Original Article:

Reconstructing the Demographic History of Orang-utans using Approximate Bayesian Computation

Alexander Nater[1], Maja P. Greminger[1], Natasha Arora[1], Carel P. van Schaik[1], Benoit Goossens[2,3,4], Ian Singleton[5,6], Ernst J. Verschoor[7], Kristin S. Warren[8], and Michael Krützen[1]

(1)   Anthropological Institute & Museum, University of Zurich, Zurich, Switzerland

(2)   Organisms and Environment Division, School of Biosciences, Cardiff University, Cardiff, UK

(3)   Danau Girang Field Centre, c/o Sabah Wildlife Department, Kota Kinabalu, Sabah, Malaysia

(4)   Sabah Wildlife Department, Kota Kinabalu, Sabah, Malaysia

(5)   Foundation for a Sustainable Ecosystem (YEL), Medan, Indonesia

(6)   Sumatran Orangutan Conservation Programme (PanEco-YEL), Medan, Indonesia

(7)   Department of Virology, Biomedical Primate Research Centre, Rijswijk, The Netherlands

(8)   College of Veterinary Medicine, School of Veterinary and Life Sciences, Murdoch University, Perth, Australia

Corresponding author:

Alexander Nater
Department of Evolutionary Biology
Evolutionary Biology Centre
Uppsala University
Norbyvägen 18D
752 36 Uppsala, Sweden
Phone number: +46 (0)18 471 2837
E-mail address: alexander.nater@ebc.uu.se

1   **Abstract**

2   Investigating how different evolutionary forces have shaped patterns of DNA variation within

3   and among species requires detailed knowledge of their demographic history. Orang-utans,

4   whose distribution is currently restricted to the Southeast Asian islands of Borneo (*Pongo*

5   *pygmaeus*) and Sumatra (*Pongo abelii*), have likely experienced a complex demographic

6   history, influenced by recurrent changes in climate and sea levels, volcanic activities and

7   anthropogenic pressures. Using the most extensive sample set of wild orang-utans to date, we

8   employed an approximate Bayesian computation (ABC) approach to test the fit of 12 different

9   demographic scenarios to the observed patterns of variation in autosomal, X-chromosomal,

10   mitochondrial and Y-chromosomal markers. In the best-fitting model, Sumatran orang-utans

11   exhibit a deep split of populations north and south of Lake Toba, probably caused by multiple

12   eruptions of the Toba volcano. In addition, we found signals for a strong decline in all

13   Sumatran populations ~24 ka, probably associated with hunting by human colonizers. In

14   contrast, Bornean orang-utans experienced a severe bottleneck ~135 ka, followed by a

15   population expansion and substructuring starting ~82 ka, which we link to an expansion from

16   a glacial refugium. Therefore, we showed that orang-utans went through drastic changes in

17   population size and connectedness, caused by the recurrent contraction and expansion of

18   rainforest habitat during Pleistocene glaciations, and probably also by the impact of hunting

19   by early humans. Our findings also emphasize the fact that important aspects of the

20   evolutionary past of species with complex demographic histories might remain obscured

21   when applying overly simplified models.

Introduction

Patterns of DNA variation are the result of both adaptive and non-adaptive processes, and the debate about the relative importance of natural selection and random genetic drift in shaping genetic diversity within and among species is still ongoing (e.g. Hahn 2008; Nei *et al.* 2010). A common approach to detect signals of selection aims at identifying genomic regions that show marked deviations in DNA variation from a neutral equilibrium model (reviewed in Nielsen 2005). However, in certain demographic scenarios, such as population size changes or population subdivision, random genetic drift can result in similar deviations as selection (e.g. Excoffier *et al.* 2009; Teshima *et al.* 2006). Therefore, confounding effects of demographic processes can only be unravelled from selective signals if the demographic history is explicitly taken into account when formulating the expectations under the neutral model against which observed patterns of DNA variation are tested (e.g. Haddrill *et al.* 2005; Stajich & Hahn 2005). Consequently, methods to reconstruct the demographic history of natural populations have recently generated great interest among evolutionary geneticists, as recent technical advances allow conducting genome-wide studies of selection in a large variety of species (reviewed in Ellegren 2014).

Orang-utans, currently restricted to two distinct species on Borneo (*Pongo pygmaeus*) and northern Sumatra (*Pongo abelii*) (Wich *et al.* 2008), are the only Asian great apes and are phylogenetically most distant to humans (Groves 2001). Their ancestral position in the lineage leading to African great apes and modern humans has evoked great interest in this taxon in the overall effort to reconstruct the adaptive evolutionary history of great apes in general and humans in particular (Locke *et al.* 2011; Prado-Martinez *et al.* 2013). However, orang-utans might have experienced a complex demographic history, as their distribution has been subject to major changes during the Pleistocene. The ancestors of extant orang-utans have sequentially colonized the islands of the Sunda archipelago arriving from the Southeast Asian mainland (Delgado & Van Schaik 2000; Rijksen & Meijaard 1999). Since then, their population history was strongly influenced by geological and climatic events: rising and falling sea levels cyclically connected and isolated the islands of Sundaland, allowing for potential terrestrial migration between the islands at certain points in time (Voris 2000).

Major volcanic eruptions, mainly on Sumatra and Java, might have led to the extinction of local orang-utan populations and subsequent re-colonisations (Muir *et al.* 2000). Of special interest here is the Toba volcano on northern Sumatra, which has seen at least four major

54   eruptions during the last 1.2 million years (Chesner *et al.* 1991). This sequence of eruptions

55   culminated in the Toba supereruption ~73 ka, which is considered to be the most powerful

56   volcanic eruption within the last 25 million years (Chesner *et al.* 1991) and is thought to have

57   had severe consequences for flora and fauna on Sundaland (Williams *et al.* 2009). In the Late

58   Pleistocene, all orang-utan populations on the mainland, southern Sumatra and Java went

59   extinct (Delgado & Van Schaik 2000; Rijksen & Meijaard 1999). Climatic changes during the

60   Pleistocene might have been responsible for the southward shift of the distribution and the

61   disappearance of orang-utans from the mainland (Jablonski 1998). Moreover, anthropogenic

62   factors, such as prehistoric hunting by hunter-gatherer societies, are likely to have played a

63   significant role in the decline and extinction of orang-utans populations on insular Southeast

64   Asia (Delgado & Van Schaik 2000).

65   Genetic signals of these past demographic changes have been found in studies of genetic

66   diversity in extant orang-utan populations on Borneo and Sumatra. Most genetic studies

67   analysing autosomal and mitochondrial DNA (mtDNA) agree that Sumatran orang-utans

68   show a higher level of sequence diversity and corresponding long-term effective population

69   size (N_e) (Locke *et al.* 2011; Muir *et al.* 2000; Prado-Martinez *et al.* 2013; Steiper 2006;

70   Zhang *et al.* 2001), even though Sumatran orang-utans have a much smaller current census

71   size and a more restricted distribution than Borneans (~6,600 vs. ~54,000 individuals, Wich *et*

72   *al.* 2008). This large N_e of the Sumatran species was interpreted as a signal of immigration

73   from multiple differentiated populations into the current Sumatran gene pool (Muir *et al.*

74   2000; Steiper 2006). However, Y-chromosomal diversity in orang-utans shows the opposite

75   pattern compared to mtDNA and autosomal data, with a smaller N_e on Sumatra than Borneo

76   (Nater *et al.* 2011). Such contrasting patterns of N_e between species and among genomic

77   regions hint at complex population dynamics that have so far not been properly investigated.

78   Recently, Locke *et al.* (2011) used extensive single-nucleotide polymorphism (SNP) data

79   from whole-genome resequencing of five Bornean and five Sumatran orang-utans to model

80   the demographic history of the two species. They found that a model with a population split

81   ~400 ka with subsequent gene flow between Borneo and Sumatra fits the observed data best.

82   Furthermore, Locke and colleagues inferred that Sumatran orang-utans underwent a

83   continuous exponential population growth since the population split, while Bornean orang-

84   utans were subject to a continuous exponential decline. Given the large amount of genetic

85   data, the study by Locke and colleagues is currently regarded as the most accurate

86   reconstruction of demographic history in orang-utans to date. However, the demographic

87   modelling approach by Locke and colleagues did not take several idiosyncrasies of orang-utan

88   biology into account, thus severely limiting the conclusions that could be drawn from their

89   findings.

90   First, it has been shown that biased sampling and disregard of population structure will

91   produce misleading results regarding N_e and its temporal changes (Chikhi *et al.* 2010; Stadler

92   *et al.* 2009). The study by Locke and colleagues incorporated data from only five captive

93   individuals each from Borneo and Sumatra without further provenance information. This

94   limited genetic sampling is unlikely to represent the entire genetic diversity present on both

95   islands. Second, given this lack of detailed sample provenance, the analyses were restricted to

96   models that treated Bornean and Sumatran orang-utans as single panmictic populations each.

97   Previous work, however, unequivocally showed that both Bornean and Sumatran orang-utans

98   are deeply structured genetically (Arora *et al.* 2010; Nater *et al.* 2011; Warren *et al.* 2001).

99   Especially on Sumatra populations north and south of Lake Toba exhibit high genetic

100   differentiation (Nater *et al.* 2013; Nater *et al.* 2011). Third, Locke and colleagues did not test

101   complex demographic models including population bottlenecks or recent declines, as

102   suggested in previous genetic studies. For example, genetic signals of a bottleneck with

103   subsequent population expansion on Borneo might be linked to a glacial refugium or the

104   impact of the Toba supereruption ~73 ka (Arora *et al.* 2010; Steiper 2006), and patterns of a

105   recent population decline in Sabah, Borneo, are most likely attributable to recent

106   anthropogenic pressures (Goossens *et al.* 2006).

107   Reconstructing the demographic history of a species has long been hindered by the fact that

108   full-likelihood methods were restricted to relatively simple demographic models (e.g. Hey &

109   Nielsen 2004; Wilson *et al.* 2003), which might not capture all relevant processes in complex

110   demographic settings. This restriction is mainly caused by the fact that the computation of the

111   likelihood function of complex demographic models with many parameters is either

112   intractable or computationally too expensive, especially for large data sets (Marjoram *et al.*

113   2003). Approximate Bayesian computation (ABC) allows circumventing these problems by

114   approximating the likelihood functions with simulations of genetic data under a given

115   demographic model (Beaumont *et al.* 2002; Marjoram *et al.* 2003). In order to estimate the

116   model parameters, parameter values are drawn from predefined prior distributions and used to

117   simulate genetic data matching the observed data in type of markers and number of loci. Both

118   observed and simulated data are then reduced to a set of summary statistics and the Euclidian

119   distance between the observed and the simulated summary statistics is calculated. Based on

120   the subset of simulations with the smallest Euclidian distance between observed and

121   simulated data, the posterior distribution of the model parameters can be approximated and

122   the relative fit of different demographic models to the data can be assessed.

123   Here we present an ABC modelling approach of the demographic history of orang-utans

124   based on autosomal and sex-linked marker systems. We aim to improve the current

125   knowledge of demographic history by applying three major improvements over previous

126   studies. First, we capitalize on the knowledge base of behavioural ecology and population

127   genetics of orang-utans in order to test realistic demographic models. Second, due to our

128   extensive set of orang-utan samples with detailed and reliable provenance, we are able to

129   investigate models incorporating population substructure in both orang-utan species, which

130   allows us to disentangle changes in population size from confounding effects due to changes

131   in population structure. Third, by combining autosomal and sex-linked markers into a

132   combined demographic analysis, we make use of the specific information content of different

133   marker systems in this species with its heavily sex-biased dispersal. Due to strong female

134   philopatry in orang-utans (Arora *et al.* 2012; Galdikas 1995; van Noordwijk *et al.* 2012),

135   mitochondrial markers contain information about population split times without the

136   confounding influence of gene flow. In contrast, Y-chromosomal loci should have more

137   power than autosomal markers to reveal low levels of male-mediated gene flow.

Materials & Methods

Sample Collection and Genetic Markers

140   A representative sampling scheme covering the whole range of a species is crucial for

141   accurate reconstruction of demographic history (Stadler *et al.* 2009). We used an extensive set

142   of samples from wild-born orang-utans from ten sampling locations, covering the entire

143   distribution of the genus (Figure 1, see Supporting Material for detailed information about

144   sample origin). Samples were analysed for several genetic marker systems with different

145   modes of inheritance and effective population sizes (Table 1), thus ensuring representation of

146   both male and female population history, an important aspect in demographic reconstructions

147   in species with strongly sex-biased dispersal (Nater *et al.* 2011; Nietlisbach *et al.* 2012).

148   The autosomal microsatellite data contained genotypes of 25 microsatellite markers from a

149   total of 237 individuals (Arora *et al.* 2010; Greminger *et al.* 2014; Nater *et al.* 2013). We also

150   included sequences from three mtDNA genes with a total length of 1,355 bp from 118

151 individuals (Nater *et al.* 2011), and Y-chromosomal haplotypes based on 11 Y-linked

152 microsatellite loci from 129 individuals (Nater *et al.* 2011). We complemented the data set by

153 additionally sequencing 8,055 bp of the non-coding X-chromosomal region Xq13.3

154 (Kaessmann *et al.* 2001) in 36 individuals and four non-coding autosomal regions (Fischer *et*

155 *al.* 2006) of a total of 8,238 bp in 22 individuals. Basic summary statistics for all marker

156 systems are provided in Table 2. The primers and cycling conditions used for PCR

157 amplification and sequencing of the X-chromosomal and autosomal regions are described in

158 the Supporting Table S1.

159 **Approximate Bayesian Computation**

160 *Model Selection Procedure*

161 We reconstructed the demographic history of orang-utans using an ABC approach

162 implemented in the software package ABCtoolbox v1.1 (Wegmann *et al.* 2010). To achieve

163 this goal, we first performed a model selection procedure, in which we used a hierarchical

164 approach to test a total of 12 different demographic models (Figure 2) with increasing levels

165 of complexity (see Supporting Tables S3 and S4 for more details about model

166 parameterisation and prior distributions).

167 We started by testing four relatively simple models that assumed a single population for each

168 of the two orang-utan species (Figure 2A). The first model in this set (I2) assumed constant

169 population sizes and no migration between the two populations. The second model (IM2)

170 incorporated asymmetric migration after the population split, up to a point in the past where

171 migration between Borneo and Sumatra ceased. Gene flow in all models with migration was

172 strictly male-mediated, as recent genetic and behavioural findings showed extreme female

173 philopatric tendencies in orang-utans (Arora *et al.* 2012; Nater *et al.* 2011; van Noordwijk *et*

174 *al.* 2012). The third model (IM2-GR) additionally allowed the two populations to change size

175 exponentially after the population split and corresponded largely to the favoured model in the

176 genomic study by Locke *et al.* (2011). In the fourth and most complex 2-population model

177 (IM2-BN-GR), both populations retained a constant size after the population split, with the

178 possibility for a sudden population size rescale followed by exponential growth or decline.

179 In order to test more biologically relevant demographic scenarios, we designed a series of 10-

180 population models that incorporated the repeatedly reported extensive population substructure

181 in extant orang-utan populations (Arora *et al.* 2010; Goossens *et al.* 2005; Kanthaswamy *et al.*

182 2006; Nater *et al.* 2013; Nater *et al.* 2011; Warren *et al.* 2001). The use of ten extant

183   population units models is justified by previously published data (Arora *et al.* 2010;

184   Greminger *et al.* 2014; Nater *et al.* 2013; Nater *et al.* 2011), and the combination of patterns

185   of population differentiation in both mtDNA and autosomal microsatellite markers, pointing

186   to six populations on Borneo, one Sumatran population south of Lake Toba, and three

187   Sumatran populations north of Lake Toba (see validation for ten population units in

188   Supporting Figures S1 and S2). For all 10-population models, we assumed equal population

189   sizes and equal symmetric migration rates among all populations within Borneo and among

190   all populations north of Lake Toba, respectively, as well as a separate population size

191   parameter for the population south of Lake Toba. We included asymmetric migration rates

192   between Borneo and south of Lake Toba, and between north of Lake Toba and south of Lake

193   Toba.

194   To assess to what extent the additional population units improve model fit, we first tested the

195   best-fitting 2-population model against two basic 10-population models (IM10 and IM10-

196   BOSU, Figure 2B). The IM10 model incorporated the population splitting sequence derived

197   from mtDNA data, i.e. the populations north and south of Lake Toba show the oldest split,

198   while Bornean populations diverged after this split (Nater *et al.* 2011). As this is in

199   discordance with the current species designation (Groves 2001), which assigns a single

200   species each to Sumatra and Borneo, we also tested this model against a model following the

201   species split pattern (IM10-BOSU), i.e. with the oldest split between Sumatra and Borneo, to

202   see if incomplete lineage sorting could be responsible for the particular phylogenetic pattern

203   observed for mtDNA.

204   We further tested for the presence of population size changes in the demographic history of

205   orang-utans, as suggested by previous studies (Arora *et al.* 2010; Goossens *et al.* 2006; Locke

206   *et al.* 2011; Steiper 2006). First, we tested for signals of recent declines on Sumatra (IM10-

207   DECSU), Borneo (IM10-DECBO) or both islands (IM10-DECALL) (Figure 2C).

208   In a second test, we evaluated the support for a bottleneck on Borneo (IM10-BNBO-DECSU),

209   possibly linked to a refugium during the penultimate glaciation (Arora *et al.* 2010) (Figure

210   2D).

211   Last, we tested for evidence for a bottleneck on Sumatra linked to the Toba supereruption,

212   either allowing for a broad prior range of the magnitude of decline (IM10-BNBO-TOBA-

213   DECSU) or restricting to a severe bottleneck of less than 100 individuals in each of the four

214   Sumatran populations (IM10-BNBO-RECOL-DECSU), resembling a founder effect after

215   local extinction and re-colonization events on Sumatra (Figure 2E).

216   *ABC Data Simulation*

217   To simulate genetic data under different demographic models, we used the software

218   FASTSIMCOAL v1.1.2 (Excoffier & Foll 2011). Simulations for the different marker

219   systems were run with the same set of parameters, whereby the effective population sizes

220   were scaled 1 to 0.75, 0.25, and 0.25 for autosomal, X-chromosomal, mitochondrial and Y-

221   chromosomal markers, respectively. We then used ARLSUMSTAT v3.5.1.3 (Excoffier &

222   Lischer 2010) to calculate a total of 259 summary statistics for each simulated data set as well

223   as for the observed data set (Supporting Table S5). The summary statistics were chosen in

224   order to capture the information in the genetic data about population differentiation, within

225   population diversity, and population size changes. To avoid problems with unreliable phasing,

226   we only used summary statistics that do not require phased sequence data for X-chromosomal

227   and autosomal loci. Since the number of simulated populations differed between the 2-

228   populuation and 10-population models, summary statistic would not be directly comparable

229   between the two sets of models. Therefore, when running the 10-population models, we

230   applied a script that pooled the simulated data into a Bornean and a Sumatran group after each

231   simulation step. Summary statistics were then also calculated island-wise, in order to be able

232   to directly compare to the 2-population models.

233   We first performed an initial run of 2×10^6 simulations with the standard rejection sampler

234   (Tavare *et al.* 1997). These simulations were used for both model selection and validation. To

235   reduce the dimensionality of the summary statistics, we performed a principal component

236   analysis (PCA) with the "prcomp" function in R version 2.12.1 (R Development Core Team

237   2010). A 100,000 random simulations from each of the two compared models were pooled

238   and standardised, and these summary statistics were used to extract the loadings of the first

239   ten principal components. We then transformed both the simulated and the observed data and

240   used it to perform a multinomial logistic regression with the R package "abc" version 1.6. For

241   this, we used the 0.1% of the simulations with the smallest Euclidean distance between the

242   transformed summary statistics and the observed data.

243   In order to assess model fit, we also calculated the marginal densities and the probability of

244   the observed data under the general linear model (GLM) used for the post sampling regression

245   for each model with ABCtoolbox (Leuenberger & Wegmann 2010). For this, we again

246   transformed both the simulated summary statistics as well as the observed data with the

       Demographic History of Orang-utans (*Pongo* spp.)

247   loadings for the first ten principal components. This time, PCA loadings were obtained for

248   each model separately by using 100,000 random simulations. The GLM was built from the

249   2,000 simulations closest to the observed data, and we assessed the goodness of fit of all

250   tested models to the observed data by calculating the p-value of the observed data under the

251   GLM (Supporting Table S6). The p-value is representing the proportion of the retained

252   simulations showing a lower or equal likelihood under the inferred GLM as compared to the

253   observed genetic data (Wegmann *et al.* 2009a). Thus, low p-values indicate that the observed

254   data is unlikely to have been generated under the inferred GLM, implying a bad model fit.

255   *Parameter Estimation*

256   To obtain good estimates of the posterior distributions of the parameters for the best-fitting

257   model (IM10-BNBO-DECSU), we used a MCMC without likelihood method (Wegmann *et*

258   *al.* 2009b). To reduce the dimensionality of the data and extract as much information as

259   possible about the model parameters, we used the first 20,000 simulations with the standard

260   sampler to define the first 12 orthogonal components of the summary statistics that maximise

261   the covariance matrix between summary statistics and model parameters. For this, we applied

262   a partial least-squares (PLS) regression approach (Boulesteix & Strimmer 2007) as

263   implemented in the "pls" R package (Mevik & Wehrens 2007) and used the R script provided

264   in the ABCtoolbox package. We defined the optimal number of PLS components by assessing

265   the drop in the root mean squared error for each parameter with the inclusion of additional

266   PLS components. This way, a large set of summary statistics is reduced to a number of

267   independent components, whereby summary statistics that are most informative about the

268   model parameters are weighted more than summary statistics that do not show much response

269   to changing parameter values (Wegmann *et al.* 2009b). The initial simulations were also used

270   to define the tolerance distance based on a tolerance level of 0.1 and to calibrate the transition

271   kernel of the MCMC run with a rangeProp setting of 1 unit of standard deviation (Wegmann

272   *et al.* 2009b; Wegmann *et al.* 2010). We then ran a total of 10^7 iterations with the MCMC

273   sampler, followed by a ABC-GLM post sampling regression (Leuenberger & Wegmann

274   2010) on the 10,000 simulations with the smallest Euclidean distance to the PLS components

275   of the observed summary statistics. Finally, we used R to plot the posterior distributions of

276   important model parameters.

277   *ABC Validation*

278   The performance of ABC in model selection and parameter estimation in complex

279   demographic settings inevitably suffers from the loss of information when the observed and

280   simulated genetic data are reduced to a set of summary statistics (Robert *et al.* 2011). This

281   necessitates a careful validation of the employed ABC procedure in order to avoid biases in

282   the approximation of posterior probabilities of evaluated models and the estimation of model

283   parameters. Accordingly, we validated our model selection and parameter estimation

284   approach with four different procedures. The first three validation approaches made use of so

285   called pseudo-observed data sets (*pods*), whereby parameter combinations are randomly

286   drawn from the prior distributions and the corresponding summary statistics were simulated

287   under a given model. These sets of summary statistics were then treated as if it were real

288   observed data, but since the model and the corresponding parameter values that generated

289   these summary statistics were known, we could use the *pods* to validate both our model

290   selection and parameter estimation procedure.

291   In the first validation step, we investigated the model misclassification rate for each pair-wise

292   model comparison by generating 100 *pods* under each model with parameters randomly

293   drawn from the prior distributions. We then performed the same model selection procedure as

294   with the real observed data and counted the number of assignments to each model. We

295   derived the model misclassification rate by counting all assignments of *pods* to a model other

296   than the one generating it (Figure 2).

297   Second, we assessed the accuracy of the parameter estimation, both in terms of different point

298   estimators (mode, average and median) and over the whole posterior distribution under

299   different tolerance levels (proportion of retained simulations). For this, we generated 1,000

300   *pods* under the best-fitting model (IM10-BNBO-DECSU) and performed the same parameter

301   estimation procedure on each *pods* as for the real data. The accuracy of the point estimators

302   was assessed using the average of the root mean squared errors (RMSE) over all 1,000 *pods*

303   (Supporting Table S7), while the root mean integrated squared error (Leuenberger &

304   Wegmann 2010) was used to assess accuracy over the whole posterior distribution

305   (Supporting Table S8). The results indicated that accuracy of the posterior distributions is

306   little affected by varying tolerance levels and that the mode of the distribution is the most

307   accurate point estimator for parameter estimation.

308   Third, to increase confidence in the parameter estimates of the best-fitting model, we checked

309   for biased posterior distributions by producing 1,000 *pods* under the best-fitting model with

310   parameter values drawn from the prior distributions. We used ABCtoolbox to calculate the

311   posterior quantiles of the true parameter values within the estimated posterior distributions for

312   each *pods* and used a Kolmogorov-Smirnov test for uniformity in R (Wegmann *et al.* 2009b).

Demographic History of Orang-utans (*Pongo* spp.)

313   Significant deviation from uniformity after sequential Bonferroni correction (Rice 1989)
314   would indicate biased posterior distributions (Cook *et al.* 2006). The distribution of posterior
315   quantiles within which the true values of the *pods* fell did not significantly deviate from the
316   expectation of uniformity for most parameters (Supporting Figure S4). In most cases where
317   the posterior quantiles were not distributed uniformly, data points were overrepresented in the
318   centre of the histogram, indicating that the posterior distributions were estimated too
319   conservatively.

320   In a last validation approach, we tested if the best-fitting model (IM10-BNBO-DECSU) and
321   the corresponding posterior distributions of the model parameters are able to reproduce the
322   summary statistics of the observed data. For this, we randomly sampled 10,000 parameter sets
323   from the inferred posterior distributions and used these to simulate genetic data under the
324   best-fitting model. We then carried out a PCA transformation of the simulated data and
325   plotted the first 16 principal components to check if the transformed observed data fell within
326   the distribution of the simulated data (Supporting Figure S5). This was the case for all the first
327   16 principal components, suggesting that the best-fitting model and its inferred parameter
328   estimates are well able to explain the observed data.

Results

Model Selection

331   We tested 12 demographic models, evaluating the impact of multiple demographic processes
332   on the current genetic makeup of orang-utan populations (Figure 2). We first compared
333   simple models that treated Bornean and Sumatran orang-utans as single populations, but
334   differed in the opportunity for migration after the population split (IM2 vs. I2, Figure 2A). We
335   found substantial support for the model allowing migration after the split (IM2, Bayes factor,
336   i.e. ratio of model posterior probabilities (BF) 5.18). However, this simple isolation with
337   migration model achieved only a very poor fit to the observed data, as shown by the
338   probability of the observed data under the general linear model used for parameter estimation
339   (GLM p-value) of 0.003, indicating that additional processes were involved in shaping the
340   gene pool of orang-utans. Of all four 2-population models tested, we observed a very strong
341   support for a model that allowed a sudden change in population size for both populations
342   followed by exponential growth (IM2-BN-GR vs. I2, IM2, IM2-GR, BF 36.79). Still, this

343   model did not achieve a good fit to the observed data, as evidenced by a p-value of the

344   observed data under the GLM of only 0.017 (Supporting Table S6).

345   The poor model fit of all tested 2-population models can be explained by the extensive

346   population substructure within the two orang-utan species (Arora *et al.* 2010; Kanthaswamy *et*

347   *al.* 2006; Nater *et al.* 2013; Nater *et al.* 2011; Warren *et al.* 2001), which differs to a great

348   extent for female- and male-mediated marker systems (Nater *et al.* 2011; Nietlisbach *et al.*

349   2012). Accordingly, the N_e for each marker system varies to a large degree and cannot be

350   described accurately with just one population size parameter per island. In agreement with

351   this notion, we found that a basic model with ten current population units (IM10) achieved a

352   better fit to the observed genetic data (GLM p-value 0.224) than all the 2-population models

353   (Supporting Table S6), and also obtained much stronger statistical support when directly

354   compared against the best 2-population model (IM10 vs. IM2-BN-GR, BF 830.21, Figure

355   2B). However, in our case, a better fit of the 10-population model compared to the 2-

356   population models was not unexpected, since part of the observed genetic data was used

357   beforehand to derive the number of population units in the 10-population models. When we

358   computed summary statistics for the IM10 model without pooling the genetic data for the

359   Sumatran populations north and south of Lake Toba, the model fit was still poor (GLM p-

360   value 0.019). In order to improve model fit, we first tested whether a population split

361   sequence following the species designation fits the data better than the pattern suggested by

362   mtDNA data (deepest split within Sumatran orang-utans north and south of Lake Toba). This

363   was strongly rejected by the observed data (IM10 vs. IM10-BOSU, BF 45.45, Figure 2B).

364   We then further tested for recent population declines on Sumatra (IM10-DECSU vs. IM10,

365   BF 57.03), on Borneo (IM10-DECBO vs. IM10, BF 0.48) or on both islands (IM10-DECALL

366   vs. IM10-DECSU, BF 0.94, Figure 2C). Incorporating a population decline on Sumatra

367   considerably improved the model fit (GLM p-value 0.553).

368   Next, we tested a model incorporating a bottleneck on Borneo together with a recent decline

369   on Sumatra (Figure 2D), which revealed substantial support for a bottleneck on Borneo

370   (IM10-BNBO-DECSU vs. IM10-DECSU, BF 3.60).

371   Finally, we evaluated the statistical support for a bottleneck on Sumatra associated with the

372   Toba supereruption (Figure 2E). We found substantial support against a bottleneck on

373   Sumatra in general (IM10-BNBO-DECSU vs. IM10-BNBO-TOBA-DECSU, BF 3.29), and

374   overwhelming support against a severe bottleneck (less than 100 individuals per population)

375   (IM10-BNBO-DECSU vs. IM10-BNBO-RECOL-DECSU, BF 10,887.60).

376   After performing a series of hierarchical model selection steps, we were able to identify a

377   demographic model (IM10-BNBO-DECSU) capable of reproducing the observed patterns of

378   DNA variation in the two current orang-utan species. Therefore, this model is likely to

379   capture the biologically most relevant processes in the demographic history of orang-utans.

380   **Parameter Estimation**

381   We estimated the model parameters for the selected 10-population model (IM10-BNBO-

382   DECSU, Figure 3) based on a total of 10 million simulations (Table 3, Figure 4). The

383   parameter estimates point to a current N_e of ~970 diploid individuals in each of the six

384   Bornean populations. We found support for a bottleneck on Borneo starting ~135 ka and

385   ending ~82 ka, during which N_e on Borneo was reduced from an ancestral N_e of ~17,000

386   individuals to ~2,600 individuals. The bottleneck on Borneo was followed by population

387   recovery and substructuring, with a current total N_e of all Bornean populations of ~6,150

388   individuals.

389   On Sumatra, the three populations north of Lake Toba suffered a decline ~24 ka from a N_e of

390   ~10,500 to currently only ~960 individuals in each of the three populations, corresponding to

391   a total N_e in the meta-population north of Lake Toba of ~38,300 and ~3,300 individuals

392   before and after the decline, respectively. We estimated that population structure north of

393   Lake Toba was established ~860 ka, with an ancestral effective population size of ~14,400

394   individuals. The population south of Lake Toba also went through a recent decline ~24 ka

395   from a N_e of ~24,200 individuals in the ancestral population to currently only ~1,030

396   individuals. Thus, Sumatran orang-utan populations first expanded during the Middle

397   Pleistocene before experiencing an island-wide population crash in the Late Pleistocene or

398   Early Holocene.

399   We inferred the population split time between Borneo and south Toba as ~1.13 Ma, and

400   between north and south of Lake Toba as ~3.39 Ma. Gene flow between Borneo and Sumatra

401   appears to have ceased ~87 ka, but this parameter was associated with a broad posterior

402   distribution. We found no evidence for asymmetric migration rates between Borneo and south

403   of Lake Toba, and between south of Lake Toba and north of Lake Toba. The migration rates

404   between the two islands were comparable to the migration rates over the Toba caldera on

405   Sumatra, while migration rates among the populations on Borneo and among those north of

406   Toba, respectively, were estimated to be about a magnitude higher.

407   ## Discussion

408   Our modelling approach capitalized on the use of multiple genetic marker systems and an

409   extensive set of geographically well-defined samples, in contrast to previous studies, which

410   based their findings on a small number of captive individuals with poorly recorded

411   provenance (Locke *et al.* 2011; Mailund *et al.* 2011; Mailund *et al.* 2012). Thus, our study

412   was able to shed light on important aspects of orang-utan demographic history that so far

413   remained unexamined due to non-representative sampling and dismissal of within-species

414   population structure. For instance, the inferred model by Locke *et al.* (2011) of a continuously

415   expanding Sumatran orang-utan population with a substantially larger current N_e as compared

416   to Bornean orang-utans was unrealistic in the light of current species distribution and

417   abundance, and did not capture recent population dynamics. Our results indicate that such

418   misleading signals are the result of a recent decline and deep divergence of orang-utan

419   populations on Sumatra, which yields a larger long-term N_e for Sumatran orang-utans as

420   compared to Bornean orang-utans in oversimplified demographic models.

421   **Inference of Best-Fitting Model**

422   We inferred that a model with comprehensive population structure, a bottleneck on Borneo

423   and a recent decline on Sumatra (IM10-BNBO-DECSU) fits the observed data significantly

424   better than a range of simplified models that treat each orang-utan species as a single

425   panmictic population. Estimation of demographic parameters under this model revealed a

426   population split time between Borneo and Sumatran populations south of Lake Toba of just

427   over a million years ago, followed by bidirectional gene flow. This species split time estimate

428   is considerably older than estimates obtained using whole genome data, suggesting a species

429   split time of between 330 and 600 ka (Locke *et al.* 2011; Mailund *et al.* 2011; Mailund *et al.*

430   2012). Such recent species split estimates are, however, in disagreement with findings based

431   on mitochondrial DNA, which yielded divergence time estimates of island specific mtDNA

432   lineages of 1–5 Ma (Nater *et al.* 2011; Steiper 2006; Warren *et al.* 2001; Xu & Arnason 1996;

433   Zhang *et al.* 2001; Zhi *et al.* 1996).

434   The discrepancy between model-based species split estimates using exclusively autosomal

435   data and mtDNA divergence time estimates from phylogenetic methods is owed to two

436   idiosyncrasies in the biology of orang-utans. First, due to the pronounced philopatric

437   tendencies of female orang-utans (Arora *et al.* 2012; Galdikas 1995; Nietlisbach *et al.* 2012;

438   van Noordwijk *et al.* 2012), mtDNA has likely experienced only little if any gene flow

439   between the two species after the species split. Therefore, the coalescent time of island-

440   specific mitochondrial lineages is expected to predate the population split between Borneo

441   and Sumatra, depending on N_e in the ancestral population (Nichols 2001). Second, due to

442   male-mediated gene flow, model-based approaches using solely autosomal data are likely to

443   underestimate species split times, as disentangling the contributions of migration and split

444   time remains challenging (Hey & Nielsen 2004). The recent split time estimates from

445   autosomal genomic data might reflect the end of an initial period of frequent, but strictly

446   male-driven gene flow after the species split. Such complex temporal fluctuations in

447   migration rates, as expected during glacial cycles for Sundaland species, are so far not

448   properly addressed in any applied demographic model. Still, combining markers with

449   different inheritance patterns as done in this study is likely to improve the estimates of both

450   migration rates and split times in species with sex-biased dispersal such as orang-utans.

451   Our findings of recent gene flow between Bornean and Sumatran orang-utans are in

452   agreement with previous observations (Becquet & Przeworski 2007; Muir *et al.* 2000;

453   Verschoor *et al.* 2004). In their genomic study, Locke and colleagues (2011) found an

454   unexpectedly high incidence of low-frequency mutations shared between Borneo and

455   Sumatra, which also hints at recent gene flow between the two islands. Contrary to studies

456   that indicated the presence of impassable dispersal barriers on the exposed Sunda shelf, either

457   due to large river systems (Harrison *et al.* 2006) or a putative savannah corridor (Bird *et al.*

458   2005; Gathorne-Hardy *et al.* 2002), it seems that habitat conditions during glacial periods did

459   at least sporadically allow male orang-utans to cross the exposed Sunda shelf. However, given

460   the strict and long lasting separation of mtDNA lineages on both islands (Nater *et al.* 2011), it

461   appears that the exposed shelf was not covered with forest able to sustain orang-utan

462   populations over prolonged periods. In fact, large parts of the Sunda shelf between Borneo

463   and Sumatra were covered with nutrient-poor sandy soils (Bird *et al.* 2005; Slik *et al.* 2011).

464   Forests on such soil types are characterized by low growth and productivity (Paoli *et al.*

465   2010). These constraints might explain why orang-utan populations on both islands could not

466   expand onto the exposed shelf to an extent where population admixture and thus exchange of

467   mtDNA lineages was possible.

Glacial Cycles and Population Size Changes

Since we also tested models that incorporated sudden population size changes, we were able to detect signals of a population bottleneck on Borneo. In contrast to Sumatra, the currently observed pattern of strong population differentiation on Borneo (Arora *et al.* 2010; Warren *et al.* 2001) seems to have been established only recently, as parameter estimation indicated that Bornean orang-utans were organized at least temporarily as a single panmictic population before ~80 ka. At ~140 ka, the ancient population on Borneo experienced a sudden drop in N_e from ~17,000 to ~2,500 individuals, which then recovered again to the current total N_e of ~6,000 for all Bornean orang-utans. Such a change in both N_e as well as population structure could be explained by a common Bornean refugium during either the penultimate (190–130 ka) or last (110-18 ka) glacial period, when the drier and more seasonal climate might have caused a drastic reduction of rainforest coverage on Borneo (Bird *et al.* 2005; Gathorne-Hardy *et al.* 2002; Morley 2000). Population contractions with subsequent expansions likely occurred multiple times on Borneo during Pleistocene glacial and interglacial cycles, but incorporating such complex population dynamics into a demographic model is currently not feasible with the data at hand.

Interestingly, a similar signal of a glacial refugium with subsequent population structuring, as observed in Bornean orang-utans, has been found in western gorillas (*Gorilla gorilla*). By using a demographic modelling approach comparable to our study, Thalmann *et al.* (2011) found that the two subspecies of western gorillas (*G. g. gorilla* and *G. g. diehli*) diverged only about ~18 ka, thus directly following the last glacial maximum (LGM) 19–26 ka (Clark *et al.* 2009). Furthermore, the ancient population of western gorillas exhibited a N_e of just ~2,500 individuals as compared to 22,000 and 17,000 individuals in the two subspecies after the population split. Therefore, it seems that western gorillas, similar to Bornean orang-utans, were constrained to a relatively small refugial population during glacial periods from which they subsequently expanded when the climate got warmer and wetter during interglacials.

Geological Processes and Population Size Changes

Linking bottleneck signals to specific environmental processes is difficult due to the large confidence intervals associated with most parameter estimates. For instance, the 90%-highest posterior interval for the estimate of the start of the bottleneck on Borneo (21–348 ka) also overlaps with the Toba supereruption on northern Sumatra ~73 ka (Chesner *et al.* 1991). It has been hypothesized that this colossal explosive eruption might have had a strong global

impact, causing a severe bottleneck in humans (Rampino & Ambrose 2000). However, evidence presented here points toward climatic changes during the glacial periods rather than the Toba supereruption as being the main cause for the detected bottleneck on Borneo, as our results showed that the supereruption did not even have a strong impact on the Sumatran populations despite their much closer geographic proximity. Models incorporating a severe bottleneck in the Sumatran populations around the time of the supereruption were clearly rejected and the signal of a recent population decline on Sumatra was considerably younger than the Toba supereruption. Studies indicate that the destruction caused by the Toba supereruption had been geographically limited, as shown by the distribution of rainforest refugia in Southeast Asia (Gathorne-Hardy *et al.* 2002), including on Mentawai Island around 350 kilometres from the Toba caldera (Gathorne-Hardy & Harcourt-Smith 2003), as well as the similar composition of Southeast Asian fossil sites before and after the date of the supereruption (Louys 2007). Given the proximity of contemporary populations of Sumatran orang-utans to the Toba caldera and the strong dependency of orang-utans on intact rain forest habitat, they are undoubtedly one of the most striking examples illustrating the limited impact of the Toba supereruption on the local flora and fauna in Southeast Asia. However, the lack of bottlenecks signals in the Sumatran populations does not imply that the activity of the Toba volcano did not influence the population history of Sumatran orang-utans at all. Rather, the results of this study, as well as previous findings (Nater *et al.* 2013; Nater *et al.* 2011), showed strikingly that the Toba eruptions must have repeatedly caused devastating damage to the local surroundings, which led to a long-lasting separation of gene pools north and south of Lake Toba.

In contrast to Toba as cause for the bottleneck on Borneo, a contraction of rainforests following the colder and drier climate during the last glacial period well explains the absence of a similar bottleneck in the Sumatran population history. During the generally drier glacial periods, large parts of Sumatra experienced considerably more rain fall compared to Borneo (Gathorne-Hardy *et al.* 2002; Whitten *et al.* 2000), because the Barisan mountain range running the length of Sumatra acted as a barrier for the wet monsoon winds, causing high precipitation along its western slopes (Whitten *et al.* 2000). This mountain ridge effect in combination with the close proximity to the sea during glacial periods, when sea levels were low, might have allowed large areas of rainforest to persist on Sumatra during glacial periods (Gathorne-Hardy *et al.* 2002). Thus, Sumatran orang-utans were almost certainly not forced into glacial refugia to the same extent as Borneans.

Demographic History of Orang-utans (*Pongo* spp.)

Anthropogenic Impacts on Orang-Utan Populations

While Sumatran orang-utans did not seem to go through glacial bottlenecks, we found evidence for recent and drastic declines in population sizes north and south of Lake Toba. These signals of population decline cannot be attributed to the large-scale human-induced habitat degradation that started in the last century (Rijksen & Meijaard 1999), of which genetic signals were found in previous studies of Bornean orang-utans (Goossens *et al.* 2006; Sharma *et al.* 2012). Rather, our results point toward an earlier decline in the Late Pleistocene or Early Holocene. In the Late Pleistocene, orang-utans went extinct on the Southeast Asian mainland as well as in many Sundaland regions (Delgado & Van Schaik 2000; Jablonski 1998; Rijksen & Meijaard 1999). Furthermore, the Pleistocene-Holocene boundary is characterized by the disappearance of many large-bodied animals world-wide (Koch & Barnosky 2006), including large parts of the megafauna in Southeast Asia (Louys *et al.* 2007). The increased occurrence of megafaunal extinctions during this period has been attributed to climatic changes following the LGM, the impact of human hunting and human-induced habitat changes, or the combination of these two factors (reviewed in Koch & Barnosky 2006).

Both climatic and anthropogenic factors might have played a role in the decline and local extinctions of orang-utan populations in the Late Pleistocene. During the LGM, the drier and more seasonal climate caused a shifting of zones of evergreen rainforest toward the equator (Flenley 1998; Jablonski 1998; Morley 2000), likely causing populations in southern China to go extinct. The warmer climate following the LGM was accompanied by rising sea levels, which drastically increased the extent of coastlines in Sundaland (Voris 2000). This enlargement of coastal habitat might have promoted an expansion of early modern humans on Sundaland, leading to increased hunting pressure on large-bodied animals, including orang-utans (Hill *et al.* 2007; Soares *et al.* 2008). Such hunting by modern humans might have caused the local extinctions of orang-utans on many Sundaland islands, and led to a strong decline in Sumatran populations north and south of Lake Toba. Bornean orang-utans did not seem to be as strongly affected by human hunting, probably because the large size and low productivity of Borneo left enough inland areas with relatively low human densities (Delgado & Van Schaik 2000).

Our modelling approach revealed that the two recognised orang-utan species experienced drastically different demographic histories. Sumatran orang-utans exhibit a deep and temporally stable population structure, including an old divergence of gene pools north and

south of Lake Toba with limited amount of gene flow over the Toba caldera. The populations on Sumatra recently suffered a strong decline, which, in combination with the strong population structure, explains the high genetic diversity found in recent genomic studies despite their low current census size (Locke *et al.* 2011; Prado-Martinez *et al.* 2013). In contrast, we find that the population structure currently observed within Bornean orang-utans has been established only recently and the population went through at least one bottleneck most likely associated with a glacial refugium.

These results strongly suggest that special consideration needs to be given to demographic factors when analysing adaptive evolutionary processes in great apes. Due to their strong dependence on intact forest habitat, most great ape taxa were severely affected by the climate shifts during glacial periods, which were accompanied by drastic changes in forest coverage in the tropics (Flenley 1998; Morley 2000). Accordingly, great ape populations experienced population bottlenecks, founder events, population expansions and population structuring as recent as 15,000 years ago (Clark *et al.* 2009). Given the long generation time of all great apes (18–30 years, Wich *et al.* 2009), great ape populations will likely not have reached an equilibrium state for most genomic regions. Thus, population expansions and substructuring caused by relatively recent climatic changes might produce erroneous signals of selective sweeps if demography is not taken into account. Our results therefore emphasize the need to further advance the development of tools to jointly estimate demography and selection in order to unravel the convoluted evolutionary history of great apes (Li *et al.* 2012).

Acknowledgements

602 **References**

603 Arora N, Nater A, van Schaik CP*, et al.* (2010) Effects of Pleistocene glaciations and rivers
604   on the population structure of Bornean orangutans (*Pongo pygmaeus*). *Proceedings of the*
605   *National Academy of Sciences* **107**, 21376-21381.
606 Arora N, van Noordwijk MA, Ackermann C*, et al.* (2012) Parentage-based pedigree
607   reconstruction reveals female matrilineal clusters and male-biased dispersal in
608   nongregarious Asian great apes, the Bornean orang-utans (*Pongo pygmaeus*). *Molecular*
609   *Ecology* **21**, 3352-3362.
610 Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in
611   population genetics. *Genetics* **162**, 2025-2035.
612 Becquet C, Przeworski M (2007) A new approach to estimate parameters of speciation
613   models with application to apes. *Genome Research* **17**, 1505-1519.
614 Bird MI, Taylor D, Hunt C (2005) Environments of insular Southeast Asia during the last
615   glacial period: a savanna corridor in Sundaland? *Quaternary Science Reviews* **24**, 2228-
616   2242.
617 Boulesteix AL, Strimmer K (2007) Partial least squares: a versatile tool for the analysis of
618   high-dimensional genomic data. *Briefings in Bioinformatics* **8**, 32-44.
619 Chesner CA, Rose WI, Deino A, Drake R, Westgate JA (1991) Eruptive history of earths
620   largest Quaternary caldera (Toba, Indonesia) clarified. *Geology* **19**, 200-203.
621 Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA (2010) The confounding effects of
622   population structure, genetic diversity and the sampling scheme on the detection and
623   quantification of population size changes. *Genetics* **186**, 983-U347.
624 Clark PU, Dyke AS, Shakun JD*, et al.* (2009) The last glacial maximum. *Science* **325**, 710-
625   714.
626 Cook SR, Gelman A, Rubin DB (2006) Validation of software for Bayesian models using
627   posterior quantiles. *Journal of Computational and Graphical Statistics* **15**, 675-692.
628 Delgado RA, Van Schaik CP (2000) The behavioral ecology and conservation of the
629   orangutan (*Pongo pygmaeus*): a tale of two islands. *Evolutionary Anthropology* **9**, 201-
630   218.
631 Ellegren H (2014) Genome sequencing and population genomics in non-model organisms.
632   *Trends in Ecology & Evolution* **29**, 51-63.
633 Excoffier L, Foll M (2011) Fastsimcoal: a continuous-time coalescent simulator of genomic
634   diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* **27**, 1332-1334.
635 Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically
636   structured population. *Heredity* **103**, 285-298.
637 Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform
638   population genetics analyses under Linux and Windows. *Molecular Ecology Resources*
639   **10**, 564-567.
640 Fischer A, Pollack J, Thalmann O, Nickel B, Paabo S (2006) Demographic history and
641   genetic differentiation in apes. *Current Biology* **16**, 1133-1138.

642 Flenley JR (1998) Tropical forests under the climates of the last 30,000 years. *Climatic*
643   *Change* **39**, 177-197.
644 Galdikas BMF (1995) Social and reproductive behavior of wild adolescent female orangutans.
645   In: *The Neglected Ape* (eds. Nadler RD, Galdikas BFM, Sheeran LK, Rosen N), pp. 163-
646   182. Plenum Press, New York.
647 Garza JC, Williamson EG (2001) Detection of reduction in population size using data from
648   microsatellite loci. *Molecular Ecology* **10**, 305-318.
649 Gathorne-Hardy FJ, Harcourt-Smith WEH (2003) The super-eruption of Toba, did it cause a
650   human bottleneck? *Journal of Human Evolution* **45**, 227-230.
651 Gathorne-Hardy FJ, Syaukani, Davies RG, Eggleton P, Jones DT (2002) Quaternary
652   rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. *Biological*
653   *Journal of the Linnean Society* **75**, 453-466.
654 Goossens B, Chikhi L, Ancrenaz M, *et al.* (2006) Genetic signature of anthropogenic
655   population collapse in orang-utans. *Plos Biology* **4**, 285-291.
656 Goossens B, Chikhi L, Jalil MF, *et al.* (2005) Patterns of genetic diversity and migration in
657   increasingly fragmented and declining orang-utan (*Pongo pygmaeus*) populations from
658   Sabah, Malaysia. *Molecular Ecology* **14**, 441-456.
659 Greminger MP, Stölting KN, Nater A, *et al.* (2014) Generation of SNP datasets for orangutan
660   population genomics using improved reduced-representation sequencing and direct
661   comparisons of SNP calling algorithms. *Bmc Genomics* **15**.
662 Groves CP (2001) *Primate taxonomy* Smithsonian Institution Press, Washington, D.C. ;
663   London.
664 Haddrill PR, Thornton KR, Charlesworth B, Andolfatto P (2005) Multilocus patterns of
665   nucleotide variability and the demographic and selection history of *Drosophila*
666   *melanogaster* populations. *Genome Research* **15**, 790-799.
667 Hahn MW (2008) Toward a selection theory of molecular evolution. *Evolution* **62**, 255-265.
668 Harrison T, Krigbaum J, Manser J (2006) Primate biogeography and ecology on the Sunda
669   shelf islands: a paleontological and zooarchaeological perspective. In: *Primate*
670   *Biogeography*. Springer US.
671 Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates
672   and divergence time, with applications to the divergence of *Drosophila pseudoobscura*
673   and *D. persimilis*. *Genetics* **167**, 747-760.
674 Hill C, Soares P, Mormina M, *et al.* (2007) A mitochondrial stratigraphy for island southeast
675   Asia. *American Journal of Human Genetics* **80**, 29-43.
676 Jablonski NG (1998) The response of catarrhine primates to Pleistocene environmental
677   fluctuations in East Asia. *Primates* **39**, 29-37.
678 Kaessmann H, Wiebe V, Weiss G, Pääbo S (2001) Great ape DNA sequences reveal a
679   reduced diversity and an expansion in humans. *Nature Genetics* **27**, 155-156.
680 Kanthaswamy S, Kurushima JD, Smith DG (2006) Inferring *Pongo* conservation units: a
681   perspective based on microsatellite and mitochondrial DNA analyses. *Primates* **47**, 310-
682   321.
683 Koch PL, Barnosky AD (2006) Late quaternary extinctions: state of the debate. *Annual*
684   *Review of Ecology Evolution and Systematics* **37**, 215-250.
685 Leuenberger C, Wegmann D (2010) Bayesian computation and model selection without
686   likelihoods. *Genetics* **184**, 243-252.
687 Li J, Li H, Jakobsson M, *et al.* (2012) Joint analysis of demography and selection in
688   population genetics: where do we stand and where could we go? *Molecular Ecology* **21**,
689   28-44.
690 Locke DP, Hillier LW, Warren WC, *et al.* (2011) Comparative and demographic analysis of
691   orang-utan genomes. *Nature* **469**, 529-533.

Louys J (2007) Limited effect of the Quaternary's largest super-eruption (Toba) on land mammals from Southeast Asia. *Quaternary Science Reviews* **26**, 3108-3117.

Louys J, Curnoe D, Tong HW (2007) Characteristics of Pleistocene megafauna extinctions in Southeast Asia. *Palaeogeography Palaeoclimatology Palaeoecology* **243**, 152-173.

Mailund T, Dutheil JY, Hobolth A, Lunter G, Schierup MH (2011) Estimating divergence time and ancestral effective population size of Bornean and Sumatran orangutan subspecies using a coalescent hidden Markov model. *Plos Genetics* **7**.

Mailund T, Halager AE, Westergaard M*, et al.* (2012) A new isolation with migration model along complete genomes infers very different divergence processes among closely related great ape species. *Plos Genetics* **8**.

Marjoram P, Molitor J, Plagnol V, Tavare S (2003) Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15324-15328.

Mevik BH, Wehrens R (2007) The pls package: principal component and partial least squares regression in R. *Journal of Statistical Software* **18**.

Morley RJ (2000) *Origin and evolution of tropical rain forests* Wiley, Chichester.

Muir CC, Galdikas BMF, Beckenbach AT (2000) mtDNA sequence diversity of orangutans from the islands of Borneo and Sumatra. *Journal of Molecular Evolution* **51**, 471-480.

Nater A, Arora N, Greminger MP*, et al.* (2013) Marked population structure and recent migration in the critically endangered Sumatran orangutan (*Pongo abelii*). *Journal of Heredity* **104**, 2-13.

Nater A, Nietlisbach P, Arora N*, et al.* (2011) Sex-biased dispersal and volcanic activities shaped phylogeographic patterns of extant orangutans (genus: *Pongo*). *Molecular Biology and Evolution* **28**, 2275-2288.

Nei M, Suzuki Y, Nozawa M (2010) The neutral theory of molecular evolution in the genomic era. *Annual Review of Genomics and Human Genetics* **11**, 265-289.

Nichols R (2001) Gene trees and species trees are not the same. *Trends in Ecology & Evolution* **16**, 358-364.

Nichols RA, Bruford MW, Groombridge JJ (2001) Sustaining genetic variation in a small population: evidence from the Mauritius kestrel. *Molecular Ecology* **10**, 593-602.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics* **39**, 197-218.

Nietlisbach P, Arora N, Nater A*, et al.* (2012) Heavily male-biased long-distance dispersal of orang-utans (genus: *Pongo*), as revealed by Y-chromosomal and mitochondrial genetic markers. *Molecular Ecology* **21**, 3173-3186.

Paoli GD, Wells PL, Meijaard E*, et al.* (2010) Biodiversity conservation in the REDD. *Carbon Balance Manag* **5**, 7.

Prado-Martinez J, Sudmant PH, Kidd JM*, et al.* (2013) Great ape genetic diversity and population history. *Nature* **499**, 471-475.

Rampino MR, Ambrose SH (2000) Volcanic winter in the Garden of Eden: the Toba supereruption and the late Pleistocene human population crash. In: *Volcanic Hazards and Disasters in Human Antiquity* (eds. McCoy FW, Heiken G). Geological Society of America, Boulder.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Rijksen HD, Meijaard E (1999) *Our vanishing relative: the status of wild orang-utans at the close of the twentieth century* Springer Netherlands.

Robert CP, Cornuet JM, Marin JM, Pillai NS (2011) Lack of confidence in approximate Bayesian computation model choice. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15112-15117.

741   Sharma R, Arora N, Goossens B, *et al.* (2012) Effective population size dynamics and the
742   demographic collapse of Bornean orang-utans. *Plos One* **7**.

743   Slik JWF, Aiba SI, Bastian M, *et al.* (2011) Soils on exposed Sunda Shelf shaped
744   biogeographic patterns in the equatorial forests of Southeast Asia. *Proceedings of the*
745   *National Academy of Sciences of the United States of America* **108**, 12343-12347.

746   Soares P, Trejaut JA, Loo JH, *et al.* (2008) Climate change and postglacial human dispersals
747   in Southeast Asia. *Molecular Biology and Evolution* **25**, 1209-1218.

748   Stadler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009) The Impact of Sampling
749   Schemes on the Site Frequency Spectrum in Nonequilibrium Subdivided Populations.
750   *Genetics* **182**, 205-216.

751   Stajich JE, Hahn MW (2005) Disentangling the effects of demography and selection in human
752   history. *Molecular Biology and Evolution* **22**, 63-73.

753   Steiper ME (2006) Population history, biogeography, and taxonomy of orangutans (Genus :
754   *Pongo*) based on a population genetic meta-analysis of multiple loci. *Journal of Human*
755   *Evolution* **50**, 509-522.

756   Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA
757   polymorphism. *Genetics* **123**, 585-595.

758   Tavare S, Balding DJ, Griffiths RC, Donnelly P (1997) Inferring coalescence times from
759   DNA sequence data. *Genetics* **145**, 505-518.

760   Team RDC (2010) R: a language and environment for statistical computing. R Foundation for
761   Statistical Computing, Vienna, Austria.

762   Teshima KM, Coop G, Przeworski M (2006) How reliable are empirical genomic scans for
763   selective sweeps? *Genome Research* **16**, 702-712.

764   Thalmann O, Wegmann D, Spitzner M, *et al.* (2011) Historical sampling reveals dramatic
765   demographic changes in western gorilla populations. *BMC Evolutionary Biology* **11**.

766   van Noordwijk MA, Arora N, Willems EP, *et al.* (2012) Female philopatry and its social
767   benefits among Bornean orangutans. *Behavioral Ecology and Sociobiology* **66**, 823-834.

768   Verschoor EJ, Langenhuijzen S, Bontjer I, *et al.* (2004) The phylogeography of orangutan
769   foamy viruses supports the theory of ancient repopulation of Sumatra. *Journal of Virology*
770   **78**, 12712-12716.

771   Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems
772   and time durations. *Journal of Biogeography* **27**, 1153-1167.

773   Warren KS, Verschoor EJ, Langenhuijzen S, *et al.* (2001) Speciation and intrasubspecific
774   variation of Bornean orangutans, *Pongo pygmaeus pygmaeus. Molecular Biology and*
775   *Evolution* **18**, 472-480.

776   Wegmann D, Leuenberger C, Excoffier L (2009a) *ABCtoolbox manual*.

777   Wegmann D, Leuenberger C, Excoffier L (2009b) Efficient Approximate Bayesian
778   computation coupled with Markov chain Monte Carlo without likelihood. *Genetics* **182**,
779   1207-1218.

780   Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a versatile
781   toolkit for approximate Bayesian computations. *Bmc Bioinformatics* **11**.

782   Whitten T, Damanik SJ, Anwar J, Hisyam N (2000) *The ecology of Sumatra* Periplus Editions
783   Ltd., Hong Kong.

784   Wich SA, de Vries H, Ancrenaz M, *et al.* (2009) Orangutan life history variation. In:
785   *Orangutans: geographic variation in behavioral ecology and conservation* (eds. Wich SA,
786   Utami Atmoko SS, Mitra Setia T, van Schaik CP). Oxford University Press.

787   Wich SA, Meijaard E, Marshall AJ, *et al.* (2008) Distribution and conservation status of the
788   orang-utan (*Pongo* spp.) on Borneo and Sumatra: how many remain? *Oryx* **42**, 329-339.

789 Williams MAJ, Ambrose SH, van der Kaars S, *et al.* (2009) Environmental impact of the 73
790   ka Toba super-eruption in South Asia. *Palaeogeography Palaeoclimatology*
791   *Palaeoecology* **284**, 295-314.
792 Wilson IJ, Weale ME, Balding DJ (2003) Inferences from DNA data: population histories,
793   evolutionary processes and forensic match probabilities. *Journal of the Royal Statistical*
794   *Society Series a-Statistics in Society* **166**, 155-188.
795 Xu XF, Arnason U (1996) The mitochondrial DNA molecule of Sumatran orangutan and a
796   molecular proposal for two (Bornean and Sumatran) species of orangutan. *Journal of*
797   *Molecular Evolution* **43**, 431-437.
798 Zhang YW, Ryder OA, Zhang YP (2001) Genetic divergence of orangutan subspecies (Pongo
799   pygmaeus). *Journal of Molecular Evolution* **52**, 516-526.
800 Zhi L, Karesh WB, Janczewski DN*, et al.* (1996) Genomic differentiation among natural
801   populations of orang-utan (Pongo pygmaeus). *Current Biology* **6**, 1326-1336.

802

Data Accessibility

Sequence data not published previously is accessible under GenBank accession numbers XXXX-YYYY.

Author Contributions

AN, MPG, CPvS, and MK designed the study; BG, IS, EJV, and KSW provided samples; AN, MPG, and NA performed laboratory procedures; AN and MPG conducted genetic data analysis; AN performed demographic modelling; AN wrote the manuscript; MPG, NA, CPvS, and MK critically revised the manuscript and provided comments at all stages; BG, EJV, and KSW edited the final manuscript.

Table 1: Sample sizes for the different marker systems in the ten geographic regions

Sampling region[a]	mtDNA	Y-STRs	Autosomal STRs	Autosomal regions	Xq13.3
North Kinabatangan (NK)	6	10	32	4	3
South Kinabatangan (SK)	13	15	76	4	3
East Kalimantan (EK)	7	9	34	4	5
Sarawak (SR)	8	2	12	2	1
Central Kalimantan (CK)	9	9	68	2	2
West Kalimantan (WK)	9	8	32	4	4
Batang Toru (BT)	8	8	18	4	3
North Aceh (NA)	7	15	32	6	3
Langkat (LK)	14	15	66	10	6
West Alas (WA)	37	38	104	4	7
Total	118	129	474	44	37

[a], sampling regions corresponding to Figure 1. Sample sizes are given as number of sampled chromosomes. The light grey shading refers to Bornean populations, middle grey to Sumatran populations north of Lake Toba, and dark grey to the Sumatran population south of Lake Toba.

Table 2: Summary statistics for the marker systems used in the ABC analysis.

Sequences	L_{Bases}[a]	Group	N_{Ind}[b]	N_{Seg}[c]	π[d]	θ_W[e]	D[f]
mtDNA (16S, ND3, CYTB)	1,355	Borneo	52	19	0.0022	0.0031	-0.92
		South Toba	8	1	0.0002	0.0003	-1.05
		North Toba	58	41	0.0100	0.0066	1.79
Autosomal regions (Chr2a_R17, Chr9_R16, Chr12_R1, Chr19_R7)	8,238	Borneo	10	19.50 ±4.56	0.0033 ±0.0011	0.0027 ±0.0006	0.68 ±0.66
		South Toba	2	13.50 ±7.79	0.0037 ±0.0020	0.0036 ±0.0020	0.08 ±0.49
		North Toba	10	28.75 ±4.66	0.0046 ±0.0012	0.0040 ±0.0006	0.51 ±0.52
Xq13.3	8,055	Borneo	18	6	0.0001	0.0002	-1.11
		South Toba	3	33	0.0027	0.0027	0.00
		North Toba	15	54	0.0020	0.0020	-0.09

Microsatellites	N_{Loci}[g]	Group	N_{Ind}	N_A[h]	H_O[i]	H_E[j]	G-W[k]
Autosomal-STR	25	Borneo	127	7.16 ±4.13	0.53 ±0.22	0.61 ±0.25	0.90 ±0.15
		South Toba	9	3.84 ±1.18	0.60 ±0.23	0.62 ±0.16	0.72 ±0.22
		North Toba	101	6.32 ±3.11	0.61 ±0.16	0.65 ±0.16	0.82 ±0.17
Y-STR	11	Borneo	53	3.18 ±2.48	-	0.31 ±0.33	0.90 ±0.14
		South Toba	8	1.27 ±0.65	-	0.08 ±0.19	0.88 ±0.18
		North Toba	68	1.91 ±1.64	-	0.12 ±0.24	0.91 ±0.17

Statistics are provided as average and standard deviation for marker systems with multiple independent loci; [a], sequence length in base pairs; [b], number of sampled individuals; [c], number of segregating sites; [d], nucleotide diversity; [e], Watterson's θ per base pair; [f], Tajima's D (Tajima 1989); [g], number of loci; [h], number of alleles; [i], observed heterozygosity; [j], expected heterozygosity; [k], Garza-Williamson index (Garza & Williamson 2001);

Table 3: Estimates of the model parameters for the selected 10-population model with a bottleneck on Borneo and a recent decline on Sumatra (IM10-BNBO-DECSU).

Parameter[a]	Prior[b]	Mode	Mean	90%-HPD[c]
$N_{NOW}BO$ [ind] (6)	logunif[100;10,000]	974	1,028	348–3,011
$N_{NOW}NT$ [ind] (3)	logunif[100;10,000]	963	933	239–3,613
$N_{NOW}ST$ [ind] (1)	logunif[100;10,000]	1,034	952	189–4,514
$N_{BN}BO$ [ind] (1)	logunif[100;10,000]	2,598	1,486	286–9,988
$N_{ANC}BO$ [ind] (1)	logunif[1,000;100,000]	17,046	12,344	2,171–89,115
$N_{STRUC}NT$ [ind] (3)	logunif[1,000;100,000]	10,508	11,278	1,886–78,264
$N_{ANC}NT$ [ind] (1)	logunif[1,000;100,000]	14,407	10,519	1,565–70,259
$N_{ANC}ST$ [ind] (1)	logunif[1,000;100,000]	24,193	13,991	2,629–99,070
$T_{BNEND}BO$ [yrs]	unif[8,750;400,000]	81,946	149,580	8,848–283,785
$T_{BNSTART}BO$ [yrs]	$T_{BNEND}BO$ + unif[250;100,000]	135,076	191,001	20,855–348,145
$T_{SPLIT}BO$ [kyrs]	unif[400;1,500]	1,128	960	497–1,436
$T_{DEC}SU$ [yrs]	unif[1.0;3.5]	23,651	36,200	4,119–67,272
$T_{STRUC}NT$ [kyrs]	unif[75;1,500]	861	820	267–1,398
$T_{SPLIT}NT$ [kyrs]	unif[1,500;4,000]	3,392	2,995	2,101–3,999
$T_{MIGSTOP}$ [yrs]	unif[2.5;4.2]	87,034	161,862	8,849–310,833
$Log(m_{BO-ST})$ [migrants/ind/gen]	unif[-5.0;-3.0]	-3.55	-3.96	-4.79–-3.09
$Log(m_{ST-BO})$ [migrants/ind/gen]	unif[-5.0;-3.0]	-3.42	-3.84	-4.61–-3.10
$Log(m_{NT-ST})$ [migrants/ind/gen]	unif[-5.0;-3.0]	-3.89	-3.98	-4.81–-3.14
$Log(m_{ST-NT})$ [migrants/ind/gen]	unif[-5.0;-3.0]	-3.65	-3.92	-4.71–-3.06
$Log(m_{BO})$ [migrants/ind/gen]	unif[-4.0;-2.0]	-2.52	-2.90	-3.66–-2.02
$Log(m_{NT})$ [migrants/ind/gen]	unif[-4.0;-2.0]	-2.51	-2.89	-3.65–-2.03

[a], BO = Borneo, NT = Sumatra north of Lake Toba, ST = Sumatra south of Lake Toba, N_{NOW} = current effective population size, N_{BN} = effective population size during population bottleneck, N_{ANC} = ancestral effective population size, N_{STRUC} = effective population size before recent decline, T_{BNEND} = time since population bottleneck ended, $T_{BNSTART}$ = time when population bottleneck started, T_{SPLIT} = population split time, T_{DEC} = time since population decline, T_{STRUC} = time since establishment of population structure, $T_{MIGSTOP}$ = time since migration between Borneo and Sumatra stopped, m = migration rate per individual per generation (an illustration of the meaning of the different model parameters can be found in Figure 3), the number in brackets next to the population size parameters refer to the number of simulated populations of this size each; [b], The prior distributions for the parameter values were either uniform or loguniform within the boundaries provided in squared brackets; [c], 90%-highest posterior density interval.
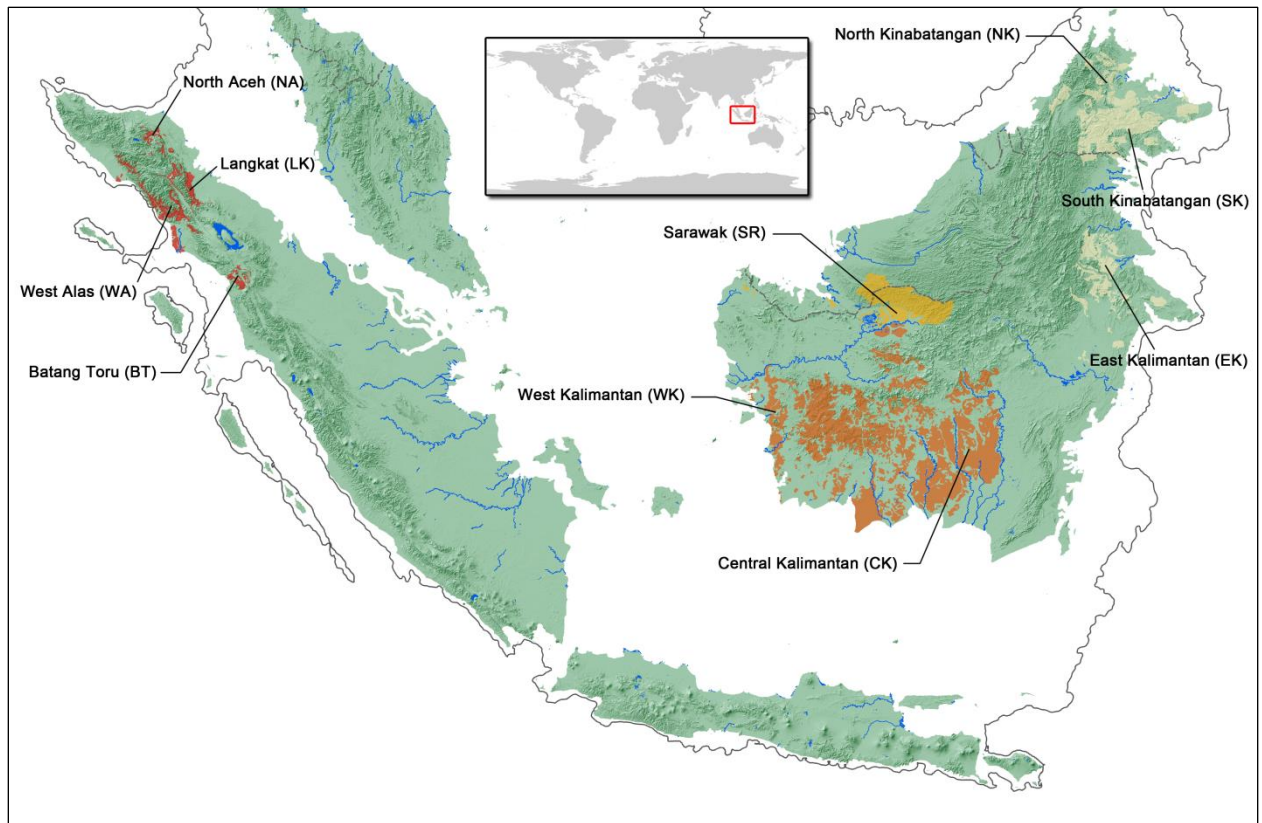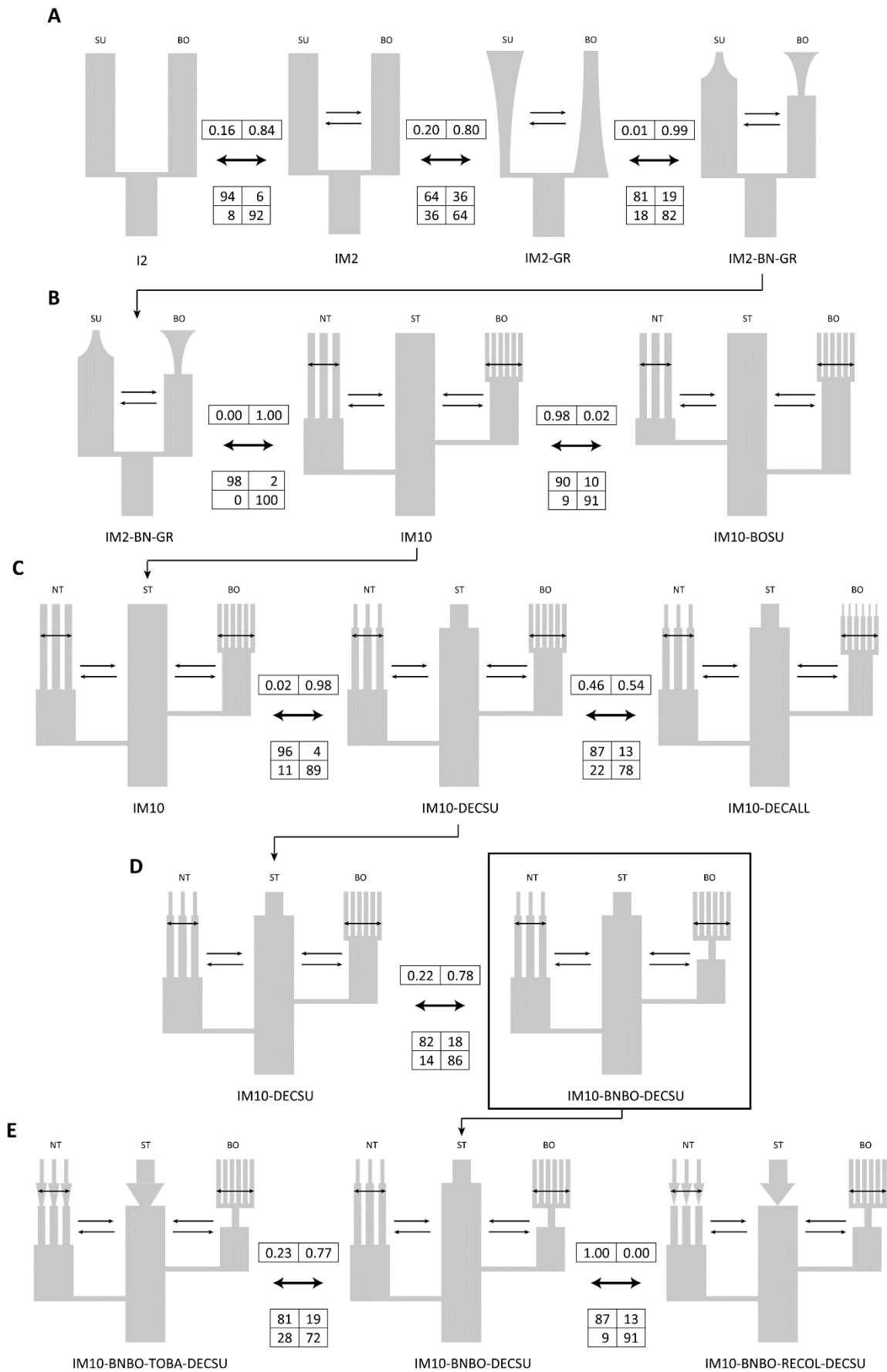
Demographic History of Orang-utans (*Pongo* spp.)

Figure 1: Map of sampling regions in Sundaland used for the demographic modelling. Shaded areas represent the current distribution of the Sumatran orang-utans and the three subspecies of Bornean orang-utans. The grey line indicates the extent of the exposed Sunda shelf during the LGM (19–26 ka, -120 meters below current sea level).

Demographic History of Orang-utans (*Pongo* spp.)

Figure 2: Schematic representation of the hierarchical model testing procedure. The 12 tested demographic models can be divided into four 2-population models and eight 10-populations models (IM10-DECBO not shown). The box above the left-right arrow shows the model posterior probabilities for each model comparison pair. The overall best-fitting model (IM10-BNBO-DECSU) is shown in a black frame. The box below the left-right arrow shows the power to distinguish between the two compared models as evaluated in a cross-validation procedure with 100 validations for each model, with the upper left and lower right boxes showing the correct model assignments for model 1 and model 2, respectively (SU = Sumatra, BO = Borneo, NT = Sumatra north of Lake Toba, ST = Sumatran south of Lake Toba). A) Comparison of four 2-population models, testing gene flow after the population split, exponential population growth or decline after the population split, and sudden population size change followed by exponential growth or decline. B) Comparison between the best fitting 2-population model and two 10-population models incorporating population structure. C) Tests of recent population declines on Sumatra, and Sumatra as well as Borneo. D) Test of population bottleneck on Borneo. E) Testing of a population bottleneck on Sumatra associated with the Toba supereruption 65–75 ka. The leftmost model implements a bottleneck in all four populations on Sumatra, followed by exponential population recovery. The rightmost model is similar, but restricts the bottleneck to a size of less than 100 surviving individuals per population, thus representing a scenario where regions devastated by the Toba eruption were recolonized from other areas after restoration of the rain forest habitat.
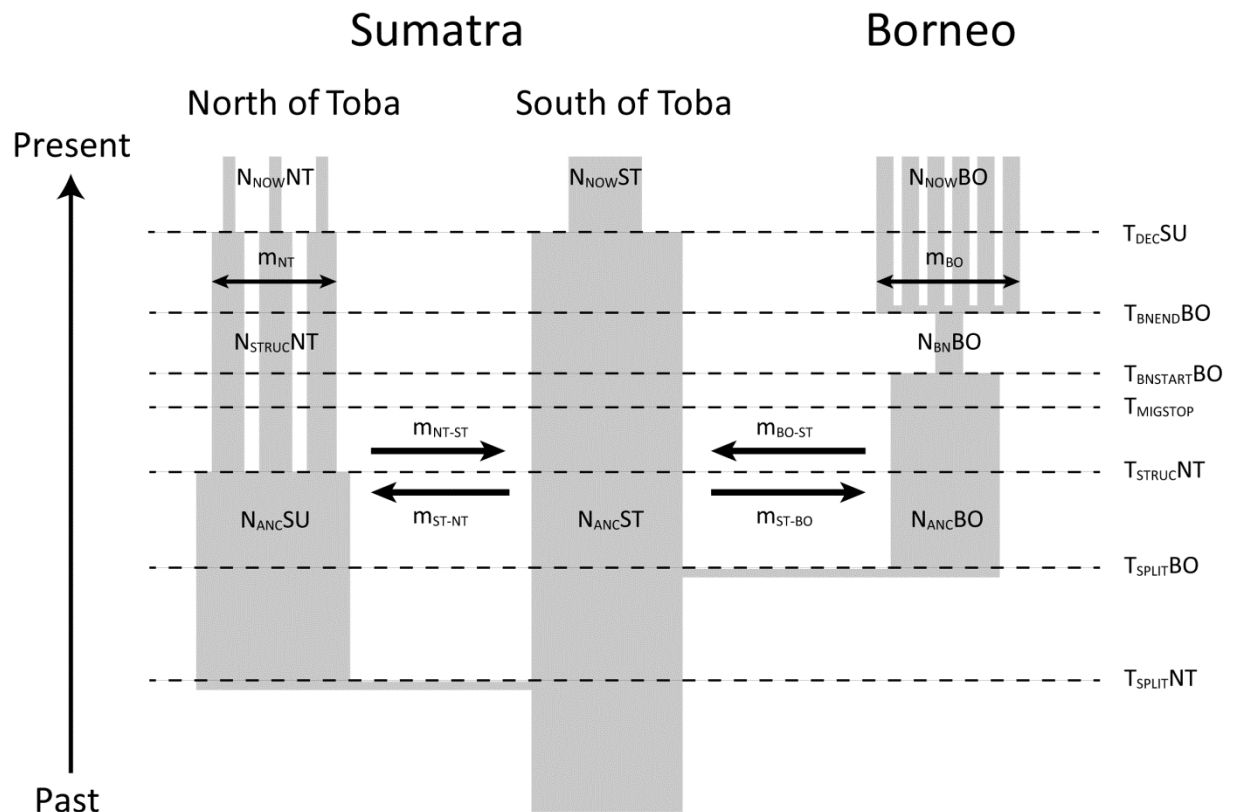
Figure 3: Schematic representation of the selected 10-population model with a bottleneck on Borneo and recent population declines in all Sumatran populations (IM10-BNBO-DECSU).
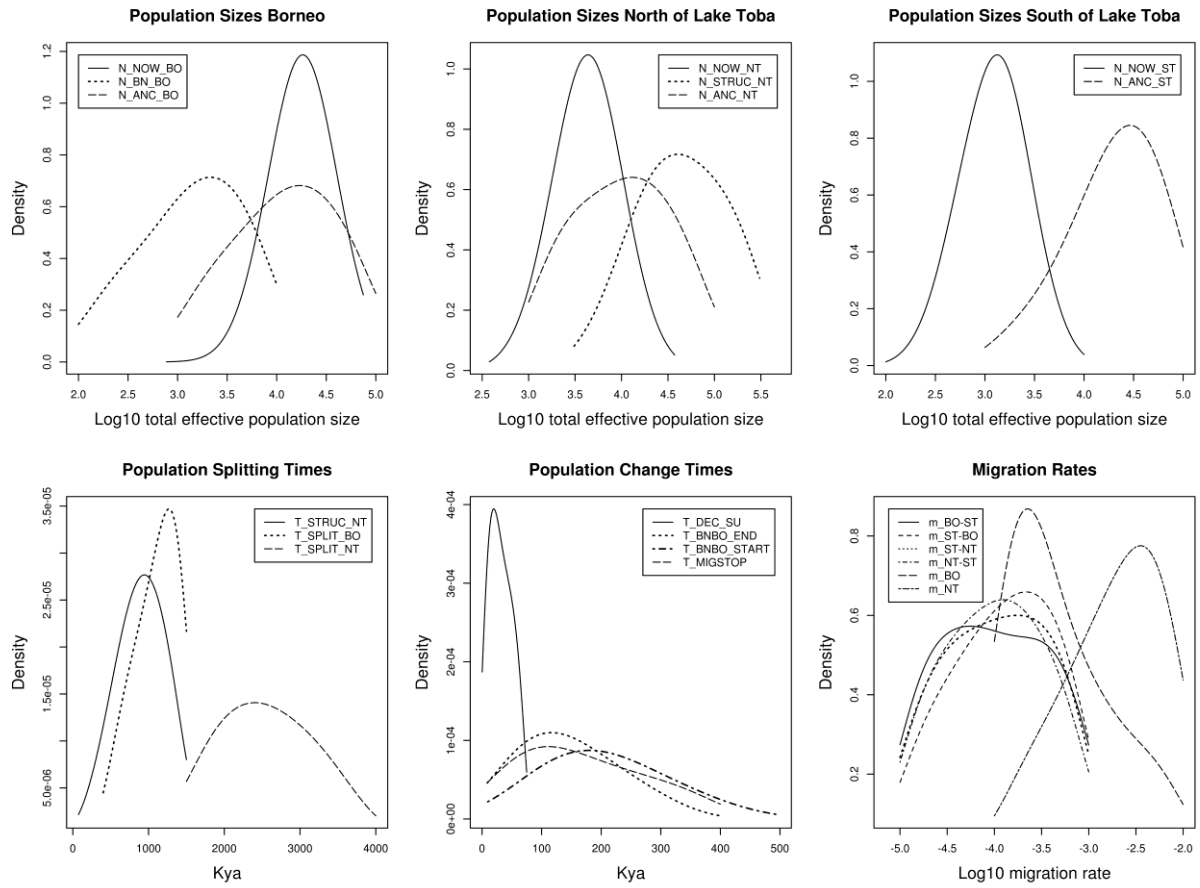
Figure 4: Posterior distributions of important model parameter under the selected 10-population model (IM10-BNBO-DECSU). The abbreviations of the model parameters correspond to the labels in Figure 3. For better comparability, the effective populations sizes of the structured meta-populations on Borneo and north of Lake Toba are given as the total effective sizes according to the formula $N_e = D \times N \times (1 + (1/(4 \times N \times m)))$, with D corresponding to the number of subpopulations, N to the mean subpopulation size and m to the total migration rate per individual per generation within the meta-population (Nichols *et al.* 2001).