On the demographic history of chimpanzees and some consequences of integrating population structure in chimpanzees and other great apes

Camille Steux, Clément Couloigner, Armando Arredondo, Willy Rodríguez, Olivier Mazet, Rémi Tournebize, Lounès Chikhi, Sanaga River

On the demographic history of chimpanzees and some consequences of integrating population structure in chimpanzees and other great apes.

Camille Steux[1, 2, *], Clément Couloigner[3, 4], Armando Arredondo[5, 6], Willy Rodríguez[5, 6], Olivier Mazet[5, 6], Rémi Tournebize[1, 3, 7], and Lounès Chikhi[1, 2, 3, *]

[1]Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Université de Toulouse, CNRS, IRD, Toulouse INP, Université Toulouse 3 – Paul Sabatier (UT3), Toulouse, France

[2]Centre for Ecology, Evolution and Environmental Changes (cE3c), Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Lisboa, Portugal

[3]Instituto Gulbenkian de Ciência, Oeiras, Portugal

[4]Centre d'Etude de la Forêt, Département de Biologie, Université Laval

[5]Université de Toulouse, Institut National des Sciences Appliquées, Institut de Mathématiques de Toulouse, Toulouse, France

[6]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse, Toulouse, France

[7]UMR DIADE - Diversité, adaptation, développement des plantes, CIRAD, CNRS, IRD, Université de Montpellier, France

[*]Authors for Correspondence: Camille Steux, camille.steux@univ-tlse3.fr ; Lounès Chikhi, lounes.chikhi@univ-tlse3.fr

Abstract

Reconstructing the evolutionary history of great apes is of particular importance for our understanding of the demographic history of humans. The reason for this is that modern humans and their hominin ancestors evolved in Africa and thus shared the continent with the ancestors of chimpanzees and gorillas. Common chimpanzees (*Pan troglodytes*) are our closest relatives with bonobos (*Pan paniscus*) and most of what we know about their evolutionary history comes from genetic and genomic studies. Most evolutionary studies of common chimpanzees have assumed that the four currently recognised subspecies can be modelled using simple tree models where each subspecies is panmictic and represented by one branch of the evolutionary tree. However, several studies have identified the existence of significant population structure, both within and between subspecies, with evidence of isolation-by-distance (IBD) patterns. This suggests that demographic models integrating population structure may be necessary to improve

our understanding of their evolutionary history. Here we propose to use *n*-island models within each subspecies to infer a demographic history integrating population structure and changes in connectivity (*i.e.* gene flow). For each subspecies, we use SNIF (structured non-stationary inference framework), a method developed to infer a piecewise stationary *n*-island model using PSMC (pairwise sequentially Markovian coalescent) curves as summary statistics. We then propose a general model integrating the four subspecies metapopulations within a phylogenetic tree. We find that this model correctly predicts estimates of within subspecies genetic diversity and differentiation, but overestimates genetic differentiation between subspecies as a consequence of the tree structure. We argue that spatial models integrating gene flow between subspecies should improve the prediction of between subspecies differentiation and IBD patterns. We also use a simple spatially structured model for bonobos and chimpanzees (without admixture) and find that it explains signals of admixture between the two species that have been reported and could thus be spurious. This may have implications for our understanding of the evolutionary history of the *Homo* genus.

Key words demographic history, population structure, chimpanzees, connectivity, fragmentation, human evolution

1 Introduction

Common chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*) are great apes found in western and central Africa, and they are the closest relatives to humans from which they diverged between 5 Mya [1, 2] and 7-8 Mya [3]. The current taxonomy of the genus *Pan* recognises bonobos as one unique species, geographically separated from common chimpanzees by the Congo river and from which it would have diverged between 0.9 and 2 Mya [2, 4, 5, 6]. Unlike bonobos, it is currently considered that common chimpanzees are further divided into four subspecies [2, 7]. Western chimpanzees, *P. t. verus*, occur in the most western part of the species geographic range, from Senegal on the west to Ghana on the east (see Fig. 1). The other three subspecies are separated from Western chimpanzees by the Dahomey gap, and their distribution ranges from Nigeria on the west to Tanzania on the east. From west to east, Nigeria-Cameroon chimpanzees (*P. t. ellioti*) are separated from Central chimpanzees (*P. t. troglodytes*) by the Sanaga river, and Eastern chimpanzees (*P. t. schweinfurthii*) are separated from Central chimpanzees by the Ubangi river (Fig. 1).

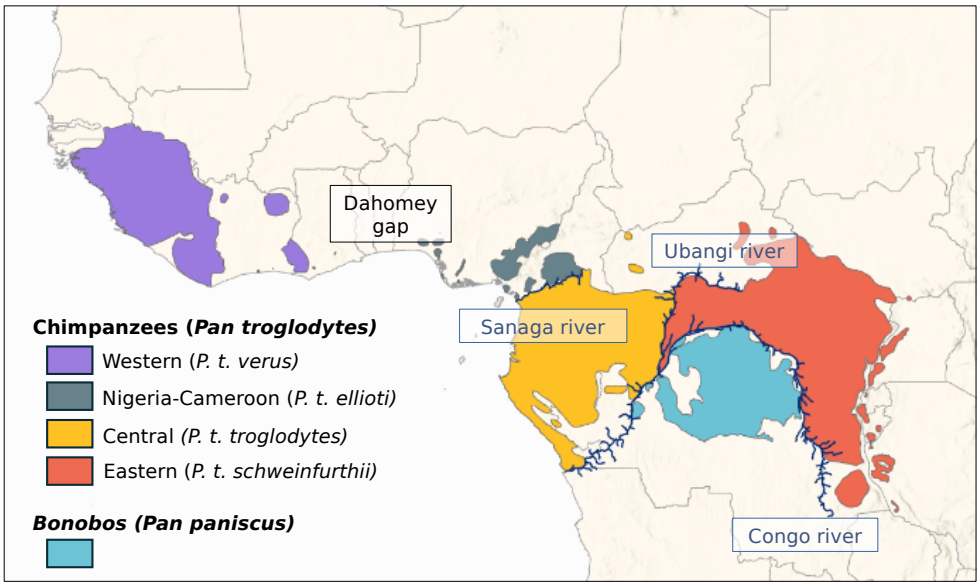


Figure 1: Distribution of the *Pan* genus. Data were extracted from the IUCN Red List of Threatened Species [8, 9].

Genetic and genomic analyses suggest that the four subspecies of common chimpanzees form two distinct monophyletic groups that split around 400-600 kya, with Western and Nigeria-Cameroon chimpanzees forming one clade and Central and Eastern chimpanzees forming the other clade [2, 10, 11, 12, 13, 4]. Sev-

eral studies estimate that Western and Nigeria-Cameroon chimpanzees probably split between 250-500 kya [2, 6], while Eastern and Central chimpanzees separated more recently, around 90-250 kya [2, 11, 12, 6]. The existence and the magnitude of gene flow between subspecies, today or in the past, remains however unclear [13, 12, 11, 2, 14]. Indeed, if the delimitation of the subspecies reflects mostly current geographic barriers (the Dahomey gap, the Sanaga and the Ubangi rivers), it is very likely that these barriers were more or less permeable in the past [15], and it has also been suggested that the ancestral population of common chimpanzees used to cover a wider and more continuous geographic range [16]. Furthermore, studies using models of isolation with migration (IM) have found signals of gene flow between subspecies, even though there is little consensus regarding the pairs of subspecies involved (see Figure 1 in Brand *et al.* [14] for a summary). For instance, Brand *et al.* [14] have identified introgressed segments from Western chimpanzees in Eastern chimpanzees, previously suggested by Hey [11], whereas Wegmann and Excoffier [13] found gene flow from Western to Central chimpanzees. Other studies have identified isolation-by-distance (IBD) patterns, where genetic distance increases with geographic distance [17, 18]. Taking the four subspecies as a single unit, Lester *et al.* [18] computed pairwise F_{ST} values between samples both within and between subspecies. They suggested that patterns of genetic diversity and differentiation could indeed correspond to a model with continuous gene flow across the whole *P. troglodytes* species geographic range, with the exceptions of a few highly isolated populations [18]. Patterns of IBD were also identified across Central and Eastern chimpanzees in an earlier study by Fünfstück *et al.* [17], questioning the classification of these two populations as distinct subspecies [17].

Additionally, several studies have pointed out the existence of structure within subspecies [17, 18, 19, 20, 21, 2, 6]. Central and Eastern chimpanzees could thus correspond to a set of populations that were connected by gene flow in the recent past, thus explaining the IBD pattern still detectable today [22, 17, 19]. In Nigeria-Cameroon chimpanzees, Mitchell *et al.* [20] identified two genetically distinct populations, one located in the forests of western Cameroon and another one in central Cameroon. Prado-Martinez *et al.* [2] also suggested the existence of substructure as they identified three Nigeria-Cameroon individuals who could belong to a distinct population than the rest of their sample. Finally, population structure has been identified as well in Western chimpanzees [21].

In parallel, much work has been done to reconstruct the demographic history of common chimpanzees using genetic and genomic data [2, 13, 7, 1, 11, 10, 21, 12, 23, 22, 19, 5]. Prado-Martinez *et al.* [2] used the Pairwise sequential Markovian coalescent (PSMC) method of Li and Durbin [24] on common chim-

panzee individual genomes to reconstruct what Prado-Martinez *et al.* [2] interpreted as a history of effective population size (N_e) changes through time. This method is particularly suited for endangered species, for which genomic data can be limited [25, 26, 27] because it only requires one single diploid genome. The interpretation of the PSMC curve is however not trivial [28, 29, 30, 31]. Indeed, whereas the interpretation of Prado-Martinez *et al.* [2] in terms of changes in N_e is potentially valid, several studies have shown that ignoring population structure can lead to the inference of spurious population size changes [32, 33]. In the case of the PSMC method, Mazet *et al.* [28] have shown that under structured population models, the PSMC curve will not only be influenced by changes in N_e, but also by population structure, and subsequently by changes in migration rates between populations [28, 29, 30, 31]. Given that population structure has been identified in common chimpanzees, both across the four subspecies and within subspecies, this means that there is currently no general model that would allow us to interpret the PSMC curves, while accounting for the observed patterns of IBD. Indeed, the current models of divergence represent the evolutionary history of the species and subspecies as successive splits of constant-size panmictic populations, which are incompatible with the PSMC curves. Altogether chimpanzees may represent interesting models to study ancient population structure and how it influences patterns of genomic diversity in present day populations. The current study benefits from work already done in humans but could also provide some interesting avenues of research for geneticists interested in ancient population structure in humans, by providing comparative data, and prompting similar work in other great apes.

Mazet *et al.* [28] introduced the IICR (inverse instantaneous coalescent rate), and showed that the PSMC is actually an estimate of the IICR and corresponds to changes in N_e under total panmixia but not necessarily under other demographic models. Mazet and colleagues also showed in several studies that the IICR can be characterized for any model of population structure under the coalescent, which opened the way to doing demographic inference using the PSMC as a summary statistic [29, 30, 34, 31]. To that purpose, Arredondo *et al.* [31] developed a method that allows to infer the parameters of a stepwise stationary n-island model [35] from a PSMC curve. For example, it allows to infer the number of islands or demes, n, their size N (in diploids), and the times, t_i, at which gene flow, $M_i = 4Nm_i$, may have changed by simply specifying a range of possible values for each one of these parameters. This method is implemented in SNIF (structured non-stationary inference framework).

In the present study, we ask whether it is possible to integrate population structure within each subspecies of common chimpanzees and infer a reasonable demographic history that explains the PSMC curves within

one single model for each subspecies and then for the species as a whole. First, we use SNIF [31] to infer non-stationary n-island models for each subspecies of common chimpanzees, assuming constant deme size, using the PSMC curves generated by Prado-Martinez *et al.* [2]. At each step we validate the inference steps by generating IICR curves from the inferred demographic models and by applying SNIF to the inferred IICR curves. From the resulting inferences, we then propose a model of demographic history for the four subspecies, integrating the n-island models and a tree model consistent with previous research. For all inferred models (for each subspecies and for the general model), we predict genetic diversity (nucleotide diversity) and differentiation (F_{ST}) both within and between subspecies and compare the predicted values to empirical estimates. We found that a model of structured populations with successive population splits and variable migration rates is sufficient to explain both the PSMC curves and several statistics of genomic diversity. We also find some discrepancies between the observed and predicted F_{ST} values between subspecies and use these to identify future directions for research. In particular we suggest that models incorporating spatial structure should be explored. As a proof of concept we use a simple example of stepping-stone model and show how signals interpreted as signatures of admixture events between chimpanzees and bonobos could actually be explained by population structure alone. These results are thus of great importance for the analysis of primate genomes in general and of humans in particular, where admixture events have been inferred, and for which population structure has also been invoked as a possible explanation [36, 37, 38].

2 Materials and Methods

2.1 Data: PSMC curves

The PSMC curves of the chimpanzees used here were retrieved from the study of Prado-Martinez *et al.* [2] who kindly shared the *.psmc* files. We only kept the PSMC curves that were computed on genomes with a coverage higher than 12X. In total, this corresponded to a total of 17 individuals, namely three Eastern chimpanzees (*P. t. schweinfurthii*), four Central chimpanzees (*P. t. troglodytes*), five Nigeria-Cameroon chimpanzees (*P. t. ellioti*) and five Western chimpanzees (*P. t. verus*). The PSMC files were used to reproduce the PSMC curves (Fig. S3).

2.2 Inference of demographic histories for each subspecies under non-stationary *n*-island models

We used SNIF (Structured Non-stationary Inferential Framework), a freely available program (`https://github.com/arredondos/snif`) based on a method developed by Arredondo *et al.* [31] to infer parameters of piecewise stationary n-island models. SNIF assumes that the number of demes (n) and their size (N diploids) are unknown and constant through time, whereas scaled migration rates ($M = 4Nm$, where m is the proportion of migrating genes at each generation) are allowed to vary over time in a piecewise manner. Note that throughout the whole paper, deme sizes are given in number of diploid individuals.

More specifically, SNIF assumes that the PSMC can be decomposed by dividing time into periods, called "components" (c), during which migration rate, M_i for component c_i, is constant. SNIF will infer the best timing (t_i) and duration ($t_{i+1} - t_i$) for these components to fit the observed PSMC/IICR curve for a given and fixed value of c provided by the user. To estimate the parameters of the model, SNIF minimizes a distance computed between the observed PSMC/IICR and the IICR simulated under the piecewise stationary n-island model (see Arredondo *et al.* [31] for details). The user must specify ω, a parameter that weights the computation of the distance between observed and simulated IICR curves by giving more weight to either recent or ancient times. The size of the parameter space explored by SNIF is defined by the user, who specifies a range of values for the parameters of the model, namely the number of islands n, their size N (in number of diploids), the scaled migration rates $M_i = 4Nm_i$ for each component c_i with $i \in \{0, ..., c-1\}$, and the times t_i (in generations) separating the components c_i and c_{i+1} (for instance t_1 separates the first component c_1 that starts at $t_0 = 0$ and the second component c_2 that ends at t_2). To scale and compare IICR curves to PSMC curves, a mutation rate (μ) and a generation time (g) are also required. The following values were used for the chimpanzee data: $\mu = 1.5 \times 10^{-8}$ per bp per generation [14] and $g = 25$ years [2].

To reduce computation time and improve consistency across runs, we first ran inferences using a wide parameter space to identify the range of parameter values that were more likely to produce IICR curves reasonably close to the observed PSMC. This allowed us to then re-run the analyses on a smaller parameter space, making the optimization algorithm more efficient for the same number of iterations. For this exploratory step, the following ranges were used: $n \in [2;100]$, $N \in [10;2 \times 10^4]$, $M_i \in [0.01;100]$, and $t_i \in [4 \times 10^2; 4 \times 10^5]$. We also identified values for c and ω that would best describe the observed PSMC curves. We tested $c \in \{4,5,6,7,8\}$ and $w \in \{0.5,1\}$ (1 being the default value). SNIF was run ten times for

each combination of c and ω value, and each run used 50 iteration steps of the optimization algorithm. The inference is expected to improve as c increases, since it allows the algorithm to add more changes in M and thus better fit the observed PSMC/IICR curve. We allowed c to vary between four and eight because the minimum number of components required to explain two humps in a PSMC is $c = 4$ [28] and because we considered that using more than eight components might lead to over-parametrization (see next section on the validation process).

This exploratory step allowed us to significantly reduce the parameter space as can be seen in Table 1, where the ranges for N and n were halved or nearly halved. We also found that ω=0.5 generated the best fit to the observed data and we identified the most likely time windows (t_i) for the changes of migration rates (see Table S1). The latter significantly reduced the time required by the optimization algorithm as the t_i and M_i values can vary over several orders of magnitude. In particular, the distance between the target and inferred IICR was on average smaller for the same number of iterations when we constrained the t_i values (based on preliminary runs) than when we did not (results not shown). We finally found that using seven components for Western, Nigeria-Cameroon and Eastern chimpanzees, and eight components for Central chimpanzees provided a good balance between model complexity and increase in fit to the observed PSMC (and validation, see next section). For Western chimpanzees, we further specified not to fit the very recent past (< 20 kya), because we noticed that SNIF would try to fit the most recent increase (forward in time) in the IICR that we wanted to ignore as it differs in magnitude between the individuals of this subspecies. Table 1 summarizes the parameter space used for the analysis of all individuals of the different subspecies and for the results shown further down. Using this final parameter space, SNIF was run ten times on each PSMC curve, each run used 50 iteration steps of the optimization algorithm.

Table 1: Parameter space used to run SNIF

Subspecies	Sample size	c	ω	n_{min}, n_{max}	N_{min}, N_{max}	$M_{i,min}, M_{i,max}$	t_i
Western	5	7	0.5	$2, 60$	$10, 10000$	$0.01, 50$	Constrained*
Nigeria-Cameroon	5	7	0.5	$2, 30$	$10, 10000$	$0.01, 50$	Constrained*
Central	4	8	0.5	$2, 60$	$10, 10000$	$0.01, 50$	Constrained*
Eastern	3	7	0.5	$2, 50$	$10, 10000$	$0.01, 50$	Constrained*

* see Supp. Table S1 for the explicit specified time windows

2.3 Validation of the inferred scenarios

Once we had inferred demographic scenarios, we performed a validation step as recommended by Arredondo *et al.* [31]. They suggested to simulate pseudo-observed data (POD) under an inferred scenario S^*, here in the form of IICR or PSMC curves, and to analyse these POD using SNIF with the same parameter ranges and c and ω values as those used to analyse the original observed PSMC curves. This procedure is somewhat similar to the validation process used in Approximate Bayesian Computation (ABC, [39]). In practice, SNIF allows the user to generate a *ms* command [40] to simulate coalescent times (T_2) under an inferred scenario S^*. This command is used by SNIF as an input to, first, produce a pseudo-observed IICR curve using simulated T_2 values, and second, infer the best stepwise stationary n-island model. This allows to quantify the discrepancy between the inferred model and the pseudo-observed underlying one. It is also possible to produce a *ms* command to simulate genomic data (by adapting the *ms* command using appropriate mutation and recombination rates and generation time) and run the PSMC method to produce a PSMC curve that can then be used as POD. Note that using genomic data instead of T_2 values is more time consuming (several hours to produce a pseudo-observed PSMC curve) and can only be applied to a limited number of simulated scenarios. Here, we performed the validation step using both simulated IICR and simulated PSMC, as explained below and in the Supplementary material.

Different individuals of the same subspecies exhibit PSMC curves that can differ more or less significantly in the recent past (see Figure S3 and Results). In addition, running SNIF on a particular PSMC curve can generate slightly different scenarios (see Results) generally characterized by similar connectivity graphs. Instead of trying to validate many similar scenarios, we arbitrarily chose an average scenario for each subspecies based on parameter values close to the median of the distribution of the inferred values. For instance, we found that the inferred values for n varied between 12 and 48 for Western chimpanzees, with 50% of inferred values between 17 and 31, and we thus selected a scenario with $n = 25$ and extracted the corresponding *ms* command. This *ms* command then served to produce as many independent IICR or PSMC curves as there were individuals in that subspecies. This allowed us to quantify the variation of inferred parameter values and compare it to the variation observed when analysing the empirical data. For each run, we simulated 10^6 T_2 values to produce an IICR curve, and simulated 10 genomic sequences of 100 Mb to produce PSMC curves, adapting the *ms* command using a mutation rate of 1.5×10^{-8} [14], a recombination rate of 0.7×10^{-8} per bp per generation [4], and a generation time of 25 years [2]. We produced

a (pseudo-observed) PSMC curve with PSMC [24] (flags -N 25 -t 15 -r 5 -p "4+25*2+4+6"). Altogether, these pseudo-observed IICR and PSMC curves were used by SNIF as POD, and the inference steps were done exactly as described in the previous section for the empirical PSMC curves.

2.4 Integration of the four subspecies *n*-island models in a general tree model

In the previous section, we inferred and validated demographic scenarios able to reproduce and fit the observed PSMC curves for each subspecies independently. Here we asked whether it was possible to integrate the four sub-subspecies into one unique demographic model that could explain the individual empirical PSMC curves, while incorporating a splitting tree based on the relationships between the subspecies as inferred by Prado-Martinez *et al.* [2] and the stepwise stationary *n*-island models within each branch of the tree, instead of assuming a panmictic population or subspecies. We constructed a scenario where an ancestral species is subdivided into n populations, and splits at time T_{CENW} into two branches which are themselves subdivided in demes (Fig. 7). One of these branches will later divides into a set of demes representing the ancestor of the Central chimpanzees and another set of demes corresponding to the Eastern chimpanzees' ancestors. The other ancestral branch becomes the ancestor to Western and Nigeria-Cameroon chimpanzees following a similar process. These splits thus generate the demes corresponding to the current four subspecies, at T_{CE} and T_{NW} for Central/Eastern chimpanzees and Western/Nigeria-Cameroon chimpanzees, respectively (see further, Figure 7).

SNIF cannot infer complex models involving both *n*-islands and tree models. Consequently, we built a general scenario manually, using the arbitrary "average" scenario used for the validation step above as a starting point. From the subspecies scenarios, we constructed a tree model where the subspecies *n*-island models merge (backward in time) in a way similar to that used by Rodríguez *et al.* [30]. For instance, the *n*-island models of Western and Nigeria-Cameroon chimpanzees were characterized by $n = 25$ and $n = 13$ islands respectively, and we thus merged the two sets of islands so as to use the largest n number for the ancestral species as suggested by Rodríguez *et al.* [30]. The same process was done for Central and Eastern chimpanzees, and then again for the ancestral branches when they merged into the most ancestral meta-population. Different values were tested for T_{CENW}, T_{NW} and T_{CE} to match the times at which the PSMC curves of the sub-populations were merging backward in time, namely $T_{CENW} \in \{900, 800, 700\}$, $T_{NW} \in \{900, 800, 700\}$ and $T_{CE} \in \{600, 500, 400\}$ in kya (thousands of years ago). For each scenario we generated the IICR plots using a script developed by W. Rodriguez [29]. We used the script to simulate 10^6

T_2 values with *ms*, sampling two haploid individuals in one deme, repeating the process for each of the four current meta-populations.

2.5 Prediction of genomic diversity and differentiation statistics

To test whether the general treee model (Figure 7) was able to predict genetic diversity and differentiation statistics in addition to the IICR, we simulated 100 segments of 1 Mb under the four subspecies models and under variations of the general model using *ms* [40], where we allowed for the splitting times (or joining times, with time going backward) to take several values (see previous section). We used a mutation rate of 1.5×10^{-8} per bp per generation [14] and a recombination rate of 0.7×10^{-8} per bp per generation [4]. We estimated genetic diversity by computing the individual observed heterozygosity (H_o) in 10 diploid individuals sampled in the present from one deme for each subspecies. We computed H_o as the number of heterozygous sites divided by the total length of the simulated genomes (100*1Mb). We also computed genetic differentiation (Hudson's F_{ST}) between demes of the same subspecies and between demes from different subspecies, sampling ten diploid individuals per deme. Pairwise F_{ST} were computed using original scripts from Tournebize and Chikhi [38].

Empirical values of genetic diversity were retrieved from Prado-Martinez *et al.* [2] (from Suppl. Table 12.4.1) and de Manuel *et al.* [6] (Table S4). Both studies used the same genomic data produced by Prado-Martinez *et al.* [2] and computed observed individual heterozygosity. We reported in Figure 9 their measures. Empirical values of F_{ST} were retrieved from Fischer *et al.* [22] (computed on autosomal sequences) and Lester *et al.* [18] (computed on microsatellites). Lester *et al.* [18] only published F'_{ST}, another estimate of genetic distance derived from F_{ST}, and they kindly shared with us the original F_{ST} values (Hudson's estimator).

3 Results

3.1 Independent demographic history of the different subspecies

We found that by using the parameter space described in Table 1 and applying SNIF to all individual PSMCs with high enough coverage within each subspecies, we were able to produce IICR curves that were similar to the observed PSMC plots, as displayed in Figure 2 for an example and Supplementary Figures S4, S5, S6, S7 for all inferences. The inferred parameters are displayed in Figure 3. Panels A and B show the distribution

of the inferred number of islands n and their size N (in number of diploid individuals), respectively (see also Table S2 and S3). The largest number of islands and the smallest deme size were inferred for Western chimpanzees, with a median n equal to 21 (50% of the inferred n values being between 17 to 31) and a median N equal to 305 (50% of the inferred N values between 239 and 335). The inferred number of islands was similar for Nigerian-Cameron and Eastern chimpanzees, with median values being 11 and 13, respectively and median N being 1154 and 800, respectively. Finally, for Central chimpanzees, 50% of the inferred n and N fall within 16-20 and 589-834, respectively.

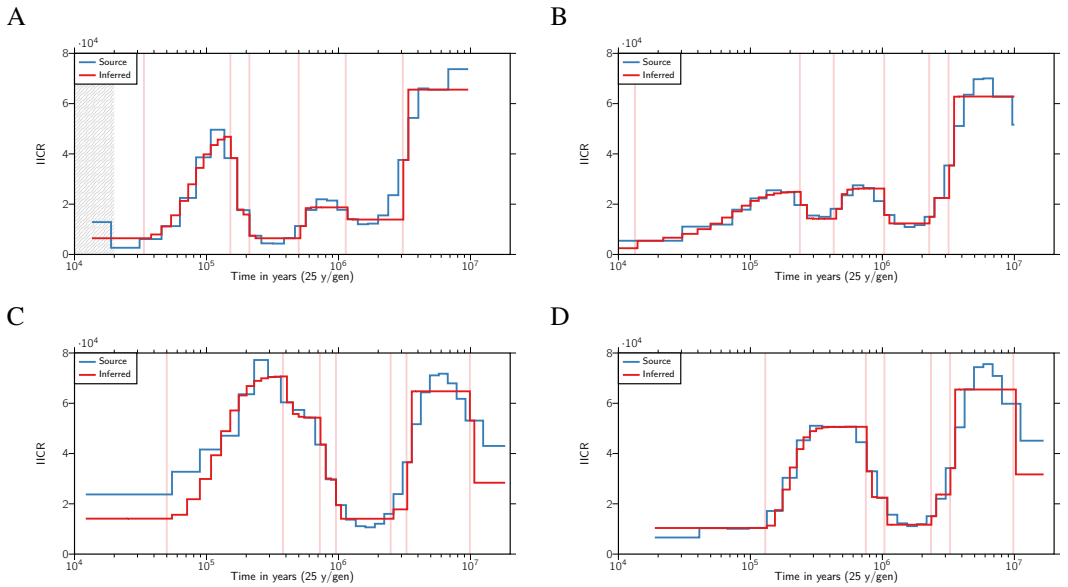


Figure 2: Inferred IICR curves and empirical PSMC curves for one individual per subspecies. The IICR curves are in red, and the empirical PSMC curves are in blue, and each panel corresponds to one individual from a different chimpanzee subspecies. Panel A . Western (Clint). Panel B. Nigerian-Cameroon (Damian). Panel C. Central (Vaillant). Panel D. Eastern (Kidongo). only one repetition of SNIF is displayed. See Supplementary Figures S4, S5, S6 and S7 for all the inferences. The vertical red lines highlight the times at which there is an inferred change in migration rate and therefore delimit the SNIF components. The grey zone in panel A corresponds to a part of the source PSMC which was not taken into account in the fitting of the curve by SNIF (see Material and Methods).

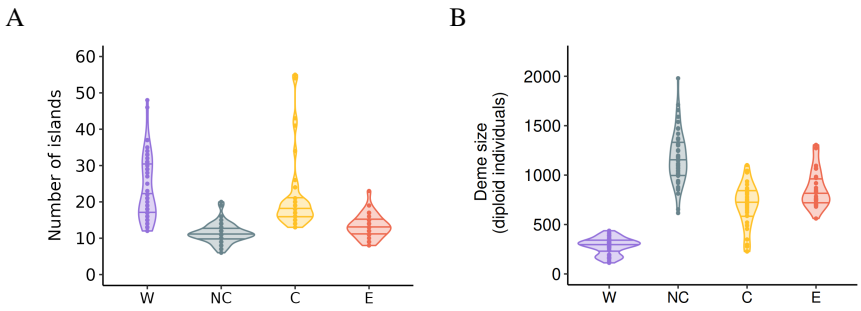


Figure 3: Distribution of the number and size of the demes inferred for the four chimpanzee subspecies. Panel A represents the inferred number of islands. Panel B represents the deme size (in number of diploid individuals). For both panels the results are plotted for each subspecies separately for Western (W), Nigeria-Cameroon (NC), Central (C) and Eastern (E) chimpanzees across the ten independent repetitions/inferences carried out per individual for all individuals, using the parameter space shown Table 1. Each dot corresponds to one repetition of SNIF done on one individual. The horizontal lines inside the violins correspond to the 25%, 50% (median) and 75% quantiles.

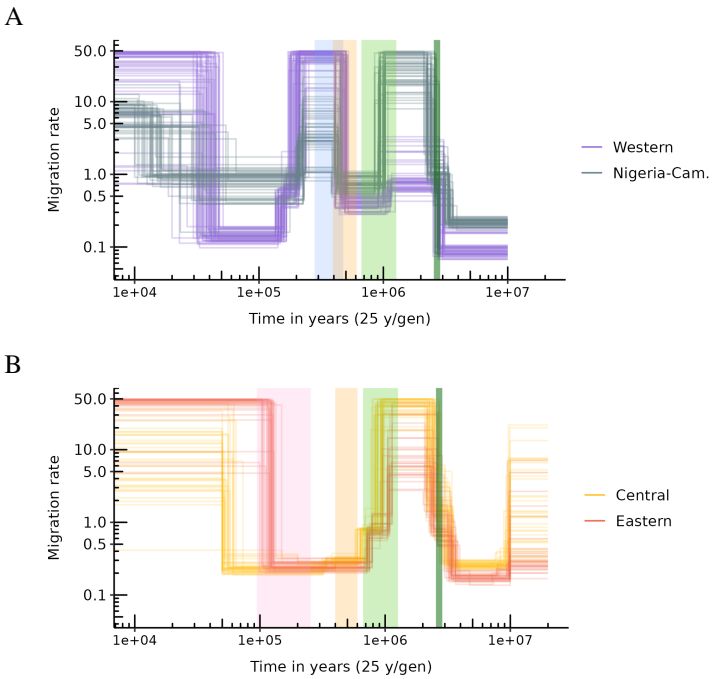


Figure 4: Connectivity graphs inferred by SNIF for the four chimpanzee subspecies. The y-axis represents scaled migration rates between demes (M_i), and the x-axis represents time in years on a logarithmic scale. Panel A shows connectivity for Western and Nigeria-Cameroon chimpanzees. Panel B shows the same for Central and Eastern chimpanzees. Each coloured curve corresponds to one inference (one repetition of SNIF) using one PSMC curve (or individual), giving therefore 50, 50, 40 and 30 curves in total for Western, Nigeria-Cameroon, Central and Eastern chimpanzees respectively. Backward in time, the vertical coloured intervals represent respectively: C-E divergence time (in pink, panel B), W-NC divergence time (in blue, panel A), C-E and W-NC ancestral populations (in yellow, both panels), the mid-Pleistocene transition (light green, both panels) and the Pliocene-Pleistocene transition (dark green, both panels).

The connectivity graphs (Figure 4) show the inferred migration rates ($M_i = 4Nm_i$) through time. For the four subspecies, we observe a significant increase in connectivity (forward in time) between 2 and 3 Mya, with a higher support for the period 2.5-3 Mya, followed by a decrease around 1 Mya. This period of increased connectivity is characterized by M_i values above 3 and up to 50 migrants per generation across the n-island for Nigeria-Cameroon, Central and Eastern chimpanzees, while values for Western chimpanzees are between 0.4 and 4. This period corresponds to a time when the four subspecies had a likely common ancestor. For Western and Nigeria-Cameroon chimpanzees, we observe a second more recent increase in connectivity between 500-600 kya and 200 kya, not observed in Central and Eastern chimpanzees, with M_i values ranging between 0.8 and 50. Finally, in the more recent past, all subspecies exhibit an increase in gene flow, occurring around 100-150 kya for Eastern chimpanzees, 50 kya for Central chimpanzees, 40 kya for Western chimpanzees and between 70 and 10 kya for Nigeria-Cameroon chimpanzees.

While the connectivity graphs are rather robust across individuals from the same subspecies and for more ancient periods for individuals from different subspecies, there is some variability in the inferred scenarios, as expected from the fact that different individuals exhibit different PSMC plots (see next section). We also observe variability in the number of islands or demes inferred for Western chimpanzees (extreme values range: 12-48 islands), and in the deme size inferred for Nigerian-Cameroon chimpanzees (extreme values range: 616-1980 diploid individuals). We also observe much variability in the most recent and most ancient parts of the connectivity graphs. For instance, there is a great variability in M_i values for Central chimpanzees in the last 50 kya and before 10 Mya, and in the M_i and t_i values for Nigeria-Cameroon chimpanzees in the last 1 Mya.

3.2 Validation step

We simulated T_2 values from an average scenarios identified for each subspecies and obtained IICR curves that were then analyzed using SNIF as a validation test of our inferential procedure (see Material and Methods). We were able to recover the original scenario with great precision, as shown in Figures 5 and 6. Black dots and lines are the pseudo-observed data, and the coloured patterns are the values inferred by SNIF. The inferred n and N are centered around the values that were used to generate the pseudo-observed IICRs, suggesting that SNIF is able to infer the complex scenarios inferred from the real data. We observe some variability in the inferences, as expected in any inference process and as we observed for the real data. Similarly, the inferred connectivity graphs were generally very good, with nearly no variability in the inferred t_i

values, and slightly more variability in the larger M_i values. The shape of the inferred connectivity graphs is however perfectly inferred for the complex scenarios that had seven or eight components.

When we simulated genomic data under inferred scenarios, generated PSMC curves and provided them to SNIF as POD (Supp. Figure S10 and S11), we also recovered the original scenarios with good precision, though SNIF inferred fewer and larger islands for Eastern chimpanzees, suggesting that we might be underestimating n and overestimating N for this particular subspecies. Inferred connectivity graphs were also generally good for the four subspecies, with higher variability in the inferred t_i and M_i values than for pseudo-observed IICRs computed on simulated T_2, which approaches the variability observed when running SNIF on empirical data. Altogether these results confirm the ability to infer a complex history of changes in connectivity under the stepwise stationary n-island model.

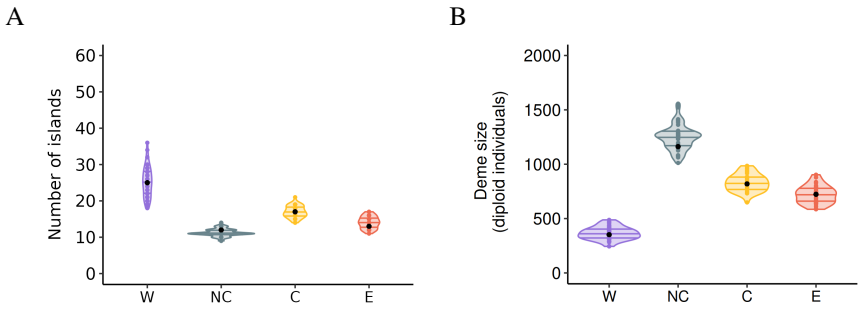


Figure 5: Distribution of the A. number of islands and B. deme size inferred by SNIF across the 10 repetitions per individual and all individuals for Western (W), Nigeria-Cameroon (NC), Central (C) and Eastern (E) common chimpanzees. Black dots are the pseudo-observed data.

3.3 General n-island model

The general tree model incorporating within its branches n-island models based on the inferences presented above for each subspecies is represented Figure 7. As explained in the Materials and Methods section, it was obtained by selecting for each subspecies an "average" inferred scenario for the different subspecies. The scenarios we kept had n = 25 islands of size N = 352 (diploid) individuals for Western chimpanzees, n = 12 islands of size N = 1162 individuals for Nigeria-Cameron chimpanzees, n = 17 islands of size N = 819 individuals for Central chimpanzees and n = 13 islands of size N = 723 individuals for Eastern chimpanzees. The sets of demes were then successively divided at times corresponding to the estimated split times for the pairs of subspecies. These split times are not known but can be approximated by using the times at which the different PSMC curves join. For instance, at time T_{CE}, the 17 demes of Central chimpanzees and

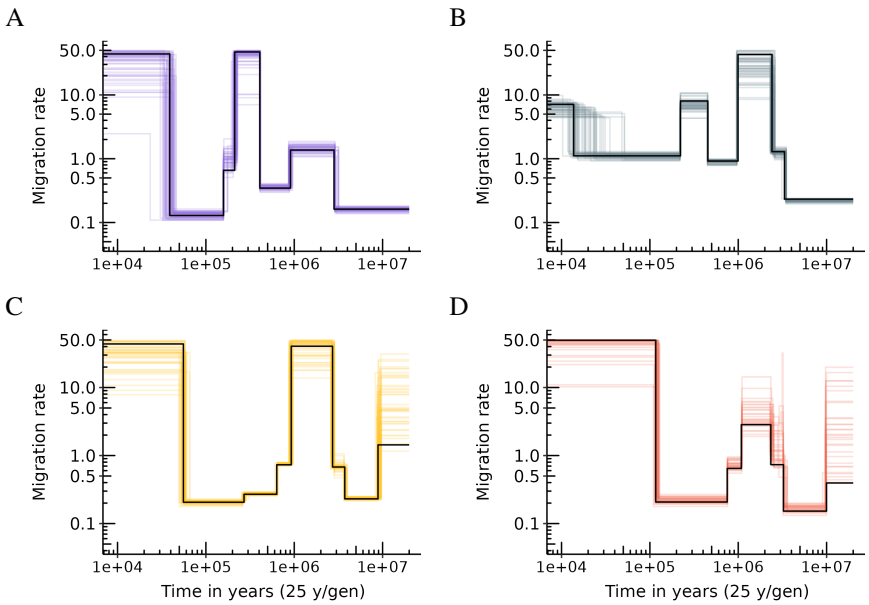


Figure 6: Inferred connectivity graph using a scenario inferred by SNIF as pseudo-observed data for each subspecies of common chimpanzee. A. Western chimpanzees, B. Nigeria-Cameroon chimpanzees, C. Central chimpanzees and D. Eastern chimpanzees. Black lines are the pseudo-observed data, and each colored line corresponds to one inference.

the 13 demes of Eastern chimpanzees are all assumed to derive (forward in time) from the demes of their ancestor. As noted in the Materials and Methods, this ancestor was assumed to have 17 demes as 17 is the largest of the two values of n. For simplicity, the 13 Eastern chimpanzee demes were assumed to derive from 13 demes rather than from all 17 demes of the ancestral meta-population. Similarly, at time T_{NW}, 12 demes of Nigeria-Cameroon chimpanzees join backward in time. 12 (out of 25) demes of the ancestor they share with Western chimpanzees. Finally, at time T_{CENW}, the 17 islands of the ancestral meta-population of Central and Eastern chimpanzees join 17 (out of 25) islands of the ancestral meta-population of Western and Nigeria-Cameroon chimpanzees. Thus, the most ancestral meta-population is represented by 25 demes.

Several values for T_{CE}, T_{NW}, and T_{CENW} were tested as it is unclear how closely splitting times of meta-populations correspond to splitting times of IICR or PSMC curves (see for instance Rodríguez *et al.* [30] and Chikhi *et al.* [29]). Empirical and simulated IICR curves presented in Figure 8 were obtained for T_{CE} = 600 kya, T_{NW} = 800 kya and T_{CENW} = 900 kya, which are the values that gave the best visual fit of the estimated IICR curve to the observed PSMC curves (in particular for Nigeria-Cameroon and Eastern chimpanzees, see below and Supp. Figure S12-S15). For each subspecies, the simulated IICR curve produced for a sample taken in a deme is represented and closely follows the corresponding observed PSMC. We found

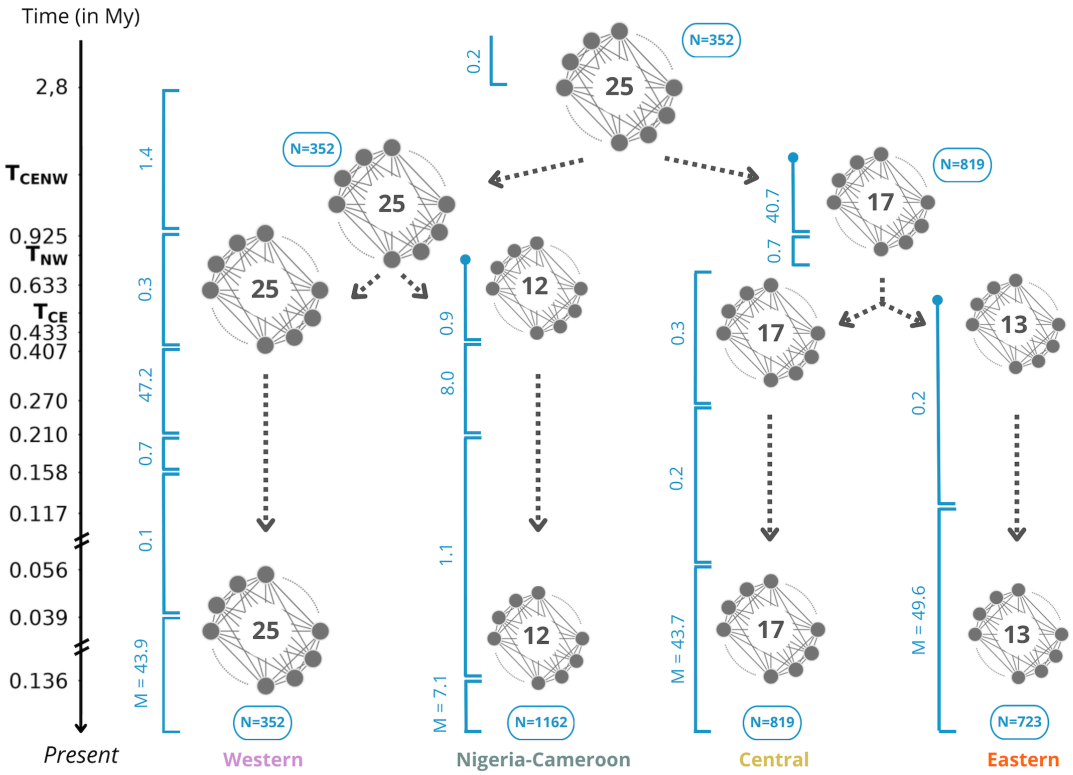


Figure 7: Proposed general n-island model for the history of the four subspecies of common chimpanzee. Dark full circles symbolize subpopulations structured in n-island, and the number in the middle corresponds to the number of demes n. Deme size N is given in number of diploid individuals. T_{CE}, T_{NW} and T_{CENW} correspond to the times at which two meta-populations join each other (backward in time), for Central/Eastern Chimpanzees, Nigeria-Cameroon/Western chimpanzees, and the two ancestral meta-populations respectively. In blue, along the vertical intervals, are the migration rates $M_i = 4Nm_i$.

that changing the splitting times (Supp. Figure S12-S15), and especially testing more recent values, did not significantly change the IICR curves, even though the estimated IICR could depart from the observed PSMC at the splitting time when the latter occurred too recently (see the case of Nigeria-Cameroon and Eastern chimpanzees in Supp. Figures S13 and S15 respectively). Altogether, we could construct a general model able to explain all the observed PSMC curves while incorporating both a tree model and intra-subspecies population structure, without any population size change within each branch.

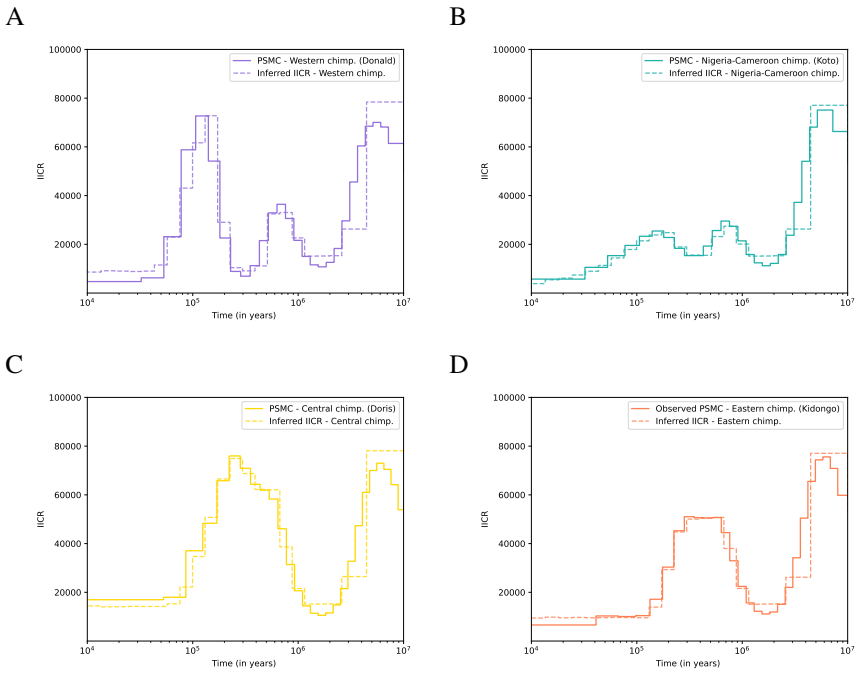


Figure 8: Empirical PSMC (solid line) and simulated IICR (dotted lines) for each subspecies of common chimpanzees under the *n*-island model displayed Figure 7, with T_{CE} = 600 Mya, T_{NW} = 800 Mya and T_{CENW} = 900 Mya. A. Western chimpanzees, B. Nigeria-Cameroon chimpanzees, C. Central chimpanzees and D. Eastern chimpanzees.

3.4 Prediction of other genomic statistics of diversity and differentiation

Though we stress that none of the models above (individual and global) should be taken at face value, the results obtained suggest that we were able to generate IICR plots that were similar to the observed PSMC plots under 1) *n*-island models for each subspecies independently and 2) a general model incorporating the four subspecies. We simulated genomic data under our general model and computed statistics representing genetic diversity and genetic differentiation to see if it could predict values close to empirical ones. As can be seen in Figure 9, we found that our simulated diversity measures were close to the observed values

computed by de Manuel *et al.* [6], and slightly lower than those computed by Prado-Martinez *et al.* [2] for Nigeria-Cameroon, Central and Eastern chimpanzees. We also found that our simulations recovered the ranking of genetic diversity for three subspecies, with Central chimpanzees being the most genetically diverse and Western chimpanzees harbouring the lowest level of genetic diversity, as observed empirically (Figure 9). However, the simulated Nigeria-Cameroon chimpanzees showed the same genetic diversity as Western chimpanzees, which is not consistent with what is found empirically [2]. These diversity estimates were identical across the different splitting times we tested.

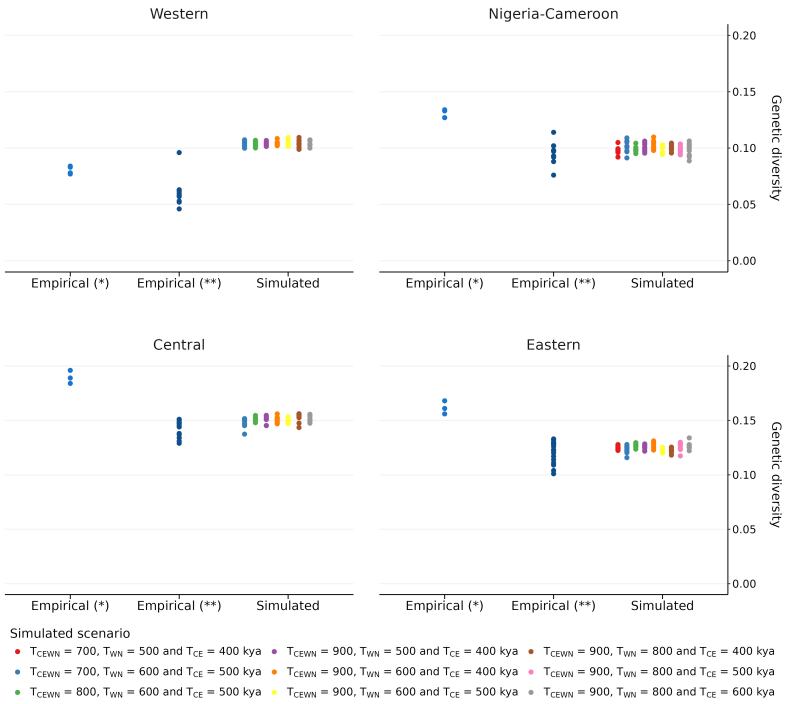


Figure 9: Genetic diversity of genomic data simulated under the model Figure 7 with several values for splitting times, sampling 10 diploid individuals in one deme. Empirical values were retrieved from (*) Prado-Martinez *et al.* [2] and (**) de Manuel *et al.* [6], each point corresponding to an estimate of individual observed heterozygosity.

Regarding the pairwise F_{ST} values, Figure 10 (and Supplementary Figures S16, S17, S18) shows that the general model predicts levels of within subspecies genetic differentiation that are within the empirical distribution, with the exception of Nigeria-Cameroon chimpanzees, a subspecies for which there are nearly no observed pairwise values. Figure 10 also shows that we strongly over-estimate between subspecies differences. As expected, the F_{ST} values between subspecies were sensitive to splitting times. For T_{CE} = 600 kya, T_{NW} = 800 kya and T_{CENW} = 900 kya, the values of splitting times that follow most closely the empirical

PSMC (see Figure 8), our model overestimated genetic differentiation in all pairs of subspecies. The F_{ST} values estimated from our model were three to five times higher than their respective empirical values for all pairs of subspecies. For instance, the median of the empirical ${F_{ST}}^{\text{W-NC}}$ was 0.12 whereas for the simulated ${F_{ST}}^{\text{W-NC}}$ it was 0.50. For the other pairs we had similar results (median empirical ${F_{ST}}^{\text{W-C}} = 0.10$ vs simulated ${F_{ST}}^{\text{W-C}} = 0.40$, median empirical ${F_{ST}}^{\text{W-E}} = 0.14$ vs simulated ${F_{ST}}^{\text{W-E}} = 0.46$, median empirical ${F_{ST}}^{\text{NC-C}} = 0.08$ vs simulated ${F_{ST}}^{\text{NC-C}} = 0.41$, median empirical ${F_{ST}}^{\text{NC-E}} = 0.08$ vs simulated ${F_{ST}}^{\text{NC-E}} = 0.47$, median empirical ${F_{ST}}^{\text{C-E}} = 0.08$ vs simulated ${F_{ST}}^{\text{C-E}} = 0.28$).

As expected we found that having more recent splitting times reduced our F_{ST} estimates which were getting closer to empirical values, though they never reached them for the parameter values we tested and which were selected to match the splitting times in the PSMC curves (Supplementary Figures S16, S17, S18). Genetic differentiation between demes of the same subspecies was much lower than between demes from different subspecies and were low and similar for three subspecies (Western, Central and Eastern) but much higher value within the Nigeria-Cameroon subspecies.

As expected as well, our general model predicted lower F_{ST} values between Western and Nigeria-Cameroon chimpanzees and between Central and Eastern chimpanzees than between other pairs of subspecies. This what is empirically observed, but we note that this is the reason why the original authors proposed the topology that we used. Finally, Eastern chimpanzees are the most differentiated subspecies to both Western and Nigeria-Cameroon chimpanzees, and Nigeria-Cameroon chimpanzees are the most distant subspecies to Central and Eastern chimpanzees, in both observed and simulated estimates.

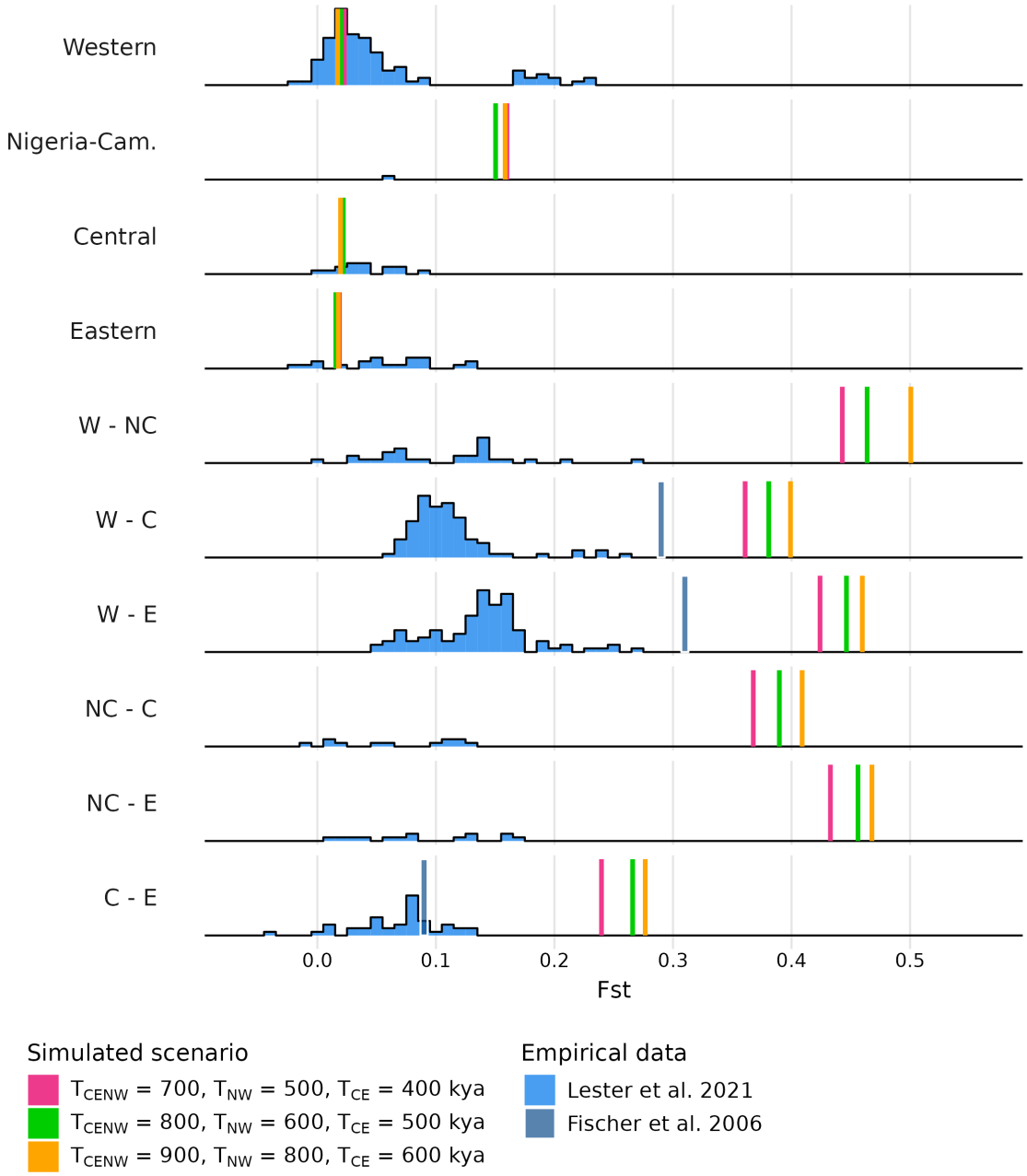


Figure 10: Genetic differentiation within and between subspecies. F_{ST} were computed on genomic data simulated under the model Figure 7 with T_{CE} = 600 kya, T_{NW} = 500 kya and $T_{CENW} \in \{700, 800, 900\}$ kya. See Supplementary Figures S16, S17, S18 for other splitting time values. In blue are the empirical values: histograms were retrieved from Lester *et al.* [18] (who used microsatellite data) and the blue vertical lines were retrieved from Fischer *et al.* [22] (who did not have samples of Nigeria-Cameroon chimpanzees in their study).

4 Discussion

Reconstructing the demographic history of species from genetic data is a complex endeavor and a major challenge because many factors have likely influenced the genetic patterns we observe today. These factors include population structure, changes in connectivity and in population size, selection, social structure (mating systems), among others [11, 41, 24, 42]. Major progress in human population genetics and genomics, including paleogenomics, have revolutionized our understanding of present-day and past genetic variation [43, 44]. Ideas and methods coming from human population genetics have influenced our understanding of the genetic diversity of other species [45, 46, 6, 14]. Here we propose to rather use what we know about other great apes to ask questions about ancient population structure and the evolutionary history of humans, and perhaps learn from such comparative analyses, as others have also suggested [47]. The work presented here is thus both an attempt at increasing our understanding of the ancient structure of chimpanzees but also on the evolutionary history of humans.

4.1 Towards a demographic history of chimpanzees incorporating population structure

We managed to explain patterns of genetic diversity within each subspecies of common chimpanzees with simple models of population structure and variable migration rates only, and we obtained demographic scenarios for each subspecies with shared periods of connectivity change. The validation step we applied (see Supplementary Material) suggests that if real genomic data had been generated under the inferred demographic scenarios, our method would have been able to infer them. While we do take inferred scenarios with a grain of salt, this validation step suggests that complex population structure can be inferred from single genomes. This is particularly notable because we recovered the simulated scenarios for each subspecies independently. Though this does not confirm that chimpanzees evolved under the inferred scenarios, it suggests that SNIF can infer different complex scenarios from the chimpanzee PSMC curves. A similar validation process is often applied in ABC studies [39, 48] but it is not that common in the literature, and we argue that it should be implemented more often to validate scenarios proposed on the basis of other inference methods.

Despite some variability in the inferences, we observed consistency across inferences within and between subspecies. We found that Western chimpanzees are characterized by a higher number of demes with a smaller size than the other three subspecies. This could be consistent with the fact that this subspecies

lives in a drier habitat, mostly in savanna [49], leading to a forest habitat that is more fragmented. However, it is unclear if it was the case throughout the evolutionary history of Western chimpanzees, and thus drawing strong conclusions is difficult at this stage. More generally, interpreting the number of islands and their size is not trivial and relating these parameters to empirical observations is not straightforward, in the same way that effective population size (N_e) values inferred in previous studies assuming panmixia are not easily interpreted. We assumed for simplicity that the number of demes and their size were constant. Allowing for the number of demes to vary would be more realistic but would also significantly increase the number of parameters in the models. At this stage, the inferred values of n and N should thus be interpreted with care.

Our models suggest that the four subspecies share a similar history of connectivity until approximately 500 kya (forward in time). We found a period of high connectivity between 2.5 and 1 Mya, followed by a decrease in connectivity until approximately 600-500 kya. The start of this period coincides with the Pliocene-Pleistocene transition boundary dated to around 2.6 Mya, whereas the drop in connectivity around 1 Mya falls within the Middle Pleistocene transition thought to have occurred between 1.2 Mya and 700 kya [50]. A second and more recent period of high connectivity, between 600-500 and 200-150 kya is also observed, although in Western and Nigeria-Cameroon chimpanzees only. The fact that only two chimpanzee subspecies were affected by this increase in migration during the most recent Middle Pleistocene period suggests that whichever environmental disturbances caused the genetic signal observed in the PSMC, these disturbances were localised mostly in Western Africa. Interestingly, Mazet *et al.* [28] and Arredondo *et al.* [31] also inferred high values of migration rates between 2.5 and around 1 Mya using PSMCs from humans, but did not identify the more recent period of high connectivity found in Nigeria-Cameroon and Western chimpanzees. This suggests that humans may share a common environmental history with Eastern and Central African chimpanzees, rather than with Western and Nigeria Cameroon chimpanzees for this period, at least. This also supports the idea that comparative analyses of genomic data from other vertebrates and primates from Africa (Eastern and Western) could improve our understanding of the history of the genus *Homo* in the last two millions of years.

We must however be careful in interpreting genomic data and environmental changes together, and more work would be needed to validate these results since they are based on a simple interpretation of the PSMC curves and on the assumption that PSMC curves infer the IICR with sufficient precision. At this stage we still lack a clear demographic model that would integrate the Pliocene-Pleistocene and Middle Pleistocene transitions and that can explain how environmental changes would have affected habitat connectivity (or

changes in effective population size), similarly impacting the PSMC curves in humans and chimpanzees. Also, we must acknowledge that reconstructing paleo-environments and habitats is still a complex endeavour [51, 52].

From a more technical point of view, we should also note that the inferred migration rates for Western chimpanzees were lower than for the other three species (Figure 4). This is surprising since all subspecies should provide similar values for the periods where there was one ancestral species to all four, and the PSMC curves overlap. We currently have no simple interpretation, and this could be due to some specificity of the Western subspecies, the quality of its genome, or the fact that we also inferred smaller deme sizes.

Similarly, we must note that the PSMC curves exhibit a large variance in the recent past with an apparent increase (forward in time) in the recent past that is interpreted in the connectivity graph as an increase in gene flow. We stress that this observed increase (forward in time) could also be due to a recent increase in the deme size, or to an uncertainty in the inference of the IICR. Indeed, the large variance in PSMC estimates in the recent past has been noted since the publication of the method of Li and Durbin [24]. Since SNIF assumes models with constant size it cannot typically fit this section of the IICR without making major changes in M. One must recall that under a model of constant size the IICR will necessarily stay at a low value (corresponding to the deme size in the recent past). IICR theory also shows that the IICR will "move up" quicker (backward in time) from the deme size to large values when M is large. Thus, under the assumption of constant size, SNIF will infer large M values in the very recent past to allow the IICR to move "up" rapidly from the inferred deme size, as explained by Mazet *et al.* [28]. At this stage, we thus considered that the recent increase in migration rate inferred for all chimpanzees should be interpreted very cautiously, if not ignored.

4.2 The general model and the limits of tree models

Using the demographic scenarios inferred for each subspecies of common chimpanzees, we successfully integrated the results of SNIF for each subspecies within a general tree model inspired by previous research on the four subspecies [2, 6, 14]. The difference with previous research was thus that we used n-island models instead of panmictic populations within each of the branches of the phylogenetic tree. We found that this model explained the PSMC curves of the four subspecies, and predicted well observed heterozygosity within demes and genetic differentiation between demes from the same subspecies. However, even using the shortest split times that would be consistent with the PSMC curves, the model led to an overestimation

of genetic differentiation between subspecies. This suggests that the statistics that depended on the SNIF inference were generally better than those that depended on the tree model. This consequently implies that using a tree topology without gene flow to define the relationships between the four subspecies and ignoring possible gene flow between them may not be appropriate. Two recent studies identified IBD patterns across the four subspecies considered as a single unit, suggesting that the four subspecies were part of a very large spatial metapopulation with gene flow between neighbouring populations, including populations currently attributed to different subspecies. Spatial models, instead of *n*-island models, might thus be necessary to represent the evolutionary history of chimpanzees.

The idea that gene exchange may have taken place between subspecies has been present in the literature [2, 14, 53, 6, 13, 11, 12, 10]. However, in most cases gene flow was seen as discrete events that could be dated, or that were limited to a pair of subspecies, which were in some cases not in geographical contact. Brand *et al.* [14] reviewed the literature on this question and found that at least 14 admixture events had been identified by eight different studies (Figure 1 of Brand *et al.* [14]) including one admixture event from a mysterious ghost species into the ancestors of bonobos [53]. Among these putative admixture events, some were identified by only one study whereas others were identified by two to five. It also appeared that some studies identified only one admixture event [7] whereas others identified as many as eight [13]. Brand *et al.* [14] themselves used an inferential method (Legofit, [54]) and a model that allowed for up to seven admixture events but only found support for two.

Altogether, this suggests that it has been difficult to find a consistent history of admixture or gene flow among previous genetic studies. We must stress that these studies are not always easy to compare, and some differences may arise from the fact that their sampling was different. For instance, several studies have no sample from one or two chimpanzee subspecies, whereas in other cases, the authors used samples from individuals with unknown geographic origin. However, one common feature of all these studies is that they consider tree models that ignore population structure below the subspecies level. They usually assume that the chimpanzee subspecies and the bonobos should be modelled as independent and panmictic lineages of an evolutionary tree where the only gene flow allowed is through these discrete admixture events.

4.3 The limits of the piecewise *n*-island model: towards spatial models

The approach used throughout the manuscript follows the theoretical and simulation-based work of several authors who found that structured stationary and non-stationary models can generate genetic signatures that

will be interpreted in terms of population size change when population structure is ignored [32, 55, 33, 56, 57, 28]. However, the *n*-island models we assumed with SNIF ignore spatial processes. As a consequence, IBD patterns observed by [17] and [18] cannot be reproduced, and suggest that the different subspecies were genetically connected in the recent past even if the chimpanzees's habitat is currently highly fragmented and discontinuous.

A difficulty that could arise from the use of spatial models is that the parameter space may increase significantly making the inference process more difficult. However, if we wish to improve our understanding of the evolutionary history of great apes, including humans, we may have no choice but integrate spatial models [58, 59]. As a simple test example and to illustrate the importance of spatial models, we developed a simple 1D stepping-stone model inspired by the demographic model proposed by de Manuel *et al.* [6] to study the demographic history of chimpanzees and bonobos. These authors assumed a tree model and allowed for the possibility of admixture events between subspecies and between bonobos and chimpanzees. They computed D-statistics and found evidence for admixture between bonobos and chimpanzees. Details of our 1D stepping-stone models and most of the results can be found in the Supplementary Material, but here we mainly wish to stress that we were able to reproduce the D-statistics with our spatial model without any introgression between bonobos and chimpanzees. By changing gene flow and deme size in the structured ancestral species, we found that we were recovering even higher D values than those observed today. This suggests that the bonobo admixture signals detected in chimpanzees might be the simple result of both ancient and recent population structure.

5 Conclusions

In this work, we showed that it is possible to propose a demographic model for common chimpanzees that accounts for population structure and gives a coherent interpretation of PSMC curves produced by previous studies. Although we stress that the general model we propose here should not be taken as face value, it manages to explain several patterns of genetic diversity within subspecies despite the limits of the *n*-island models (e.g. lack of spatial attributes). We noted the importance of using spatial models to account for the genetic differentiation between the subspecies and also showed that spatial models might also explain possibly spurious signatures of admixture with bonobos. This work is a first step towards more complex models, though we recognise the difficulty of such endeavour. There is an increasing recognition that ig-

noring population structure and spatial processes may lead to the inference of events that may never have happened during the evolutionary history of the species studied [33, 36, 37, 60, 38]. This has implications for many species but also important consequences on the understanding of human evolutionary history.

Ethics Statement

Not applicable.

Consent for Publication

Not applicable.

Availability of Data and Material

All code used for this paper is available at https://github.com/camillesteux/structurechimpanzees.

Funding

We acknowledge the financial support of a PhD studentship from the Ministère de l'Enseignement Supérieur et de la Recherche to CS. LC's research was supported by the DevOCGen project, funded by the Occitanie Regional Council's "Key challenges BiodivOc" program. This work was also supported by the LABEX entitled TULIP (ANR-10-558 LABX-41 and ANR-11-IDEX-0002-02) as well as the Investissement d'Avenir grant of the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01). We thank the IRP BEEG-B (International Research Project Bioinformatics, Ecology, Evolution, Genomics and Behaviour) for facilitating travel and collaboration between Toulouse (EDB, IMT and INSA) and Lisbon (IGC and cE3c).

Competing Interests

Not applicable.

Author Contributions

Conceptualization: LC; Methodology: LC, CS, CC, AA, WR, OM, RT; Investigation: LC, CS; Visualization: LC, CS, RT; Supervision: LC; Writing—original draft: LC, CS, RT ; Writing—review: CC, AA, WR, OM.

Acknowledgments

We thank all the members of the Population and Conservation Genetic group at the Instituto Gulbenkian de Ciência (IGC) for their support and for insightful discussions on the topic. We also thank the Bioinformatics Unit and the Informatics Team of the IGC for their help and support with computational resources, and the Genotoul bioinformatics platform. Finally, we thank Mimi Arandjelovic, Jack Lester and Javier Prado-Martinez for sharing their data with us.

Supplementary data

Supplementary Information S1: On the validation step.

Supplementary Information S2: Spatial structure confounds ancient admixture estimates.

Table S1: Prior ranges for the t_i.

Table S2: Distribution of inferred n.

Table S3: Distribution of inferred N.

Figure S1: Model and statistics calculated under a no-admixture structured population model.

Figure S2: PSMC curves provided by Prado-Martinez *et al.*.

Figures S3-S6: Inferred IICR curves for each subspecies of chimpanzees.

Figure S7: Distribution of inferred n and N coloured by individual.

Figure S8: Inferred connectivity coloured by individual.

Figure S9: Pseudo-observed data given to SNIF for the validation step.

Figure S10: Results of the validation step when giving to SNIF simulated PSMC curves as pseudo-observed data.

Figures S11-S14: Empirical PSMC and simulated IICR under the general tree model for each species of chimpanzees.

Figures S15-S17: Genetic distances within and between subspecies computed on genomic data simulated under the general tree model for different splitting times.

References

[1] Chen FC, Li WH. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. The American Journal of Human Genetics. 2001;68(2):444-56.

[2] Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, et al. Great ape genetic diversity and population history. Nature. 2013;499(7459):471-5.

[3] Langergraber KE, Prüfer K, Rowney C, Boesch C, Crockford C, Fawcett K, et al. Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. Proceedings of the National Academy of Sciences. 2012;109(39):15716-21.

[4] Kuhlwilm M, de Manuel M, Nater A, Greminger MP, Krützen M, Marques-Bonet T. Evolution and demography of the great apes. Current Opinion in Genetics and Development. 2016;41:124-9.

[5] Lobon I, Tucci S, de Manuel M, Ghirotto S, Benazzo A, Prado-Martinez J, et al. Demographic history of the genus Pan inferred from whole mitochondrial genome reconstructions. Genome Biology and Evolution. 2016;8(6):2020-30.

[6] De Manuel M, Kuhlwilm M, Frandsen P, Sousa VC, Desai T, Prado-Martinez J, et al. Chimpanzee genomic diversity reveals ancient admixture with bonobos. Science. 2016;354(6311):477-81.

[7] Caswell JL, Mallick S, Richter DJ, Neubauer J, Schirmer C, Gnerre S, et al. Analysis of chimpanzee history based on genome sequence alignments. PLoS genetics. 2008;4(4):e1000057.

[8] Humle T, Maisels F, Oates JF, Plumptre A, Williamson EA. Pan troglodytes. The IUCN Red List of Threatened Species. 2016. Accessed on 10 October 2023. Available from: `https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15933A17964454.en`.

[9] Fruth BI, Hickey J, André C, Furuichi T, Hart J, Hart T, et al. Pan paniscus. The IUCN Red List of Threatened Species. 2016. Accessed on 31 May 2024. Available from: https://www.iucnredlist.org/species/15932/102331567.

[10] Won YJ, Hey J. Divergence population genetics of chimpanzees. Molecular Biology and Evolution. 2005;22(2):297-307.

[11] Hey J. The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. Molecular Biology and Evolution. 2010;27(4):921-33.

[12] Becquet C, Przeworski M. A new approach to estimate parameters of speciation models with application to apes. Genome research. 2007;17(10):1505-19.

[13] Wegmann D, Excoffier L. Bayesian inference of the demographic history of chimpanzees. Molecular Biology and Evolution. 2010;27(6):1425-35.

[14] Brand CM, White FJ, Rogers AR, Webster TH. Estimating bonobo (Pan paniscus) and chimpanzee (Pan troglodytes) evolutionary history from nucleotide site patterns. Proceedings of the National Academy of Sciences of the United States of America. 2022;119(17):1-12.

[15] Barratt CD, Lester JD, Gratton P, Onstein RE, Kalan AK, McCarthy MS, et al. Quantitative estimates of glacial refugia for chimpanzees (Pan troglodytes) since the Last Interglacial (120,000 BP). American Journal of Primatology. 2021;83(10):e23320.

[16] McBrearty S, Jablonski NG. First fossil chimpanzee. Nature. 2005;437(7055):105-8.

[17] Fünfstück T, Arandjelovic M, Morgan DB, Sanz C, Reed P, Olson SH, et al. The sampling scheme matters: Pan troglodytes troglodytes and P. t. schweinfurthii are characterized by clinal genetic variation rather than a strong subspecies break. American Journal of Physical Anthropology. 2015;156(2):181-91.

[18] Lester JD, Vigilant L, Gratton P, McCarthy MS, Barratt CD, Dieguez P, et al. Recent genetic connectivity and clinal variation in chimpanzees. Communications Biology. 2021;4(1):1-11. Available from: http://dx.doi.org/10.1038/s42003-021-01806-x.

[19] Fontsere C, Kuhlwilm M, Morcillo-Suarez C, Alvarez-Estape M, Lester JD, Gratton P, et al. Population dynamics and genetic connectivity in recent chimpanzee history. Cell Genomics. 2022;2(6).

[20] Mitchell MW, Locatelli S, Ghobrial L, Pokempner AA, Clee PRS, Abwe EE, et al. The population genetics of wild chimpanzees in Cameroon and Nigeria suggests a positive role for selection in the evolution of chimpanzee subspecies. BMC Evolutionary Biology. 2015;15(1):1-15.

[21] Becquet C, Patterson N, Stone AC, Przeworski M, Reich D. Genetic structure of chimpanzee populations. PLoS genetics. 2007;3(4):e66.

[22] Fischer A, Pollack J, Thalmann O, Nickel B, Pääbo S. Demographic history and genetic differentiation in apes. Current Biology. 2006;16(11):1133-8.

[23] Fischer A, Wiebe V, Pääbo S, Przeworski M. Evidence for a complex demographic history of chimpanzees. Molecular biology and evolution. 2004;21(5):799-808.

[24] Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011;475(7357):493-6.

[25] Fan H, Wu Q, Wei F, Yang F, Ng BL, Hu Y. Chromosome-level genome assembly for giant panda provides novel insights into Carnivora chromosome evolution. Genome biology. 2019;20(1):1-12.

[26] Bursell MG, Dikow RB, Figueiró HV, Dudchenko O, Flanagan JP, Aiden EL, et al. Whole genome analysis of clouded leopard species reveals an ancient divergence and distinct demographic histories. Iscience. 2022;25(12):105647.

[27] Liu S, Westbury MV, Dussex N, Mitchell KJ, Sinding MHS, Heintzman PD, et al. Ancient and modern genomes unravel the evolutionary history of the rhinoceros family. Cell. 2021;184(19):4874-85.

[28] Mazet O, Rodríguez W, Grusea S, Boitard S, Chikhi L. On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? Heredity. 2016;116(4):362-71.

[29] Chikhi L, Rodríguez W, Grusea S, Santos P, Boitard S, Mazet O. The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. Heredity. 2018;120(1):13-24.

[30] Rodríguez W, Mazet O, Grusea S, Arredondo A, Corujo JM, Boitard S, et al. The IICR and the non-stationary structured coalescent: towards demographic inference with arbitrary changes in population structure. Heredity. 2018;121(6):663-78. Available from: http://dx.doi.org/10.1038/s41437-018-0148-0.

[31] Arredondo A, Mourato B, Nguyen K, Boitard S, Rodríguez W, Mazet O, et al. Inferring number of populations and changes in connectivity under the n-island model. Heredity. 2021;126(6):896-912.

[32] Wakeley J. Nonequilibrium migration in human history. Genetics. 1999;153(4):1863-71.

[33] Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA. The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. Genetics. 2010;186(3).

[34] Grusea S, Rodríguez W, Pinchon D, Chikhi L, Boitard S, Mazet O. Coalescence times for three genes provide sufficient information to distinguish population structure from population size changes. Journal of Mathematical Biology. 2019;78:189-224.

[35] Wright S. Evolution in Mendelian populations. Genetics. 1931;16(2):97.

[36] Eriksson A, Manica A. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. Proceedings of the National Academy of Sciences. 2012;109(35):13956-60.

[37] Eriksson A, Manica A. The doubly conditioned frequency spectrum does not distinguish between ancient population structure and hybridization. Molecular Biology and Evolution. 2014;31(6):1618-21.

[38] Tournebize R, Chikhi L. Questioning Neanderthal admixture: on models, robustness and consensus in human evolution. bioRxiv. 2023. Available from: https://www.biorxiv.org/content/early/2023/04/05/2023.04.05.535686.

[39] Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian computation in population genetics. Genetics. 2002;162(4):2025-35.

[40] Hudson RR. Generating samples under a Wright–Fisher neutral model of genetic variation. Bioinformatics. 2002;18(2):337-8.

[41] Wang K, Mathieson I, O'Connell J, Schiffels S. Tracking human population structure through time from whole genome sequences. PLoS genetics. 2020;16(3):e1008552.

[42] Parreira BR, Chikhi L. On some genetic consequences of social structure, mating systems, dispersal, and sampling. Proceedings of the National Academy of Sciences. 2015;112(26):E3318-26.

[43] Speidel L, Forest M, Shi S, Myers SR. A method for genome-wide genealogy estimation for thousands of samples. Nature genetics. 2019;51(9):1321-9.

[44] Yang MA. A genetic history of migration, diversification, and admixture in Asia. Human Population Genetics and Genomics. 2022;2(1).

[45] Kozak KM, Joron M, McMillan WO, Jiggins CD. Rampant Genome-Wide Admixture across the Heliconius Radiation. Genome Biology and Evolution. 2021 05;13(7):evab099. Available from: https://doi.org/10.1093/gbe/evab099.

[46] Teixeira H, Salmona J, Arredondo A, Mourato B, Manzi S, Rakotondravony R, et al. Impact of model assumptions on demographic inferences: the case study of two sympatric mouse lemurs in northwestern Madagascar. BMC Ecology and Evolution. 2021;21:1-18.

[47] Fischer J, Higham JP, Alberts SC, Barrett L, Beehner JC, Bergman TJ, et al. The Natural History of Model Organisms: Insights into the evolution of social systems and species from baboon studies. eLife. 2019 nov;8:e50989. Available from: https://doi.org/10.7554/eLife.50989.

[48] Beaumont MA. Approximate Bayesian computation in evolution and ecology. Annual Review of Ecology, Evolution, and Systematics. 2010;41:379-406.

[49] Humle T, Boesch C, Campbell G, Junker J, Koops, K, et al. Pan troglodytes ssp. verus. The IUCN Red List of Threatened Species. 2016. Accessed on 30 May 2023.

[50] Clark PU, Archer D, Pollard D, Blum JD, Rial JA, Brovkin V, et al. The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO2. Quaternary Science Reviews. 2006;25(23-24):3150-84.

[51] Berner N, Trauth MH, Holschneider M. Bayesian inference about Plio-Pleistocene climate transitions in Africa. Quaternary Science Reviews. 2022;277:107287.

[52] Demenocal PB. Plio-pleistocene African climate. Science. 1995;270(5233):53-9.

[53] Kuhlwilm M, Han S, Sousa VC, Excoffier L, Marques-Bonet T. Ancient admixture from an extinct ape lineage into bonobos. Nature Ecology and Evolution. 2019;3(6):957-65. Available from: http://dx.doi.org/10.1038/s41559-019-0881-7.

[54] Rogers AR. Legofit: estimating population history from genetic data. BMC bioinformatics. 2019;20:1-10.

[55] Beaumont MA. Recent developments in genetic data analysis: what can they tell us about human demographic history? Heredity. 2004;92(5):365-79.

[56] Heller R, Chikhi L, Siegismund HR. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. PloS one. 2013;8(5):e62992.

[57] Paz-Vinas I, Quéméré E, Chikhi L, Loot G, Blanchet S. The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. Molecular ecology. 2013;22(12):3279-91.

[58] Barbujani G, Sokal RR, Oden NL. Indo-European origins: A computer-simulation test of five hypotheses. American Journal of Physical Anthropology. 1995;96(2):109-32.

[59] Ray N, Currat M, Berthier P, Excoffier L. Recovering the geographic origin of early modern humans by realistic and spatially explicit simulations. Genome Research. 2005;15(8):1161-7.

[60] Boitard S, Arredondo A, Chikhi L, Mazet O. Heterogeneity in effective size across the genome: effects on the inverse instantaneous coalescence rate (IICR) and implications for demographic inference under linked selection. Genetics. 2022;220(3):iyac008.

[61] Durand EY, Patterson N, Reich D, Slatkin M. Testing for ancient admixture between closely related populations. Molecular biology and evolution. 2011;28(8):2239-52.

[62] Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, Li WH. Nucleotide diversity in gorillas. Genetics. 2004;166(3):1375-83.

[63] Baumdicker F, Bisschop G, Goldstein D, Gower G, Ragsdale AP, Tsambos G, et al. Efficient ancestry and mutation simulation with msprime 1.0. Genetics. 2022;220(3):iyab229.

[64] Kelleher J, Etheridge AM, McVean G. Efficient coalescent simulation and genealogical analysis for large sample sizes. PLoS computational biology. 2016;12(5):e1004842.

[65] Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, et al. Ancient admixture in human history. Genetics. 2012;192(3):1065-93.

Supplementary Information S1: On the validation process

The validation process that we applied in this study is similar to that used in ABC studies. We started with a PSMC curve for which we inferred ten scenarios corresponding to ten runs of SNIF, as explained in the Materials and Methods section. When these ten scenarios were similar we chose one of them, say scenario S^*, and generated a new IICR curve using the corresponding *ms* command. This IICR curve served as input to a new SNIF inference and we obtained ten newly inferred scenarios , $S_1^{**}, S_2^{**}, ...S_{10}^{**}$ from the ten independent runs. If these scenarios are different from each other, this may suggest that the optimization has not reached equilibrium and the number of iterations should be increased. Assuming now that the $S_1^{**}, S_2^{**}, ...S_{10}^{**}$ do not differ significantly from each other but differ from S^*, this suggests that the inferred scenario S^* should not be trusted at this stage since SNIF was not able to infer it despite using ten independent runs and reaching equilibrium. This was never the case with the chimpanzee data and we come back to this below.

In such a case where S^* should not be trusted, we suggest several solutions. One could explore more simple n-island models (with smaller c values for instance) to determine if S^* is too parameter-rich to be inferred by SNIF. If this fails too, it could also indicate that the real evolutionary history may be too complex to be approximated by a piecewise stationary n-island model such as S^* or simpler versions of S^*. In such a case, other demographic models should be explored, perhaps involving population size changes or spatial structure (stepping stone models for instance), or tree models.

If, on the contrary, S^* can be inferred reasonably well (i.e. the S_i^{**} scenarios are similar to S^*) as we observe with the chimpanzees, this could suggest that the scenario S^* is not only able to explain the original observed data but it can be inferred by SNIF if it were true. In other words, if the real species had evolved under S^* we would be able to infer S^* with SNIF. This does not prove that the species evolved under S^* but that S^* might be a reasonable approximation of reality to explain the PSMC computed for that species, at least until we find better alternatives.

These (or other) validation steps are fundamental and we suggest that they should be applied more often in demographic inference studies. It is more commonly applied in ABC studies, but there are still studies which infer a scenario without demonstrating that if real data had been generated under the inferred scenario, the authors would indeed be able to infer it back again, even approximately. While this validation process may be difficult to apply to some methods using genomic data, the current study shows that it is possible.

Supplementary Information S2: Spatial structure confounds ancient admixture estimates

Several studies have inferred ancient admixture events between lineages of the *Pan* genus [6, 13, 11, 21] (cf. Brand *et al.* [14] for a review). The mode and tempo of these putative events vary greatly with little consistency across studies (see the Discussion section in the main manuscript). This could suggest a complex admixture history involving the different subspecies and the bonobos, of which previous research works identified only some elements. Alternatively, the inferred admixture events might also be caused by the confounding effect of population structure (or other departures from model assumptions) which could generate different results across studies, depending on the models, statistics, samples, etc. used by the authors.

For instance, de Manuel *et al.* [6] inferred introgression between chimpanzees and bonobos using different approaches including the *D*-statistic [61], `TreeMix` method, and *SFS*-based demographic inference (Site Frequency Spectrum). However, like most previous studies, they assumed panmixia within each subspecies, and thus did not test for ancient population structure, which is an increasingly recognised confounder for the detection of admixture [61, 36, 37, 38].

Here, we implemented a simple linear stepping-stone model to test the hypothesis that population structure alone, without any gene flow between bonobos and chimpanzees, could replicate the observed *D*-statistics. We found that non-zero *D*-statistics could indeed be generated, following a gradient similar to what is observed on the empirical data. Noteworthy, our model not only predicted the empirical levels of *D* but it also predicted realistic values of the nucleotide diversity and the F_{ST} among chimpanzee subspecies. Altogether, this suggests that population structure can reproduce several signals of present-day genetic diversity, including purported admixture ones. It thus calls for more caution regarding the strength of the evidence favoring ancient admixture over population structure in the extant *Pan* genus, in a way that is similar to that suggested in recent research on the genus *Homo* [36, 37, 38].

S2.1 Methods

S2.1.1 Demographic model

We implemented a simple linear (one-dimensional) stepping-stone model (Fig. S1A), where each subspecies of chimpanzee is treated as a "metapopulation" of five connected demes of respective size N_i diploids. Bonobos are modelled in a similar way, except that no bonobo subspecies is currently recognised. We will abbreviate henceforth: Western chimpanzees as "W", Nigeria-Cameroon as "NC", Central as "C", Eastern as "E", and Bonobos as "P" (for *paniscus*). Thus, the full model is composed of five metapopulations of five demes each.

For simplicity, we set the deme size in each metapopulation using the θ_W estimate produced by de Manuel *et al.* [6] (Table S2 in the original article), with $N = \theta_W/(4n\mu)$, where n is the number of demes we consider in this study for each metapopulation ($n = 5$). We used the same mutation rate as de Manuel *et al.*, *i.e.* $\mu = 1.2 \times 10^{-8}$ per bp per generation. This led us to set $N_{NC} = 5,559$ diploids, $N_E = 6,498$, $N_C = 9,462$. For Western chimpanzees, we found that the original θ (leading to $N_W = 3,475$) produced an excess of nucleotide diversity compared to empirical values and we thus set it to a lower value ($N_W = 600$). We fixed the deme size in bonobos to $N_P = 4,000$ (about half N_C [62]), in the absence of the θ_W estimate in de Manuel *et al.* [6].

Within all metapopulations, demes are connected to their neighbours with a per-generation symmetric migration rate of $m_w = 10^{-2}$ (*i.e.* $M_{i,j} = m_w N_i$ diploid individuals migrating from deme i to deme j at each generation, backward). The different chimpanzee metapopulations are connected to the neighbouring metapopulation by a low-level (symmetric) gene flow of $m_b = 8 \times 10^{-4}$. The migration rate within the bonobo metapopulation was fixed at $m_{w,P} = 1.1 \times 10^{-4}$.

We acknowledge that several of these parameter values were fixed arbitrarily, since our purpose was not to infer parameters. We intended to showcase how a simple model can explain empirical data (when accounting for known population structure, without requiring unknown ancient admixture events), especially in producing apparent signatures of putative ancient admixture events.

Metapopulations split times were implemented in the model using the average of the divergence times reported in de Manuel *et al.* [6] (Figure 3 in the original article). Specifically, NC started to expand from W at 250 kya; E from C at 160.5 kya, the two ancestral chimpanzee lineages at 588.5 kya, bonobos from chimpanzees at 1.88 Mya. Each expansion is modelled, forward in time, by successive founding of each

within-metapopulation deme, every 200 years. We note that these founding events correspond exclusively to the instantaneous creation of a deme of size N_i (*i.e.* no bottlenecks).

For the time period prior to the first split between chimpanzee metapopulations (588.5 kya), the deme sizes in the ancestral chimpanzees were set to $N_{anc} = 7{,}000$ and the migration rate between the demes of this ancestral metapopulation was set to the same value as the present-day bonobo migration rate (*i.e.* $m_{anc} = m_{w,P} = 1.1 \times 10^{-4}$). All parameters for the remaining bonobo metapopulation were kept the same.

It should be clear that this model assumes that bonobos and chimpanzees never exchange gene flow at any time of their history after they separated 1.88 Mya.

The total number of non-redundant parameters in our model is 15, making it less parameter-rich than the panmictic model of de Manuel *et al.* [6]:

- 1 parameter for the number of demes per metapopulation (set to 5),
- 2 parameters for within-metapopulation migration rate in chimpanzees and bonobos (set to $m_w = 10^{-2}$ and 1.1×10^{-4}, respectively),
- 1 parameter for between-metapopulation migration rate in chimpanzees (set to $m_b = 8 \times 10^{-4}$),
- 5 parameters for deme sizes (N_i),
- 1 parameter for ancestral deme size (N_{anc}),
- 4 expansion times (*i.e.* the founding of the different metapopulations),
- 1 parameter for the delay between successive founding of demes (set to 200 years).

S2.1.2 Simulations

Using `msprime v1.1` with Hudson's coalescent algorithm [63, 64], we simulated genetic data for 20 chromosomes of 20 Mbp each ($G = 400$ Mbp), sampling 15 diploid individuals in the central deme of each metapopulation. This resulted in a total of 75 diploid genotypes per segregating site. Mutations were generated using a binary model, with two flipping alleles, at a rate of 1.2×10^{-8} per bp per generation. Recombination was assumed uniform, with rate 0.7×10^{-8} per bp per generation. We used a generation time of 25 years for both chimpanzees and bonobos [6].

S2.1.3 Statistics

Based on the simulated biallelic genetic data, we calculated, using the `scikit-allel v1.3.5` package:

- The nucleotide diversity π for each population sample (average number of pairwise differences).
- The differentiation index F_{ST} (Hudson's formula based on expected heterozygosities) between all pairs of chimpanzee populations.
- The D-statistic, as $D(X, Y; bonobo, outgroup)$ with X and Y being any chimpanzee population sample, and $outgroup$ being a virtual diploid genotype homozygous for the ancestral allele. To calculate the standard error (SE), we used a block-weighted jackknife with a typical block size of around 5 Mbp [65]. Confidence intervals at 95% were calculated as $D \pm 1.96 \times SE$.

S2.1.4 Observed data

We retrieved the empirical statistical values from previously published papers:

- π (average pairwise differences): from Fischer *et al.* [22], calculated on 26 intergenic sequences totalling 22.4 kbp for around 10 diploids in each sampled population.
- F_{ST} (Hudson's formula): from Fischer *et al.* [22], calculated on the same dataset as π.
- D-statistic: from de Manuel *et al.* [6], calculated as $D(X, Y, bonobo, humans)$ using whole-genome sequences on a set of 68 *Pan* samples. We used the D values estimated by the authors when aligning the *Pan* sequences on the human *hg19* assembly.

S2.2 Results

Our 1D stepping-stone structured population model replicated the observed gradient of D-statistics across focal chimpanzee subspecies (Fig. S1D). It further predicted the empirical values of between-chimpanzee F_{ST} (with the Central-Eastern being slightly underestimated in our simulations) (Fig. S1C), as well as the nucleotide diversity, with a slight excess of diversity for Eastern chimpanzees (Fig. S1A) (which likely results from the N_E value that we extracted from [6]).

In conclusion, these results show that the ancient admixture inferred in the *Pan* genus might not be robust to ancestral population structure.

We further investigated the variation of the *D*-statistic as a function of the ancestral connectivity, i.e. the connectivity within the metapopulation ancestral to extant chimpanzees (and further, extant *Pan* species) from 588.5 kya towards the past. To this end, we sampled $N_{anc}m_{anc}$ values from 10^{-3} to 10, every 0.2 on a $\log_{10}$-scale. Keeping N_{anc} fixed at 7,000 (cf. previous model), we derived m_{anc} according to the composite parameter $N_{anc}m_{anc}$. All other demographic and simulation parameters were kept the same as in the previous model, except for genome size, restricted to 10×10 Mbp chromosomes for computational reasons. We estimated *D* between present-day Central and Western chimpanzees: $D(C, W, bonobo, outgroup)$.

Our results (Fig. S2) show that the *D*-statistic values follow a sigmoid curve, "saturating" at zero for the highest $N_{anc}m_{anc}$ values (towards the right). This is expected, since with such high connectivity, the population model becomes nearly panmictic, and it does not incorporate admixture from bonobos into C lineages. The curve also plateaus around 0.6 for the lowest $N_{anc}m_{anc}$ range. Interestingly, we note that the *D*-statistic values start to become significant, in our model, around $N_{anc}m_{anc} = 1$ migrant per generation. A steep variation in *D* is observed for connectivity values ranging between around 0.06 and 0.6 migrants per generation. These results confirm that very significant values of the *D*-statistic can be produced under a structured population model in the absence of admixture, and that they depend mostly on the level of the ancestral connectivity within the metapopulation ancestral to the tested samples. We show that *D* can reach very high values (0.6) compared to the empirical values reported here, confirming that the level and significance of the *D* statistic cannot be used as indisputable evidence for admixture.

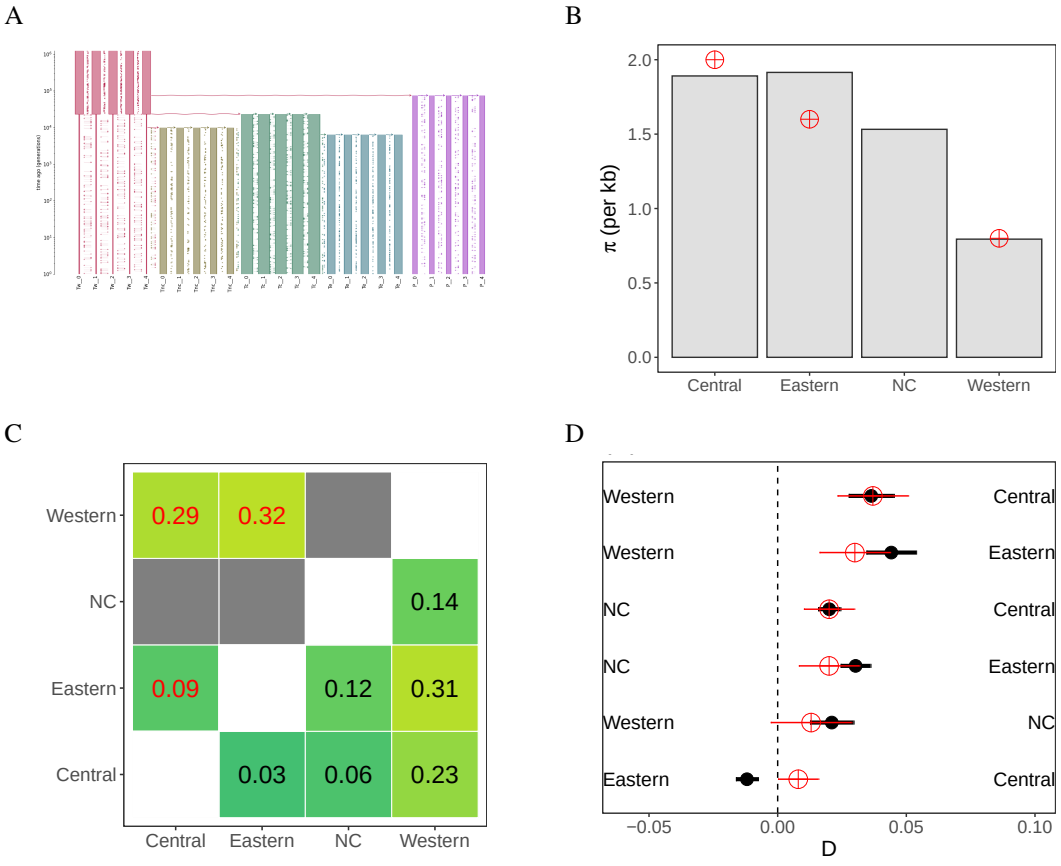


Figure S1: Model and statistics calculated under a no-admixture structured population model. A. Representation of the simulated demographic model. Each color corresponds to a chimpanzee subspecies or to the bonobos. Red: W, olive: NC, green: C, blue: E, purple: bonobos. B. Nucleotide diversity per kbp estimated for each population, for the data simulated under the structured population model (gray bars), and on the empirical data (red points). No NC samples were available in the [22] study. C. Pairwise F_{ST} obtained on the simulated data (lower-right triangle, black text) and on the empirical data (upper-left triangle, red text). We note that in the original article, the NC population was not studied [22], thus appearing here as empty gray cells. D. D-statistics from the simulated data (black points with 95% confidence interval) and empirical data (red points). The D-statistic values were calculated as $D(X,Y,P,O)$ with "P" the bonobos. On the plot, X populations are labeled on the right and Y on the left.

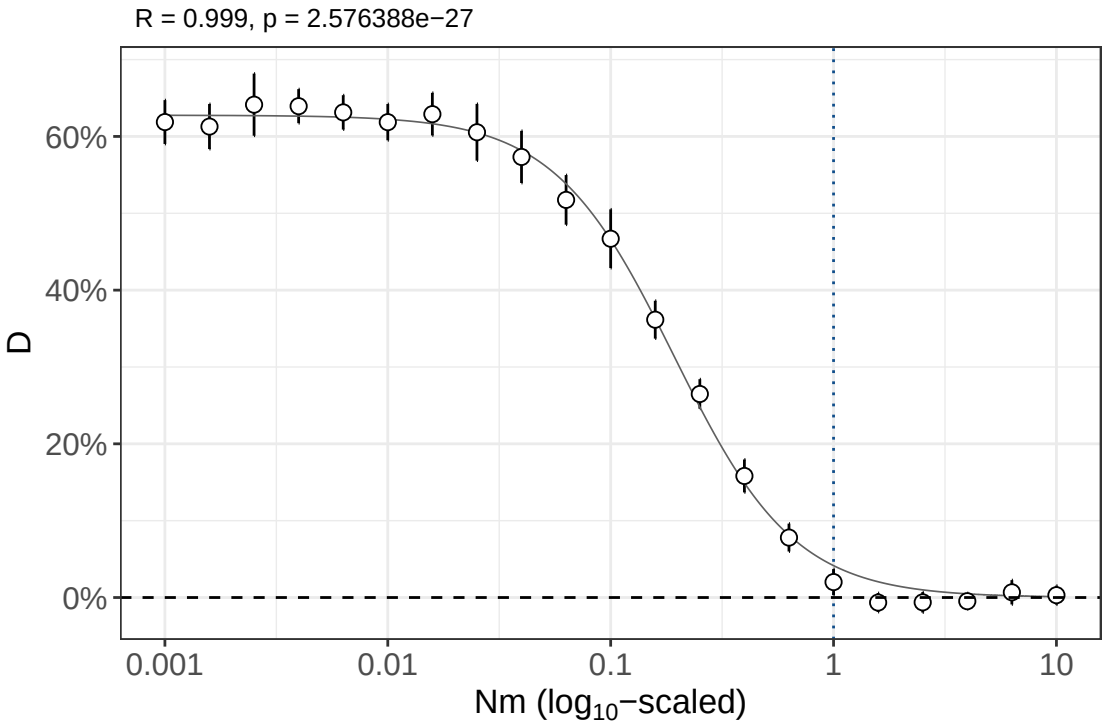


Figure S2: Distribution of the D-statistic as a function of ancestral migration rates. The D-statistic was computed as $D(C,W,P,O)$ for our no-admixture structured model, where we allowed the migration rate in the ancestral metapopulation to vary. This ancestral metapopulation corresponds to the period prior to 588.5 kya, *i.e.* before the foundation of the bonobos (cf. model description) We allowed $N_{anc} \times m_{anc}$ to vary from 10^{-3} to 10. The error bars represent the confidence intervals at 95%. The x-axis is $\log_{10}$-scaled. The vertical blue dotted line represents the first tested $N_{anc}m_{anc}$ value from decreasing order with significant D-statistics. The gray curve is a logistic fit to the empirical scatter plot. The Pearson's correlation coefficient representing the fit of the curve to the empirical data is reported in the title.

Table S1: Prior ranges of the t_i (times at which migrations rates are allowed to change, delimiting the components) given to SNIF for the final analysis. Note that here the times are given in years, but they have to be divided by the generation time when given to SNIF.

Subspecies	c	Priors of t_i (in years)
Western	7	(2e4, 1e5), (1e5, 3e5), (1e5, 3e5), (3e5, 7e5), (7e5, 1.5e6), (1.5e6, 5e6)
Nigeria-Cam.	7	(1e4, 1e5), (1.5e5, 3e5), (3.5e5, 5e5), (8e5, 1.1e6), (2e6, 3e6), (4e6, 7e6)
Central	8	(5e4, 2e5), (2e5, 4e5), (5e5, 2e6), (5e5, 2e6), (2e6, 5e6), (2e6, 5e6), (6e6, 1e7)
Eastern	7	(5e4, 2e5), (4e5, 1.5e6), (4e5, 1.5e6), (2e6, 6e6), (2e6, 6e6), (7e6, 1e7)

Table S2: Distribution of inferred n, the number of demes of the n-island models.

Subspecies	**Min**	**25% quantile**	**Median**	**Mean**	**75% quantile**	**Max**
Western	12	17	21	23	31	48
Nigeria-Cameroon	6	10	11	11	13	20
Central	13	16	18	21	20	55
Eastern	8	11	13	13	15	23

Table S3: Distribution of inferred N, the deme size of the n-island models (in number of diploids).

Subspecies	**Min**	**25% quantile**	**Median**	**Mean**	**75% quantile**	**Max**
Western	112	239	305	285	335	437
Nigeria-Cameroon	616	1009	1154	1174	1311	1980
Central	227	589	737	694	834	1101
Eastern	561	728	801	863	975	1306

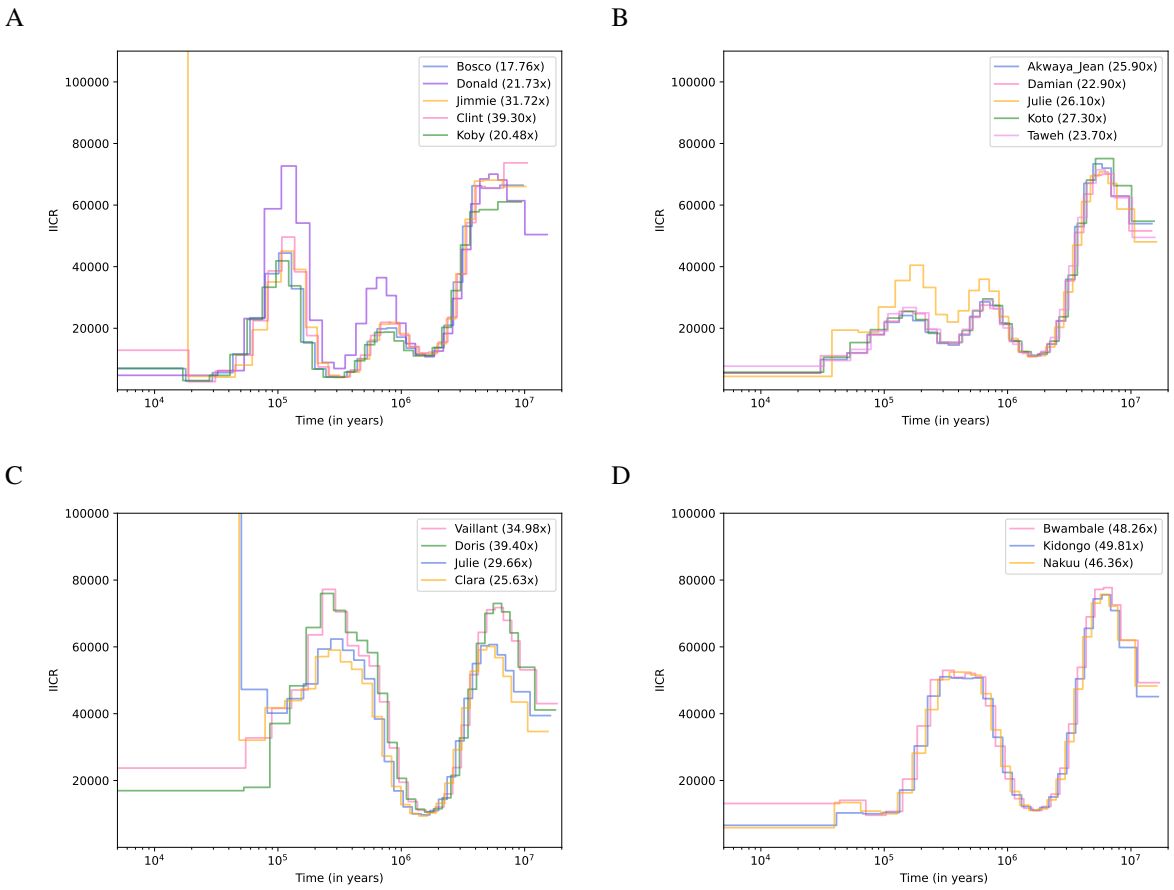


Figure S3: PSMC curves for A. Western chimpanzees, B. Nigeria-Cameroon chimpanzees, C. Central chimpanzees and D. Eastern chimpanzees, computed using PSMC files provided to us by Prado-Martinez *et al.* [2]

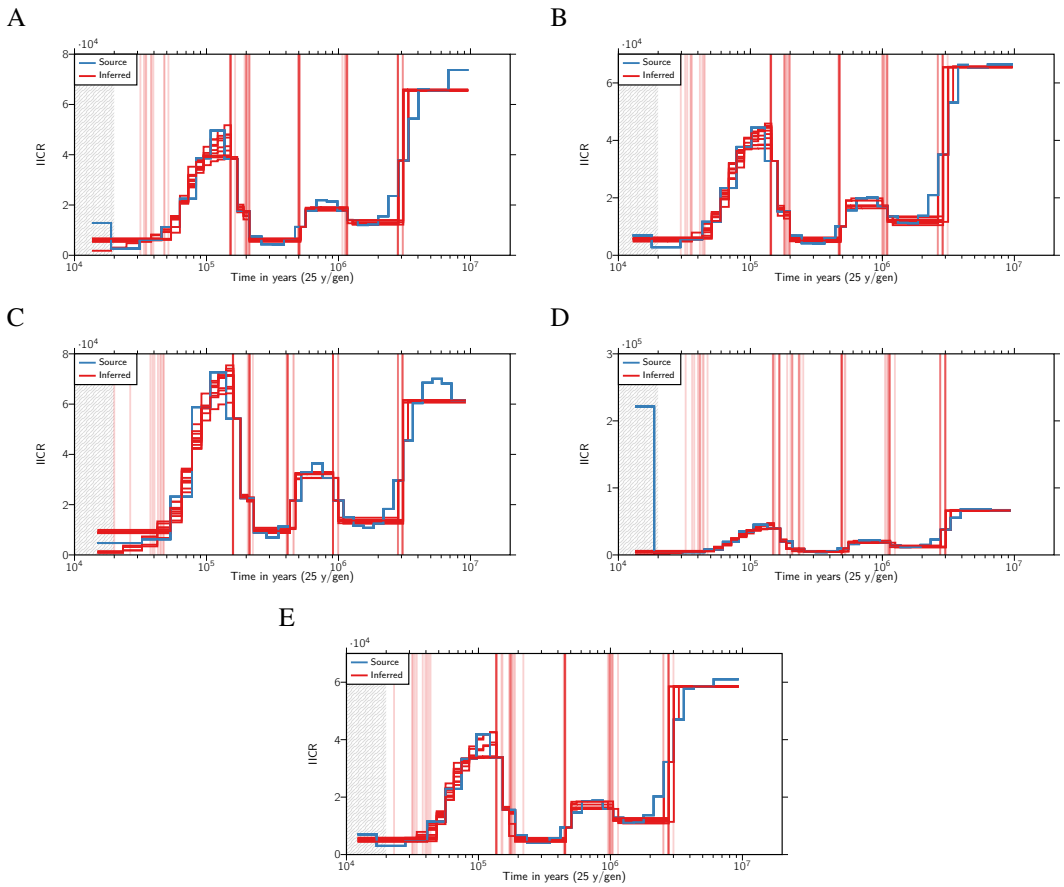


Figure S4: Inferred IICR curves (in red) and empirical PSMC (in blue) for the five Western common chimpanzees. Each red curve is a repetition of SNIF. The vertical red lines highlight the times at which there is an inferred change in migration rate and therefore delimit the components. A. Clint, B. Bosco, C. Donald, D. Jimmie and E. Koby. The grey zone corresponds to a part of the source PSMC which was not taken into account for the fitting of the curve by SNIF (see Material and Methods).

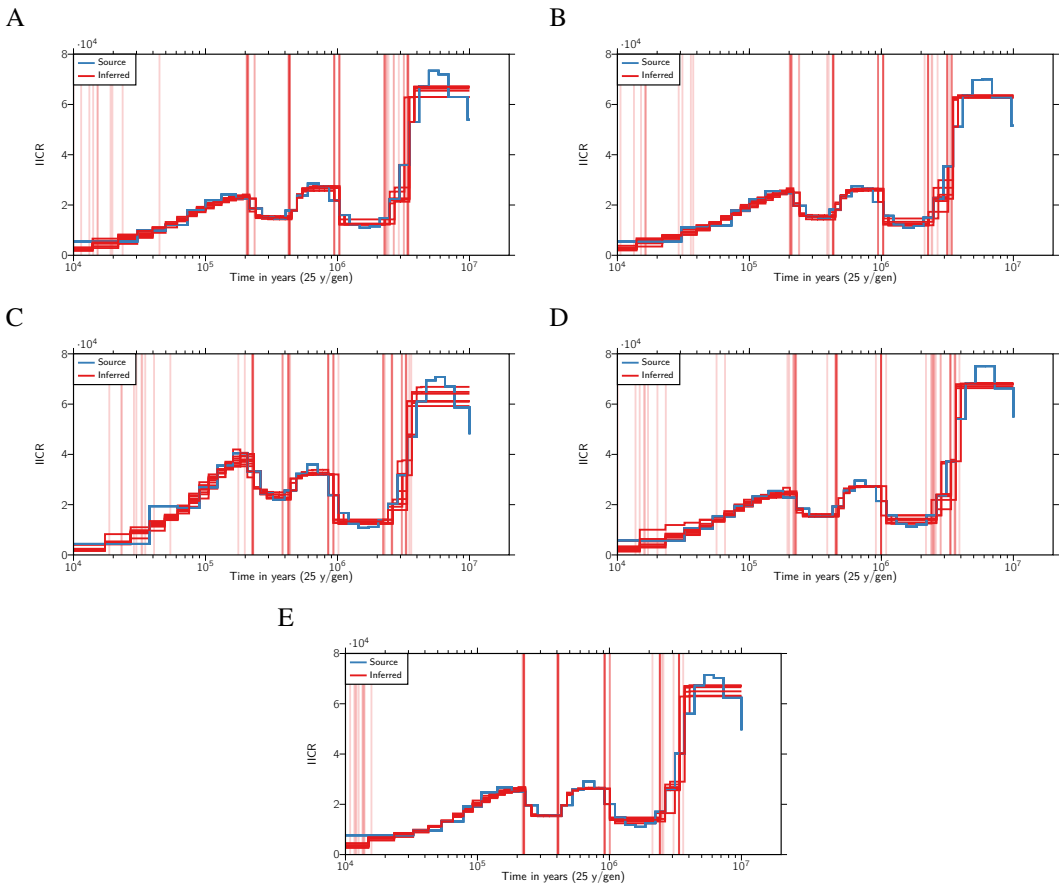


Figure S5: Inferred IICR curves (in red) and empirical PSMC (in blue) for the five Nigeria-Cameroon common chimpanzees. Each red curve is a repetition of SNIF. The vertical red lines highlight the times at which there is an inferred change in migration rate and therefore delimit the components. A. Akwaya-Jean, B. Damian, C. Julie, D. Koto and E. Taweh.

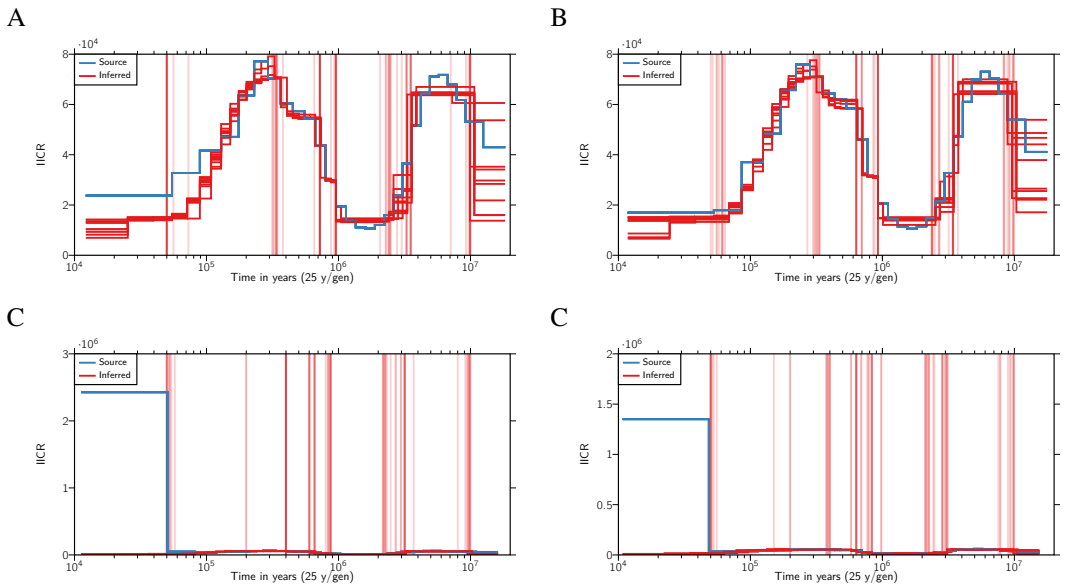


Figure S6: Inferred IICR curves (in red) and empirical PSMC (in blue) for the four Central common chimpanzees. Each red curve is a repetition of SNIF. The vertical red lines highlight the times at which there is an inferred change in migration rate and therefore delimit the components. A. Vaillant, B. Doris, C. Julie and D. Clara.

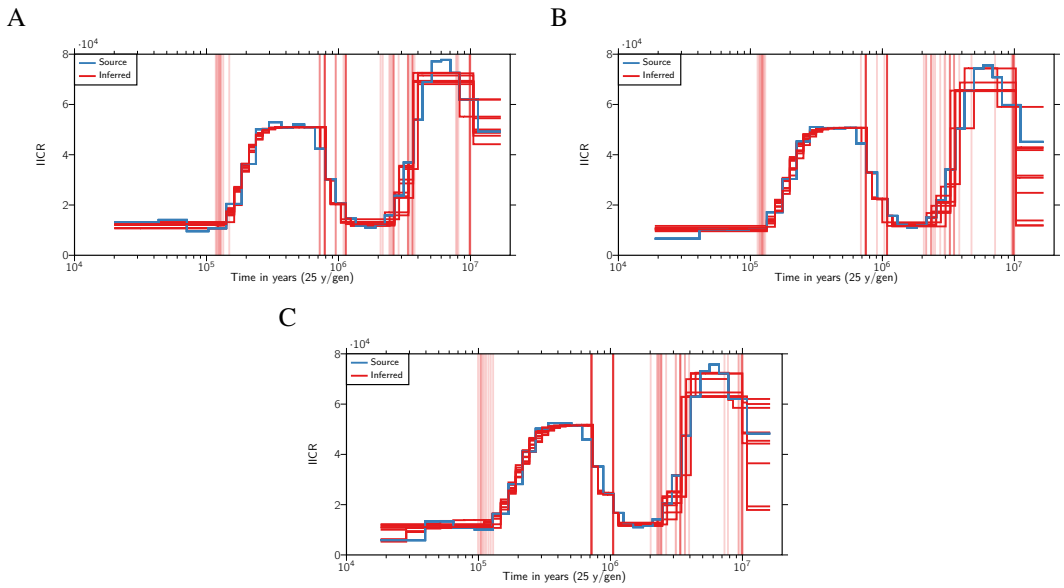


Figure S7: Inferred IICR curves (in red) and empirical PSMC (in blue) for the three Eastern common chimpanzees. Each red curve is a repetition of SNIF. The vertical red lines highlight the times at which there is an inferred change in migration rate and therefore delimit the components. A. Bwambale, B. Kidongo and C. Nakuu.

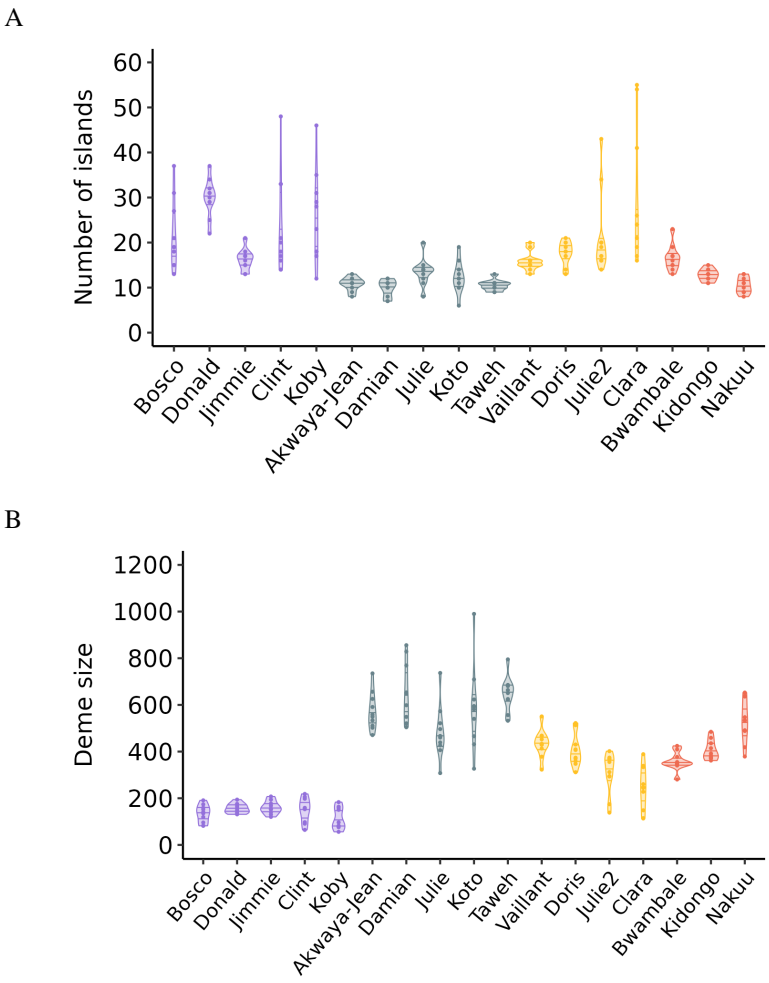


Figure S8: Distribution of the inferred A. number of islands and B. deme size (given in number of diploid individuals) across the 10 repetitions for each individual, using the parameter space shown Table 1. Horizontal lines in the violins represent the 25%, 50% (median) and 75% quantiles.

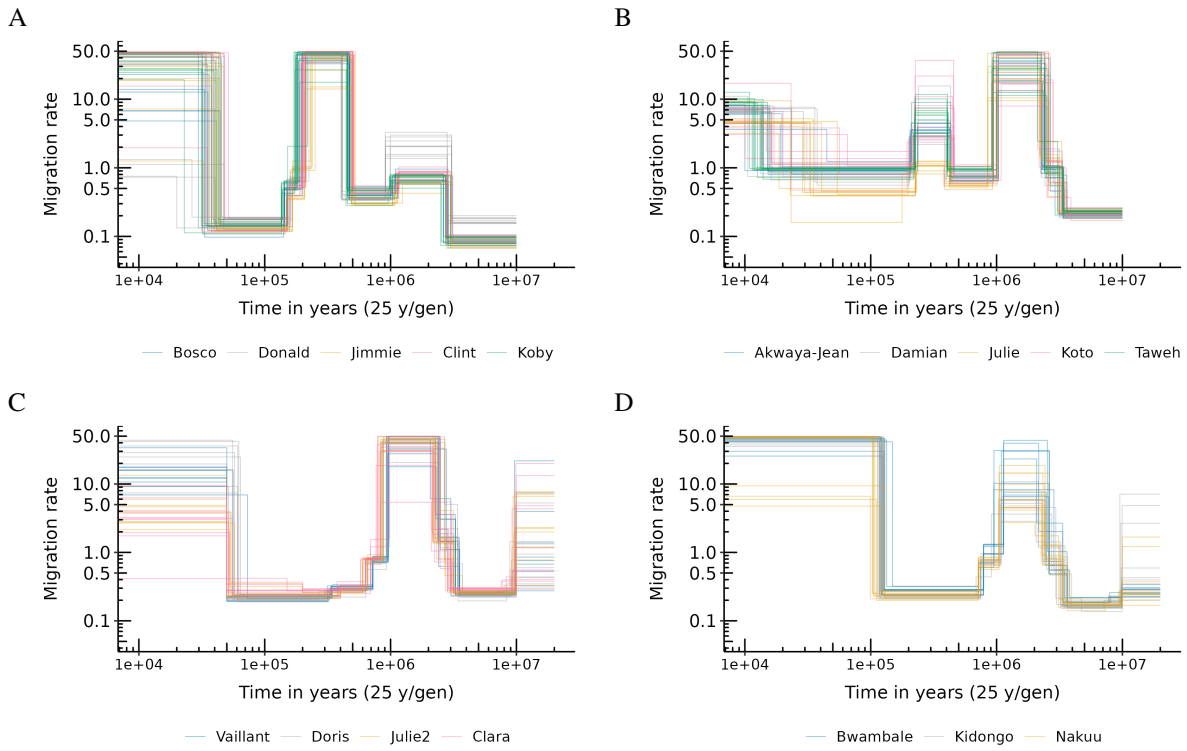


Figure S9: Connectivity graph (migration rates along successive time components) inferred by SNIF coloured by individual for A. Western, B. Nigeria-Cameroon, C. Central and D. Eastern chimpanzees. Each line corresponds to one inference (one repetition of SNIF) using one PSMC curve (or individual) as observed data.

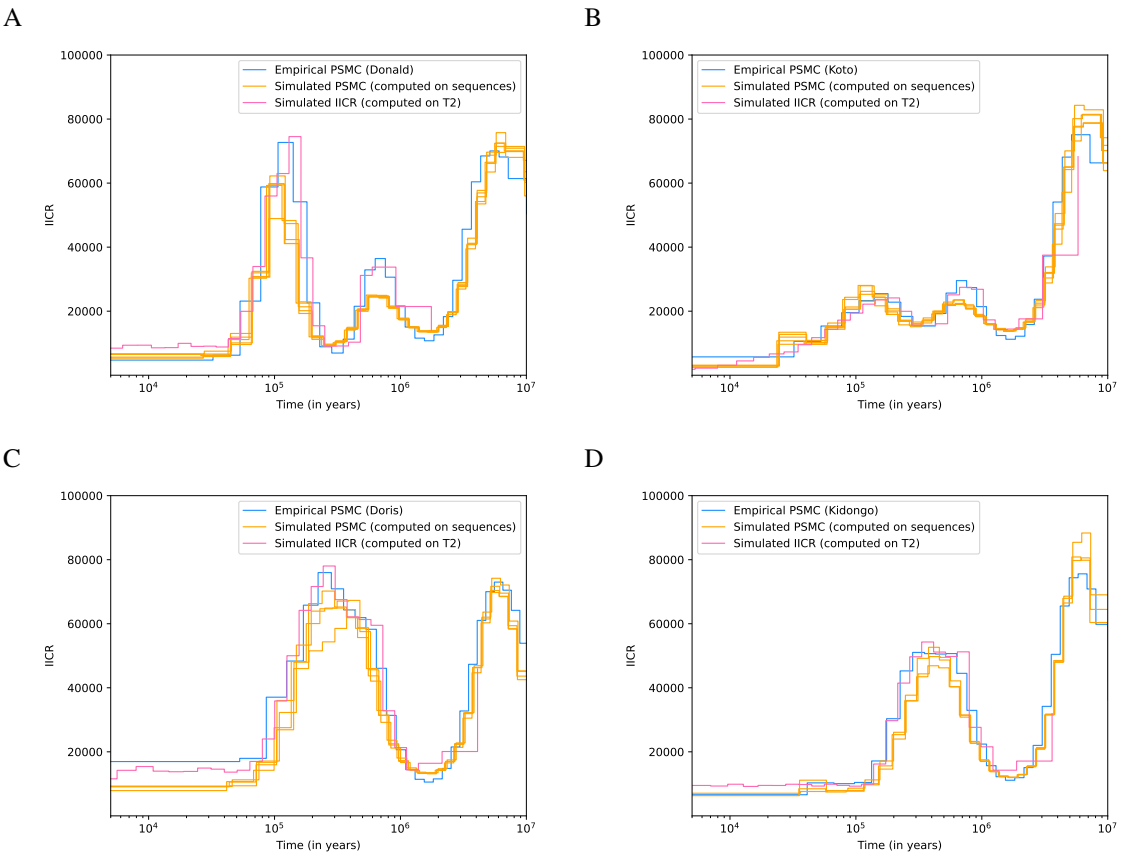


Figure S10: Simulated PSMC computed on simulated sequences (10x100Mb) (in blue) and simulated IICR computed on simulated coalescent times (T_2) (in orange), both given to SNIF as pseudo-observed data for the validation procedure. A. Western chimpanzees, B. Nigeria-Cameroon chimpanzees, C. Central chimpanzees and D. Eastern chimpanzees.

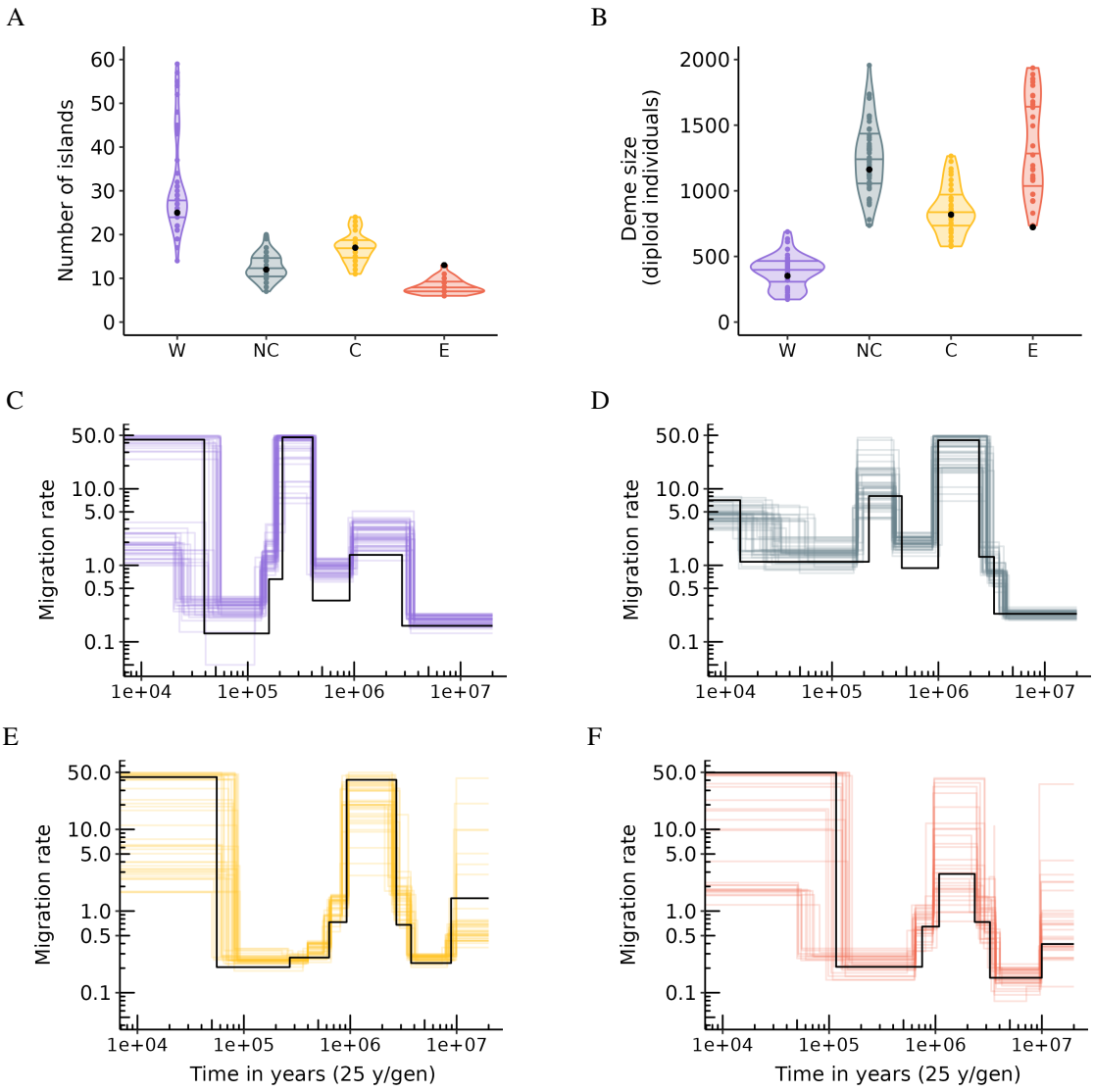
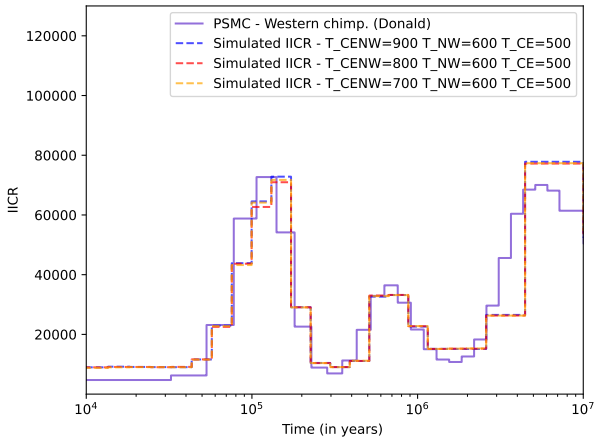


Figure S11: Results of the validation procedure when giving to SNIF simulated PSMC (see orange curves in Figure S10) as pseudo-observed data. A. Inferred number of islands (n), B. Inferred deme sizes (N), C. Inferred connectivity graph for Western chimpanzees, D. Inferred connectivity graph for Nigeria-Cameroon chimpanzees, E. Inferred connectivity graph for Central chimpanzees and F. Inferred connectivity graph for Eastern chimpanzees. Horizontal lines in the violins on panels A and B represent the 25%, 50% (median) and 75% quantiles.

A



B

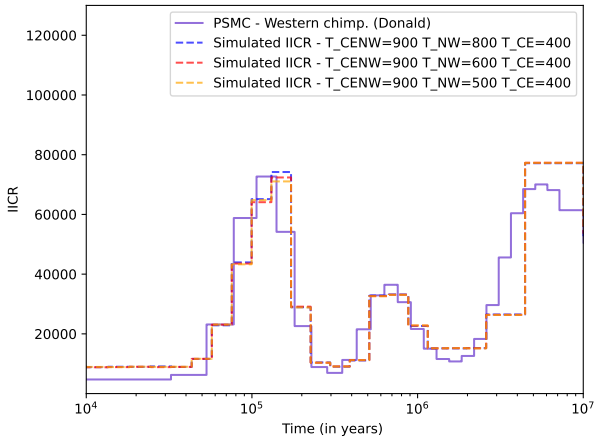
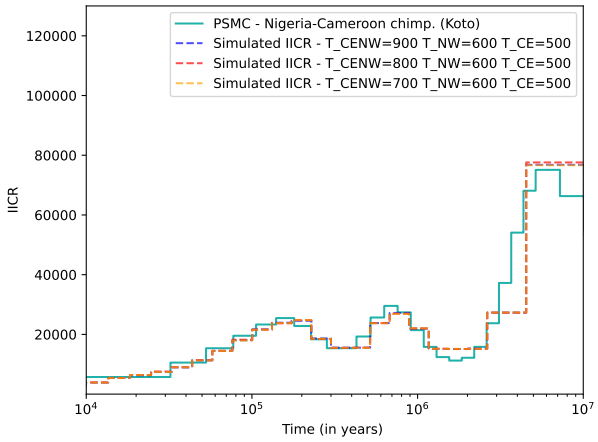


Figure S12: Empirical PSMC (solid line) and simulated IICR (dotted lines) for Western common chimpanzees under the general n-island model Figure 7 for different values of splitting times. (A) $T_{CENW} \in \{700, 800, 900\}$, $T_{NW} = 600$ and $T_{CE} = 500$ and (B) $T_{CENW} = 900$, $T_{NW} \in \{500, 600, 800\}$ and $T_{CE} = 400$ (in kya).

A



B

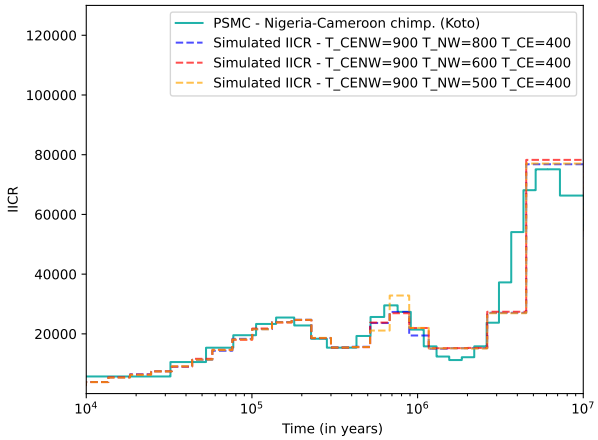
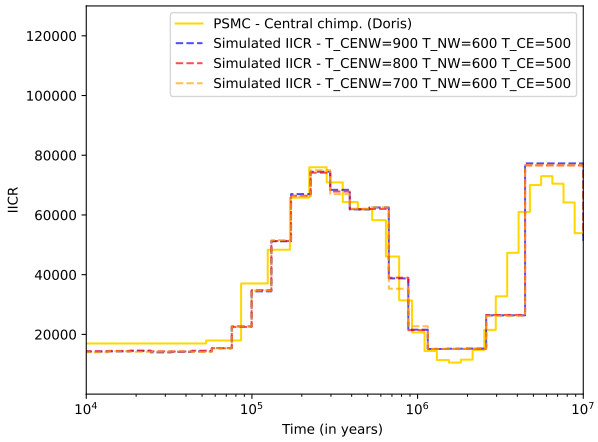


Figure S13: Empirical PSMC (solid line) and simulated IICR (dotted lines) for Nigeria-Cameroon common chimpanzees under the general n-island model Figure 7 for different values of splitting times. (A) $T_{CENW} \in \{700, 800, 900\}$, $T_{NW} = 600$ and $T_{CE} = 500$ and (B) $T_{CENW} = 900$, $T_{NW} \in \{500, 600, 800\}$ and $T_{CE} = 400$ (in kya)

A



B

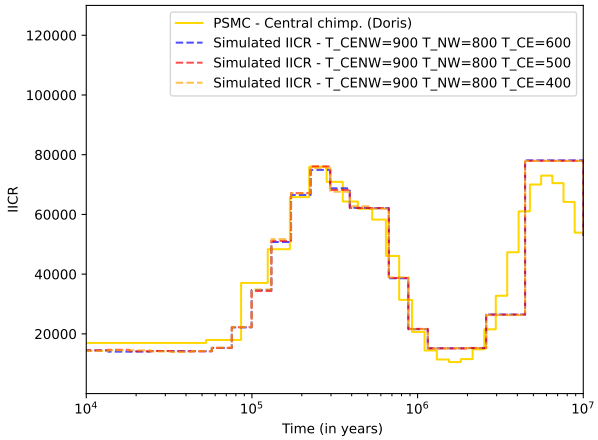
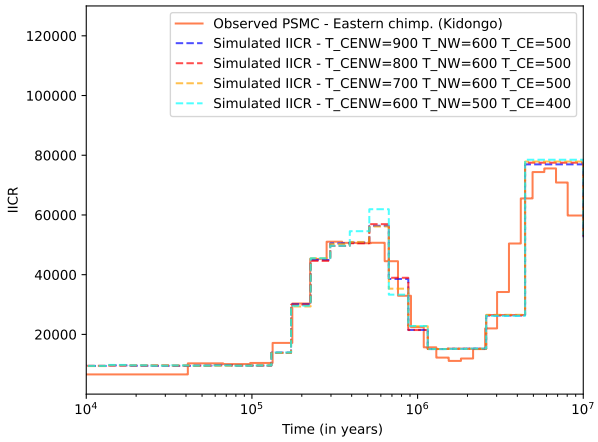


Figure S14: Empirical PSMC (solid line) and simulated IICR (dotted lines) for Central common chimpanzees under the general n-island model Figure 7 for different values of splitting times. (A) $T_{CENW} \in \{700, 800, 900\}$, $T_{NW} = 600$ and $T_{CE} = 500$ and (B) $T_{CENW} = 900$, $T_{NW} = 800$ and $T_{CE} \in \{400, 500, 600\}$ (in kya).

A



B

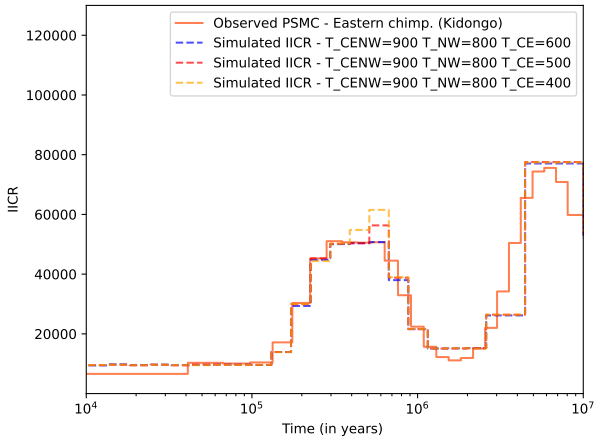


Figure S15: Empirical PSMC (solid line) and simulated IICR (dotted lines) for Eastern common chimpanzees under the general n-island model Figure 7 for different values of splitting times. (A) $T_{CENW} \in \{700, 800, 900\}$, $T_{NW} = 600$ and $T_{CE} = 500$ and (B) $T_{CENW} = 900$, $T_{NW} = 800$ and $T_{CE} \in \{400, 500, 600\}$ (in kya).

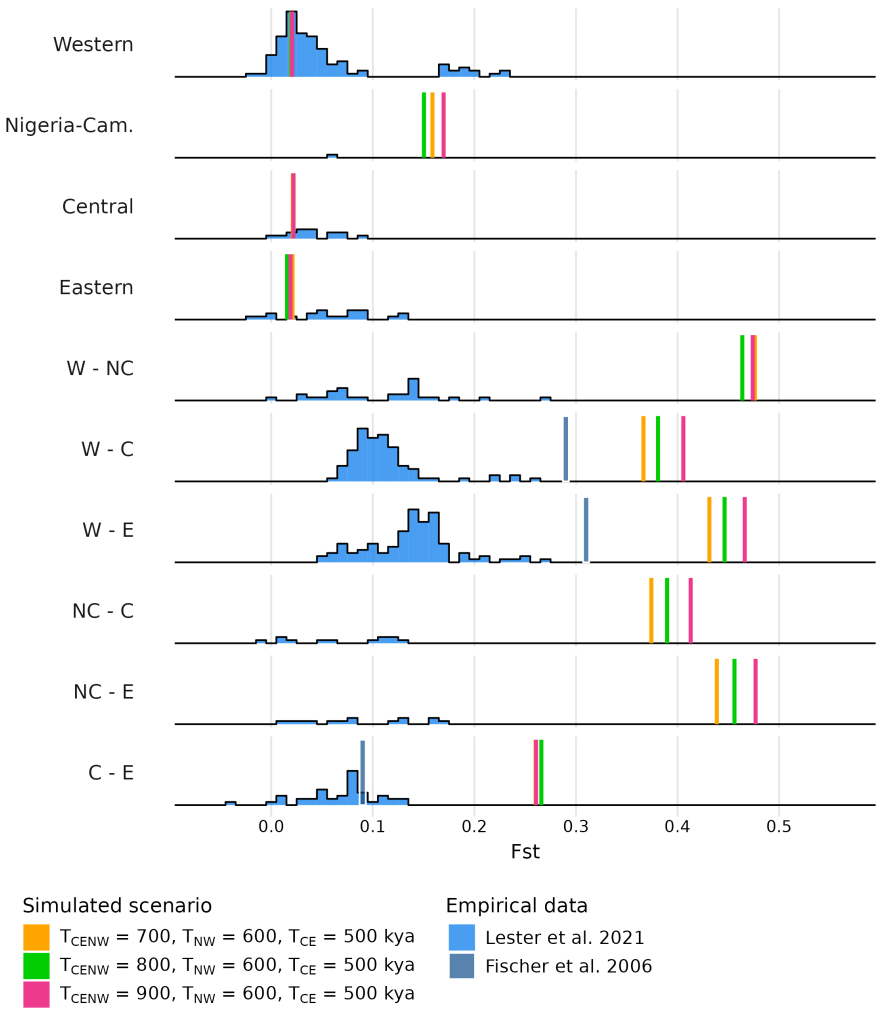


Figure S16: Genetic distance (F_{ST}) between demes of the same subspecies or between demes from different subspecies computed on genomic data simulated under the model Figure 7 with T_{CE} = 500 kya, T_{NW} = 800 kya and $T_{CENW} \in \{700, 800, 900\}$ kya. In blue are the empirical values: histograms were retrieved from Lester *et al.* [18] and the darker blue vertical lines were retrieved from Fischer *et al.* [22]

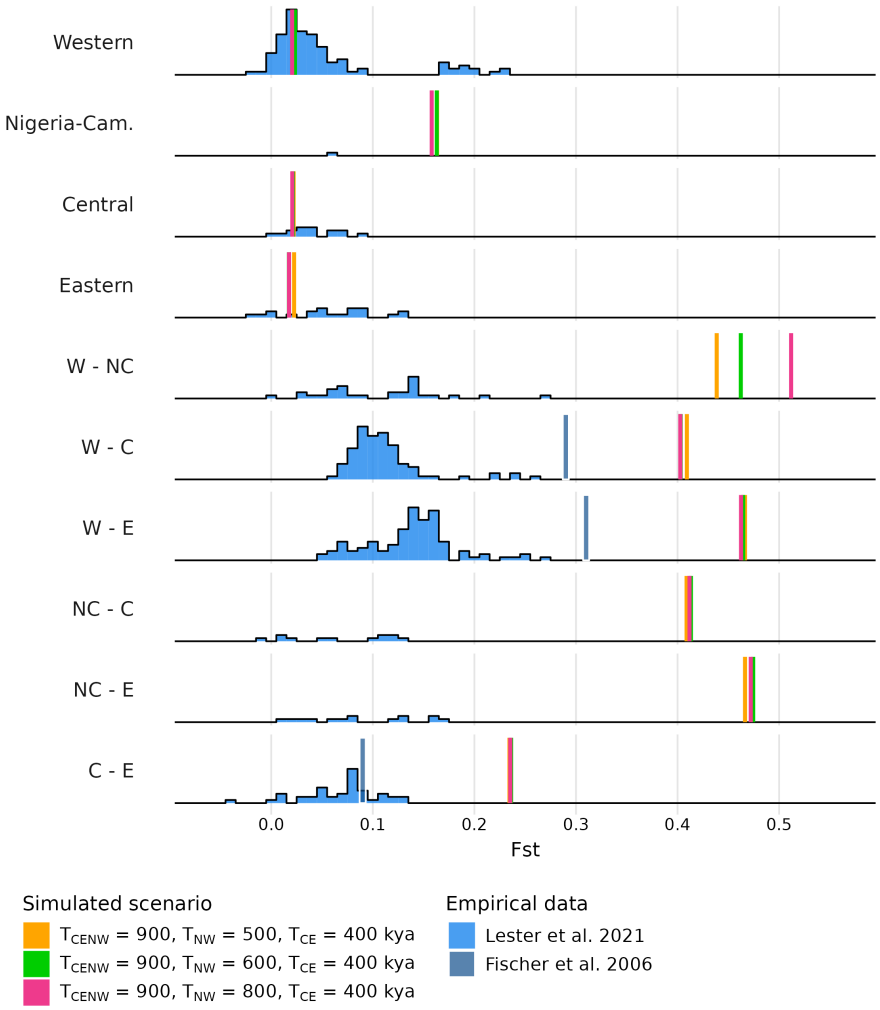


Figure S17: Genetic distance (F_{ST}) between demes of the same subspecies or between demes from different subspecies computed on genomic data simulated under the model Figure 7 with T_{CE} = 500 kya, T_{CENW} = 900 kya and $T_{NW} \in \{500, 600, 800\}$ kya. In blue are the empirical values: histograms were retrieved from Lester *et al.* [18] and the blue vertical lines were retrieved from Fischer *et al.* [22]

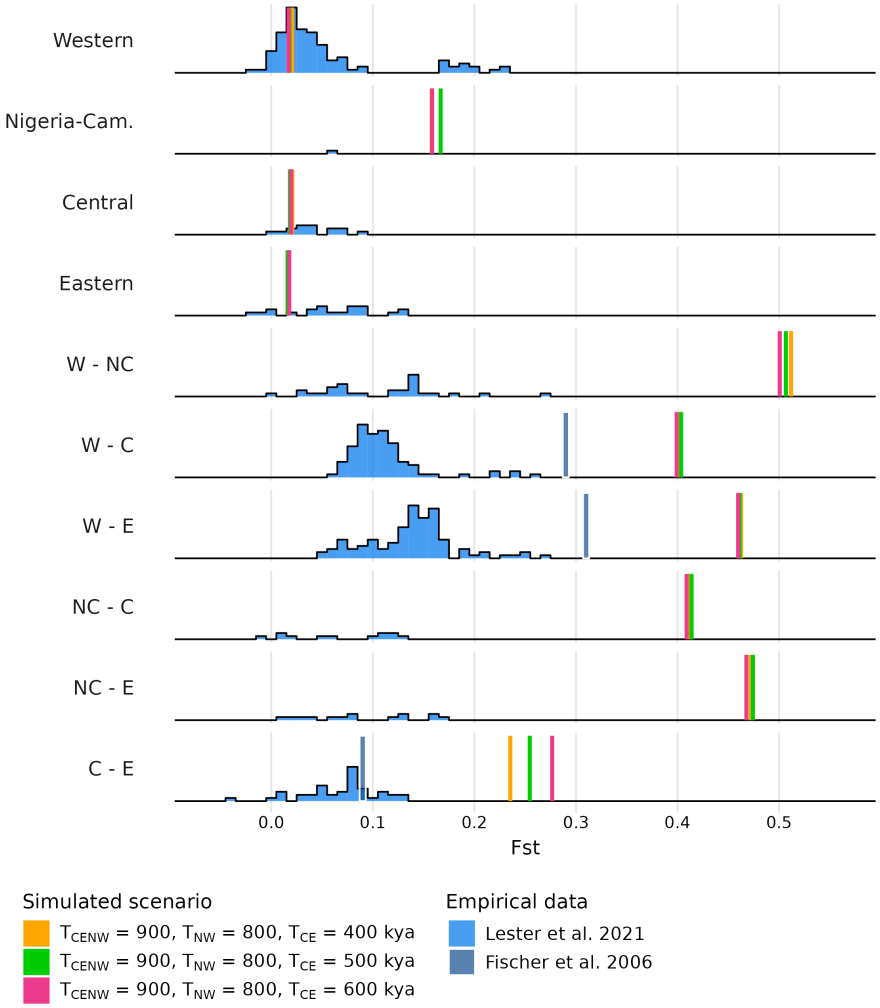


Figure S18: Genetic distance (F_{ST}) within and between subspecies computed on genomic data simulated under the model Figure 7 with T_{CENW} = 900 kya, T_{NW} = 800 kya and $T_{CE} \in \{400, 500, 600\}$ kya. In blue are the empirical values: histograms were retrieved from Lester *et al.* [18] and the blue vertical lines were retrieved from Fischer *et al.* [22]