

# Reconstructing the demographic history of orang-utans using Approximate Bayesian Computation

ALEXANDER NATER,\* MAJA P. GREMINGER,\* NATASHA ARORA,\* CAREL P. VAN SCHAİK,\* BENOIT GOOSSENS,†‡§ IAN SINGLETON,¶\*\* ERNST J. VERSCHOOR,†† KRISTIN S. WARREN‡‡ and MICHAEL KRÜTZEN\*

\*Anthropological Institute & Museum, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland, †Organisms and Environment Division, School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK, ‡c/o Sabah Wildlife Department, Danau Girang Field Centre, Kota Kinabalu, Sabah, Malaysia, §Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia, ¶Foundation for a Sustainable Ecosystem (YEL), Jl. Wahid Hasyim No. 51/74, Medan 20154, Indonesia, \*\*Sumatran Orangutan Conservation Programme (PanEco-YEL), Jl. Wahid Hasyim No. 51/74, Medan 20154, Indonesia, ††Department of Virology, Biomedical Primate Research Centre, PO Box 3306, 2280 GH, Rijswijk, The Netherlands, ‡‡College of Veterinary Medicine, School of Veterinary and Life Sciences, Murdoch University, 90 South Street, Murdoch, WA 6150, Australia

## Abstract

Investigating how different evolutionary forces have shaped patterns of DNA variation within and among species requires detailed knowledge of their demographic history. Orang-utans, whose distribution is currently restricted to the South-East Asian islands of Borneo (*Pongo pygmaeus*) and Sumatra (*Pongo abelii*), have likely experienced a complex demographic history, influenced by recurrent changes in climate and sea levels, volcanic activities and anthropogenic pressures. Using the most extensive sample set of wild orang-utans to date, we employed an Approximate Bayesian Computation (ABC) approach to test the fit of 12 different demographic scenarios to the observed patterns of variation in autosomal, X-chromosomal, mitochondrial and Y-chromosomal markers. In the best-fitting model, Sumatran orang-utans exhibit a deep split of populations north and south of Lake Toba, probably caused by multiple eruptions of the Toba volcano. In addition, we found signals for a strong decline in all Sumatran populations ~24 ka, probably associated with hunting by human colonizers. In contrast, Bornean orang-utans experienced a severe bottleneck ~135 ka, followed by a population expansion and substructuring starting ~82 ka, which we link to an expansion from a glacial refugium. We showed that orang-utans went through drastic changes in population size and connectedness, caused by recurrent contraction and expansion of rainforest habitat during Pleistocene glaciations and probably hunting by early humans. Our findings emphasize the fact that important aspects of the evolutionary past of species with complex demographic histories might remain obscured when applying overly simplified models.

**Keywords:** Approximate Bayesian Computation, demographic history, *Pongo* spp., population structure

Received 24 April 2014; revision received 24 November 2014; accepted 27 November 2014

## Introduction

Patterns of DNA variation are the result of both adaptive and nonadaptive 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, under certain demographic scenarios, such as population size

Correspondence: Alexander Nater, Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 752 36, Uppsala Sweden.  
Fax: +46 18 4716310; E-mail: alexander.nater@ebc.uu.se

changes or population subdivision, random genetic drift can result in similar deviations as selection (e.g. Teshima *et al.* 2006; Excoffier *et al.* 2009). 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 genomewide 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 South-East Asian mainland (Rijksen & Meijaard 1999; Delgado & Van Schaik 2000). 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 recolonizations (Muir *et al.* 2000). Of special interest here is the Toba volcano on northern Sumatra, which has seen at least four major eruptions during the last 1.2 million years (Chesner *et al.* 1991). This sequence of eruptions culminated in the Toba supereruption ~73 ka, which is considered to be the most powerful volcanic eruption within the last 25 million years (Chesner *et al.* 1991) and is thought to have had severe consequences for flora and fauna on Sundaland (Williams *et al.* 2009). In the Late Pleistocene, all orang-utan populations on the mainland, southern Sumatra and Java went extinct (Rijksen & Meijaard 1999; Delgado & Van Schaik 2000). Climatic changes during the Pleistocene might have been responsible for the southward shift of the distribution and the

disappearance of orang-utans from the mainland (Jablonski 1998). Moreover, anthropogenic factors, such as prehistoric hunting by hunter-gatherer societies, are likely to have played a significant role in the decline and extinction of orang-utans populations on insular South-East Asia (Delgado & Van Schaik 2000).

Genetic signals of these past demographic changes have been found in studies of genetic diversity in extant orang-utan populations on Borneo and Sumatra. Most genetic studies analysing autosomal and mitochondrial DNA (mtDNA) agree that Sumatran orang-utans show a higher level of sequence diversity and corresponding long-term effective population size ( $N_e$ ) (Muir *et al.* 2000; Zhang *et al.* 2001; Steiper 2006; Locke *et al.* 2011; Prado-Martinez *et al.* 2013), even though Sumatran orang-utans have a much smaller current census size and a more restricted distribution than Borneans (~6600 vs. ~54 000 individuals, Wich *et al.* 2008). The large  $N_e$  of the Sumatran species was interpreted as a signal of immigration from multiple differentiated populations into the current Sumatran gene pool (Muir *et al.* 2000; Steiper 2006). However, Y-chromosomal diversity in orang-utans shows the opposite pattern compared to mtDNA and autosomal data, with a smaller  $N_e$  on Sumatra than Borneo (Nater *et al.* 2011). Such contrasting patterns of  $N_e$  between species and among genomic regions hint at complex population dynamics that have so far not been properly investigated.

Recently, Locke *et al.* (2011) used extensive single nucleotide polymorphism (SNP) data from whole-genome resequencing of five Bornean and five Sumatran orang-utans to model the demographic history of the two species. They found that a model with a population split ~400 ka with subsequent gene flow between Borneo and Sumatra fits the observed data best. Furthermore, Locke and colleagues inferred that Sumatran orang-utans underwent a continuous exponential population growth since the population split, while Bornean orang-utans were subject to a continuous exponential decline. Given the large amount of genetic data, the study by Locke and colleagues is currently regarded as the most accurate reconstruction of demographic history in orang-utans to date. However, the demographic modelling approach by Locke and colleagues did not take several idiosyncrasies of orang-utan biology into account, thus severely limiting the conclusions that could be drawn from their findings.

First, it has been shown that biased sampling and disregard of population structure will produce misleading results regarding  $N_e$  and its temporal changes (Stadler *et al.* 2009; Chikhi *et al.* 2010). The study by Locke and colleagues incorporated data from only five captive individuals from Borneo and Sumatra each without further provenance information. This limited genetic

sampling is unlikely to represent the entire genetic diversity present on both islands. Second, given the lack of detailed sample provenance, analyses were restricted to models treating Bornean and Sumatran orang-utans as single panmictic populations each. Previous work, however, unequivocally showed that both Bornean and Sumatran orang-utans are genetically deeply structured (Warren *et al.* 2001; Arora *et al.* 2010; Nater *et al.* 2011). In particular, on Sumatra, populations north and south of Lake Toba exhibit high genetic differentiation (Nater *et al.* 2011, 2013). Third, Locke and colleagues did not test complex demographic models including population bottlenecks or recent declines, as suggested in previous genetic studies. For example, genetic signals of a bottleneck with subsequent population expansion on Borneo might be linked to a glacial refugium or the impact of the Toba supereruption ~73 ka (Steiper 2006; Arora *et al.* 2010), and patterns of a recent population decline in Sabah, Borneo, are most likely attributable to recent anthropogenic pressures (Goossens *et al.* 2006).

Reconstructing the demographic history of a species has long been hindered by the fact that full-likelihood methods were restricted to relatively simple demographic models (e.g. Wilson *et al.* 2003; Hey & Nielsen 2004), which might not capture all relevant processes in complex demographic settings. This restriction is mainly caused by the fact that the computation of the likelihood function of complex demographic models with many parameters is either intractable or computationally too expensive, especially for large data sets (Marjoram *et al.* 2003). Approximate Bayesian Computation (ABC) allows circumventing these problems by approximating the likelihood functions with simulations of genetic data under a given demographic model (Beaumont *et al.* 2002; Marjoram *et al.* 2003). To estimate the model parameters, parameter values are drawn from predefined prior distributions and used to simulate genetic data matching the observed data in type of markers and number of loci. Both observed and simulated data are then reduced to a set of summary statistics, and the Euclidian distance between the observed and the simulated summary statistics is calculated. Based on the subset of simulations with the smallest Euclidian distance between observed and simulated data, the posterior distribution of the model parameters can be approximated and the relative fit of different demographic models to the data can be assessed.

Here, we present an ABC modelling approach of the demographic history of orang-utans based on autosomal and sex-linked marker systems. We aim to improve the current knowledge of demographic history by applying three major improvements over previous studies. First, we capitalize on the knowledge base of behavioural ecology and population genetics of

orang-utans to test realistic demographic models. Second, due to our extensive set of orang-utan samples with detailed and reliable provenance, we are able to investigate models incorporating population substructure in both orang-utan species, which allows us to disentangle changes in population size from confounding effects due to changes in population structure. Third, by combining autosomal and sex-linked markers into a single demographic analysis, we take advantage of the specific information content of different marker systems in this species with its heavily sex-biased dispersal. Due to strong female philopatry in orang-utans (Galdikas 1995; Arora *et al.* 2012; van Noordwijk *et al.* 2012), mitochondrial markers contain information about population split times without the confounding influence of gene flow. In contrast, Y-chromosomal loci should have more power than autosomal markers to reveal low levels of male-mediated gene flow.

## Materials and methods

### *Sample collection and genetic markers*

A representative sampling scheme covering the whole range of a species is crucial for accurate reconstruction of demographic history (Stadler *et al.* 2009). We used an extensive set of samples from wild-born orang-utans from 10 sampling locations, covering the entire distribution of the genus (Table 1 and Fig. 1, see Supporting information for detailed information about sample origin). Samples were analysed for several genetic marker systems with different modes of inheritance and effective population sizes (Table 2), thus ensuring representation of both male and female population history, an important aspect in demographic reconstructions in species with strongly sex-biased dispersal (Nater *et al.* 2011; Nietlisbach *et al.* 2012).

The autosomal microsatellite data contained genotypes of 25 microsatellite markers from a total of 237 individuals (Arora *et al.* 2010; Nater *et al.* 2013; Greminger *et al.* 2014). We also included sequences from three mtDNA genes with a total length of 1355 bp from 118 individuals (Nater *et al.* 2011), and Y-chromosomal haplotypes based on 11 Y-linked microsatellite loci from 129 individuals (Nater *et al.* 2011; Nietlisbach *et al.* 2012). We complemented the data set by additionally sequencing 8055 bp of the noncoding X-chromosomal region Xq13.3 (Kaessmann *et al.* 2001) in 36 individuals and four noncoding autosomal regions (Fischer *et al.* 2006) of a total of 8238 bp in 22 individuals. Basic summary statistics for all marker systems are provided in Table 2. The primers and cycling conditions used for PCR amplification and sequencing of the autosomal and X-chromosomal regions are

**Table 1** Sample sizes for the different marker systems in the 10 geographic regions

Sampling region*	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

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.

\*Sampling regions corresponding to Fig. 1.

described in the Supporting Table S1, Supporting information.

#### Approximate Bayesian Computation

**Model selection procedure.** We reconstructed the demographic history of orang-utans using an ABC approach implemented in the software package ABC-TOOLBOX v1.1 (Wegmann *et al.* 2010). To achieve this goal, we first performed a model selection procedure, in which we used a hierarchical approach to test a total of 12 different demographic models (Fig. 2) with increasing levels of complexity (see Tables S5 and S6, Supporting information, for more details about model parameterization and prior distributions).

We started by testing four relatively simple models assuming a single population for each of the two orang-utan species (Fig. 2A). The first model in this set (I2) posited constant population sizes and no migration between the two populations. The second model (IM2) incorporated asymmetric migration after the population split, up to a point in the past where migration between Borneo and Sumatra ceased. Gene flow in all models with migration was strictly male-mediated, as recent genetic and behavioural findings showed extreme female philopatric tendencies in orang-utans (Nater *et al.* 2011; Arora *et al.* 2012; van Noordwijk *et al.* 2012). The third model (IM2-GR) additionally allowed the two populations to change size exponentially after the population split and corresponded largely to the favoured model in the genomic study by Locke *et al.* (2011). In the fourth and most complex 2-population model (IM2-BN-GR), both populations retained a constant size after the population split, with the possibility for a sudden population size rescale followed by exponential growth or decline.

To test more biologically relevant demographic scenarios, we designed a series of 10-population models incorporating the repeatedly reported extensive population substructure in extant orang-utan populations (Warren *et al.* 2001; Goossens *et al.* 2005; Kanthaswamy *et al.* 2006; Arora *et al.* 2010; Nater *et al.* 2011, 2013). The use of 10 extant population units models is justified by previously published data (Arora *et al.* 2010; Nater *et al.* 2011, 2013; Greminger *et al.* 2014). The combination of patterns of population differentiation in both mtDNA and autosomal microsatellite markers points to six populations on Borneo, one Sumatran population south of Lake Toba and three Sumatran populations north of Lake Toba (see validation of population units in Supporting information). For all 10-population models, we assumed equal population sizes and equal symmetric migration rates among all populations within Borneo and among all populations north of Lake Toba, respectively, as well as a separate population size parameter for the population south of Lake Toba. We included asymmetric migration rates between Borneo and south of Lake Toba, and between north of Lake Toba and south of Lake Toba.

To assess to what extent the additional population units improve model fit, we first tested the best-fitting 2-population model against two basic 10-population models (IM10 and IM10<sub>BO-NT</sub>, Fig. 2B). The IM10 model incorporated the population splitting sequence derived from mtDNA data, that is the populations north and south of Lake Toba show the oldest split, while Bornean populations diverged after this split (Nater *et al.* 2011). As this is in discordance with the current species designation (Groves 2001), which assigns a single species each to Sumatra and Borneo, we also tested this model against a model following the species split pattern (IM10<sub>BO-NT</sub>), that is with the oldest split between

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

Sequences	$L_{\text{Bases}}^*$	Group	$N_{\text{Ind}}^\dagger$	$N_{\text{Seg}}^\ddagger$	$\pi^§$	$\theta_W^\P$	$D^{**}$
mtDNA (16S, ND3, CYTB)	1355	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)	8238	Borneo	10	$19.50 \pm 4.56$	$0.0033 \pm 0.0011$	$0.0027 \pm 0.0006$	$0.68 \pm 0.66$
		South Toba	2	$13.50 \pm 7.79$	$0.0037 \pm 0.0020$	$0.0036 \pm 0.0020$	$0.08 \pm 0.49$
		North Toba	10	$28.75 \pm 4.66$	$0.0046 \pm 0.0012$	$0.0040 \pm 0.0006$	$0.51 \pm 0.52$
Xq13.3	8055	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	$N_{\text{Loci}}^{\dagger\dagger}$	Group	$N_{\text{Ind}}$	$N_A^{\ddagger\dagger}$	$H_O^{\S\S}$	$H_E^{\P\P}$	$G-W^{***}$
Autosomal STR	25	Borneo	127	$7.16 \pm 4.13$	$0.53 \pm 0.22$	$0.61 \pm 0.25$	$0.90 \pm 0.15$
		South Toba	9	$3.84 \pm 1.18$	$0.60 \pm 0.23$	$0.62 \pm 0.16$	$0.72 \pm 0.22$
		North Toba	101	$6.32 \pm 3.11$	$0.61 \pm 0.16$	$0.65 \pm 0.16$	$0.82 \pm 0.17$
Y-STR	11	Borneo	53	$3.18 \pm 2.48$	–	$0.31 \pm 0.33$	$0.90 \pm 0.14$
		South Toba	8	$1.27 \pm 0.65$	–	$0.08 \pm 0.19$	$0.88 \pm 0.18$
		North Toba	68	$1.91 \pm 1.64$	–	$0.12 \pm 0.24$	$0.91 \pm 0.17$

Statistics are provided as average and standard deviation for marker systems with multiple independent loci.

\*Sequence length in base pairs.

<sup>†</sup>Number of sampled individuals.

<sup>‡</sup>Number of segregating sites

<sup>§</sup>Nucleotide diversity

<sup>¶</sup>Watterson's  $\theta$  per base pair.

\*\*Tajima's D (Tajima 1989).

<sup>††</sup>Number of loci.

<sup>‡‡</sup>Number of alleles.

<sup>§§</sup>Observed heterozygosity.

<sup>¶¶</sup>Expected heterozygosity.

\*\*\*Garza–Williamson index (Garza & Williamson 2001).

Sumatra and Borneo, to see whether incomplete lineage sorting could be responsible for the particular phylogenetic pattern observed for mtDNA.

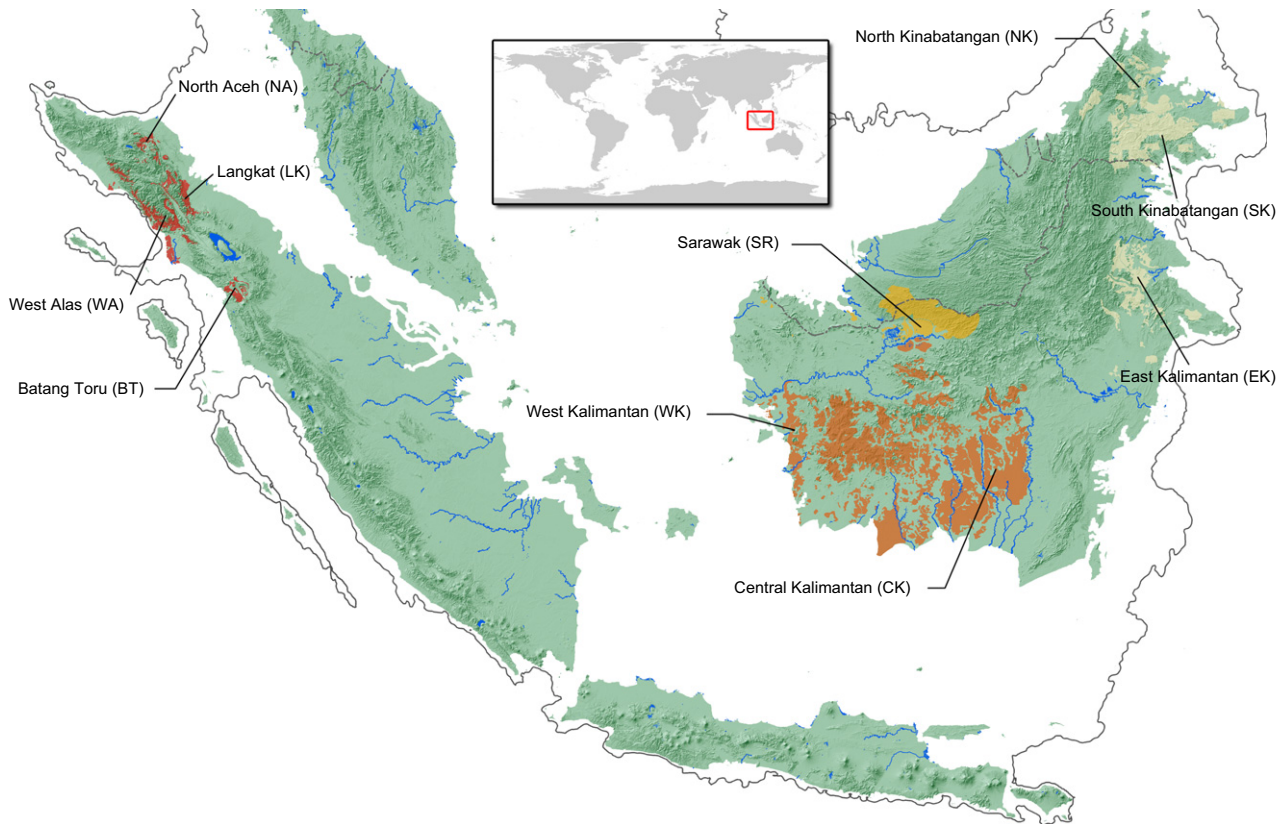
We further tested for the presence of population size changes in the demographic history of orang-utans, as suggested by previous studies (Goossens *et al.* 2006; Steiper 2006; Arora *et al.* 2010; Locke *et al.* 2011). First, we tested for signals of recent declines in Sumatra (IM10-DEC<sub>SU</sub>), Borneo (IM10-DEC<sub>BO</sub>) or both islands (IM10-DEC<sub>ALL</sub>) (Fig. 2C).

In a second test, we evaluated the support for a bottleneck on Borneo (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>), possibly linked to a refugium during the penultimate glaciation (Arora *et al.* 2010) (Fig. 2D).

Last, we tested for evidence for a bottleneck on Sumatra linked to the Toba supereruption, either allowing for a broad prior range of the magnitude of decline (IM10-BN<sub>BO</sub>-TOBA-DEC<sub>SU</sub>) or restricting to a severe bottleneck of <100 individuals in each of the four Sumatran populations (IM10-BN<sub>BO</sub>-RECOL-DEC<sub>SU</sub>), resembling a founder effect after local extinction and recolonization events on Sumatra (Fig. 2E).

**ABC data simulation.** To simulate genetic data under different demographic models, we used the software FASTSIMCOAL v1.1.2 (Excoffier & Foll 2011). Simulations for the different marker systems were run with the same set of parameters, whereby the effective population sizes were scaled 1 to 0.75, 0.25 and 0.25 for autosomal, X-chromosomal, mitochondrial and Y-chromosomal markers, respectively. We then used ARLSTAT v3.5.1.3 (Excoffier & Lischer 2010) to calculate a total of 259 summary statistics for each simulated data set as well as for the observed data set (Table S7, Supporting information). The summary statistics were chosen to capture the information in the genetic data about population differentiation, within population diversity, and population size changes. To avoid problems with unreliable phasing, we only used summary statistics that do not require phased sequence data for X-chromosomal and autosomal loci. As the number of simulated populations differed between the 2-population and 10-population models, summary statistic would not be directly comparable between the two sets of models. Therefore, when running the 10-population models, we applied a script





**Fig. 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 last glacial maximum (19–26 ka, 120 m below current sea level).

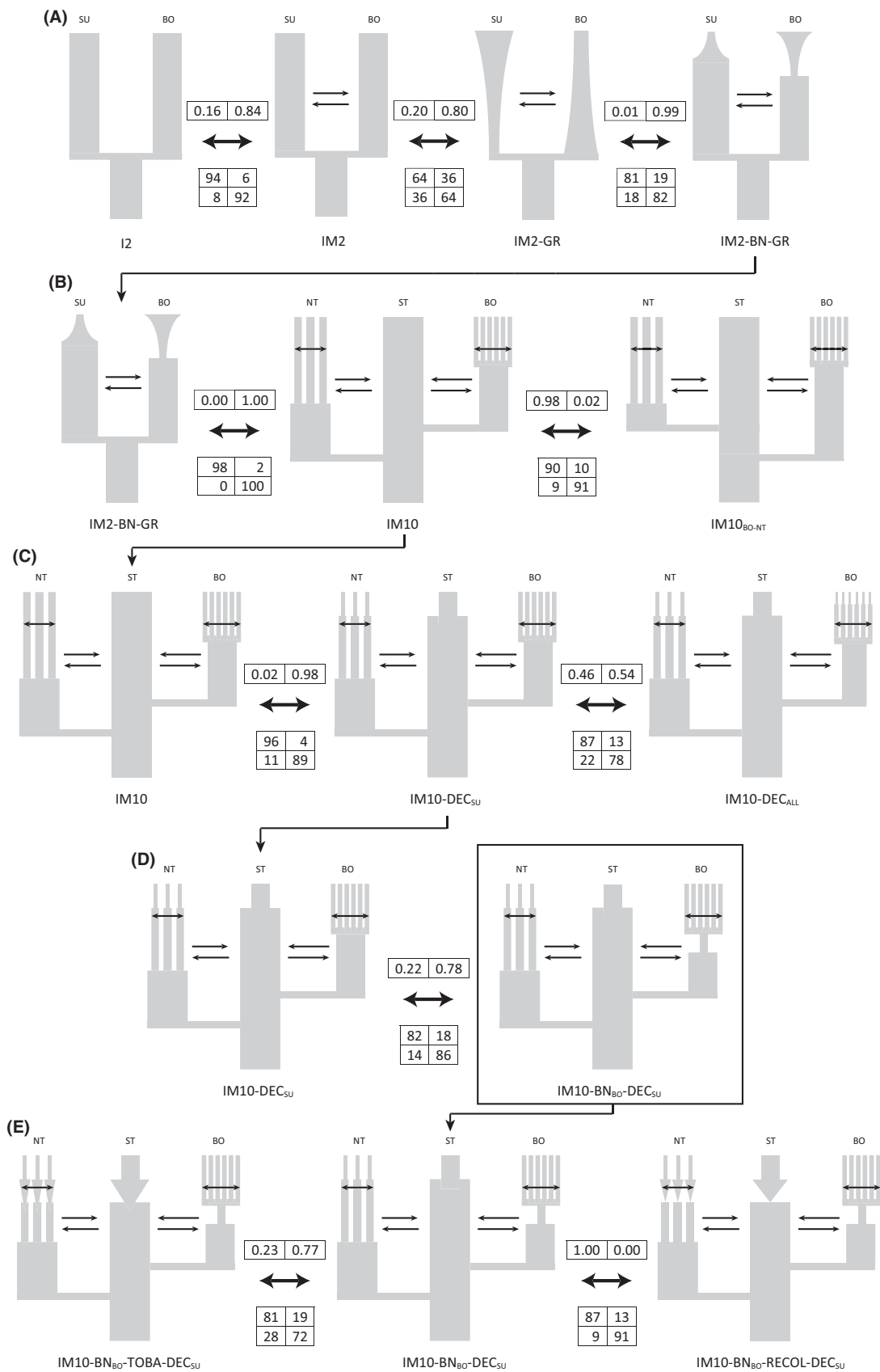
pooling the simulated data into a Bornean and a Sumatran group after each simulation step. Summary statistics were then also calculated islandwise, in order to be able to directly compare to the 2-population models.

We first performed an initial run of  $2 \times 10^6$  simulations with the standard rejection sampler (Tavare *et al.* 1997). These simulations were used for both model selection and validation. To reduce the dimensionality of the summary statistics, we performed a principal component analysis (PCA) with the 'prcomp' function in R version 2.12.1 (R Development Core Team 2010). We pooled and standardized 100 000 random simulations from each of the two compared models and used these summary statistics to extract the loadings of the first 10 principal components. We then transformed both the simulated and the observed data to perform a multinomial logistic regression with the R package 'abc' version 1.6. For this, we used the 0.1% of the simulations with the smallest Euclidean distance between the transformed summary statistics and the observed data.

To assess model fit, we also calculated the marginal densities and the probability of the observed data under

the general linear model (GLM) used for the postsampling regression for each model with ABCTOOLBOX (Leuenberger & Wegmann 2010). For this, we again transformed both the simulated summary statistics and the observed data with the loadings for the first 10 principal components. This time, PCA loadings were obtained for each model separately using 100 000 random simulations. The GLM was built from the 2000 simulations closest to the observed data, and we assessed the goodness of fit of all tested models to the observed data by calculating the *P*-value of the observed data under the GLM (Supporting Table S8, Supporting information). The *P*-value is representing the proportion of the retained simulations showing a lower or equal likelihood under the inferred GLM as compared to the observed genetic data (Wegmann *et al.* 2009b). Thus, low *P*-values indicate that the observed data is unlikely to have been generated under the inferred GLM, implying a bad model fit.

**Parameter estimation.** To obtain good estimates of the posterior distributions of the parameters for the best-fitting model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>), we used a MCMC



without likelihood method (Wegmann *et al.* 2009a). To reduce the dimensionality of the data and extract as much information as possible about the model parameters, we used the first 20 000 simulations with the standard sampler to define the first 12 orthogonal components of the summary statistics that maximize the covariance matrix between summary statistics and model parameters. For this, we applied a partial least-squares (PLS) regression approach (Boulesteix & Strimmer 2007) as implemented in the 'pls' R package (Mevik & Wehrens 2007) and used the R script provided in the ABCTOOLBOX package. We defined the optimal number of PLS components by assessing the drop in the root-mean-squared error for each parameter with the inclusion of additional PLS components. This way, a large set of summary statistics is reduced to a number of independent components, whereby summary statistics that are most informative about the model parameters are weighted more than summary statistics that do not show much response to changing parameter values (Wegmann *et al.* 2009a). The initial simulations were also used to define the tolerance distance based on a tolerance level of 0.1 and to calibrate the transition kernel of the MCMC run with a rangeProp setting of 1 unit of standard deviation (Wegmann *et al.* 2009a, 2010). We then ran a total of  $10^7$  iterations with the MCMC sampler, followed by a ABC-GLM postsampling regression (Leuenberger & Wegmann 2010) on the 10 000 simulations with the smallest Euclidean distance to the PLS components of the observed summary statistics. Finally, we used R to plot the posterior distributions of important model parameters.

**ABC validation.** The performance of ABC in model selection and parameter estimation in complex demographic settings inevitably suffers from the loss of information when the observed and simulated genetic data are reduced to a set of summary statistics (Robert *et al.* 2011). This necessitates a careful validation of the employed

ABC procedure to avoid biases in the approximation of posterior probabilities of evaluated models and the estimation of model parameters. Accordingly, we validated our model selection and parameter estimation approach with four different procedures. The first three validation approaches made use of so-called pseudo-observed data sets (*pods*), whereby parameter combinations are randomly drawn from the prior distributions and the corresponding summary statistics were simulated under a given model. These sets of summary statistics were then treated as if it were real observed data, but as the model and the corresponding parameter values that generated these summary statistics were known, we could use the *pods* to validate both our model selection and parameter estimation procedure.

In the first validation step, we investigated the model misclassification rate for each pairwise model comparison by generating 100 *pods* under each model with parameters randomly drawn from the prior distributions. We then performed the same model selection procedure as with the real observed data and counted the number of assignments to each model. We derived the model misclassification rate by counting all assignments of *pods* to a model other than the one generating it (Fig. 2).

Second, we assessed the accuracy of the parameter estimation, in terms of both different point estimators (mode, average and median) and over the whole posterior distribution under different tolerance levels (proportion of retained simulations). For this, we generated 1000 *pods* under the best-fitting model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>) and performed the same parameter estimation procedure on each *pods* as for the real data. The accuracy of the point estimators was assessed using the average of the root-mean-squared errors (RMSE) over all 1000 *pods* (Table S9, Supporting information), while the root-mean-integrated-squared error (Leuenberger & Wegmann 2010) was used to assess accuracy over the whole posterior distribution (Table S10, Supporting information). The results indicated that accuracy of the posterior distributions

**Fig. 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-DEC<sub>BO</sub> 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-BN<sub>BO</sub>-DEC<sub>SU</sub>) 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 in 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 <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.



is little affected by varying tolerance levels and that the mode of the distribution is the most accurate point estimator for parameter estimation.

Third, to increase confidence in the parameter estimates of the best-fitting model, we checked for biased posterior distributions by producing 1000 *pods* under the best-fitting model with parameter values drawn from the prior distributions. We used **ABCTOOLBOX** to calculate the posterior quantiles of the true parameter values within the estimated posterior distributions for each *pods* and used a Kolmogorov–Smirnov test for uniformity in **R** (Wegmann *et al.* 2009a). Significant deviation from uniformity after sequential Bonferroni correction (Rice 1989) would indicate biased posterior distributions (Cook *et al.* 2006). The distribution of posterior quantiles within which the true values of the *pods* fell did not significantly deviate from the expectation of uniformity for most parameters (Fig. S4, Supporting information). In most cases where the posterior quantiles were not distributed uniformly, data points were overrepresented in the centre of the histogram, indicating that the posterior distributions were estimated too conservatively.

In a last validation approach, we tested whether the best-fitting model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>) and the corresponding posterior distributions of the model parameters are able to reproduce the summary statistics of the observed data. For this, we randomly sampled 10 000 parameter sets from the inferred posterior distributions and used these to simulate genetic data under the best-fitting model. We then carried out a PCA transformation of the simulated data and plotted the first 16 principal components to check whether the transformed observed data fell within the distribution of the simulated data (Fig. S5, Supporting information). This was the case for all the first 16 principal components, suggesting that the best-fitting model and its inferred parameter estimates are well able to explain the observed data.

## Results

### Model selection

We tested 12 demographic models, evaluating the impact of multiple demographic processes on the current genetic makeup of orang-utan populations (Fig. 2). We first compared simple models that treated Bornean and Sumatran orang-utans as single populations, but differed in the opportunity for migration after the population split (IM2 vs. I2, Fig. 2A). We found substantial support for the model allowing migration after the split [IM2, Bayes factor, i.e. ratio of model posterior probabilities (BF) 5.18]. However, this simple isolation with migration model achieved only a

very poor fit to the observed data, as shown by the probability of the observed data under the GLM used for parameter estimation (GLM *P*-value) of 0.003, indicating that additional processes were involved in shaping the gene pool of orang-utans. Of all four 2-population models tested, we observed a very strong support for a model that allowed a sudden change in population size for both populations followed by exponential growth (IM2-BN-GR vs. I2, IM2, IM2-GR, BF 36.79). Still, this model did not achieve a good fit to the observed data, as evidenced by a *P*-value of the observed data under the GLM of only 0.017 (Table S8, Supporting information).

The poor model fit of all tested 2-population models can be explained by the extensive population substructure within the two orang-utan species (Warren *et al.* 2001; Kanthaswamy *et al.* 2006; Arora *et al.* 2010; Nater *et al.* 2011, 2013), which differs to a great extent for female- and male-mediated marker systems (Nater *et al.* 2011; Nietlisbach *et al.* 2012). Accordingly, the  $N_e$  for each marker system varies to a large degree and cannot be described accurately with just one population size parameter per island. In agreement with this notion, we found that a basic model with 10 current population units (IM10) achieved a better fit to the observed genetic data (GLM *P*-value 0.224) than all the 2-population models (Table S8, Supporting information), and also obtained much stronger statistical support when directly compared against the best 2-population model (IM10 vs. IM2-BN-GR, BF 830.21, Fig. 2B). However, in our case, a better fit of the 10-population model compared to the 2-population models was not unexpected, as part of the observed genetic data was used beforehand to derive the number of population units in the 10-population models. When we computed summary statistics for the IM10 model without pooling the genetic data for the Sumatran populations north and south of Lake Toba, the model fit was still poor (GLM *P*-value 0.019). To improve model fit, we first tested whether a population split sequence following the species designation fits the data better than the pattern suggested by mtDNA data (deepest split within Sumatran orang-utans north and south of Lake Toba). This was strongly rejected by the observed data (IM10 vs. IM10<sub>BO-NT</sub>, BF 45.45, Fig. 2B).

We then further tested for recent population declines in Sumatra (IM10-DEC<sub>SU</sub> vs. IM10, BF 57.03), on Borneo (IM10-DEC<sub>BO</sub> vs. IM10, BF 0.48) or in both islands (IM10-DEC<sub>ALL</sub> vs. IM10-DEC<sub>SU</sub>, BF 0.94, Fig. 2C). Incorporating a population decline in Sumatra considerably improved the model fit (GLM *P*-value 0.553).

Next, we tested a model incorporating a bottleneck on Borneo together with a recent decline in Sumatra (Fig. 2D), which revealed substantial support for a bot-

tleneck on Borneo (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub> vs. IM10-DEC<sub>SU</sub>, BF 3.60).

Finally, we evaluated the statistical support for a bottleneck on Sumatra associated with the Toba supereruption (Fig. 2E). We found substantial support against a bottleneck on Sumatra in general (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub> vs. IM10-BN<sub>BO</sub>-TOBA-DEC<sub>SU</sub>, BF 3.29) and overwhelming support against a severe bottleneck (less than 100 individuals per population) (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub> vs. IM10-BN<sub>BO</sub>-RECOL-DEC<sub>SU</sub>, BF 10 887.60).

After performing a series of hierarchical model selection steps, we were able to identify a demographic model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>) capable of reproducing the observed patterns of DNA variation in the two current orang-utan species. Therefore, this model is likely to capture the biologically most relevant processes in the demographic history of orang-utans.

### Parameter estimation

We estimated the model parameters for the selected 10-population model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>, Fig. 3) based on a total of 10 million simulations (Table 3, Fig. 4). The parameter estimates point to a current  $N_e$  of ~970 diploid individuals in each of the six Bornean populations. We found support for a bottleneck on Borneo starting ~135 ka and ending ~82 ka, during which  $N_e$  on Borneo was reduced from an ancestral  $N_e$  of ~17 000 individuals to ~2600 individuals. The bottleneck on Borneo was followed by population recovery and substructuring, with a current total  $N_e$  of all Bornean populations of ~6150 individuals.

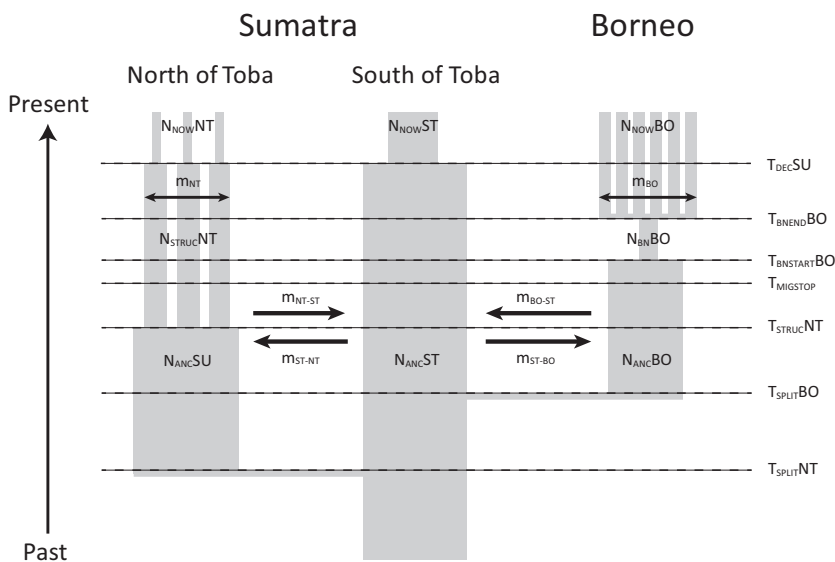
On Sumatra, the three populations north of Lake Toba suffered a decline ~24 ka from a  $N_e$  of ~10 500 to

currently only ~960 individuals in each of the three populations, corresponding to a total  $N_e$  in the metapopulation north of Lake Toba of ~38 300 and ~3300 individuals before and after the decline, respectively. We estimated that population structure north of Lake Toba was established ~860 ka, with an ancestral effective population size of ~14 400 individuals. The population south of Lake Toba also went through a recent decline ~24 ka from a  $N_e$  of ~24 200 individuals in the ancestral population to currently only ~1030 individuals. Thus, Sumatran orang-utan populations first expanded during the Middle Pleistocene before experiencing an islandwide population crash in the Late Pleistocene or Early Holocene.

We inferred the population split time between Borneo and south Toba as ~1.13 Ma, and between north and south of Lake Toba as ~3.39 Ma. Gene flow between Borneo and Sumatra appears to have ceased ~87 ka, but this parameter was associated with a broad posterior distribution. We found no evidence for asymmetric migration rates between Borneo and south of Lake Toba, and between south of Lake Toba and north of Lake Toba. The migration rates between the two islands were comparable to the migration rates over the Toba caldera on Sumatra, while migration rates among the populations on Borneo and among those north of Toba, respectively, were estimated to be about a magnitude higher.

### Discussion

Our modelling approach capitalized on the use of multiple genetic marker systems and an extensive set of geographically well-defined samples, in contrast to



**Fig. 3** Schematic representation of the selected 10-population model with a bottleneck on Borneo and recent population declines in all Sumatran populations (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>).

**Table 3** Estimates of the model parameters for the selected 10-population model with a bottleneck on Borneo and a recent decline in Sumatra (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>)

Parameter	Prior*	Mode	Mean	90%-HPD†
N <sub>NOW</sub> BO [ind] (6)	logunif[100;10 000]	974	1028	348–3011
N <sub>NOW</sub> NT [ind] (3)	logunif[100;10 000]	963	933	239–3613
N <sub>NOW</sub> ST [ind] (1)	logunif[100;10 000]	1034	952	189–4514
N <sub>BN</sub> BO [ind] (1)	logunif[100;10 000]	2598	1486	286–9988
N <sub>ANC</sub> BO [ind] (1)	logunif[1000;100 000]	17 046	12 344	2171–89 115
N <sub>STRUC</sub> NT [ind] (3)	logunif[1000;100 000]	10 508	11 278	1886–78 264
N <sub>ANC</sub> NT [ind] (1)	logunif[1000;100 000]	14 407	10 519	1565–70 259
N <sub>ANC</sub> ST [ind] (1)	logunif[1000;100 000]	24 193	13 991	2629–99 070
T <sub>BNEND</sub> BO [years]	unif[8750;400 000]	81 946	149 580	8848–283 785
T <sub>BNSTART</sub> BO [years]	T <sub>BNEND</sub> BO + unif[250;100 000]	135 076	191 001	20 855–348 145
T <sub>SPLIT</sub> BO [ka]	unif[400;1500]	1128	960	497–1436
T <sub>DEC</sub> SU [years]	unif[1.0;3.5]	23 651	36 200	4119–67 272
T <sub>STRUC</sub> NT [ka]	unif[75;1500]	861	820	267–1398
T <sub>SPLIT</sub> NT [ka]	unif[1500;4000]	3392	2995	2101–3999
T <sub>MIGSTOP</sub> [years]	unif[2.5;4.2]	87 034	161 862	8849–310 833
Log(m <sub>BO-ST</sub> ) [migrants/ind/gen]	unif[−5.0; −3.0]	−3.55	−3.96	−4.79 to −3.09
Log(m <sub>ST-BO</sub> ) [migrants/ind/gen]	unif[−5.0; −3.0]	−3.42	−3.84	−4.61 to −3.10
Log(m <sub>NT-ST</sub> ) [migrants/ind/gen]	unif[−5.0; −3.0]	−3.89	−3.98	−4.81 to −3.14
Log(m <sub>ST-NT</sub> ) [migrants/ind/gen]	unif[−5.0; −3.0]	−3.65	−3.92	−4.71 to −3.06
Log(m <sub>BO</sub> ) [migrants/ind/gen]	unif[−4.0; −2.0]	−2.52	−2.90	−3.66 to −2.02
Log(m <sub>NT</sub> ) [migrants/ind/gen]	unif[−4.0; −2.0]	−2.51	−2.89	−3.65 to −2.03

BO, Borneo, NT, Sumatra north of Lake Toba; ST, Sumatra south of Lake Toba; N<sub>NOW</sub>, current effective population size; N<sub>BN</sub>, effective population size during population bottleneck; N<sub>ANC</sub>, ancestral effective population size; N<sub>STRUC</sub>, effective population size before recent decline; T<sub>BNEND</sub>, time since population bottleneck ended; T<sub>BNSTART</sub>, time when population bottleneck started; T<sub>SPLIT</sub>, population split time; T<sub>DEC</sub>, time since population decline; T<sub>STRUC</sub>, time since establishment of population structure; T<sub>MIGSTOP</sub>, 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 Fig. 3), the number in parentheses next to the population size parameters refer to the number of simulated populations of this size each.

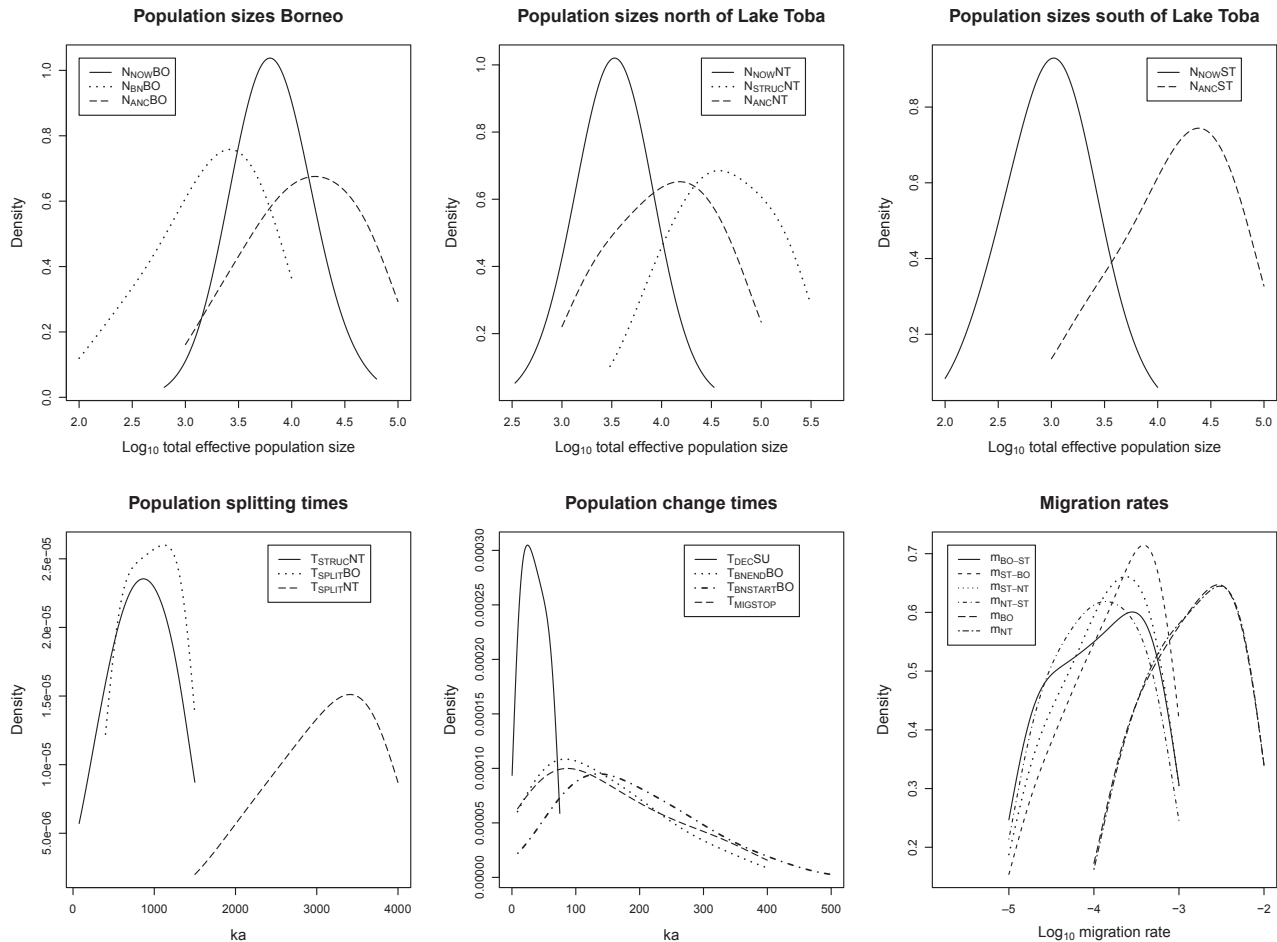
\*The prior distributions for the parameter values were either uniform or loguniform within the boundaries provided in squared brackets

†Ninety percent highest posterior density interval.

previous studies, which based their findings on a small number of captive individuals with poorly recorded provenance (Locke *et al.* 2011; Mailund *et al.* 2011, 2012). Thus, our study was able to shed light on important aspects of orang-utan demographic history that so far remained unexamined due to nonrepresentative sampling and dismissal of within-species population structure. For instance, the inferred model by Locke *et al.* (2011) of a continuously expanding Sumatran orang-utan population with a substantially larger current  $N_e$  as compared to Bornean orang-utans was unrealistic in the light of current species distribution and abundance and did not capture recent population dynamics. Our results indicate that such misleading signals are the result of a recent decline and deep divergence of orang-utan populations on Sumatra, which yields a larger long-term  $N_e$  for Sumatran orang-utans as compared to Bornean orang-utans in oversimplified demographic models.

### Inference of best-fitting model

We inferred that a model with comprehensive population structure, a bottleneck on Borneo and a recent decline in Sumatra (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>), fits the observed data significantly better than a range of simplified models that treat each orang-utan species as a single panmictic population. Estimation of demographic parameters under this model revealed a population split time between Borneo and Sumatran populations south of Lake Toba of just over a million years ago, followed by bidirectional gene flow. This species split time estimate is considerably older than estimates obtained using whole-genome data, suggesting a species split time of between 330 and 600 ka (Locke *et al.* 2011; Mailund *et al.* 2011, 2012). Such recent species split estimates are, however, in disagreement with findings based on mitochondrial DNA, which yielded divergence time estimates of island-specific mtDNA lineages of 1–5 Ma (Xu & Arnason 1996; Zhi *et al.* 1996;



**Fig. 4** Posterior distributions of important model parameter under the selected 10-population model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>). The abbreviations of the model parameters correspond to the labels in Fig. 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).

Warren *et al.* 2001; Zhang *et al.* 2001; Steiper 2006; Nater *et al.* 2011).

The discrepancy between model-based species split estimates using exclusively autosomal data and mtDNA divergence time estimates from phylogenetic methods is owed to two idiosyncrasies in the biology of orang-utans. First, due to the pronounced philopatric tendencies of female orang-utans (Galdikas 1995; Arora *et al.* 2012; Nietlisbach *et al.* 2012; van Noordwijk *et al.* 2012), mtDNA has likely experienced only little if any gene flow between the two species after the species split. Therefore, the coalescent time of island-specific mitochondrial lineages is expected to predate the population split between Borneo and Sumatra, depending on  $N_e$  in the ancestral population (Nichols 2001). Second, due to male-mediated gene flow, model-based approaches using solely autosomal data are likely to underestimate

species split times, as disentangling the contributions of migration and split time remains challenging (Hey & Nielsen 2004). The recent split time estimates from autosomal genomic data might reflect the end of an initial period of frequent, but strictly male-driven gene flow after the species split. Such complex temporal fluctuations in migration rates, as expected during glacial cycles for Sundaland species, are so far not properly addressed in any applied demographic model. Still, combining markers with different inheritance patterns as carried out in this study is likely to improve the estimates of both migration rates and split times in species with sex-biased dispersal such as orang-utans.

Our findings of recent gene flow between Bornean and Sumatran orang-utans are in agreement with previous observations (Muir *et al.* 2000; Verschoor *et al.* 2004; Becquet & Przeworski 2007). In their genomic study,

Locke *et al.* (2011) found an unexpectedly high incidence of low-frequency mutations shared between Borneo and Sumatra, which also hints at recent gene flow between the two islands. Contrary to studies indicating the presence of impassable dispersal barriers on the exposed Sunda shelf, due to either large river systems (Harrison *et al.* 2006) or a putative savannah corridor (Gathorne-Hardy *et al.* 2002; Bird *et al.* 2005), it seems that habitat conditions during glacial periods did at least sporadically allow male orang-utans to cross the exposed Sunda shelf. However, given the strict and long-lasting separation of mtDNA lineages on both islands (Nater *et al.* 2011), it appears that the exposed shelf was not covered with forest able to sustain orang-utan populations over prolonged periods. In fact, large parts of the Sunda shelf between Borneo and Sumatra were covered with nutrient-poor sandy soils (Bird *et al.* 2005; Slik *et al.* 2011). Forests on such soil types are characterized by low growth and productivity (Paoli *et al.* 2010). These constraints might explain why orang-utan populations on both islands could not expand onto the exposed shelf to an extent where population admixture and thus exchange of mtDNA lineages was possible.

#### *Glacial cycles and population size changes*

As 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 (Warren *et al.* 2001; Arora *et al.* 2010) 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 ~2500 individuals, which then recovered again to the current total  $N_e$  of ~6000 for all Bornean orang-utans. Such a change in both  $N_e$  and 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 (Morley 2000; Gathorne-Hardy *et al.* 2002; Bird *et al.* 2005). 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*). 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 ~2500 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 density 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 towards 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 in 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 South-East Asia (Gathorne-Hardy *et al.* 2002), including on Mentawai Island around 350 km from the Toba caldera (Gathorne-Hardy & Harcourt-Smith 2003), as well as the similar composition of South-East 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 South-East



Asia. However, the lack of bottleneck 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.* 2011, 2013), highlighted 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 explains the absence of a similar bottleneck in the Sumatran population history well. During the generally drier glacial periods, large parts of Sumatra experienced considerably more rain fall compared to Borneo (Whitten *et al.* 2000; Gathorne-Hardy *et al.* 2002), 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.

#### *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 towards an earlier decline in the Late Pleistocene or Early Holocene. In the Late Pleistocene, orang-utans went extinct on the South-East Asian mainland as well as in many Sundaland regions (Jablonski 1998; Rijksen & Meijaard 1999; Delgado & Van Schaik 2000). Furthermore, the Pleistocene–Holocene boundary is characterized by the disappearance of many large-bodied animals worldwide (Koch & Barnosky 2006), including large parts of the megafauna in South-East 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 towards 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 recognized 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 went recently through a strong decline, which, in combination with 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

We are indebted to Pirmin Nietlisbach, Nicole Ponta, Livia Gerber, Corinne Ackermann and Kai Ansmann for providing valuable laboratory work for this study. We thank Erik Willems for the Sundaland map. Laurentius N. Ambu, Maria A. van Noordwijk, Helen Morrogh-Bernard, Cheryl Knott, Noko Kuze, Tomoko Kanamori, Joko Pamungkas, Dyah Perwitasari-Farajallah and Muhammad Agil provided orang-utan samples that were analysed for this study or helped with administration in Indonesia and Malaysia. Special thanks goes to Daniel Wegmann for his technical help with the ABCtoolbox software. We are grateful to three anonymous reviewers for their valuable comments. This project was financially supported by the Swiss National Science Foundation (Grant no. 3100A-116848 to MK and CPvS), Forschungskredit of the University of Zurich (Grant no. 57020601 to MPG), Messerli Foundation, A.H.-Schultz Foundation and Claraz Schenkung. Furthermore, we thank the following institutions for supporting our research: Primate Research Center of the Bogor Agricultural University (IPB), Indonesian State Ministry of Research and Technology (RISTEK), Indonesian Institute of Sciences (LIPI), Sabah Wildlife Department, Taman Nasional Gunung Leuser (TNGL), Borneo Orangutan Survival Foundation (BOSF), Leuser International Foundation (LIF) and Badan Pengelola Kawasan Ekosistem Leuser (BPKEK).

## References

- Arora N, Nater A, van Schaik CP *et al.* (2010) Effects of Pleistocene glaciations and rivers on the population structure of Bornean orangutans (*Pongo pygmaeus*). *Proceedings of the National Academy of Sciences*, **107**, 21376–21381.
- Arora N, van Noordwijk MA, Ackermann C *et al.* (2012) Parentage-based pedigree reconstruction reveals female matrilineal clusters and male-biased dispersal in nongregarious Asian great apes, the Bornean orang-utans (*Pongo pygmaeus*). *Molecular Ecology*, **21**, 3352–3362.
- Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.
- Becquet C, Przeworski M (2007) A new approach to estimate parameters of speciation models with application to apes. *Genome Research*, **17**, 1505–1519.
- Bird MI, Taylor D, Hunt C (2005) Environments of insular Southeast Asia during the last glacial period: a savanna corridor in Sundaland? *Quaternary Science Reviews*, **24**, 2228–2242.
- Boulesteix AL, Strimmer K (2007) Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Briefings in Bioinformatics*, **8**, 32–44.
- Chesner CA, Rose WI, Deino A, Drake R, Westgate JA (1991) Eruptive history of Earth's largest Quaternary caldera (Toba, Indonesia) clarified. *Geology*, **19**, 200–203.
- Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA (2010) The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics*, **186**, 983–995.
- Clark PU, Dyke AS, Shakun JD *et al.* (2009) The last glacial maximum. *Science*, **325**, 710–714.
- Cook SR, Gelman A, Rubin DB (2006) Validation of software for Bayesian models using posterior quantiles. *Journal of Computational and Graphical Statistics*, **15**, 675–692.
- Delgado RA, Van Schaik CP (2000) The behavioral ecology and conservation of the orangutan (*Pongo pygmaeus*): a tale of two islands. *Evolutionary Anthropology*, **9**, 201–218.
- Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, **29**, 51–63.
- Excoffier L, Foll M (2011) FASTSIMCOAL: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, **27**, 1332–1334.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
- Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.
- Fischer A, Pollack J, Thalmann O, Nickel B, Paabo S (2006) Demographic history and genetic differentiation in apes. *Current Biology*, **16**, 1133–1138.
- Flenley JR (1998) Tropical forests under the climates of the last 30,000 years. *Climatic Change*, **39**, 177–197.
- Galdikas BMF (1995) Social and reproductive behavior of wild adolescent female orangutans. In: *The Neglected Ape* (eds Nader RD, Galdikas BFM, Sheeran LK, Rosen N), pp. 163–182. Plenum Press, New York.
- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.
- Gathorne-Hardy FJ, Harcourt-Smith WEH (2003) The supereruption of Toba, did it cause a human bottleneck? *Journal of Human Evolution*, **45**, 227–230.
- Gathorne-Hardy FJ, Syaukani Davies RG, Eggleton P, Jones DT (2002) Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. *Biological Journal of the Linnean Society*, **75**, 453–466.
- Goossens B, Chikhi L, Jalil MF *et al.* (2005) Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (*Pongo pygmaeus*) populations from Sabah, Malaysia. *Molecular Ecology*, **14**, 441–456.
- Goossens B, Chikhi L, Ancenaz M *et al.* (2006) Genetic signature of anthropogenic population collapse in orang-utans. *PLoS Biology*, **4**, 285–291.
- Greminger MP, Stölting KN, Nater A *et al.* (2014) Generation of SNP datasets for orangutan population genomics using improved reduced-representation sequencing and direct comparisons of SNP calling algorithms. *BMC Genomics*, **15**, 16. doi:10.1186/1471-2164-15-16.

- Groves CP (2001) *Primate Taxonomy*. Smithsonian Institution Press, Washington, District of Columbia; London.
- Haddrill PR, Thornton KR, Charlesworth B, Andolfatto P (2005) Multilocus patterns of nucleotide variability and the demographic and selection history of *Drosophila melanogaster* populations. *Genome Research*, **15**, 790–799.
- Hahn MW (2008) Toward a selection theory of molecular evolution. *Evolution*, **62**, 255–265.
- Harrison T, Krigbaum J, Manser J (2006) Primate biogeography and ecology on the Sunda shelf islands: a paleontological and zooarchaeological perspective. In: *Primate Biogeography* (eds Lehman SM, Fleagle JG), pp. 331–372. Springer, New York, USA.
- Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.
- Hill C, Soares P, Mormina M *et al.* (2007) A mitochondrial stratigraphy for island southeast Asia. *American Journal of Human Genetics*, **80**, 29–43.
- Jablonski NG (1998) The response of catarrhine primates to Pleistocene environmental fluctuations in East Asia. *Primates*, **39**, 29–37.
- Kaessmann H, Wiebe V, Weiss G, Pääbo S (2001) Great ape DNA sequences reveal a reduced diversity and an expansion in humans. *Nature Genetics*, **27**, 155–156.
- Kanhaswamy S, Kurushima JD, Smith DG (2006) Inferring *Pongo* conservation units: a perspective based on microsatellite and mitochondrial DNA analyses. *Primates*, **47**, 310–321.
- Koch PL, Barnosky AD (2006) Late quaternary extinctions: state of the debate. *Annual Review of Ecology Evolution and Systematics*, **37**, 215–250.
- Leuenberger C, Wegmann D (2010) Bayesian computation and model selection without likelihoods. *Genetics*, **184**, 243–252.
- Li J, Li H, Jakobsson M *et al.* (2012) Joint analysis of demography and selection in population genetics: where do we stand and where could we go? *Molecular Ecology*, **21**, 28–44.
- Locke DP, Hillier LW, Warren WC *et al.* (2011) Comparative and demographic analysis of orang-utan genomes. *Nature*, **469**, 529–533.
- Louys J (2007) Limited effect of the Quaternary's largest supereruption (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**, e1001319. doi:10.1371/journal.pgen.1001319.
- 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**, e100312. doi:10.1371/journal.pgen.1003125.
- Marjoram P, Molitor J, Plagnol V, Tavaré 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**, 1–28.
- 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, 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.
- 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.
- 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.
- van Noordwijk MA, Arora N, Willems EP *et al.* (2012) Female philopatry and its social benefits among Bornean orangutans. *Behavioral Ecology and Sociobiology*, **66**, 823–834.
- Paoli GD, Wells PL, Meijaard E *et al.* (2010) Biodiversity conservation in the REDD. *Carbon Balance and Management*, **5**, 7.
- Prado-Martinez J, Sudmant PH, Kidd JM *et al.* (2013) Great ape genetic diversity and population history. *Nature*, **499**, 471–475.
- R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- 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), pp. 71–82. 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*. Kluwer Academic Publishers, Dordrecht, the 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.
- Sharma R, Arora N, Goossens B *et al.* (2012) Effective population size dynamics and the demographic collapse of Bornean orang-utans. *PLoS ONE*, **7**, e49429. doi:10.1371/journal.pone.0049429.



- Slik JWF, Aiba SI, Bastian M *et al.* (2011) Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of Southeast Asia. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 12343–12347.
- Soares P, Trejaut JA, Loo JH *et al.* (2008) Climate change and postglacial human dispersals in Southeast Asia. *Molecular Biology and Evolution*, **25**, 1209–1218.
- Stadler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009) The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics*, **182**, 205–216.
- Stajich JE, Hahn MW (2005) Disentangling the effects of demography and selection in human history. *Molecular Biology and Evolution*, **22**, 63–73.
- Steiper ME (2006) Population history, biogeography, and taxonomy of orangutans (Genus: *Pongo*) based on a population genetic meta-analysis of multiple loci. *Journal of Human Evolution*, **50**, 509–522.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Tavare S, Balding DJ, Griffiths RC, Donnelly P (1997) Inferring coalescence times from DNA sequence data. *Genetics*, **145**, 505–518.
- Teshima KM, Coop G, Przeworski M (2006) How reliable are empirical genomic scans for selective sweeps? *Genome Research*, **16**, 702–712.
- Thalmann O, Wegmann D, Spitzner M *et al.* (2011) Historical sampling reveals dramatic demographic changes in western gorilla populations. *BMC Evolutionary Biology*, **11**, 85. doi:10.1186/1471-2148-11-85.
- Verschoor EJ, Langenhuijzen S, Bontjer I *et al.* (2004) The phylogeography of orangutan foamy viruses supports the theory of ancient repopulation of Sumatra. *Journal of Virology*, **78**, 12712–12716.
- Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153–1167.
- Warren KS, Verschoor EJ, Langenhuijzen S *et al.* (2001) Speciation and intrasubspecific variation of Bornean orangutans, *Pongo pygmaeus pygmaeus*. *Molecular Biology and Evolution*, **18**, 472–480.
- Wegmann D, Leuenberger C, Excoffier L (2009a) Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, **182**, 1207–1218.
- Wegmann D, Leuenberger C, Excoffier L (2009b) Using ABCTOOLBOX. [http://cmpg.unibe.ch/software/ABCToolbox/ABCToolbox\\_manual.pdf](http://cmpg.unibe.ch/software/ABCToolbox/ABCToolbox_manual.pdf).
- Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCTOOLBOX: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, **11**, 116. doi:10.1186/1471-2105-11-116.
- Whitten T, Damanik SJ, Anwar J, Hisyam N (2000) *The Ecology of Sumatra*. Periplus Editions Ltd., Hong Kong.
- Wich SA, Meijaard E, Marshall AJ *et al.* (2008) Distribution and conservation status of the orang-utan (*Pongo* spp.) on Borneo and Sumatra: how many remain? *Oryx*, **42**, 329–339.
- Wich SA, deVries H, Ancrenaz M *et al.* (2009) Orangutan life history variation. In: *Orangutans: Geographic Variation in Behavioral Ecology and Conservation* (eds Wich SA, Utami Atmoko SS, Mitra Setia T, van Schaik CP), pp. 65–75. Oxford University Press, Oxford.
- Williams MAJ, Ambrose SH, van der Kaars S *et al.* (2009) Environmental impact of the 73 ka Toba super-eruption in South Asia. *Palaeogeography Palaeoclimatology Palaeoecology*, **284**, 295–314.
- Wilson IJ, Weale ME, Balding DJ (2003) Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. *Journal of the Royal Statistical Society Series A-Statistics in Society*, **166**, 155–188.
- Xu XF, Arnason U (1996) The mitochondrial DNA molecule of Sumatran orangutan and a molecular proposal for two (Bornean and Sumatran) species of orangutan. *Journal of Molecular Evolution*, **43**, 431–437.
- Zhang YW, Ryder OA, Zhang YP (2001) Genetic divergence of orangutan subspecies (*Pongo pygmaeus*). *Journal of Molecular Evolution*, **52**, 516–526.
- Zhi L, Karesh WB, Janczewski DN *et al.* (1996) Genomic differentiation among natural populations of orang-utan (*Pongo pygmaeus*). *Current Biology*, **6**, 1326–1336.

---

A.N., M.P.G., C.P.vS. and M.K. designed the study; B.G., I.S., E.J.V. and K.S.W. provided samples; A.N., M.P.G. and N.A. performed laboratory procedures; A.N. and M.P.G. conducted genetic data analysis; A.N. performed demographic modelling; A.N. wrote the manuscript; M.P.G., N.A., C.P.vS. and M.K. critically revised the manuscript and provided comments at all stages; B.G., E.J.V. and K.S.W. edited the final manuscript.

---

## Data accessibility

Sequence alignments, microsatellite genotypes, summary statistics, input files and custom-made scripts used in this study are available on Dryad: doi:10.5061/dryad.1jv55. Sequence data published previously are accessible under GenBank Accession nos HQ912716–HQ912752 (Table S2, Supporting information).

## Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Primers used for amplification and sequencing of four autosomal regions and one X-chromosomal region.

**Table S2** List of sequence loci used in the ABC analysis.

**Table S3** List of microsatellite loci used in the ABC analysis.

**Table S4** Mutation rate estimates of sequence loci.

**Table S5** Parameterisation and parameter prior distributions for all 2-population models.

**Table S6** Parameterisation and parameter prior distributions for all 10-population models.

**Table S7** Summary statistics used for Approximate Bayesian Computation.

**Table S8** Model fits of all tested demographic models.

**Table S9** Accuracy of different point estimators in parameter estimation.

**Table S10** Accuracy of parameter estimation under different tolerance levels.

**Fig. S1**  $\Pr(\text{Data} | K)$  and  $\Delta K$  statistics for all STRUCTURE runs.

**Fig. S2** Structure plot for 25 microsatellite markers used for the demographic modelling.

**Fig. S3** Gene trees based on sequence data of six different loci.

**Fig. S4** Cross validation of the parameter estimation.

**Fig. S5** First 16 principal components of the posterior predictive distribution for the selected model (IM10-BN<sub>BO</sub>-DEC<sub>SU</sub>).