

1 **Report**

2 **Title: Morphometric, behavioral, and genomic evidence for a new**
3 **orangutan species**

4 **Authors:** Alexander Nater^{1,2,3§*}, Maja P. Mattle-Greminger^{1,2§}, Anton Nurcahyo^{4§}, Matthew G.
5 Nowak^{5,6§}, Marc de Manuel⁷, Tariq Desai⁸, Colin Groves⁴, Marc Pybus⁷, Tugce Bilgin Sonay¹,
6 Christian Roos⁹, Adriano R. Lameira^{10,11}, Serge A. Wich^{12,13}, James Askew¹⁴, Marina Davila-Ross¹⁵,
7 Gabriella Fredriksson^{5,13}, Guillem de Valles⁷, Ferran Casals¹⁶, Javier Prado-Martinez¹⁷, Benoit
8 Goossens^{18,19,20,21}, Ernst J. Verschoor²², Kristin S. Warren²³, Ian Singleton^{5,24}, David A. Marques^{1,25},
9 Joko Pamungkas^{26,27}, Dyah Perwitasari-Farajallah^{26,28}, Puji Rianti^{28,26,1}, Augustine Tuuga²⁰, Ivo G.
10 Gut^{29,30}, Marta Gut^{29,30}, Pablo Orozco-terWengel¹⁸, Carel P. van Schaik¹, Jaume Bertranpetit^{7,31}, Maria
11 Anisimova^{32,33}, Aylwyn Scally⁸, Tomas Marques-Bonet^{7,29,34}, Erik Meijaard^{4,35*} and Michael
12 Krützen^{1*}

13 §These authors contributed equally to this work.

14 *Correspondence to: alexander.nater@uzh.ch, emeijaard@gmail.com (EM),
15 michael.krutzen@aim.uzh.ch (MK)

16 **Affiliations:**

17 ¹Evolutionary Genetics Group, Department of Anthropology, University of Zurich,
18 Winterthurerstrasse 190, 8057 Zürich, Switzerland.

19 ²Department of Evolutionary Biology and Environmental Studies, University of Zurich,
20 Winterthurerstrasse 190, 8057 Zürich, Switzerland.

21 ³Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz,
22 Universitätsstrasse 10, 78457 Konstanz, Germany.

23 ⁴School of Archaeology and Anthropology, Australian National University, Canberra, Australia.

24 ⁵Sumatran Orangutan Conservation Programme (PanEco-YEL), Jalan Wahid Hasyim 51/74, Medan
25 20154, Indonesia.

26 ⁶Department of Anthropology, Southern Illinois University, 1000 Faner Drive, Carbondale, IL 62901,
27 USA.

28 ⁷Institut de Biologia Evolutiva (UPF-CSIC), Universitat Pompeu Fabra, Doctor Aiguader 88,
29 Barcelona 08003, Spain.

A NEW SPECIES OF ORANGUTAN

- 30 ⁸Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK.
- 31 ⁹Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute
32 for Primate Research, 37077 Göttingen, Germany.
- 33 ¹⁰Department of Anthropology, Durham University, Dawson Building, South Road, Durham, DH1
34 3LE, UK.
- 35 ¹¹School of Psychology & Neuroscience, St. Andrews University, St Mary's Quad, South Street, St.
36 Andrews, Fife, KY16 9JP, Scotland, United Kingdom.
- 37 ¹²School of Natural Sciences and Psychology, Liverpool John Moores University, James Parsons
38 Building, Byrom Street, L33AF Liverpool, UK.
- 39 ¹³Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Sciencepark 904,
40 Amsterdam 1098, Netherlands.
- 41 ¹⁴Department of Biological Sciences, University of Southern California, 3616 Trousdale Parkway,
42 Los Angeles, CA 90089, USA.
- 43 ¹⁵Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1st Street,
44 Portsmouth, PO1 2DY, UK.
- 45 ¹⁶Servei de Genòmica, Universitat Pompeu Fabra, Doctor Aiguader 88, Barcelona 08003, Spain.
- 46 ¹⁷Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK.
- 47 ¹⁸School of Biosciences, Cardiff University, Sir Martin Evans Building, Museum Avenue, Cardiff
48 CF10 3AX, UK.
- 49 ¹⁹Danau Girang Field Centre, c/o Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu,
50 Sabah, Malaysia.
- 51 ²⁰Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia.
- 52 ²¹Sustainable Places Research Institute, Cardiff University, 33 Park Place, Cardiff CF10 3BA, UK.
- 53 ²²Department of Virology, Biomedical Primate Research Centre, Lange Kleiweg 161, 2288GJ
54 Rijswijk, The Netherlands.
- 55 ²³Conservation Medicine Program, College of Veterinary Medicine, Murdoch University, South
56 Street, Murdoch 6150, Australia.
- 57 ²⁴Foundation for a Sustainable Ecosystem (YEL), Medan, Indonesia.
- 58 ²⁵Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland.
- 59 ²⁶Primate Research Center, Bogor Agricultural University, Bogor 16151, Indonesia.

60 ²⁷Faculty of Veterinary Medicine, Bogor Agricultural University, Darmaga Campus, Bogor 16680,
61 Indonesia.

62 ²⁸Animal Biosystematics and Ecology Division, Department of Biology, Bogor Agricultural
63 University, Jalan Agatis, Dramaga Campus, Bogor 16680, Indonesia.

64 ²⁹CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and
65 Technology (BIST), Baldri i Reixac 4, Barcelona 08028, Spain.

66 ³⁰Universitat Pompeu Fabra (UPF), Plaça de la Mercè, 10, 08002 Barcelona, Spain.

67 ³¹Leverhulme Centre for Human Evolutionary Studies, Department of Archaeology and
68 Anthropology, University of Cambridge, Cambridge, UK.

69 ³²Institute of Applied Simulations, School of Life Sciences and Facility Management, Zurich
70 University of Applied Sciences ZHAW, Einsiedlerstrasse 31a, 8820 Wädenswil, Switzerland.

71 ³³Swiss Institute of Bioinformatics, Quartier Sorge - Batiment Genopode, 1015 Lausanne,
72 Switzerland.

73 ³⁴Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona 08010, Spain.

74 ³⁵Borneo Futures, Bandar Seri Begawan, Brunei Darussalam.

75 **Summary**

76 Six extant species of non-human great apes are currently recognized: Sumatran and Bornean
77 orangutans, eastern and western gorillas, and chimpanzees and bonobos [1]. However, large gaps
78 remain in our knowledge of the fine-scale variation in hominoid morphology, behavior, and genetics,
79 and aspects of great ape taxonomy remain in flux. This is particularly true for orangutans (genus:
80 *Pongo*), the only Asian great apes, and phylogenetically our most distant relatives among extant
81 hominids [1]. Definitive designation of Bornean and Sumatran orangutans, *P. pygmaeus* (Linnaeus
82 1760) and *P. abelii* (Lesson 1827), as distinct species took place as recently as 2001 [1, 2]. Based on
83 morphological, behavioral and environmental data, and corroborated by population genetic analyses
84 of 37 orangutan genomes, we show that an isolated population of the Sumatran orangutan south of
85 Lake Toba, Batang Toru, is highly distinct from the northern Sumatran and Bornean populations. The
86 deepest split in the evolutionary history of extant orangutans occurred ~3.38 Ma between this remnant
87 population south of Lake Toba and those to the north, while both currently recognized species
88 separated much later about 674 ka. A morphometric analysis based on cranio-mandibular and dental
89 characters as well as behavioral data revealed consistent differences between individuals from Batang
90 Toru and other extant Ponginae. Our combined analyses support a new classification of orangutans
91 into three extant species. One of them, *P. tapanuliensis*, encompasses the Batang Toru population, of
92 which fewer than 800 individuals survive.

93 **Results and Discussion**

94 Despite decades of field studies [3] our knowledge of variation among orangutans remains limited as
 95 many populations occur in isolated and inaccessible habitats, leaving questions regarding their
 96 evolutionary history and taxonomic classification largely unresolved. In particular, Sumatran
 97 populations south of Lake Toba had long been overlooked, even though a 1939 review of the species'
 98 range mentioned that orangutans had been reported in several forest areas in that region [4]. Based on
 99 diverse sources of evidence, we describe a new orangutan species, *Pongo tapanuliensis*, which
 100 encompasses a geographically and genetically isolated population found in the Batang Toru area at
 101 the southernmost range of extant Sumatran orangutans, south of Lake Toba, Indonesia.

102 **Systematics**

103 Genus *Pongo* Lacépède, 1799

104 *Pongo tapanuliensis* sp. nov. Nurcahyo, Meijaard, Nowak, Fredriksson & Groves

105 Tapanuli Orangutan

106 **Etymology.** The species name refers to three North Sumatran districts (North, Central, and South
 107 Tapanuli) to which *P. tapanuliensis* is endemic.

108 **Holotype.** The complete skeleton of an adult male orangutan that died from wounds sustained by local
 109 villagers in November 2013 near Sugi Tonga, Marancar, Tapanuli (Batang Toru) Forest Complex
 110 (1°35'54.1"N, 99°16'36.5"E), South Tapanuli District, North Sumatra, Indonesia. Skull and
 111 postcranium are lodged in the Museum Zoologicum Bogoriense, Indonesia, accession number
 112 MZB39182. High-resolution 3D reconstructions of the skull and mandible are available as
 113 supplementary material.

114 **Paratypes.** Adult individuals of *P. tapanuliensis* (P2591-M435788 – P2591-M435790) photographed
 115 by Tim Laman in the Batang Toru Forest Complex (1°41'9.1"N, 98°59'38.1"E), North Tapanuli
 116 District, North Sumatra, Indonesia. Paratypes are available from <http://www.morphobank.org> (Login:
 117 2591 / Password: tapanuliorangutan).

118 **Differential diagnosis.** Unless otherwise stated, all units are [mm]. Summary statistics for all
 119 measurements are listed in Tables S1–3. *Pongo tapanuliensis* differs from all extant orangutans in the
 120 breadth of the upper canine (21.5, vs. <20.86); the shallow face depth (6.0 vs. >8.4); the narrower
 121 interpterygoid distance (at posterior end of pterygoids 33.8 vs. >43.9; at anterior end of pterygoids,
 122 33.7 vs. >43.0); the shorter tympanic tube (23.9 vs. >28.4, mostly >30); the shorter
 123 temporomandibular joint (22.5 vs. >24.7); the narrower maxillary incisor row (28.3 vs. >30.1); the
 124 narrower distance across the palate at the first molars (62.7 vs. >65.7); the shorter horizontal length of

125 the mandibular symphysis (49.3 vs. >53.7); the smaller inferior transverse torus (horizontal length
126 from anterior surface of symphysis 31.8 compared to >36.0); and the width of the ascending ramus of
127 the mandible (55.9 vs. >56.3).

128 *Pongo tapanuliensis* differs specifically from *P. abelii* by its deep suborbital fossa, triangular
129 pyriform aperture, and angled facial profile; the longer nuchal surface (70.5 vs. <64.7); the wider
130 rostrum, posterior to the canines (59.9 vs. <59); the narrower orbits (33.8 vs. <34.6); the shorter (29.2
131 vs. >30.0) and narrower foramen magnum (23.2 vs. >23.3); the narrower bicondylar breadth (120.0
132 vs. >127.2); the narrower mandibular incisor row (24.4 vs. >28.3); the greater mesio-distal length of
133 the upper canine (19.44 vs. <17.55). The male long call has a higher maximum frequency range of the
134 roar pulse type (> 800 Hz vs. <747) with a higher ‘shape’ (>952 Hz/s vs. <934).

135 *Pongo tapanuliensis* differs from *P. pygmaeus* by possessing a nearly straight zygomaxillary suture;
136 the lower orbit (orbit height 33.4 vs. >35.3); the male long call has a longer duration (>111 seconds
137 vs. <90) with a greater number of pulses (>52 pulses vs. <45), and is delivered at a greater rate (>0.82
138 pulses per 20 seconds vs. <0.79).

139 *Pongo tapanuliensis* differs specifically from *Pongo ‘pygmaeus’ palaeosumatrensis* in the smaller
140 size of the first upper molar (mesio-distal length 13.65 vs. >14.0, buccolingual breadth 11.37 vs.
141 >12.10, crown area 155.2 mm² vs. >175.45).

142 **Description.** Craniometrically, the type skull of *P. tapanuliensis* (Fig. 1B) is significantly smaller than
143 any skull of comparable developmental stage of other orangutans; it falls outside of the interquartile
144 ranges of *P. abelii* and *P. pygmaeus* for 24 of 39 cranio-mandibular measurements (Table S1). A PCA
145 of 26 cranio-mandibular measurements commonly used in primate taxonomic classification [5, 6]
146 shows consistent differences between *P. tapanuliensis* and the two currently recognized species (Fig.
147 1C).

148 The external morphology of *P. tapanuliensis* is more similar to *P. abelii* in its linear body build and
149 more cinnamon pelage than *P. pygmaeus*. The hair texture of *P. tapanuliensis* is frizzier, contrasting
150 in particular with the long, loose body hair of *P. abelii*. *Pongo tapanuliensis* has a prominent
151 moustache and flat flanges covered in downy hair in dominant males, while flanges of older males
152 resemble more those of Bornean males. Females of *P. tapanuliensis* have beards, unlike *P. pygmaeus*.

153 **Distribution.** *Pongo tapanuliensis* occurs only in a small number of forest fragments in the districts of
154 Central, North, and South Tapanuli, Indonesia (Fig. 1A). The total distribution covers approximately
155 1,000 km², with an estimated population size of fewer than 800 individuals [7]. The current
156 distribution of *P. tapanuliensis* is almost completely restricted to medium elevation hill and
157 submontane forest (~300–1300 m asl) [7–9]. While densities are highest in primary forest, it does

158 occur at lower densities in mixed agroforest at the edge of primary forest areas [10, 11]. Until
159 relatively recently, *P. tapanuliensis* was more widespread to the south and west of the current
160 distribution, although evidence for this is largely anecdotal [12, 13].

161 To corroborate our morphological analysis, we investigated the evolutionary history of the genus
162 *Pongo*, particularly the relationships between *P. tapanuliensis* and its extant congeners, using whole-
163 genome data of 37 orangutans (Table S4). Our dataset is based on wild-born individuals with known
164 provenance, covering the entire range of extant orangutans including areas never sampled before (Fig.
165 2A). This dataset therefore provides a valuable resource for future studies of great ape biology and
166 genomics.

167 A principal component analysis (PCA; Fig. 2B) of genomic diversity highlighted the divergence
168 between individuals from Borneo and Sumatra (PC1), but also separated *P. tapanuliensis* from *P.*
169 *abelii* (PC2). The same clustering pattern was also found in a model-based analysis of population
170 structure (Fig. 2C), and is consistent with an earlier genetic study analyzing a larger number of non-
171 invasively collected samples using microsatellite markers [14]. However, while powerful in detecting
172 extant population structure, population history and speciation cannot be inferred, as they are not suited
173 to distinguish between old divergences with gene flow and cases of recent divergence with isolation
174 [15, 16]. To address this problem and further investigate the timing of population splits and gene flow,
175 we therefore employed different complementary modeling and phylogenetic approaches.

176 We applied an Approximate Bayesian Computation (ABC) approach, which allows to infer and
177 compare arbitrarily complex demographic modes based on the comparison of the observed genomic
178 data to extensive population genetic simulations [17]. Our analyses revealed three deep evolutionary
179 lineages in extant orangutans (Figs. 3A and B). Colonization scenarios in which the earliest split
180 within *Pongo* occurred between the lineages leading to *P. abelii* and *P. tapanuliensis* were much
181 better supported than scenarios in which the earliest split was between Bornean and Sumatran species
182 (combined posterior probability: 99.91%, Fig. 3A). Of the two best scenarios, a model postulating
183 colonization of both northern Sumatra and Borneo from an ancestral population likely situated south
184 of Lake Toba on Sumatra, had the highest support (posterior probability 97.56%, Fig. 3A). Our results
185 supported a scenario in which orangutans from mainland Asia first entered Sundaland south of what is
186 now Lake Toba on Sumatra, the most likely entry point based on paleogeographic reconstructions
187 [18]. This ancestral population, of which *P. tapanuliensis* is a direct descendant, then served as a
188 source for the subsequent different colonization events of what is now Borneo, Java and northern
189 Sumatra.

190 We estimated the split time between populations north and south of Lake Toba at ~3.4 Ma (Fig. 3B,
191 Table S5). Under our best-fitting model, we found evidence for post-split gene flow across Lake Toba

192 (~0.3–0.9 migrants per generation, Table S5), which is consistent with highly significant signatures of
193 gene flow between *P. abelii* and *P. tapanuliensis* using D-statistics (CK, BT, WA, *Homo sapiens*: $D =$
194 -0.2819 , $p\text{-value} < 0.00001$; WK, BT, LK, *Homo sapiens*: $D = -0.2967$, $p\text{-value} < 0.00001$)[19]. Such
195 gene flow resulted in comparatively high autosomal affinity of *P. tapanuliensis* to *P. abelii* in the
196 PCA (Fig. 2B), explaining the smaller amount of variance captured by PC2 (separating *P.*
197 *tapanuliensis* from all other populations) compared to PC1 (separating *P. pygmaeus* from the
198 Sumatran populations). The parameter estimates from a Bayesian full-likelihood analysis
199 implemented in the software G-PhoCS were in good agreement with those obtained by the ABC
200 analysis, although the split time between populations north and south of Lake Toba was more recent
201 (2.28 Ma, 95%-HPD: 2.21–2.35, Table S5). The G-PhoCS analysis revealed highly asymmetric gene
202 flow between populations north and south of the Toba caldera, with much lower levels from north to
203 south than vice versa (Table S5).

204 The existence of two deep evolutionary lineages among extant Sumatran orangutans was corroborated
205 by phylogenetic analyses based on whole mitochondrial genomes (Fig. 4A), in which the deepest split
206 occurred between populations north of Lake Toba and all other orangutans at ~3.97 Ma (95%-HPD:
207 2.35–5.57). Sumatran orangutans formed a paraphyletic group, with *P. tapanuliensis* being more
208 closely related to the Bornean lineage from which it diverged ~2.41 Ma (1.26–3.42 Ma). In contrast,
209 Bornean populations formed a monophyletic group with a very recent mitochondrial coalescence
210 at ~160 ka (94–227 ka).

211 Due to strong female philopatry [20], gene flow in orangutans is almost exclusively male-mediated
212 [21]. Consistent with these pronounced differences in dispersal behavior, phylogenetic analysis of
213 extensive Y-chromosomal sequencing data revealed a comparatively recent coalescence of Y
214 chromosomes of all extant orangutans ~430 ka (Fig. 4B). The single available Y-haplotype from *P.*
215 *tapanuliensis* was nested within the other Sumatran sequences, pointing at the occurrence of recent
216 male-mediated gene flow across the Toba divide. Thus, in combination with our modeling results, the
217 sex-specific data highlighted the impact of extraordinarily strong male-biased dispersal in the
218 speciation process of orangutans.

219 Our analyses revealed significant divergence between *P. tapanuliensis* and *P. abelii* (Figs. 3B and
220 4A), but also low levels of male-mediated gene flow until recently between both species (Figs. 3B and
221 4B). Populations north and south of Lake Toba on Sumatra had been in genetic contact for most of the
222 time since their split, but there was a marked reduction in gene flow after ~100 ka (Fig. 3C),
223 consistent with habitat destruction caused by the Toba supereruption 73 ka ago [22]. *Pongo*
224 *tapanuliensis* and *P. abelii* have been on independent evolutionary trajectories at least since the late
225 Pleistocene/early Holocene, as gene flow between these populations has ceased completely 10–20 ka
226 (Fig. 3C) and is now impossible because of habitat loss in areas between the species' ranges [7].

227 Nowadays, most biologists would probably adopt an operational species definition such as: ‘a species
228 is a population (or group of populations) with fixed heritable differences from other such populations
229 (or groups of populations)’ [23]. With totally allopatric populations, a ‘reproductive isolation’
230 criterion, such as is still espoused by adherents of the biological species concept, is not possible [24,
231 25]. Notwithstanding a long-running debate about the role of gene flow during speciation and genetic
232 interpretations of the species concept [26, 27], genomic studies have found evidence for many
233 instances of recent or ongoing gene flow between taxa which are recognized as distinct and well-
234 established species. This includes examples within each of the other three hominid genera. A recent
235 genomic study using comparable methods to ours revealed extensive gene flow between *Gorilla*
236 *gorilla* and *G. beringei* until ~20–30 ka [28]. Similar, albeit older and less extensive, admixture
237 occurred between *Pan troglodytes* and *P. paniscus* [29], and between *Homo sapiens* and *H.*
238 *neanderthalensis* [30]. *Pongo tapanuliensis* and *P. abelii* appear to be further examples, showing
239 diagnostic phenotypic and other distinctions that have persisted despite gene flow between them.

240 With a census size of fewer than 800 individuals [7], *P. tapanuliensis* is the least numerous of all
241 great ape species [31]. Its range is located around 200 km from the closest population of *P. abelii* to
242 the north (Fig. 2A). A combination of small population size and geographic isolation is of particular
243 high conservation concern, as it may lead to inbreeding depression [32] and threaten population
244 persistence [33]. Highlighting this, we discovered extensive runs of homozygosity in the genomes of
245 both *P. tapanuliensis* individuals (Fig. S3), pointing at the occurrence of recent inbreeding.

246 To ensure long-term survival of *P. tapanuliensis*, conservation measures need to be implemented
247 swiftly. Due to the rugged terrain, external threats have been primarily limited to road construction,
248 illegal clearing of forests, hunting, killings during crop conflict and trade in orangutans [7, 11]. A
249 hydro-electric development has been proposed recently in the area of highest orangutan density,
250 which could impact up to 8% of *P. tapanuliensis*’ habitat. This project might lead to further genetic
251 impoverishment and inbreeding, as it would jeopardize chances of maintaining habitat corridors
252 between the western and eastern range (Fig. 1A), and smaller nature reserves, all of which maintain
253 small populations of *P. tapanuliensis*.

254 **References**

- 255 1. Groves, C.P. (2001). Primate taxonomy, (Washington, D.C. ; London: Smithsonian Institution
256 Press).
- 257 2. Xu, X., and Arnason, U. (1996). The mitochondrial DNA molecule of Sumatran orangutan
258 and a molecular proposal for two (Bornean and Sumatran) species of orangutan. *J. Mol. Evol.*
259 *43*, 431-437.
- 260 3. Wich, S.A., Utami Atmoko, S.S., Mitra Setia, T., and van Schaik, C.P. (2009). Orangutans:
261 geographic variation in behavioral ecology and conservation, (Oxford University Press).
- 262 4. Nederlandsch-Indische Vereeniging tot Natuurbescherming (1939). *Natuur in Zuid- en Oost-*
263 *Borneo. Fauna, flora en natuurbescherming in de Zuider- en Ooster-Afdeeling van Borneo.* In
264 *3 Jaren Indisch natuur leven. Opstellen over landschappen, dieren en planten, tevens elfde*
265 *verslag (1936-1938), Nederlandsch-Indische Vereeniging tot Natuurbescherming, ed.*
266 *(Batavia, Indonesia), pp. 334-411.*
- 267 5. Groves, C.P., Westwood, C., and Shea, B.T. (1992). Unfinished business - Mahalanobis and a
268 clockwork orang. *J. Hum. Evol.* *22*, 327-340.
- 269 6. Groves, C.P. (1986). Systematics of the great apes. In *Comparative primate biology, Vol.1:*
270 *Systematics, evolution, and anatomy*, D.R. Swindler and J. Erwin, eds. (New York: Alan R.
271 Liss), pp. 187-217.
- 272 7. Wich, S.A., Singleton, I., Nowak, M.G., Utami Atmoko, S.S., Nisam, G., Arif, S.M., Putra,
273 R.H., Ardi, R., Fredriksson, G., Usher, G., et al. (2016). Land-cover changes predict steep
274 declines for the Sumatran orangutan (*Pongo abelii*). *Sci. Adv.* *2*, e1500789.
- 275 8. Laumonier, Y., Uryu, Y., Stüwe, M., Budiman, A., Setiabudi, B., and Hadian, O. (2010). Eco-
276 floristic sectors and deforestation threats in Sumatra: identifying new conservation area
277 network priorities for ecosystem-based land use planning. *Biodivers. Conserv.* *19*, 1153-1174.
- 278 9. Wich, S.A., Usher, G., Peters, H.H., Khakim, M.F.R., Nowak, M.G., and Fredriksson, G.M.
279 (2014). Preliminary data on the highland Sumatran orangutans (*Pongo abelii*) of Batang Toru.
280 In *High Altitude Primates*, B.N. Grow, S. Gursky-Doyen and A. Krzton, eds. (New York,
281 NY: Springer New York), pp. 265-283.
- 282 10. Meijaard, E. (1997). A survey of some forested areas in South and Central Tapanuli, North
283 Sumatra; new chances for orangutan conservation. (Wageningen: Tropenbos and the Golden
284 Ark).
- 285 11. Wich, S.A., Fredriksson, G.M., Usher, G., Peters, H.H., Priatna, D., Basalamah, F., Susanto,
286 W., and Kuhl, H. (2012). Hunting of Sumatran orang-utans and its importance in determining
287 distribution and density. *Biol. Conserv.* *146*, 163-169.
- 288 12. Kramm, W. (1879). Tochtjes in Tapanoeli. *Sumatra-Courant* *20*, 1-2.
- 289 13. Miller, G.S. (1903). Mammals collected by Dr. W.L. Abbott on the coast and islands of
290 northwest Sumatra. *Proceedings US National Museum, Washington* *26*, 437-484.
- 291 14. Nater, A., Arora, N., Greminger, M.P., van Schaik, C.P., Singleton, I., Wich, S.A.,
292 Fredriksson, G., Perwitasari-Farajallah, D., Pamungkas, J., and Krützen, M. (2013). Marked
293 population structure and recent migration in the critically endangered Sumatran orangutan
294 (*Pongo abelii*). *J. Hered.* *104*, 2-13.
- 295 15. Nielsen, R., and Wakeley, J. (2001). Distinguishing migration from isolation: a Markov chain
296 Monte Carlo approach. *Genetics* *158*, 885-896.
- 297 16. Palsboll, P.J., Berube, M., Aguilar, A., Notarbartolo-Di-Sciara, G., and Nielsen, R. (2004).
298 Discerning between recurrent gene flow and recent divergence under a finite-site mutation
299 model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*)
300 populations. *Evolution* *58*, 670-675.
- 301 17. Beaumont, M.A., Zhang, W.Y., and Balding, D.J. (2002). Approximate Bayesian
302 computation in population genetics. *Genetics* *162*, 2025-2035.
- 303 18. Meijaard, E. (2004). Solving mammalian riddles: a reconstruction of the Tertiary and
304 Quaternary distribution of mammals and their palaeoenvironments in island South-East Asia.
305 (Australian National University), p. 2 v.

- 306 19. Durand, E.Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture
307 between closely related populations. *Mol. Biol. Evol.* 28, 2239-2252.
- 308 20. Arora, N., van Noordwijk, M.A., Ackermann, C., Willems, E.P., Nater, A., Greminger, M.,
309 Nietlisbach, P., Dunkel, L.P., Utami Atmoko, S.S., Pamungkas, J., et al. (2012). Parentage-
310 based pedigree reconstruction reveals female matrilineal clusters and male-biased dispersal in
311 nongregarious Asian great apes, the Bornean orang-utans (*Pongo pygmaeus*). *Mol. Ecol.* 21,
312 3352-3362.
- 313 21. Nater, A., Nietlisbach, P., Arora, N., van Schaik, C.P., van Noordwijk, M.A., Willems, E.P.,
314 Singleton, I., Wich, S.A., Goossens, B., Warren, K.S., et al. (2011). Sex-biased dispersal and
315 volcanic activities shaped phylogeographic patterns of extant orangutans (genus: *Pongo*).
316 *Mol. Biol. Evol.* 28, 2275-2288.
- 317 22. Chesner, C.A., Rose, W.I., Deino, A., Drake, R., and Westgate, J.A. (1991). Eruptive history
318 of earths largest Quaternary caldera (Toba, Indonesia) clarified. *Geology* 19, 200-203.
- 319 23. Groves, C.P., and Grubb, P. (2011). *Ungulate taxonomy*, (Baltimore, Md.: Johns Hopkins
320 University Press).
- 321 24. Coyne, J.A., and Orr, H.A. (2004). *Speciation*, (Sunderland, MA: Sinauer Associates, Inc.).
- 322 25. Mayr, E. (1963). *Animal species and evolution*, (Cambridge,: Belknap Press of Harvard
323 University Press).
- 324 26. Arnold, M.L. (2016). *Divergence with Genetic Exchange*, (Oxford, UK: Oxford University
325 Press).
- 326 27. Reznick, D.N., and Ricklefs, R.E. (2009). Darwin's bridge between microevolution and
327 macroevolution. *Nature* 457, 837-842.
- 328 28. Scally, A., Dutheil, J.Y., Hillier, L.W., Jordan, G.E., Goodhead, I., Herrero, J., Hobolth, A.,
329 Lappalainen, T., Mailund, T., Marques-Bonet, T., et al. (2012). Insights into hominid
330 evolution from the gorilla genome sequence. *Nature* 483, 169-175.
- 331 29. de Manuel, M., Kuhlwilm, M., Frandsen, P., Sousa, V.C., Desai, T., Prado-Martinez, J.,
332 Hernandez-Rodriguez, J., Dupanloup, I., Lao, O., Hallast, P., et al. (2016). Chimpanzee
333 genomic diversity reveals ancient admixture with bonobos. *Science* 354, 477.
- 334 30. Kuhlwilm, M., Gronau, I., Hubisz, M.J., de Filippo, C., Prado-Martinez, J., Kircher, M., Fu,
335 Q., Burbano, H.A., Lalueza-Fox, C., de la Rasilla, M., et al. (2016). Ancient gene flow from
336 early modern humans into Eastern Neanderthals. *Nature* 530, 429-433.
- 337 31. IUCN (2016). *IUCN Red List of Threatened Species*. Version 2016.2.
- 338 32. Hedrick, P.W., and Kalinowski, S.T. (2000). Inbreeding depression in conservation biology.
339 *Annu. Rev. Ecol. Syst.* 31, 139-162.
- 340 33. Allendorf, F.W., Luikart, G., and Aitken, S.N. (2013). *Conservation and the genetics of*
341 *populations*, 2nd Edition, (Hoboken: John Wiley & Sons).

342

343

344 **Acknowledgments**

345 We thank the following institutions and organizations for supporting our research: Indonesian State
346 Ministry for Research and Technology, Sabah Wildlife Department, Ministry of Environment and
347 Forestry of the Republic of Indonesia, Indonesian Institute of Sciences, Leuser International
348 Foundation, Gunung Leuser National Park, Borneo Orangutan Survival Foundation, Agisoft,
349 NVIDIA, and the 10 museums where we measured the specimens. This work was financially
350 supported by University of Zurich (UZH) Forschungskredit grants FK-10-50 (MPMG), FK-15-103
351 (AN), and FK-14-094 (TBS), Swiss National Science Foundation grant 3100A-116848 (MK, CvS),
352 Leakey Foundation (MPMG), A.H. Schultz Foundation grants (MK, MPMG), UZH Research Priority
353 Program ‘Evolution in Action’ (MK), the Arcus Foundation (EM), Australian National University
354 (ANU) research fund (AnN), ANU Vice Chancellor Travel Grant (AnN), Australia Awards
355 Scholarship-DFAT (AnN), ERC Starting Grant 260372 (TMB), EMBO YIP 2013 (TMB), MINECO
356 BFU2014-55090-P, BFU2015-7116-ERC, BFU2015-6215-ERCU01, and MH106874 (TMB),
357 Fundacio Zoo Barcelona (TMB), Julius–Klaus Foundation (MK), MINECO/FEDER BFU2016-
358 77961-P (JB, MP), Gates Cambridge Trust (TD), and the Department of Anthropology at the
359 University of Zurich. Novel raw sequencing data have been deposited into the European Nucleotide
360 Archive (ENA; <http://www.ebi.ac.uk/ena>) under study accession number PRJEB19688.

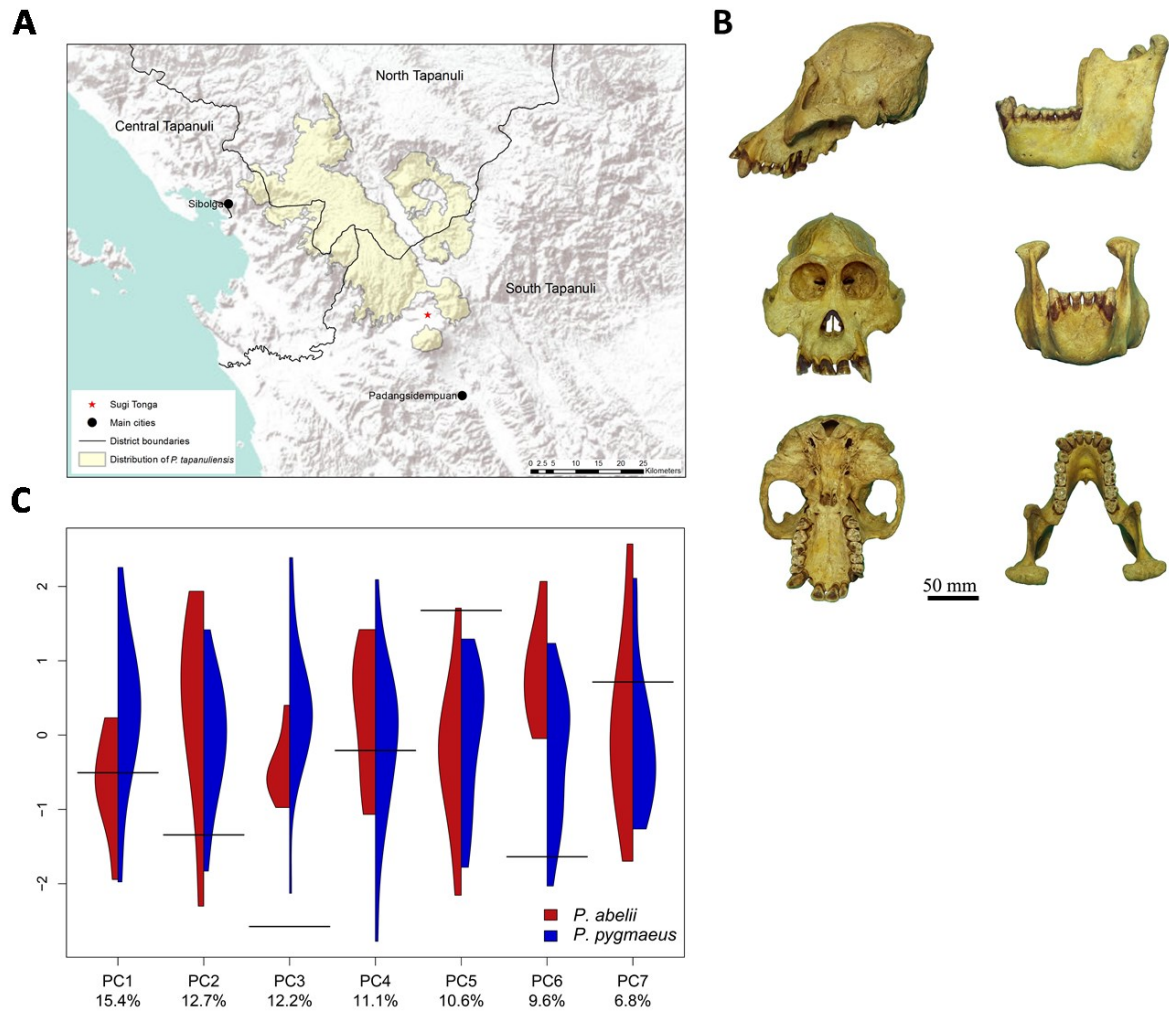
361 **Supplemental Information**

362 Figures S1–S4

363 Tables S1–S4

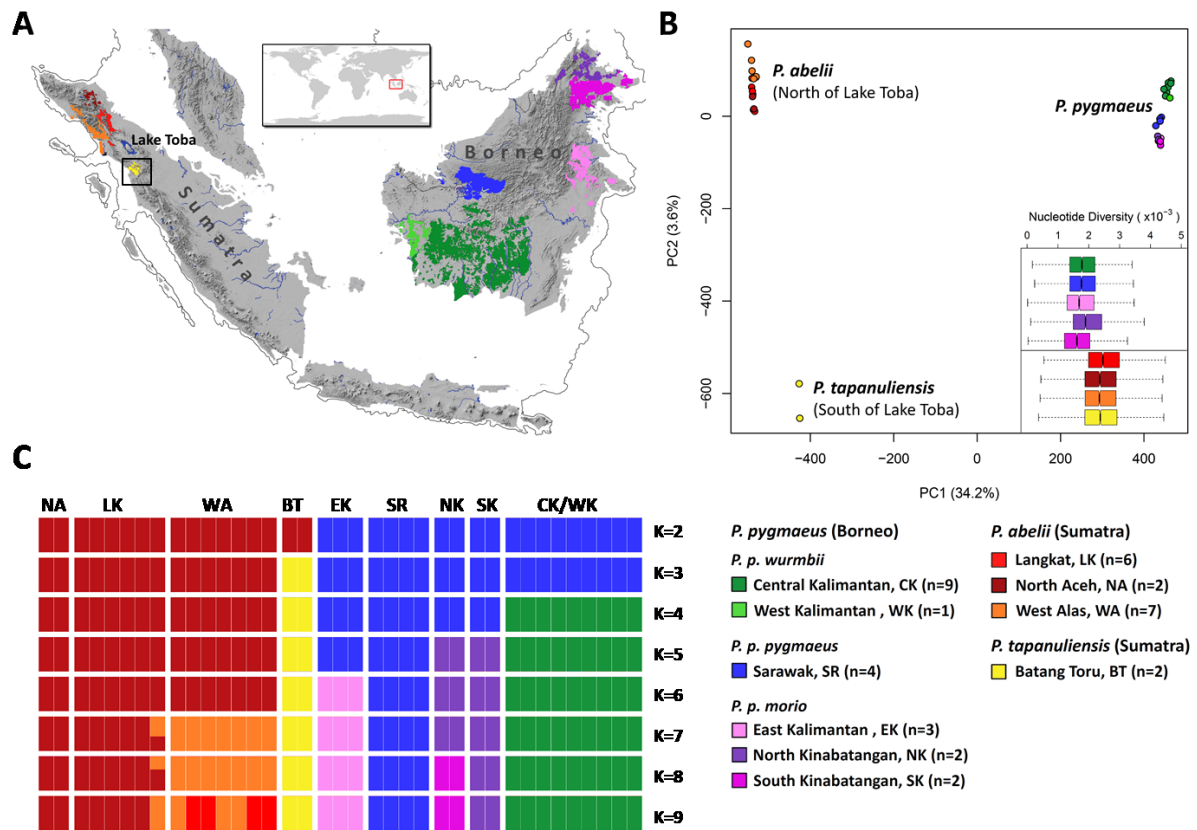
364 **Author Contributions**

365 Conceived the study and wrote the paper: MPMG, AIN, EM, MK, MGN, CG. Edited the manuscript:
366 SW, GF, CvS, AS, TMB, DAM, TBS, TD, BG, FC, KSW, EV, POtW, PR, JB, MA, AnN. Carried out
367 statistical analyses: MPMG, AIN, MGN, AnN, CG, MdM, TD, JA, MDR, AL, MP, JPM, MK, EM,
368 AS, TMB. Provided samples, and behavioral and ecological data: MGN, MPMG, AnN, AIN, GF, JA,
369 AL, MDR, BG, EJV, KSW, IS, JP, DPF, PR, WB. Performed sequencing: MPMG, IGG, MG, CR.



370

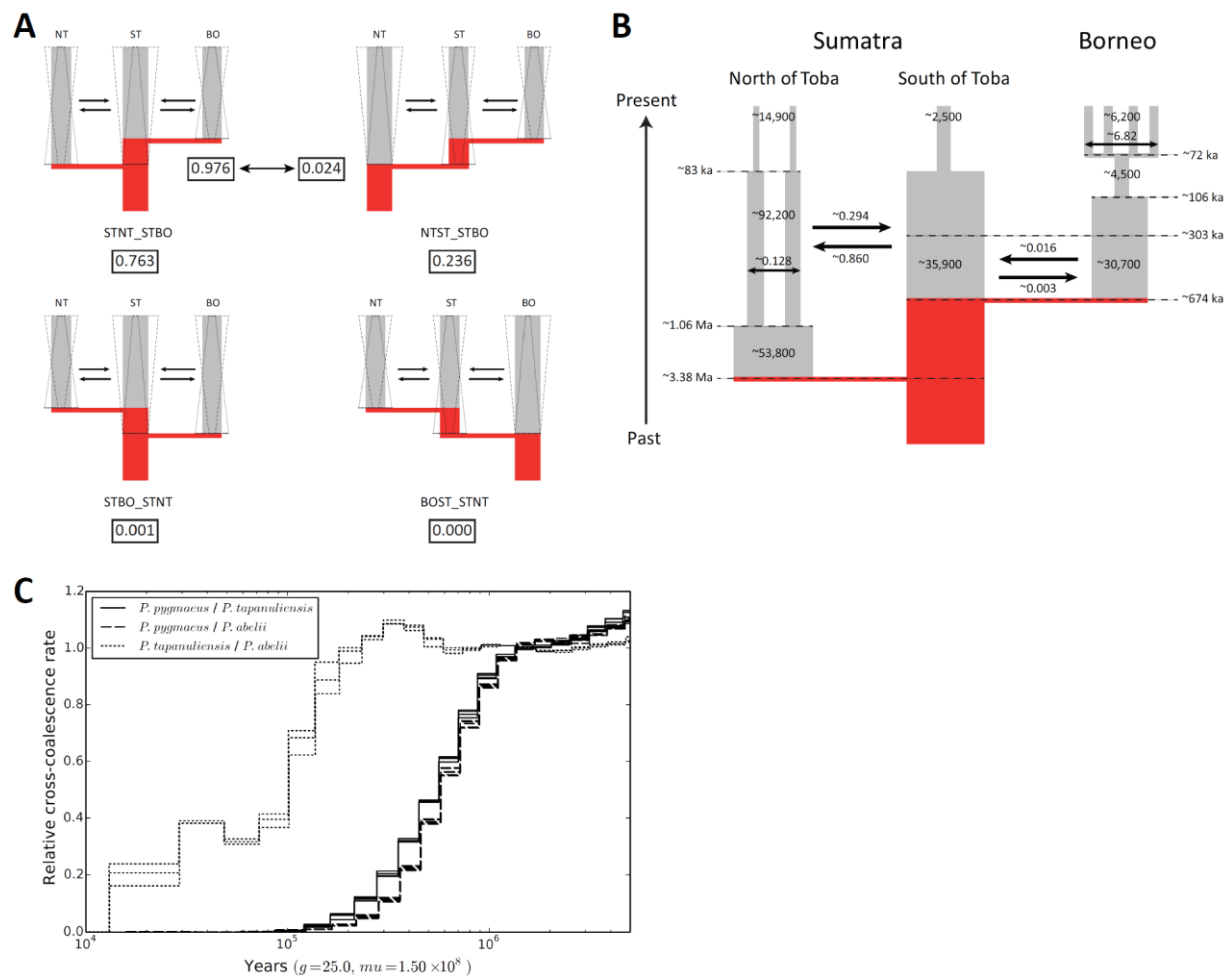
371 **Figure 1. Morphological evidence supporting a new orangutan species.** A) Current distribution of
 372 *Pongo tapanuliensis* on Sumatra. The holotype locality is marked with a red star. The area shown in
 373 the map is indicated in Fig. 2A. B) Holotype skull and mandible of *P. tapanuliensis* from a recently
 374 deceased individual from Batang Toru. C) Violin plots of the first seven principal components of 26
 375 cranio-mandibular morphological variables of 8 north Sumatran *P. abelii* and 19 Bornean *P.*
 376 *pygmaeus* individuals of similar developmental state as the holotype skull (black horizontal lines).



377

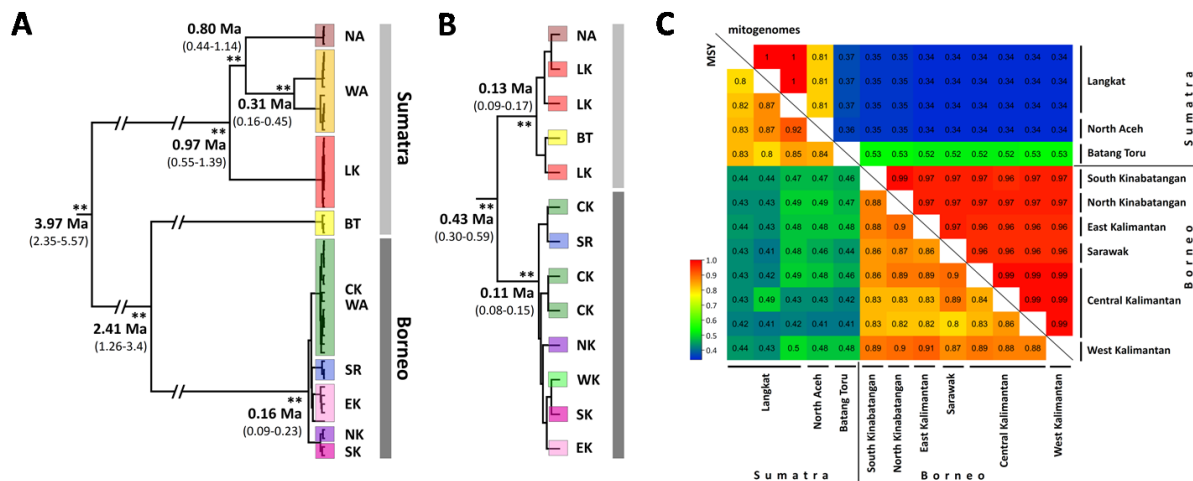
378 **Figure 2. Distribution, genomic diversity, and population structure of the genus *Pongo*.** A)
 379 Sampling areas across the current distribution of orangutans. The contour indicates the extent of the
 380 exposed Sunda Shelf during the last glacial maximum. The black rectangle delimits the area shown in
 381 Fig. 1A. n = numbers of sequenced individuals. B) Principal component analysis of genomic diversity
 382 in *Pongo*. Axis labels show the percentages of the total variance explained by the first two principal
 383 components. Colored bars in the insert represent the distribution of nucleotide diversity in genome-
 384 wide 1-Mb windows across sampling areas. (C) Bayesian clustering analysis of population structure
 385 using the program ADMIXTURE. Each vertical bar depicts an individual, with colors representing the
 386 inferred ancestry proportions with different assumed numbers of genetic clusters (K, horizontal
 387 sections).

A NEW SPECIES OF ORANGUTAN



388

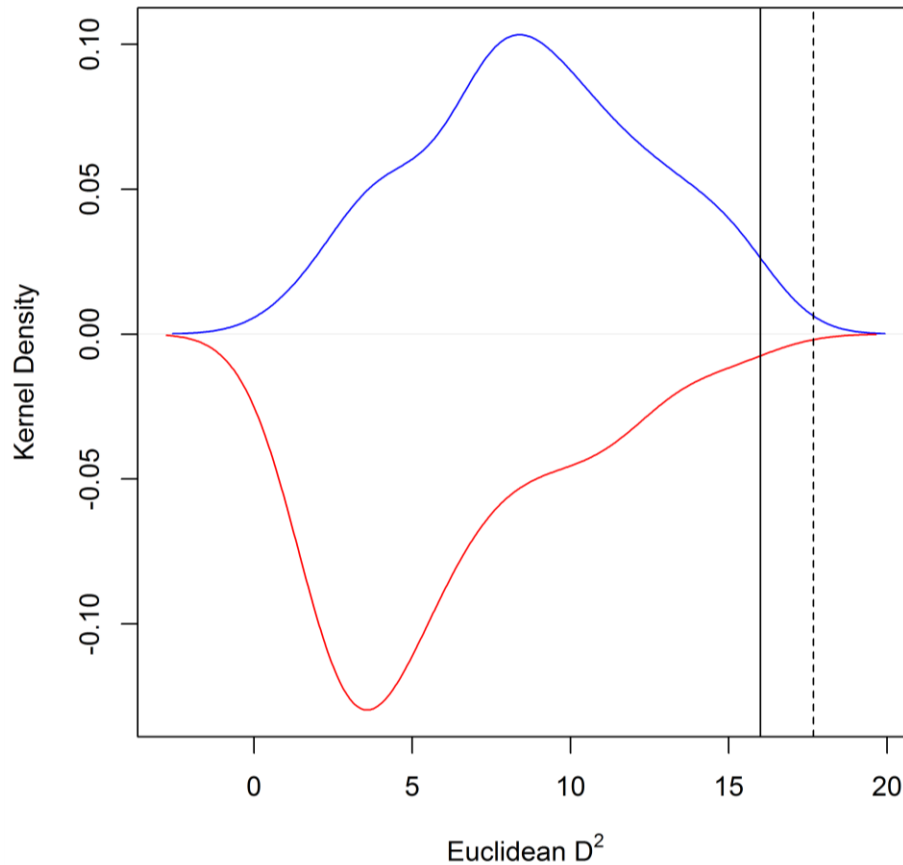
389 **Figure 3. Demographic history and gene flow in *Pongo*.** A) Model selection by Approximate
 390 Bayesian Computation (ABC) of plausible colonization histories of orangutans on Sundaland. The
 391 ABC analyses are based on the comparison of ~3,000 non-coding 2-kb loci randomly distributed
 392 across the genome with corresponding data simulated under the different demographic models. The
 393 numbers in the black boxes indicate the model's posterior probability. NT = Sumatran populations
 394 north of Lake Toba, ST = the Sumatran population of Batang Toru south of Lake Toba, BO =
 395 Bornean populations. B) ABC parameter estimates based on the full demographic model with
 396 colonization pattern inferred in panel A. Numbers in grey rectangles represent point estimates of
 397 effective population size (N_e). Arrows indicate gene flow among populations, numbers above the
 398 arrows represent point estimates of numbers of migrants per generation. C) Relative cross-coalescent
 399 rate (RCCR) analysis for between-species pairs of phased high-coverage genomes. A RCCR close to
 400 1 indicates extensive gene flow between species, while a ratio close to 0 indicates genetic isolation
 401 between species pairs. The x-axis shows time scaled in years, assuming a generation time of 25 years
 402 and an autosomal mutation rate of 1.5×10^{-8} per site per generation.



403

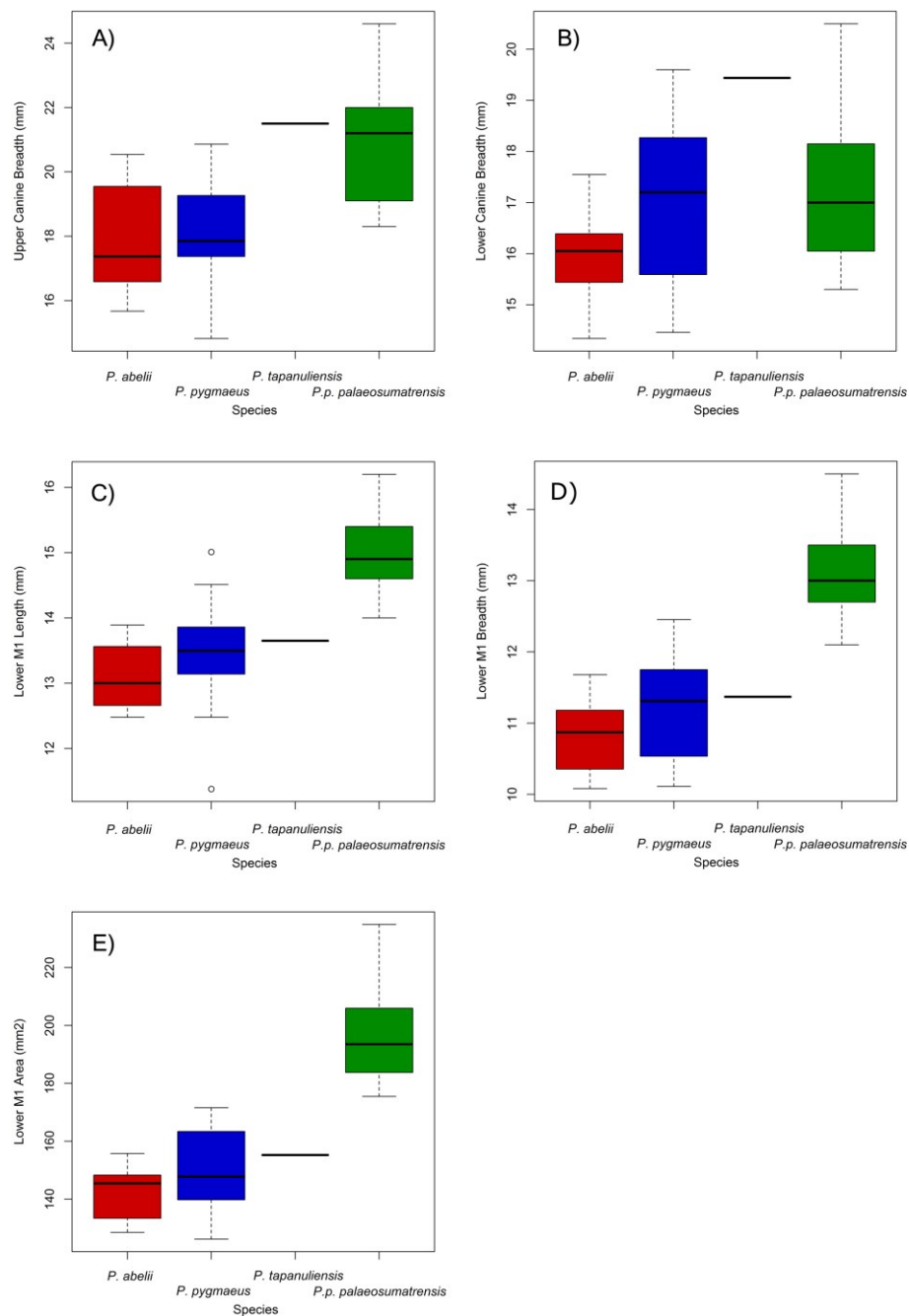
404 **Figure 4. Sex-specific evolutionary history of orangutans.** Bayesian phylogenetic trees for (A)
 405 mitochondrial genomes and (B) Y chromosomes. The mitochondrial tree is rooted with a human and a
 406 central chimpanzee sequence, the Y chromosome tree with a human sequence (not shown). **
 407 Posterior probability = 1.00. (C) Genotype-sharing matrix for mitogenomes (above the diagonal) and
 408 Y chromosomes (below the diagonal) for all analyzed male orangutans. A value of 1 indicates that
 409 two males have identical genotypes at all polymorphic sites; a value of 0 means that they have
 410 different genotypes at all variable positions.

411



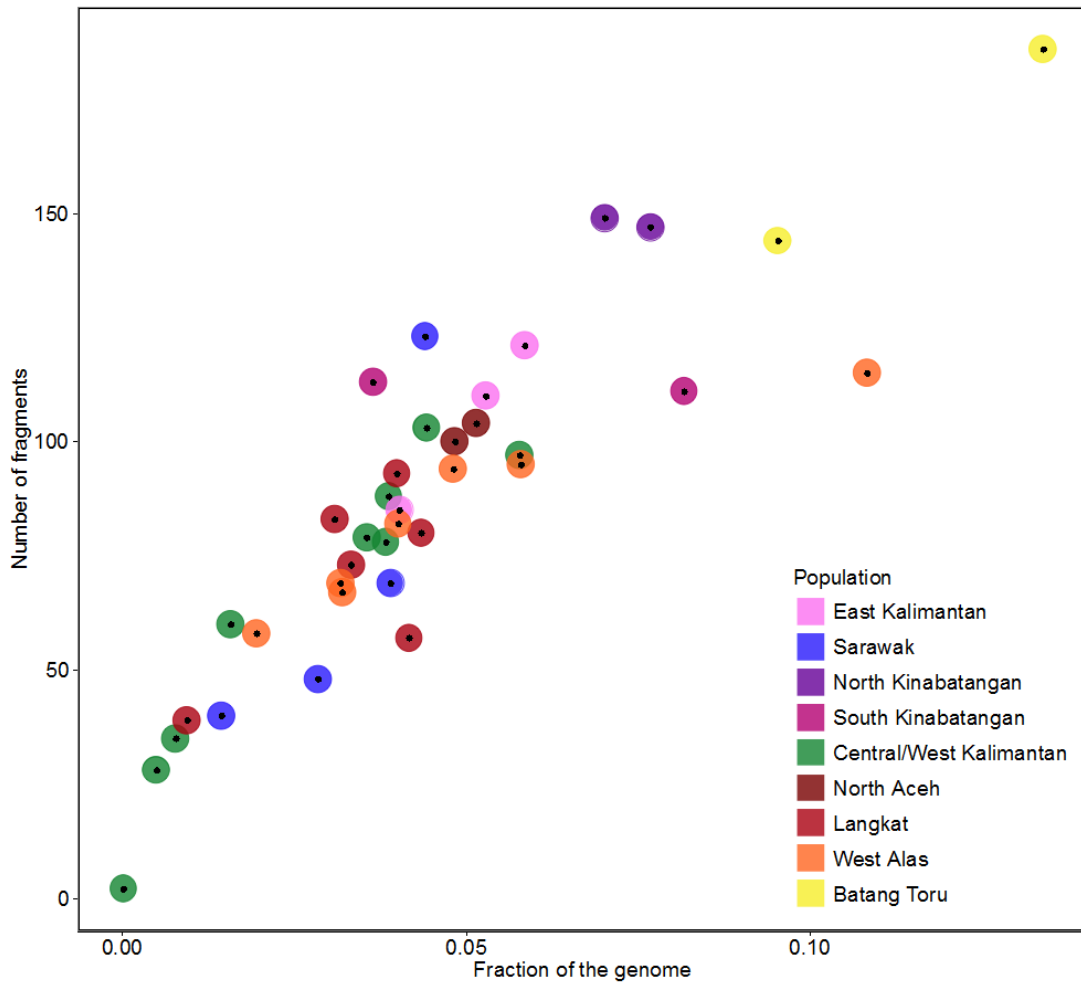
412

413 **Supplemental Fig. S1, related to Fig. 1C.** Kernel density mirror plots showing the results of
 414 Euclidean D^2 analyses of six principal components calculated from 26 cranio-mandibular
 415 morphological variables (Table S1). The between-species distribution (blue line) was calculated as the
 416 distances of all *Pongo abelii* samples to the *P. pygmaeus* centroid plus all of the *P. pygmaeus* samples
 417 to the *P. abelii* centroid, whereas the within-species distribution (red line) was calculated as the
 418 distances of all *P. abelii* samples to the *P. abelii* centroid plus all of the *P. pygmaeus* samples to the *P.*
 419 *pygmaeus* centroid. The dotted line represents the distance of the *P. tapanuliensis* sample to the *P.*
 420 *abelii* centroid (exact permutation test; within-species: p-value<0.001; between-species:
 421 p-value<0.001), whereas solid line represents the distance of the *P. tapanuliensis* samples to the *P.*
 422 *pygmaeus* centroid (within-species: p-value<0.001; between-species: p-value<0.001).



423

424 **Supplemental Fig. S2, related to Fig. 1C.** Comparisons of five dental variables across *P. abelii*
 425 (red), *P. pygmaeus* (blue), *P. tapanuliensis* (black horizontal line), and *P. p. palaeosumatrensis*
 426 (green). Variables include upper canine breadth (A), lower canine breadth (B), lower M1 length (C),
 427 lower M1 breadth (D), and lower M1 area (E). For each boxplot, the middle line is the median value
 428 of the distribution, with the box representing the first (lower extreme) and third (upper extreme)
 429 quartile values (*i.e.*, the interquartile range [IQR]), and the whiskers representing the lower and upper
 430 extreme values that are within 1.5 x IQR of the first and third quartile values.



431

432 **Supplemental Fig. S3, related to Fig. 3C.** Number of genomic fragments that are autozygous (y-
 433 axis) plotted against the total fraction of the genome covered by such fragments (x-axis). Each dot
 434 represents an individual, with sample origins represented by colors corresponding to those in Fig.
 435 2A.