of the specimen's identity, genus, sex, and age as presented in the text: C, sternum wooden replica; D, cranium; E, molar teeth; F, vestigial tusks; G, pelvis. Scale bar in panel A, in white, equals approximately 50 cm.

possible because nuclear loci could not be sequenced (Ishida *et al.*, 2011). Despite the short length of target amplicons and attempts to optimize the PCR reactions/conditions, the degraded DNA of the foetus failed to amplify for two nuclear loci with sites diagnostic between forest and savannah elephants.

IDENTIFICATION OF A LECTOTYPE

Having established that the Seba foetus was *Loxodonta*, we exhaustively scrutinized the references cited by Linnaeus (1758). In one of these references, John Ray's '*Synopsis Methodica Animalium Quadrupedum et Serpentini Generis*' (1693: 131), we found a detailed description of the skeleton of an elephant that Ray had observed while visiting Florence between 14 July and 1 September 1664 (Supporting Information S4). His description (Supporting Information S5), and that of his travelling companion Phillip Skippon (1732), do not contain information sufficient to identify the specimen as either Asian or African elephant. However, anatomical details, together with other historical sources, clearly indicate that the skeleton is the one still on display at the Natural History Museum of the University

of Florence (Fig. 3A). Records indicate that these remains belong to an elephant that died in Florence on 9 November 1655 (Supporting Information S6), as reported in a drawing by the artist Stefano Della Bella (Fig. 3B). This was probably the itinerant performing elephant known as 'Hansken', born in Ceylon (today Sri Lanka), and portrayed in 1637 by the Dutch artist Rembrandt van Rijn (Supporting Information S7; Slatkes, 1980). After dissection, the skeleton was mounted and exhibited in a room at the Uffizi Gallery (Targioni Tozzetti, 1780) until, in September 1771 the Grand Duke of Tuscany ordered the relocation of all specimens of scientific interest into the newly constituted Museum of Physics and Natural History, now Natural History Museum. The continuous presence of the elephant remains at the Uffizi Gallery first, then at the Natural History Museum, is documented by several bibliographical sources (Ray, 1673, 1693; Skippon, 1732) and museum inventories (Supporting Information S6). In addition, the current wooden replica of the sternum (Fig. 3C) conforms to Ray's indication that this element was missing, whereas further anatomical peculiarities mentioned by Skippon (1732) can be observed today in the Florence skeleton: fusion of the second and third

cervical vertebrae, and fusion of the spinous processes of 18th and 19th dorsal vertebrae.

Morphologically, the skeleton is clearly that of an Asian elephant. Features include the high cranial apex with a double parietal dome (Fig. 3D), and the characteristic form of the enamel loops of the molars, all in distinction to *Loxodonta* (Shoshani, 2000). The jaws preserve the second molars (M2) in wear and the third (M3) close to eruption (M5–6 of some authors), indicating (Fig. 3E) an adult animal of *c*. 25–30 years at death (Roth & Shoshani, 1988). The specimen's relatively small size (shoulder height *c*. 230 cm), vestigial tusks – tushes – (Fig. 3F), and wide pelvic aperture (Fig. 3G), indicate that the individual was female (Lister, 1996).

Analysis of a 741-bp region of mitochondrial sequence (using DNA derived from a small fragment of the left humerus) also confirmed the identity of the Florence specimen as *E. maximus*. The sequence was a novel haplotype, closest to the previously reported haplotype BP (= MDL2), from which it differed at one nucleotide site (Vidya, Sukumar & Melnick, 2009; Lei *et al*., 2012). Although particularly common in Sri Lanka, haplotype BP is found widely across Asia. Thus, further resolution of the specimen's geographical origin was not possible with current molecular data, but for nomenclatural stability, the type locality of *E. maximus* should continue to be understood to be the island of Ceylon ('Zeylonae' of Linnaeus, 1758).

CONCLUSION

We established that the Seba foetus was *Loxodonta*, and thus that Linnaeus's original conception of the taxon *E. maximus* is composite. Consequently, designation of a lectotype as the single name-bearing specimen is necessary to fix the identity of the species for all future work on its biology, phylogeny, and conservation.

Exhaustive examination of references cited by Linnaeus (1758) led us to a detailed description of the skeleton of an elephant in Florence. We identified the described skeleton as corresponding to an extant specimen that proved to be on display at the Natural History Museum of the University of Florence, and which both morphological and ancient DNA identified as an Asian elephant.

Based on this evidence, and in conformity with Article 74.7 of the ICZN, we hereby designate the elephant skeleton in Florence, catalogue no. MZUF-734, as the lectotype of *E. maximus* Linnaeus, 1758, to preserve the traditional understanding and application of this name to the Asian elephant. Not only was it amongst the type series but its anatomical completeness makes it a suitable name-bearing speci-men from the original type series for this iconic mammal (Supporting Information S8).

This resolution of nomenclatural inconsistencies was made possible, 320 years after the first published description of the hereby designated lectotype specimen and 255 years after the naming of the species, by integrating state-of-the-art experimental research, based on high-throughput ancient protein and DNA sequencing, and investigation of the historical literature, streamlined by digitalization and online availability of historical texts. The electronic version of this article, published online ahead of print, is registered in the ZooBank database (zoobank.org) with the following LSID: urn:lsid:zoobank.org:pub:1E9CC99D-6037-41DA-AFE0-EB9BDF603749.

REFERENCES

Aldrovandi U. 1616. *De quadrupedibus solidipedibus, volumen integrum.* Bologna.

Artimo PP, Jonnalagedda MM, Arnold KK, Baratin DD, Csardi GG, de Castro EE, Duvaud SS, Flegel VV, Fortier AA, Gasteiger EE, Grosdidier AA, Hernandez CC, Ioannidis VV, Kuznetsov DD, Liechti RR, Moretti SS, Mostaguir KK, Redaschi NN, Rossier GG, Xenarios II, Stockinger HH. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research* **40:** W597–W603.

Barnes I, Shapiro B, Lister A, Kuznetsova T, Sher A, Guthrie D, Thomas MG. 2007. Genetic structure and extinction of the woolly mammoth. *Mammuthus primigenius. Current Biology* **17:** 1072–1075.

Boeseman M. 1970. The vicissitudes and dispersal of Albertus Seba's zoological specimens. *Zoologische Mededelingen* **44:** 177–206.

Cappellini E, Jensen LJ, Szklarczyk D, Ginolhac A, da Fonseca RAR, Stafford TW, Holen SR, Collins MJ, Orlando L, Willerslev E, Gilbert MTP, Olsen JV. 2012. Proteomic analysis of a Pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *Journal of Proteome Research* **11:** 917–926.

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26:** 1367–1372.

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of Proteome Research* **10:** 1794–1805.

Engel H. 1961. The sale-catalogue of the Cabinets of Natural History of Albertus Seba (1752). A curious document from the period of the Naturae Curiosi. *Bulletin of the Research Council of Israel, Section B (Zoology)* **10B:** 119–131.

Gesner C. 1551. *Historiae animalium lib. I de quadrupedibus viviparis* . . . Zurich: Froschover. 1104.

International Commission on Zoological Nomenclature. 1999. *International Code of Zoological Nomenclature.* London: International Trust for Zoological Nomenclature.

Ishida Y, Demeke Y, van Coeverden de Groot PJ, Georgiadis NJ, Leggett KEA, Fox VE, Roca AL. 2011. Distinguishing forest and savanna African elephants using short nuclear DNA sequences. *The Journal of Heredity* **102:** 610–616.

Ishida Y, Georgiadis NJ, Hondo T, Roca AL. 2013. Triangulating the provenance of African elephants using mitochondrial DNA. *Evolutionary Applications* **6:** 253–265.

Jonston J. 1650. *Historiae naturalis de quadrupedibus liber 1.* De quadrupedibus Solipedibus. Marianus: Frankfurt.

Kelstrup CD, Young C, Lavallee R, Nielsen ML, Olsen JV. 2012. Optimized fast and sensitive acquisition methods for shotgun proteomics on a Quadrupole Orbitrap Mass Spectrometer. *Journal of Proteome Research* **11:** 3487–3497.

Lei R, Brenneman RA, Schmitt DL, Louis EEJ. 2012. Genetic diversity in North American captive Asian elephants. *Journal of Zoology* **286:** 38–47.

Linnaeus C. 1758. *Systema naturae*, Ed. 10. Salvius: Holmiae.

Linnaeus C. 1764. *Museum S:ae R:ae M:tis Adolphi Friderici, Regis* . . . Salvius: Holmiae.

Lister AM. 1996. Sexual dimorphism in the mammoth pelvis: an aid to gender determination. In: Shoshani J, Tassy P, eds. *The Proboscidea.* Oxford: Oxford University Press, 254–259.

Lönnberg E. 1904. *Demonstration eines Fötus vom westafrikanischen Elephant, Elephas cyclotis Matschie.* Comptes rendus du 6e Congrès international de Zoologie, Session de Berne. 3me Section: Vertébrés (Anatomie).

Lovén S. 1887. On the species of Echinoidea described by Linnaeus in his Museum Ludovicae Ulricae. *Bihang till Kongl. Svenska Vetenskaps-Akademiens Handlingar* **13**(4,5): 1–183.

Opazo JC, Sloan AM, Campbell KL, Storz JF. 2009. Origin and ascendancy of a chimeric fusion gene: the beta/delta-globin gene of paenungulate mammals. *Molecular Biology and Evolution* **26:** 1469–1478.

Rappsilber J, Ishihama Y, Mann M. 2003. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical Chemistry* **75:** 663–670.

Ray J. 1673. *Observations Topographical, Moral, and Physiological Made in a Journey Through Part of the Low-Countries, Germany, Italy, and France.* London: Martyn.

Ray J. 1693. *Synopsis methodica animalium quadrupedum et serpentini generis.* London: Smith & Walford.

Roth VL, Shoshani J. 1988. Dental identification and age-determination in *Elephas maximus. Journal of Zoology* **214:** 567–588.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Misener S, Krawetz SA, eds. *Bioinformatics methods and protocols.* Totowa: Humana Press, 365–386.

Sanders WS, Wang N, Bridges SM, Malone BM, Dandass YS, McCarthy FM, Nanduri B, Lawrence ML, Burgess SC. 2011. The proteogenomic mapping tool. *BMC Bioinformatics* **12:** 115–121.

Seba A. 1734, 1759, 1765. *Locupletissimi rerum naturalium thesauri accurata descriptio et iconibus artificiosissimis expressio per universam physices historiam.* Amsterdam.

Shoshani J. 2000. *Elephants: majestic creatures of the wild.* New York: Checkmark Books.

Skippon P. 1732. *An Account of a Journey Made Thro' Part of the Low-Countries, Germany, Italy, and France.* London: Walthoe *et al*. 359–738.

Slatkes LJ. 1980. Rembrandt's elephant. *Simiolus: Netherlands Quarterly for the History of Art* **11:** 7–13.

Sundevall CJ. 1857. Mammalia in Spiritu Vini asservata Mus[eum] Stockh[olmiensis]: Unpublished handwritten inventory, NRM archives.

Targioni Tozzetti G. 1780. *Atti e memorie inedite dell' Accademia del cimento e notizie aneddote dei progressi delle scienze in Toscana*. Firenze.

Vidya TNCT, Sukumar RR, Melnick DJD. 2009. Range-wide mtDNA phylogeography yields insights into the origins of Asian elephants. *Proceedings. Biological Sciences/The Royal Society* **276:** 893–902.

Vizcaíno JA, Côté RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O'Kelly G, Schoenegger A, Ovelleiro D, Pérez-Riverol Y, Reisinger F, Ríos D, Wang R, Hermjakob H. 2013. The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Research* **41:** D1063–D1069.

Wisniewski JR, Zougman A, Nagaraj N, Mann M. 2009. Universal sample preparation method for proteome analysis. *Nature Methods* **6:** 359–363.

Yang DY, Eng B, Waye JS, Dudar JC, Saunders SR. 1998. Improved DNA extraction from ancient bones using silica-based spin columns. *American Journal of Physical Anthropology* **105:** 539–543.

Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology* **7:** 203–214.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

S1. Iconography of the sources, other than Seba and Ray, cited by Linnaeus (1758) in the description of *Elephas maximus*.

Figure S1.1. Gesner, C. 1551. *Historiae Animalium Lib. 1 de quadrupedibus viviparis*. [39], 1104. [11] pp. Froschover, Zurich.

Figure S1.2. Aldrovandi, U. 1616. *De quadrupedibus solidipedibus volumen integrum*. 495 pp. Bologna.

Figure S1.3. Jonston, J. 1650. Historiae naturalis de quadrupedibus liber 1. De quadrupedibus Solipedibus. 231 pp., 80 pls. Marianus, Frankfurt.

S2. Albertus Seba, his collection and the fate of his elephant foetus.

S3. Ancient protein sequencing and identification of genus-diagnostic peptides: materials and methods.

Figure S3.1. Tandem spectrum, generated by nanoLC-MS/MS analysis of the ancient foetus sample, supporting identification of the *L. africana* version of the genus-diagnostic tryptic peptide $E_{86}AALVDVV_{93}NDGVEDLR_{101}$.

Figure S3.2. Tandem spectrum, generated from the analysis of the Seba elephant sample, supporting identification of the *L. africana* version of the genus-diagnostic tryptic peptide $L_{627}YLI_{630}NSPVVR_{636}$.

Figure S3.3. Tandem spectrum, generated from the analysis of the Seba elephant sample, supporting identification of the *L. africana* version of the genus-diagnostic tryptic peptide $N_{184}M_{185}MFQVLAAEEPTVR_{198}$.

Figure S3.4. Tandem spectrum, generated from the analysis of the Seba elephant sample, supporting identification of the *L. africana* version of the genus-diagnostic tryptic peptide $V_{202}APLQGV_{208}LPSLLAPLR_{217}$.

Table S3.1. Protein recoveries, before and after digestion, for each of the two Seba foetus extracted samples.

Table S3.2. Statistics for MaxQuant MS/MS spectra matching against the *Loxodonta africana* reference proteome.

Table S3.3. Statistics for MaxQuant MS/MS spectrometry spectra matching against the *Elephas maximus* complete protein list.

Table S3.4. Primer sets used to amplify *Loxodonta/Elephas* DNA regions coding for genus-diagnostic peptides.

Table S3.5. Variable and diagnostic sites across geographically diverse Asian, African forest and African savannah elephants.

Table S3.6. Primers used to amplify *Loxodonta/Elephas* diagnostic sites in mitochondrial DNA.

Table S3.7. PCR primers and M13 tails used in this study to attempt amplification of nuclear nucleotide sites diagnostic between African forest and African savannah elephants (Ishida et al., 2011).

Table S3.8. Evidence supporting the identities of proteins bearing genus-diagnostic tryptic peptides.

Table S3.9. Clade-informative sites overlapping mtDNA sequences generated for the Seba elephant foetus.

S4. John Ray and his tour across part of Europe.

S5. Translation of Ray's Latin text mentioning the elephant he observed in Florence.

S6. Reconstruction of the fate of the Florence elephant.

Table S6.1. Inventory numbers documenting the presence of the elephant skeleton observed by John Ray in the Florence Royal Museum since the end of the 18th century.

Table S6.2. Comparison of the measures of the elephant skeleton reported by Targioni Tozzetti (1763) and those of specimen 734 as measurable today.

S7. Possible identification of the Florence elephant as the performing itinerant elephant known as 'Hansken'.

Figure S7.1. Rembrandt, An elephant, 1637 (drawing). Vienna, Albertina.

S8. Another extant syntype specimen of *Elephas maximus*: a partial elephant tooth in Uppsala.

Figure S8.1. *Elephas maximus* molar fragment (catalogue number UUZM 370) in the Uppsala University Zoological Museum collection.