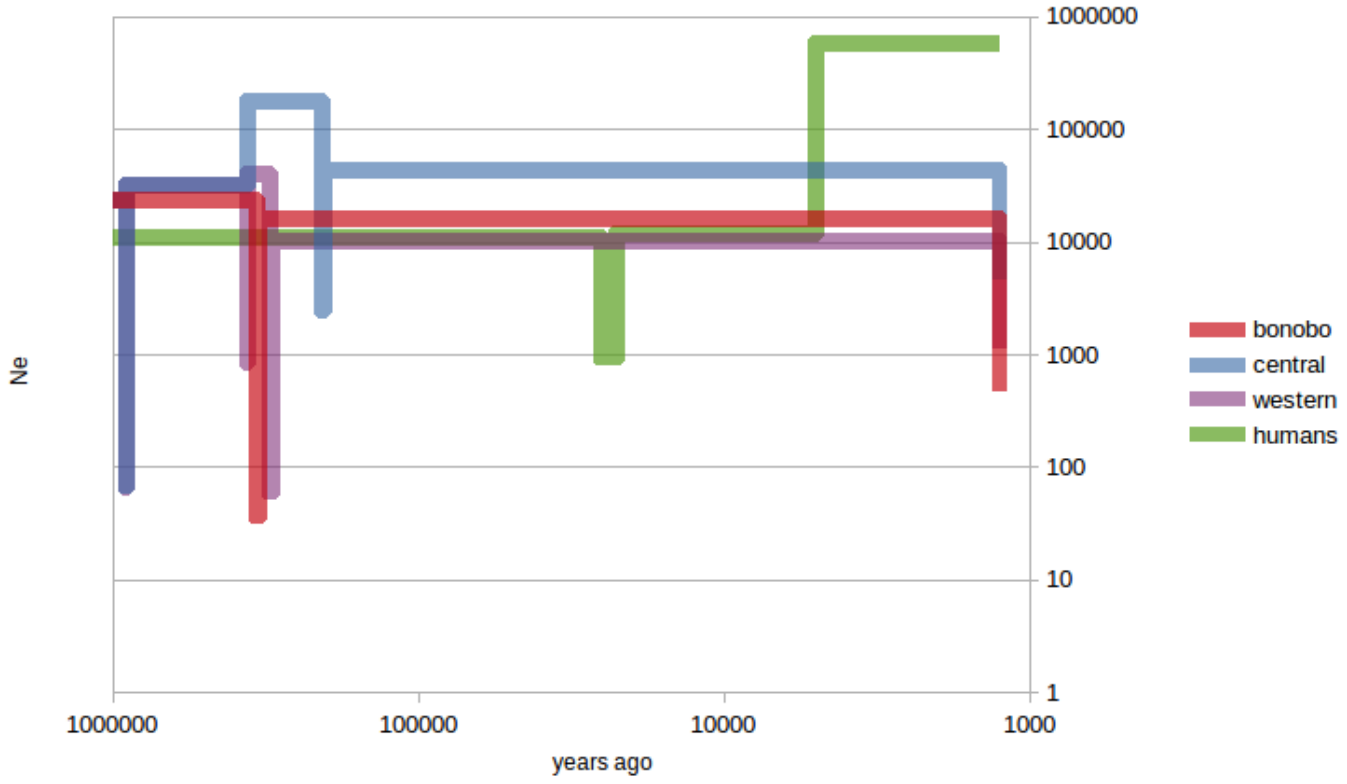
Supplementary Material - Forward Simulations

We perform simulations to check that the nuisance (*r*) parameters of polyDFE (Tataru *et al.* 2017) accurately account for demography and other distorters of the uSFS such as linkage. We use the forward-simulator SFS_CODE (Hernandez 2008) to generate three populations diverging from one source population, following the demographic parameters of bonobos, central and western chimpanzees (de Manuel *et al.* 2016), together with an independent (European) human population (Kim *et al.* 2017). We chose these four scenarios because they show remarkably different population histories and we believe they are good representatives of the demographic heterogeneity found in the great apes. We simulated samples consisting in 4 diploid individuals per population, for twenty loci of 1,440 base pair length (matching the mean protein length in humans). We run 450 iterations of each scenario. These numbers yield 9,000 protein-coding genes to approximately match the size of our data set. To account for background selection at fine-scale we assume there is no recombination within a given locus but there is free recombination between loci. We simulated coding regions where synonymous mutations are fully neutral and non-synonymous mutations are deleterious. The selection coefficients of non-synonymous mutations are drawn from a gamma distribution with shape 0.16. We assume that ~1,000,000 years ago the ancestral population of chimpanzees and bonobos was composed of 23,450 diploid individuals while during the same period of time the human population was composed by 10,907 diploid individuals (values retrieved from de Manuel *et al.* 2016; Kim *et al.* 2017, respectively). The simulated mean population scaled effect size of deleterious mutations ($S_d = 2N_e s_d$) for the ancestral population of chimpanzees and bonobos is $S_d = -1500$ and for the ancestral human population is $S_d = -698$. This matches approximately the S_d estimated in humans under model M3S (Supplementary Table 4 C). Hence, we assume that the mean s_d is the same for human, bonobos and the two chimpanzee subpopulations ($s_d = S_d/2N_e$ = 1,500/(2 x 23,450) = 698/(2 x 10,907) = 0.032). The ratio of nonsynonymous to synonymous sites is 2.93 and the mutation rate per site and generation is assumed to be $\mu = 1.65 \times 10^{-8}$ per generation

and per site. This mutation rate matches current levels of synonymous diversity in all four populations assuming that the true N_e is the harmonic mean of the simulated population sizes. For simplicity, we assume 25 years per generation in all populations. To speed up the forward simulations, we assumed an ancestral effective population size of 4000 (diploids) and scaled mutation rate, selection coefficients, and demographic parameters accordingly (Aberer and Stamatakis 2013). This ancestral population was simulated for a burn-in period of 10N generations.

We used the complex demography estimated in the *Pan* clade (de Manuel *et al.* 2016) (demographic parameters extracted from Supplementary Figure 51 and Supplementary Table 15, assuming no migration between populations), with a divergence of 36,240 generations between chimpanzees and bonobos, and 14,520 generations between central and western chimpanzees (Supplementary Figure A).



Supplementary Figure A. Schematic representation of the simulated demographic histories of bonobos, central chimpanzees, western chimpanzees, and humans. The exact value of each parameter can be found above.

This is an example of one iteration for bonobo and the two chimpanzee subpopulations:

sfs_code 3 1 -N 4000 -TS 0 0 1 -Td 0 P 0 1.0220895522 -Td 0 P 1 0.0028997868 -Td 0.0125 P 1 478.5735294118 -TS 2.715 1 2 -Td 2.715 P 1 5.5324954675 -Td 2.715 P 2 0.0267338598 -Td 2.7275 P 2 47.0942528736 -Td 2.84 P 0 0.001543725 -Td 2.8525 P 0 438.7027027027 -Td 3 P 2 0.0014644147 -Td 3.0125 P 2 173.2 -Td 3.495 P 1 0.014035458 -Td 3.5075 P 1 17.4317372378 -Td 4.52375 P 0 0.0287703302 -Td 4.52375 P 1 0.1045629966 -Td 4.52375 P 2 0.1086412625 -TE 4.53 -W 2 0 1 4 0.16 0.0002133333 --length 20 1440 --linkage p -1 -n 4 -a C --theta 0.0015517162 --rho 0

For the human population we simulated a bottleneck 1005 generations ago that lasted 2500 generations, then a recent expansion 201 generations ago (Kim *et al.* 2017) (parameters retrieved from Supplementary Table 1 for the 1000G data set with 864 chromosomes). Example of the command line used to simulate the human demography:

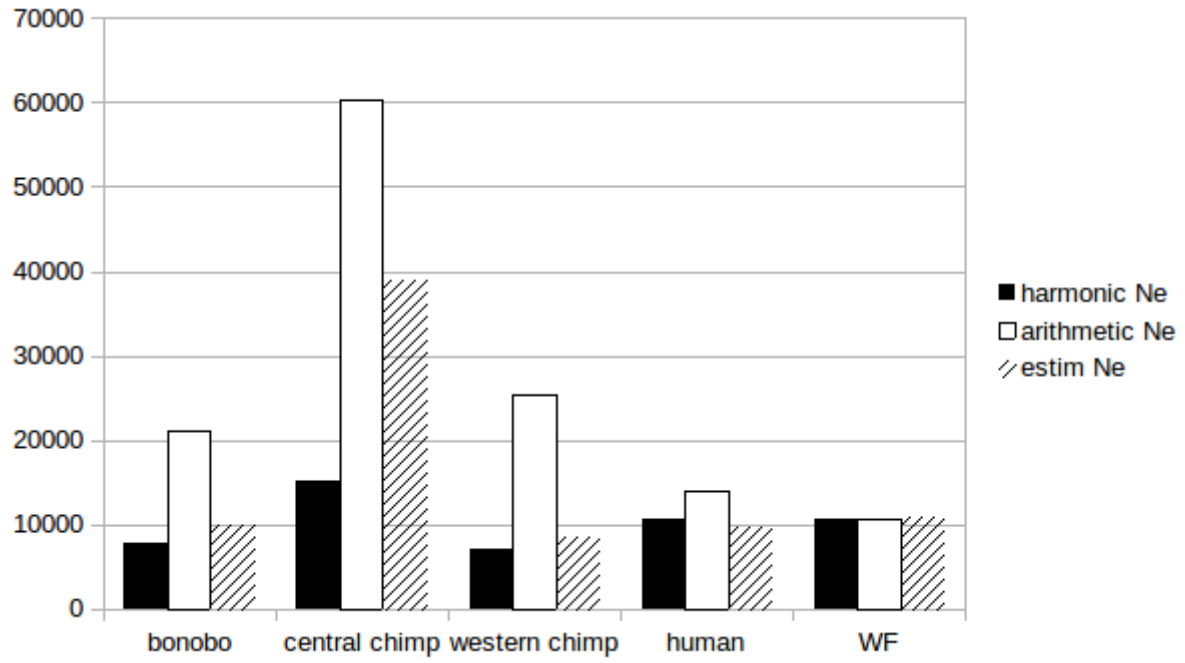
sfs_code 1 1 -N 4000 -Td 4.40435 P 0 0.08469 -Td 4.41685 P 0 12.9968119022 -Td 4.5048875 P 0 48.4082856364 -TE 4.53 -W 2 0 1 4 0.16 0.0004586657 --length 20 1440 --linkage p -1 -n 4 -a C --theta 0.00072173 --rho 0

Finally, the true uSFS for synonymous and nonsynonymous mutations for each population is extracted using the program SFS_convert (Hernandez 2008). Then we apply polyDFE as described in the main text including the bootstrap approach. We report inference performance using log2(estim/true) on a log-modulus scale. Here, estim is the estimated value, while true is the simulated value. We assess the inference accuracy of polyDFE estimates of the shape parameter (b) of the deleterious DFE (gamma distributed) and the mean population scaled effect size of new deleterious mutations ($S_d = 2N_e s_d$). To estimate the expected or true value of S_d we first computed the realized N_e by dividing the levels of synonymous diversity by four times the mutation rate ($\mu = 1.65 \times 10^{-8}$). Then the expected $S_d = 2N_e s_d$, where s_d is 3.2%.

Simulation Results

Inference Quality

As expected the value of the estimated N_e is close to the harmonic population size (computed from the demographic parameters described above) (Supplementary Figure B). One exception is central chimpanzees, where the old and large population expansion (that lasted 156,000 years) has increased the estimated N_e relative to the expected harmonic mean N_e. It is likely then that this expansion was overestimated (and/or the bottlenecks were stronger) given current levels of diversity. Nonetheless, these simulations are still valid to assess the ability of polyDFE to estimate the underlying DFE parameters.

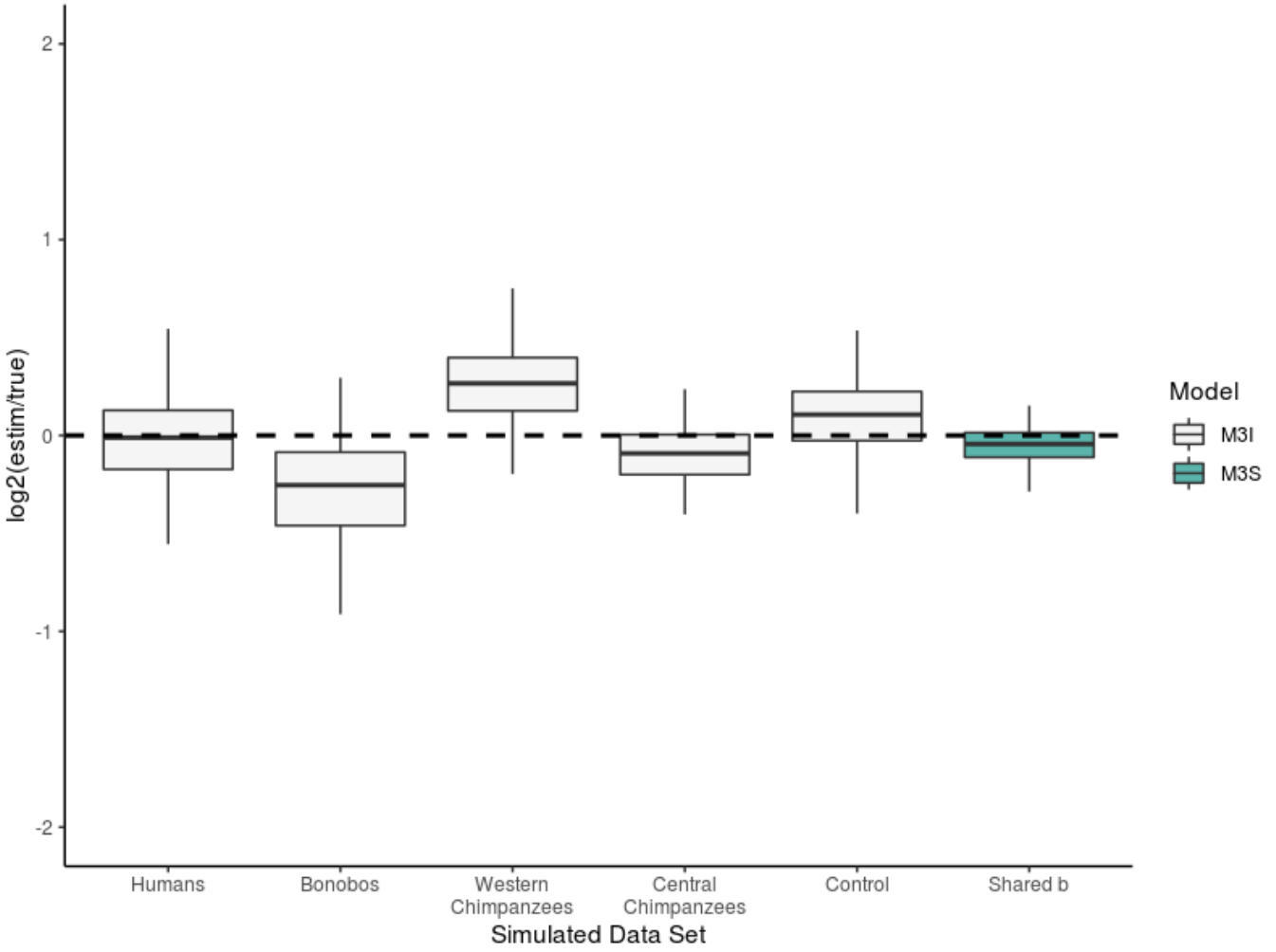


Supplementary Figure B. Recovering the simulated population size. Harmonic (in black) and arithmetic (in white) diploid *population size* in the last million years for the simulated demographies of bonobos, central chimpanzees, western chimpanzees and humans. For comparison, there is also a constant Wright-Fisher population (WF with N_e = 10,907). To estimate the N_e from synonymous diversity (estim Ne, hatching pattern) we divided Θ by 4μ, where $\mu = 1.65 \times 10^{-8}$.

When ranking populations by the estimated N_e (or equivalently by the levels of neutral diversity at the end of the simulation) we realize that bonobos show a larger N_e than humans. This is because the estimated N_e in bonobos is slightly above the expected harmonic mean, while in humans it is slightly below the harmonic mean. This can be explained by the fact that we are simulating a European population

while in our real sample we have 3 African individuals that have not undergone the recent out of Africa bottleneck.

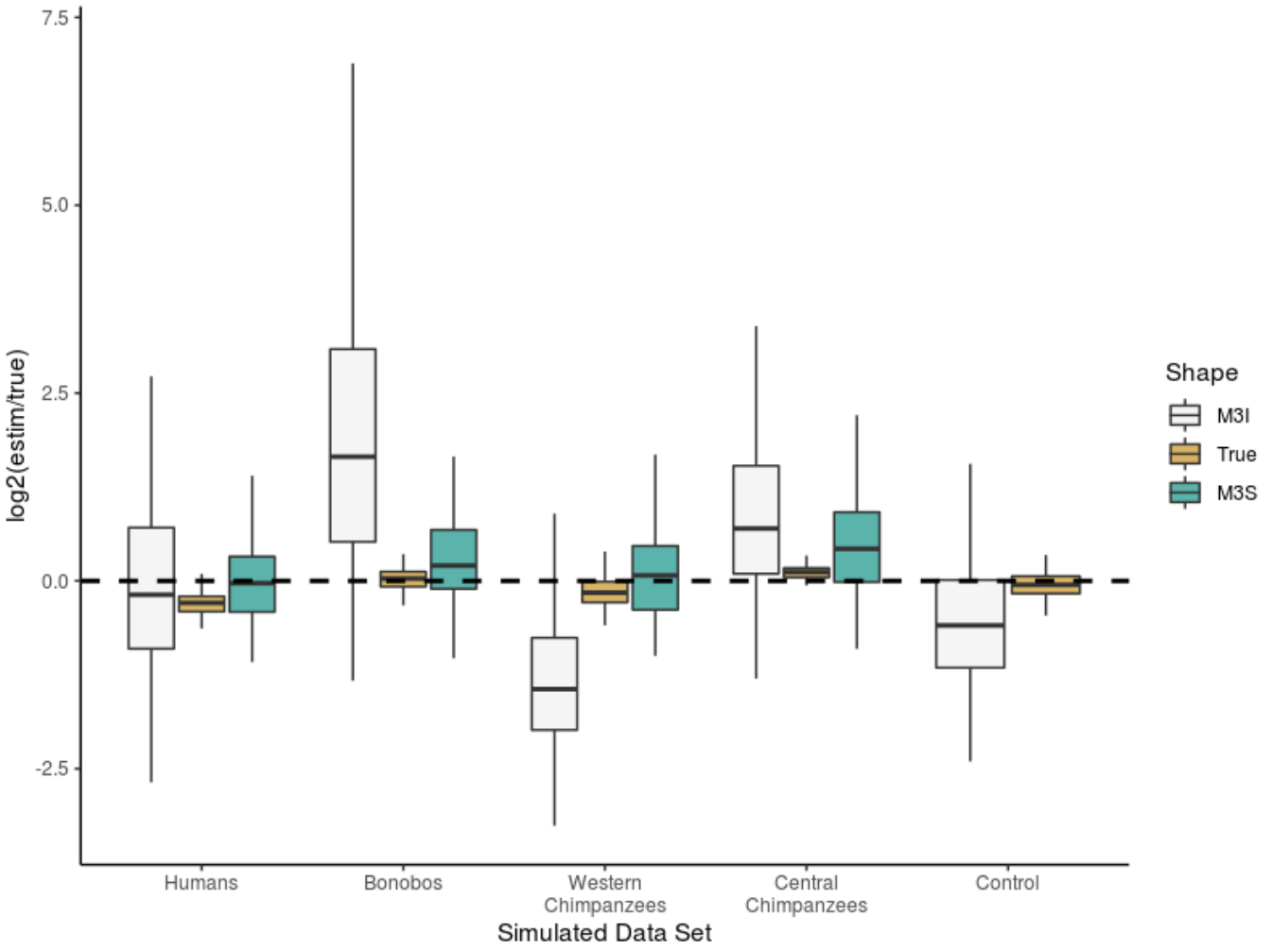
The true shape parameter is in general successfully recovered under M3I for the simulated great ape populations and for a control population evolving under a constant size (Supplementary Figure C). The shape parameter under model M3S is successfully recovered but less noisy than with model M3I.



Supplementary Figure C. Inference quality of the shape parameter of the deleterious DFE with model M3I and M3S (Table 2 for details). For comparison, we show the inference quality of a control Wright-Fisher population evolving under a constant population size. Note that the control population is not included when inferring the shared shape parameter with model M3S. Each boxplot represents 100 bootstrap replicates.

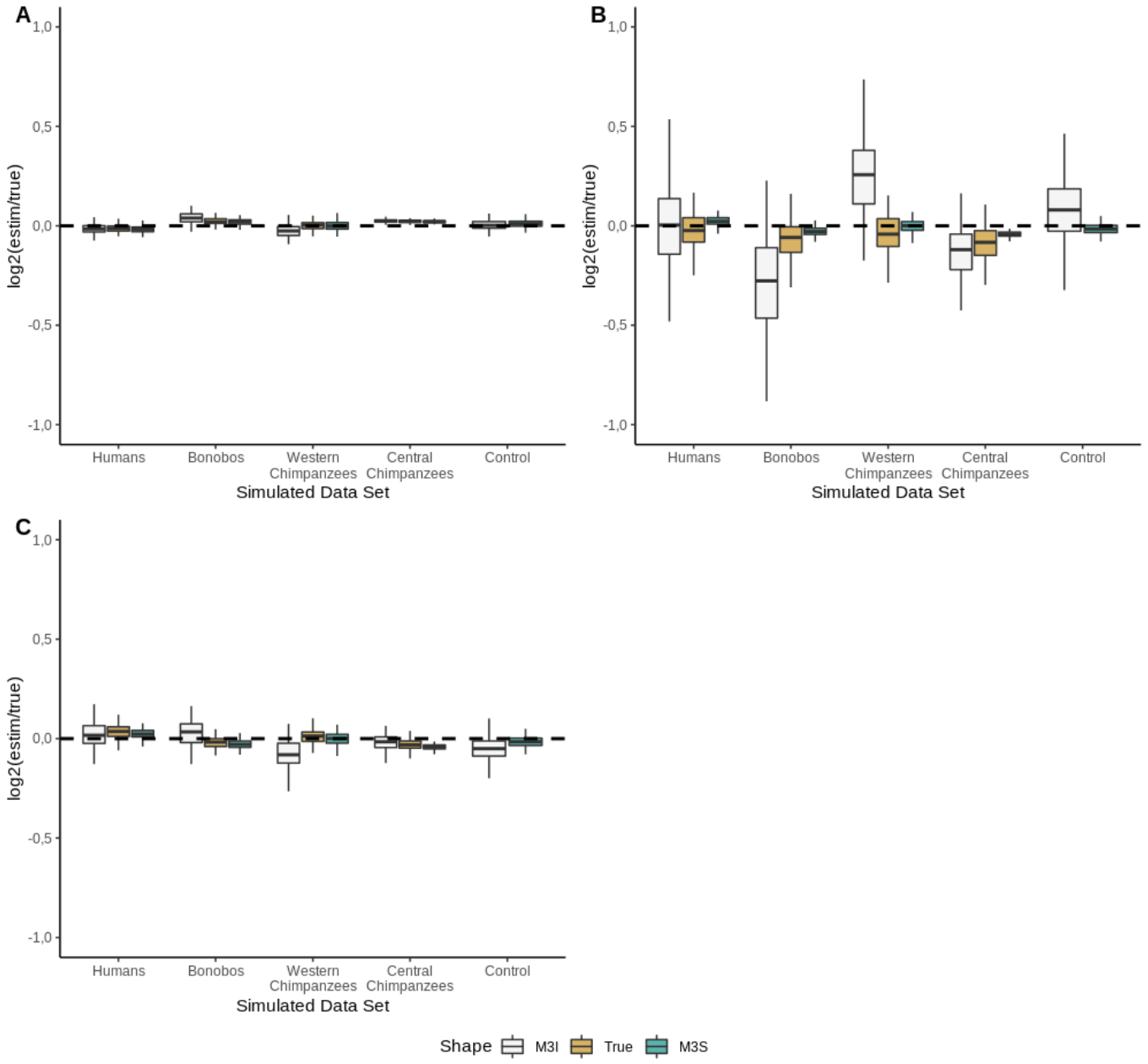
S_d estimates with model M3I are noisy and in general not accurate, except for humans. S_d estimates with model M3S are unbiased, but noisier than S_d estimates under the true shape parameter

(Supplementary Figure D). Note that the inference quality for S_d in bonobo and western chimpanzee is very similar to that for humans. This is an important result because we observe an increase in S_d estimates in bonobos and western chimpanzees relative to humans in the natural populations. Hence, our simulations show that this finding is not an artefact resulting from demography.



Supplementary Figure D. Inference quality of the mean population scaled effect size of new deleterious mutations, S_d, under model M3I and M3S (Table 2 for details). For comparison, we show the inference quality of a control Wright-Fisher population evolving under a constant population size and S_d estimated under the true shape parameter (b = 0.16) in gold. Note that the control population is not included when inferring the shared shape parameter with model M3S. Each boxplot represents 100 bootstrap replicates.

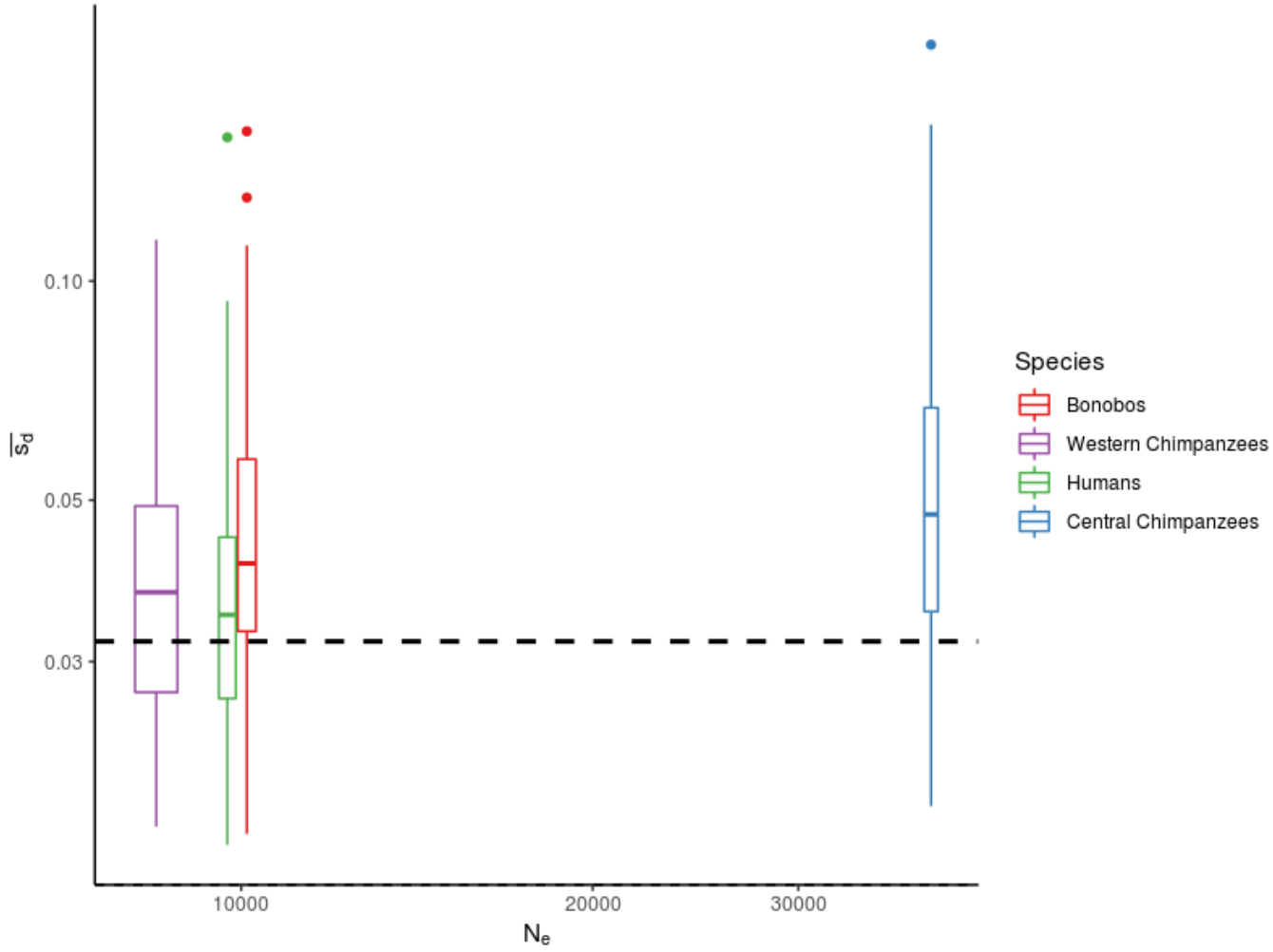
As reported before, the inference quality of the discretized S_d ranges is very high despite the statistical uncertainty surrounding the estimation of S_d and shape parameter (Supplementary Figure E). In other words, the discretized S_d ranges are fairly robust to model assumptions.



Supplementary Figure E. Inference quality of the proportion of strongly deleterious (A), weakly deleterious (B) and effectively neutral (A) mutations under model M3I and M3S (Table 2 for details). For comparison, we show the inference quality of a control Wright-Fisher population evolving under a constant population size and S_d estimated under the true shape parameter (b = 0.16) in gold. Note that the control population is not included when inferring the shared shape parameter with model M3S. Each boxplot represents 100 bootstrap replicates.

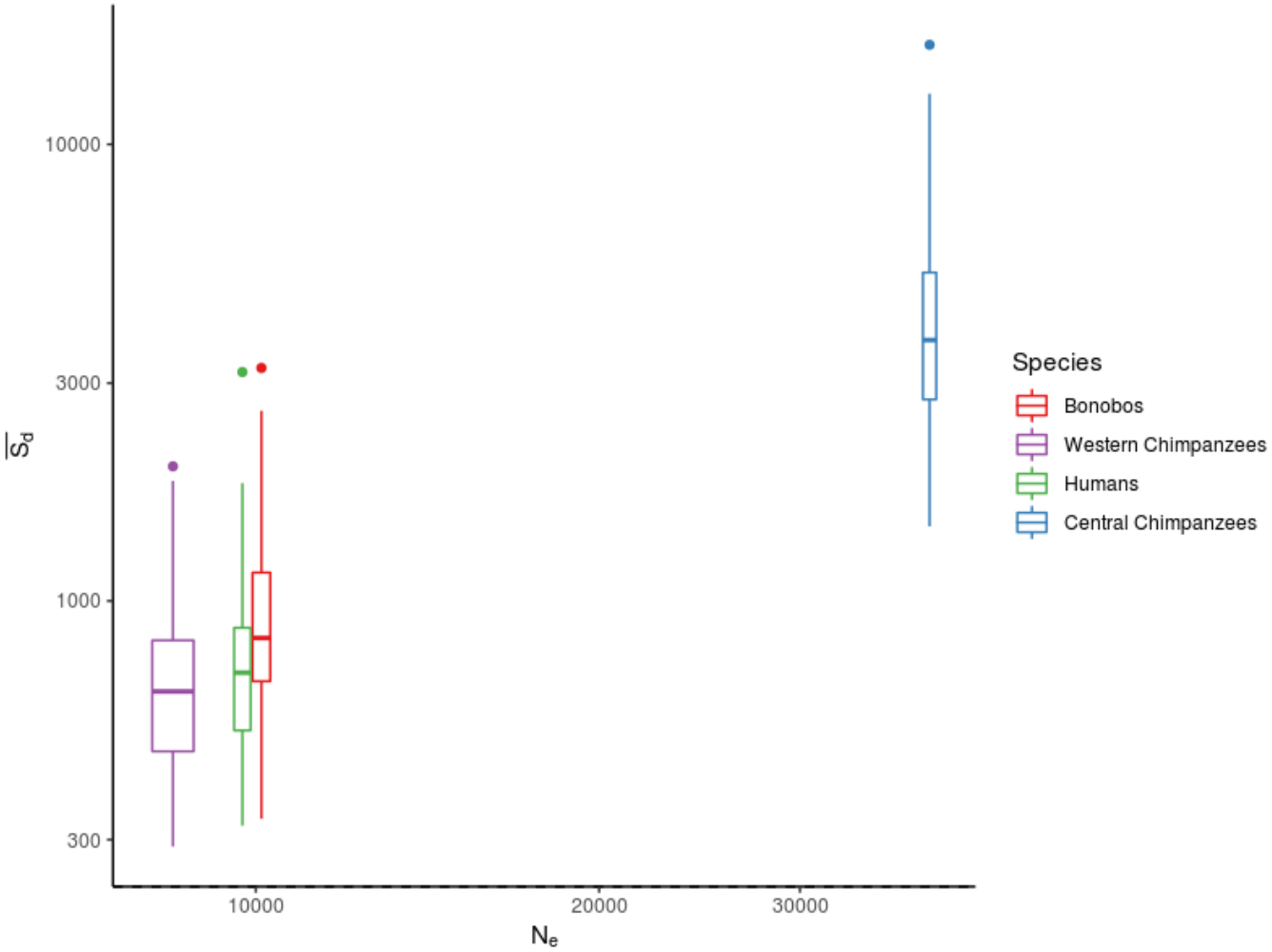
Relationship between N_e and the mean effect size of new deleterious mutations

We explored whether S_d estimates vary more than expected in great apes given their differences in population size. To do so, we investigate the slope in our simulated data set. Either polyDFE is introducing spurious co-variation between s_d and N_e or there is genuine co-variation in the data we analyzed. The slope between s_d estimates and N_e with the simulated data set is not significant (log-log slope = $7e^{-17}$ [$-2e^{-16}$, $3e^{-16}$]) while the observed slope within these four species is positive and significant (Table 3). This suggests that the reported co-variation is driven by genuine co-variation between the effect of new deleterious mutations.



Supplementary Figure F. Expected relationship between the estimated absolute mean s_d and N_e in a log-log scale. Parameters estimated under model M3S. The dashed line shows the expected true value of s_d. Each boxplot represents 100 bootstrap replicates.

Equivalently, the direct comparison of the slope between S_d and N_e in the simulated and real data confirms the variation in s_d in the great ape populations (simulated log-log slope = 1.17 [0.95, 1.27]; real log-log slope = 1.43 [1.29, 1.58]; Table 3) (Supplementary Figure G). Note that this is a conservative analysis because for the estimation of the real slope we only use bonobos, western chimpanzees, central chimpanzees and humans, and for this subset of species the slope is shallower due to the increase in S_d in bonobos and western chimpanzees. The slope between log(S_d) and log(N_e) for the natural populations excluding bonobos and western chimpanzees is 3.05 (2.94, 3.16) and for all the great apes is 1.90 (1.64, 2.16) (Table 3).



Supplementary Figure G. Expected relationship between the estimated absolute population scaled mean S_d and N_e in a log-log scale. Parameters estimated under model M3S. Each boxplot represents 100 bootstrap replicates.

References

Aberer A. J., and A. Stamatakis, 2013 Rapid forward-in-time simulation at the chromosome and genome level. BMC Bioinformatics 14: 216.

Hernandez R. D., 2008 A flexible forward simulator for populations subject to selection and demography. Bioinformatics 24: 2786–2787.

Kim B. Y., C. D. Huber, and K. E. Lohmueller, 2017 Inference of the Distribution of Selection Coefficients for New Nonsynonymous Mutations Using Large Samples. Genetics 206: 345–361.

Manuel M. de, M. Kuhlwilm, P. Frandsen, V. C. Sousa, T. Desai, *et al.*, 2016 Chimpanzee genomic diversity reveals ancient admixture with bonobos. Science 354: 477–481.

Tataru P., M. Mollion, S. Glémin, and T. Bataillon, 2017 Inference of Distribution of Fitness Effects and Proportion of Adaptive Substitutions from Polymorphism Data. Genetics 207: 1103–1119.