HUMAN EVOLUTION

Genomic inference of a severe human bottleneck during the Early to Middle Pleistocene transition

Wangjie Hu[1,2]†‡, Ziqian Hao[3]†, Pengyuan Du[1,3], Fabio Di Vincenzo[4], Giorgio Manzi[5], Jialong Cui[2], Yun-Xin Fu[6,7], Yi-Hsuan Pan[2]*, Haipeng Li[1,8]*

Population size history is essential for studying human evolution. However, ancient population size history during the Pleistocene is notoriously difficult to unravel. In this study, we developed a fast infinitesimal time coalescent process (FitCoal) to circumvent this difficulty and calculated the composite likelihood for present-day human genomic sequences of 3154 individuals. Results showed that human ancestors went through a severe population bottleneck with about 1280 breeding individuals between around 930,000 and 813,000 years ago. The bottleneck lasted for about 117,000 years and brought human ancestors close to extinction. This bottleneck is congruent with a substantial chronological gap in the available African and Eurasian fossil record. Our results provide new insights into our ancestry and suggest a coincident speciation event.

Although the lineage of humans is estimated to have separated from that of chimpanzees and bonobos more than 6 million years ago, anatomically modern humans (*Homo sapiens*) are estimated to have originated around 300 thousand to 200 thousand years before the present (kyr BP) in Africa (*1*–*3*). On the basis of present-day human genomic sequences, the recent population size histories (i.e., the dynamics of population size since the emergence of modern humans) have been intensively studied, revealing the worldwide spread of our ancestors (*4*–*8*). However, ancient population size history of the genus *Homo* during the Pleistocene is still poorly known, although it is essential for understanding the origin of the human lineage. It is likely to be very difficult or impossible to obtain ancient DNA from African *Homo* samples dated before the emergence of *H. sapiens*. It would be particularly notable if present-day human genomic sequences could be used to robustly infer both the recent and ancient population size histories of humankind. Thus, a new approach is needed to improve the inference accuracy of population size history.

Population size changes that occurred hundreds of thousands of years ago affected the rates of coalescence and thus have left their signatures in the site frequency spectrum (SFS) of genomic sequences. The SFS is the distribution of allele frequencies in the sequences, randomly collected from the present-day human population. Each SFS category contains a certain number of mutations of the same size. Because SFS is crucial for demographic inference (*6*, *8*–*12*) and construction of key summary statistics (*13*), many efforts have been devoted to deriving its analytical formulas (*14*–*17*). However, these formulae may not achieve the required accuracy because of propagation and accumulation of numerical errors resulting from their dependence on the joint probability density function of coalescent times (*16*–*18*).

In this study, to circumvent this numerical problem, we developed the fast infinitesimal time coalescent process (FitCoal) (Fig. 1) that analytically derives expected branch length for each SFS category under arbitrary demographic models. FitCoal does not need phased haplotype data and prior information on demography. The effects of sequencing errors or hitchhiking due to positive selection can be circumvented, largely by focusing on a subset of SFS that are less influenced by those factors. We used FitCoal to analyze a large number of present-day human genomic sequences from 10 African and 40 non-African populations. Results showed that our ancestors experienced a severe population bottleneck between about 930 and 813 kyr BP, most likely because of climatic changes. The average number of breeding individuals was only about 1280 during the bottleneck period. Our findings indicate that the severe bottleneck brought the ancestral human population close to extinction and completely reshaped present-day human genetic diversity.

Results

FitCoal

We developed FitCoal to determine the expected branch lengths for an SFS (Fig. 1). During FitCoal analysis of a sample, the time period in which the most recent common ancestor

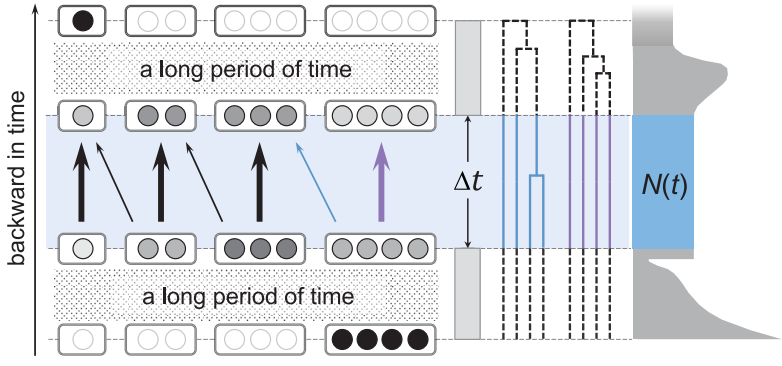

Fig. 1. Illustration of FitCoal. (**Left**) The backward process in which four lineages (represented by the four solid black circles at the bottom) coalesce into one (represented by the single solid black circle at the top) after passing through millions of infinitesimal time intervals (Δt). The area highlighted in blue shows the backward transformation process of different coalescent states with tiny probability changes in an infinitesimal time interval. Thick arrows indicate high transformation probabilities, and thin arrows indicate low transformation probabilities. The blue and purple arrows correlate to the two events in the middle pane represented by blue- and purple-colored lines. Each state is indicated with a box, in which one circle indicates one lineage. The boxes with solid black circles represent the states with the probability of 1. The boxes with empty circles represent the states with the probability of 0. The probabilities between 0 and 1 are represented by gray circles. (**Middle**) Hypothetical coalescent trees with branches of different states, indicating the number of lineages. Blue branches represent a transformation from four to three lineages. Purple branches indicate that no coalescent event occurred. (**Right**) The size of a theoretical population over time. The width of shadowed area denoted as N(t) indicates the effective population size (i.e., the number of breeding individuals) at time t. It is assumed that the effective population size remains unchanged within a Δt.

[1]CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China. [2]Key Laboratory of Brain Functional Genomics of Ministry of Education, School of Life Science, East China Normal University, Shanghai, China. [3]College of Artificial Intelligence and Big Data for Medical Sciences, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, China. [4]Natural History Museum, University of Florence, Florence, Italy. [5]Department of Environmental Biology, Sapienza University of Rome, Rome, Italy. [6]Department of Biostatistics and Data Science, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA. [7]Key Laboratory for Conservation and Utilization of Bioresources, Yunnan University, Kunming, China. [8]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China.
***Corresponding author. Email: yxpan@sat.ecnu.edu.cn (Y.-H.P.); lihaipeng@sinh.ac.cn (H.L.)**
†These authors contributed equally to this work.
‡Present address: Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

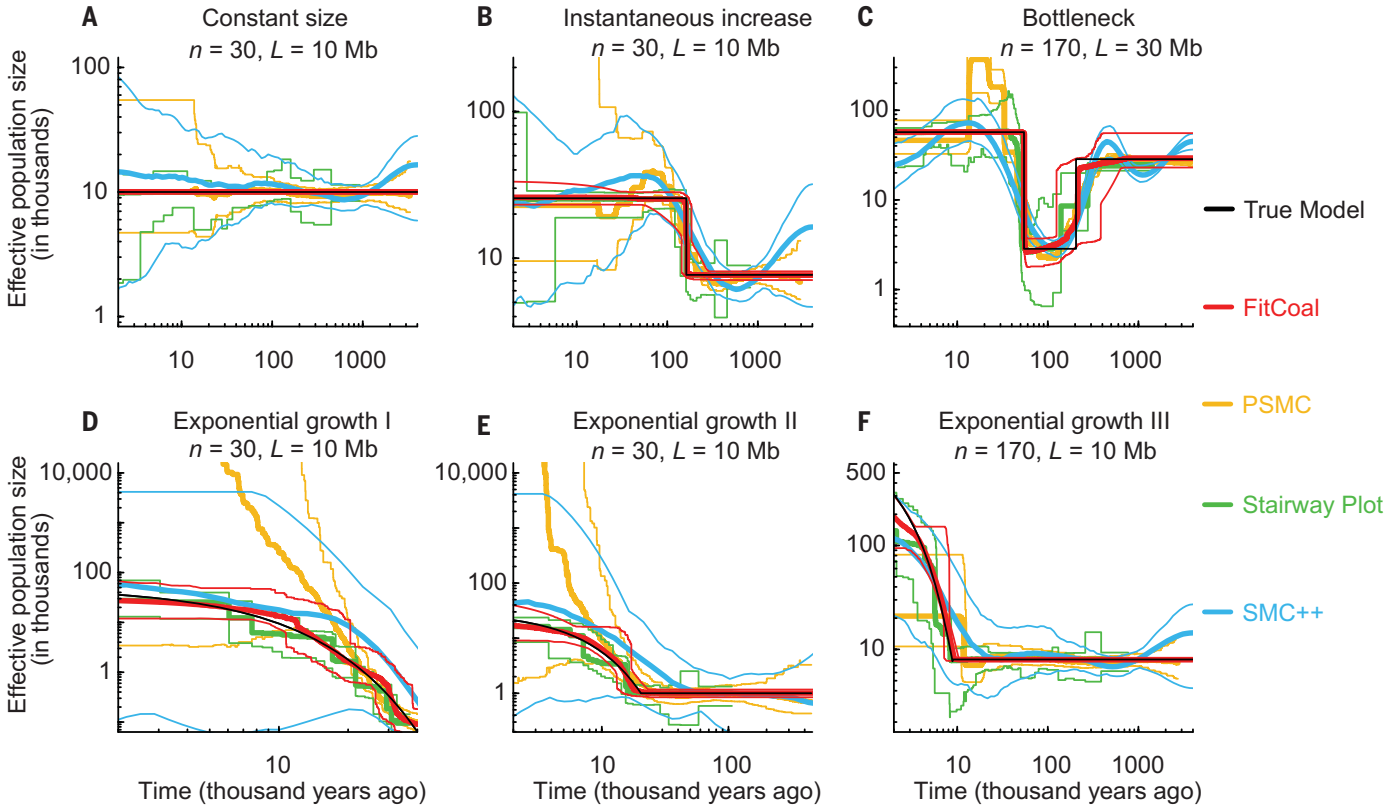

Fig. 2. Population size histories inferred by FitCoal, PSMC, Stairway Plot, and SMC++ with simulated samples. (**A**) Constant size model. (**B**) Instantaneous increase model. (**C**) Bottleneck model. (**D**) Exponential growth I model. (**E**) Exponential growth II model. (**F**) Exponential growth III model. In all panels, thin black lines indicate the true models. Thick red lines indicate the medians of FitCoal-inferred population size histories; thin red lines represent 2.5 and 97.5 percentiles of FitCoal-inferred population size histories. Yellow, green, and blue lines indicate results obtained with PSMC, Stairway Plot, and SMC++, respectively. The mutation rate is assumed to be 1.2×10^{-8} per base per generation, and a generation time is assumed to be 24 years. *n* is the number of simulated sequences, and *L* is the length in Mb of each simulated sequence.

originated was partitioned into as many time intervals as needed, such that each time interval (Δt) was very small (e.g., 1 month or 1 year). During each time interval, the population size was assumed to be constant. The probabilities of all coalescent states (i.e., all possible ancestral lineages) were calculated backward in time. For each state, the branch length during a time interval was calculated by multiplying its probability with population size and then transformed to determine the expected branch lengths. Because the expected branch length of an SFS category during a time interval was precalculated, FitCoal could be very fast.

FitCoal demographic inference

After the expected branch lengths were determined, the composite likelihood of an SFS (*6*, *9*, *19*) was calculated. FitCoal is effective for a wide range of sample sizes in the calculation of the composite likelihood of a given SFS and is much more accurate than simulation approaches (fig. S1). When inferring population size history, the likelihood was maximized in a wide range of demographic scenarios. Moreover, both instantaneous and long-term exponential population changes were considered. Similar to previous studies (*6*, *19*), the likelihood of the SFS was first maximized with the constant size model, followed by repeated maximization of the likelihood with increased number of inference time intervals until the best model was found.

Demographic inference from simulated data

The accuracy of FitCoal demographic inference was evaluated by simulation and comparison of results with those of PSMC (pairwise sequentially Markovian coalescent), Stairway Plot, and SMC++ methods (*6*–*8*) (Fig. 2). To ensure fair comparisons, we tested six demographic models by simulating 200 independent datasets for each model, as described previously (*6*), with the assumption that a generation time is 24 years (*6*, *20*) and that the mutation rate is 1.2×10^{-8} per site per generation (*6*, *21*).

Results showed that the medians of FitCoal-inferred population size histories were almost identical to the true models, and the 95% confidence intervals of FitCoal inference were narrower than those of PSMC, Stairway Plot, and SMC++ (Fig. 2). FitCoal inference accuracy could be further improved by increasing sample size and lengths of sequences (fig. S2). The proportion of the most recent changes in population size inferred from the six models showed that FitCoal could distinguish between instantaneous and exponential changes (table S1). Overall, our results confirmed that SFSs could be used to estimate population size histories (*22*).

It has been suggested that a population size history could be inferred by using a subset of SFS or a collapsed SFS (*6*, *19*); the latter is an SFS with high frequency mutations combined into one category. Results of simulations showed that the FitCoal could still accurately determine a population size history even when a portion (10 to 90%) of an SFS was truncated (i.e., excluded for analysis) (figs. S3 to S5), thus reducing the impact of confounding factors, such as hitchhiking effect due to positive selection (fig. S6) or sequencing errors on FitCoal analyses.

Demographic inference of African populations

To infer population size histories of African populations, seven African populations in the 1000 Genomes Project (1000GP) (*23*) and three African populations in the Human Genome Diversity Project–Centre d'Etude du Polymorphisme Humain (HGDP-CEPH) panel (*24*) were analyzed by the FitCoal (tables S2 to S4). Only autosomal noncoding regions were used to partially avoid the effect of purifying selection. To avoid hitchhiking effect due to positive

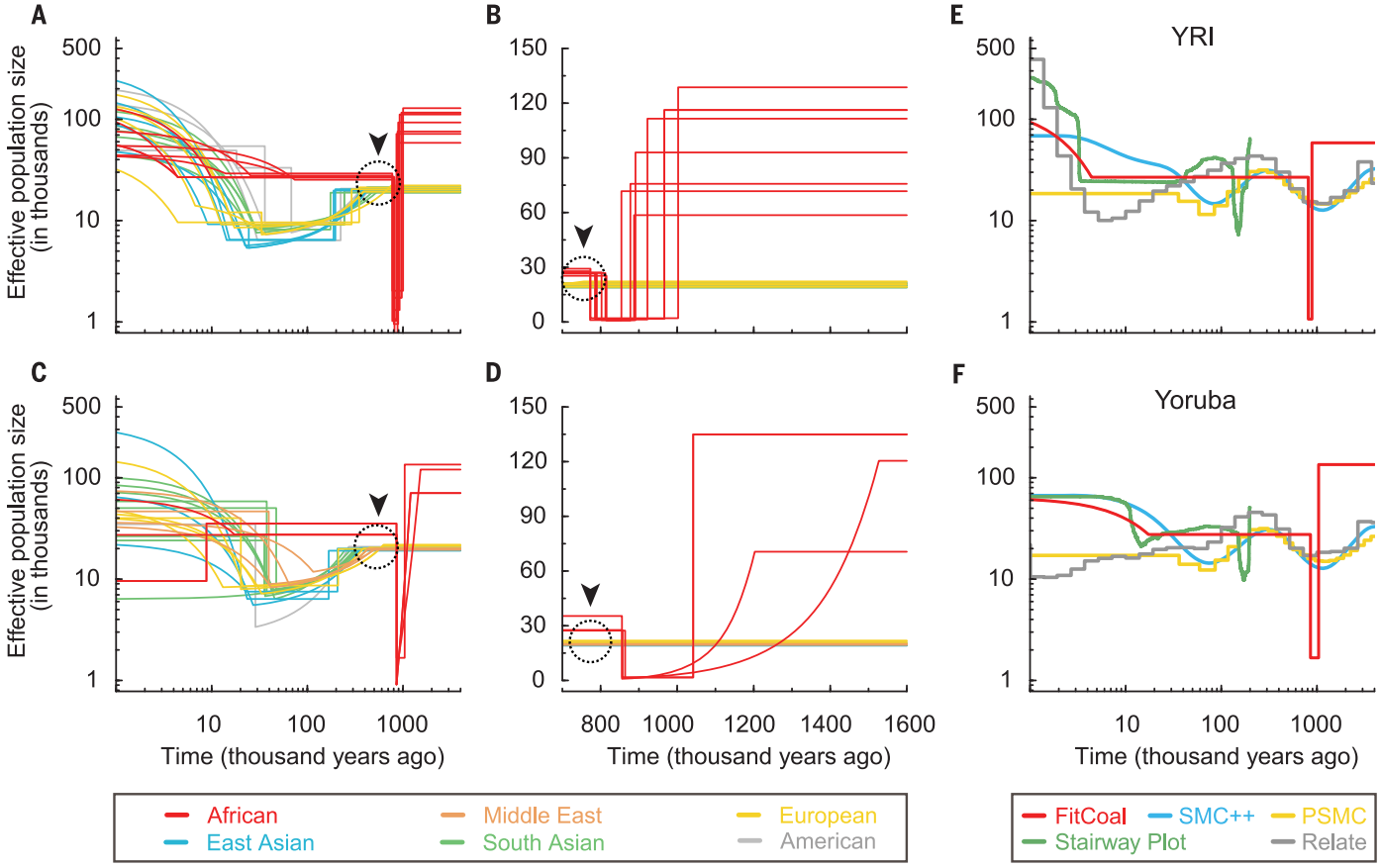

Fig. 3. Histories of human populations in 1000GP and HGPD-CEPH genomic datasets inferred by FitCoal, SMC++, Stairway Plot, PSMC, and Relate. The mutation rate is assumed to be 1.2×10^{-8} per base per generation, and a generation time is assumed to be 24 years. (**A**) Inferred population size histories of 26 populations in 1000GP. (**B**) Linear-scaled estimation of sizes over time of populations in 1000GP during the severe bottleneck period. (**C**) Inferred population size histories of 24 populations in the HGPD-CEPH panel. (**D**) Linear-scaled estimation of sizes over time of populations in the HGPD-CEPH panel during the severe bottleneck period. (**E** and **F**) Comparison of population size histories of African YRI and Yoruba populations inferred by FitCoal, SMC++, Stairway Plot, PSMC, and Relate. Only the population size histories up to 200 kyr BP were analyzed by Stairway Plot. In (A) to (D), colored lines indicate the following: red, African populations; brown, Middle East populations; yellow, European populations; blue, East Asian populations; green, Central or South Asian populations; and gray, American populations. Black dashed circles with arrow heads represent the discrepancy in the population size between African agriculturalist and non-African populations. In (E) and (F), red, blue, yellow, green, and gray lines indicate results obtained with FitCoal, SMC++, PSMC, Stairway Plot, and Relate, respectively.

selection (*25*), high-frequency mutations were excluded from the analysis. Results showed that all 10 African populations went through a severe bottleneck (Fig. 3 and figs. S7 and S8). The bottleneck was estimated to persist for 117 kyr, from 930 ± 23.52 (SEM) (range, 854 to 1042) to 813 ± 11.02 (SEM) (range, 772 to 864) kyr BP. The average effective population size (i.e., the number of breeding individuals) (*26*) during the bottleneck period was determined to be 1280 ± 131 (SEM) (range, 770 to 2030), which was only 1.3% of its ancestral size (98,130 ± 8720; range, 58,600 to 135,000). To evaluate the impact of the bottleneck on current human genetic diversity, we analyzed the expected pairwise nucleotide diversity. Results showed that 65.85% of current human genetic diversity was lost because of the bottleneck.

Demographic inference of non-African populations

The severe bottleneck was not directly detected in all 19 non-African populations in 1000GP and the 21 non-African populations in the HGDP-CEPH panel (Fig. 3, A to D; figs. S7 and S8; and tables S2, S5, and S6). The average ancestral population sizes of the populations in the two datasets were 20,260 (range, 18,850 to 22,220) and 20,030 (range, 19,060 to 21,850), respectively, similar to those determined in previous studies (*7*, *8*, *24*). The estimated population size started to decline around 368 (range, 175 to 756) and 367 (range, 167 to 628) kyr BP, respectively, which is consistent with previous findings that African and non-African divergence occurred much earlier than the out-of-Africa dispersal (*7*, *8*, *23*, *24*). The inferred out-of-Africa dispersal and the recent population size expansion and reduction are consistent with those of previous studies (*5*–*8*, *23*, *24*).

Severe bottleneck during the Early to Middle Pleistocene transition

The ancient severe bottleneck was directly detected in each of the 10 African populations but in none of the 40 non-African populations. To investigate this discrepancy, we performed simulations with three demographic models, designated bottlenecks I, II, and III (Fig. 4, A to C, and figs. S9 and S10). Bottleneck I simulated the population size history of African agriculturalist populations with the ancient severe bottleneck, and bottlenecks II and III simulated that of non-African populations without and with the ancient severe bottleneck, respectively. Both bottlenecks I and II were inferred precisely in all simulated data sets (tables S7 to S9). However, no ancient severe bottleneck was detected in bottleneck III simulations, which indicates that the out-of-Africa dispersal hinders the chance of discovering the ancient severe bottleneck. Furthermore, the ancient severe bottleneck was found to cause a discrepancy in the estimation of the population size between the bottleneck III model and the inferred population size history after the bottleneck was relieved, which suggests a hidden effect

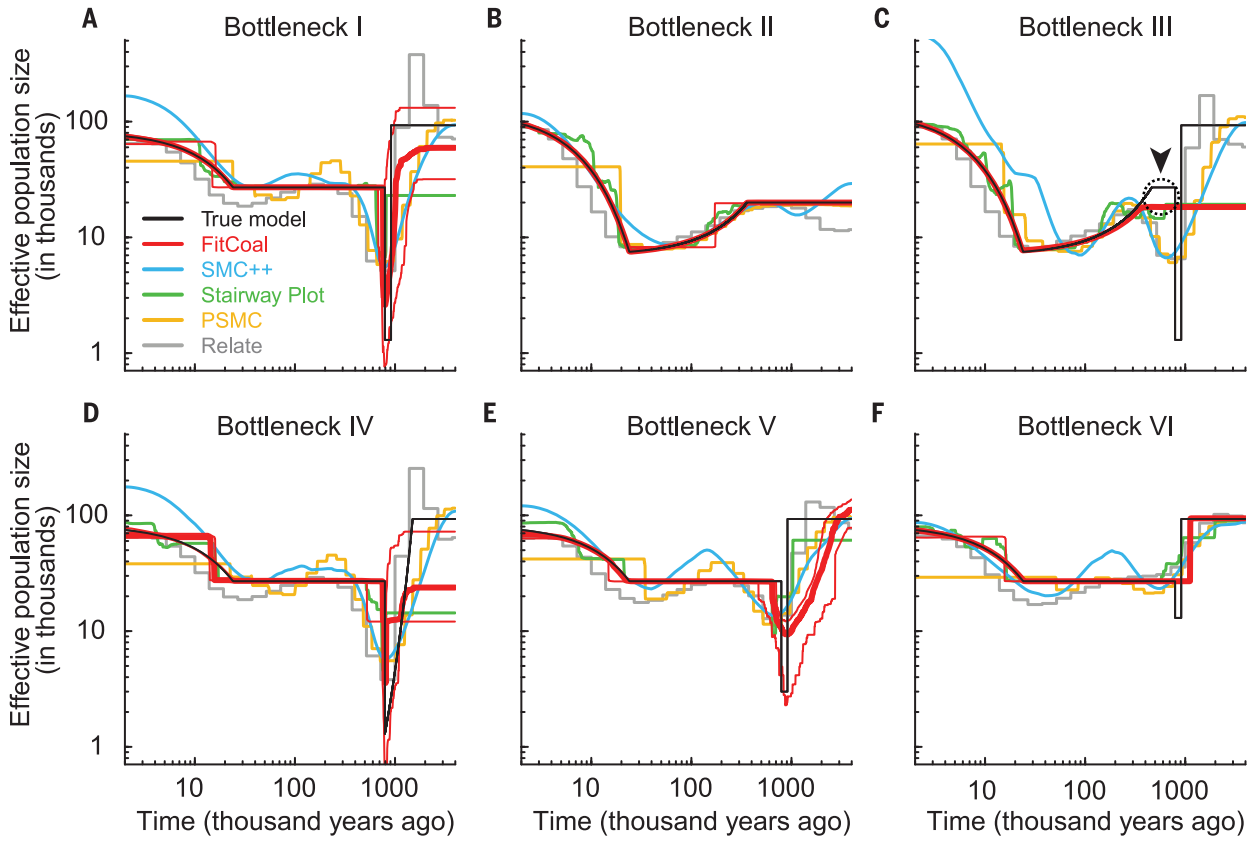

Fig. 4. Verification of the severe bottleneck. (**A**) Bottleneck I model, mimicking the true population size history of African agriculturalist populations. (**B**) Bottleneck II model, mimicking the inferred population size history of non-African populations. (**C**) Bottleneck III model, mimicking the true population size history of non-African populations. The gap between actual and FitCoal estimated population size is indicated by the black dashed circle and arrowhead. (**D**) Bottleneck IV model, mimicking a population with an exponential reduction in size 1.5 million years ago. (**E**) Bottleneck V model, mimicking a population with a moderate bottleneck. (**F**) Bottleneck VI model, mimicking a population with a weak bottleneck. Black lines represent the true models. Thick red lines represent the medians of FitCoal estimated population sizes over time; thin red lines represent 2.5 and 97.5 percentiles of FitCoal estimated population sizes over time. Blue, green, yellow, and gray lines represent the medians of 10 runs each of SMC++, Stairway Plot, PSMC, and Relate. The mutation rate is assumed to be 1.2×10^{-8} per base per generation, and a generation time is assumed to be 24 years. Numbers of simulated sequences are 188 in bottlenecks I, IV, V, and VI and 194 in bottlenecks II and III. The lengths of simulated sequences are 800 Mb each.

of the ancient severe bottleneck on non-African populations (Fig. 4C and figs. S9C and S10C).

After the bottleneck was relieved, the average population size of non-African populations in 1000GP was 20,260, and that of those in HGDP-CEHP was 20,030. For African agriculturalist populations, the average population size in 1000GP was 27,080, and that of those in HGDP-CEHP was 27,440. This population size difference of 7020 (Fig. 3, A and C) is likely due to the hidden effect of the ancient severe bottleneck on non-African populations. Because the out-of-Africa dispersal existed in non-African populations but not in African populations, African populations had more lineages remaining to be traced back to the ancient severe bottleneck (fig. S11). In the analysis of the African YRI (Yoruba in Ibadan, Nigeria) population, the minimum sample size was three individuals for detection of the severe bottleneck (fig. S12).

Because the signal for the existence of the severe bottleneck was too weak to be detected in non-African populations by FitCoal, we performed an extended FitCoal analysis. To eliminate noise effects resulting from problems such as sequencing or overfitting errors on the inference of population size history, we used the FitCoal-inferred recent population size history as a starting point for size estimation of an ancient population. With this modification, all 19 non-African populations in 1000GP were found to have gone through the severe bottleneck with approximately 1450 individuals between 921 and 785 kyr BP (fig. S13 and table S10). This result is consistent with that obtained with African populations.

To further examine the severe bottleneck, we simulated a slow population reduction starting 1.5 million years ago (Fig. 4D). The FitCoal-inferred population size histories were different from those observed in 1000GP and HDGP-CEHP populations, which supports the hypothesis that a sudden size reduction occurred at the beginning of the bottleneck. Results of simulations were similar to that of the observed cases (Fig. 3, E and F) in that they also showed that PSMC, SMC++, and Relate methods (*7*, *8*, *27*) underestimated the severity of the ancient bottleneck (Fig. 4 and figs. S14 to S17). Moreover, the inferred population declines shown in Fig. 4A were more severe than those inferred from the real data because the simulated data were generated under the assumption of a neutrally evolved single population and a homogeneous recombination rate, whereas humans evolved with subpopulations and a heterogeneous recombination rate (*28*, *29*). FitCoal was not found to overestimate such severity (Fig. 4) or to falsely detect a bottleneck when examining the effects of continuous and pulsed introgressions among existing and ghost populations (figs. S18 to S21). Therefore, the discovery of the ancient severe bottleneck was not due to overfitting the data in FitCoal analyses.

Discussion

In this study, we developed FitCoal, a model-flexible method for demographic inference. One key feature of FitCoal is that the expected

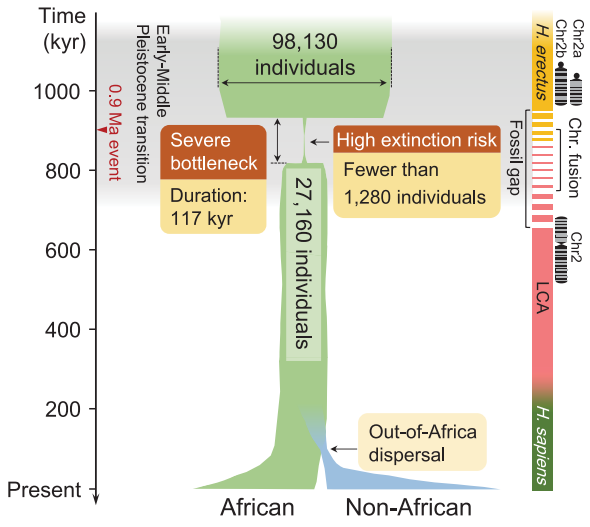

Fig. 5. Schematic diagram of human population size history. Both African (light green) and non-African (light blue) populations are presented. The width of the boxes represents the effective population size (i.e., the number of breeding individuals) with naturally occurred fluctuations. The occurrence time of the out-of-Africa dispersal and the divergence between African and non-African populations are indicated. The gray-shaded time duration indicates the Early to Middle Pleistocene transition between 1250 and 700 kyr BP. The red arrow indicates the peak of glaciation during the transition (i.e., the 0.9 Ma event). The ancient severe bottleneck inferred in this study is highlighted. The gap in the available African hominin fossil record and an indicative chronology for *H. erectus*, the LCA, and *H. sapiens* are shown. The estimated time period in which two ancestral chromosomes (chromosome, Chr.) fused to become one is also shown on the right.

branch lengths can be accurately determined for an SFS under an arbitrary demographic model, which allows precise calculation of the likelihood. Analyses by FitCoal are, in most cases, less time consuming than those by other methods such as PSMC, SMC++, Stairway Plot, and Relate (28/32 = 87.5%) (tables S11 and S12). By discarding rare and high-frequency mutations, FitCoal can avoid the effects of sequencing errors or hitchhiking due to positive selection without losing its inference accuracy. Because both instantaneous and exponential changes are allowed within each inference time interval, FitCoal can reveal the dynamic of population size precisely. Because coalescent events become rare when tracing backward in time, the length of inference time interval is usually set to increase progressively (*6*–*8*). Although this strategy can capture recent population size changes, it may miss ancient ones. Therefore, FitCoal inference time intervals are allowed to vary during demographic inference, and FitCoal can make a fast and accurate inference of recent and ancient population size histories.

The most important discovery with FitCoal is that human ancestors went through a severe bottleneck in the late Lower Pleistocene (Fig. 5). This ancient severe bottleneck was directly found in all 10 African populations, but only a weak signal of the existence of such was detected in all 40 non-African populations. This observation is consistent with the coalescent theory and the occurrence of the out-of-Africa dispersal. Results of our large-scale simulations demonstrated that FitCoal did not falsely infer the bottleneck because of positive selections (figs. S6 and S22) or population structure (fig. S18 to S21). Because we observed no overfitting cases and results obtained by examining different sets of genomic regions (Fig. 3, A and C, and fig. S23) were similar, the existence of the ancient severe bottleneck was ascertained.

Our results indicate that the ancient severe bottleneck lasted for approximately 117 kyr (Fig. 5), and that about 98.7% of human ancestors were lost at the beginning of the bottleneck, thus threatening our ancestors with extinction. The estimated effective population size during the bottleneck period was only 1280 breeding individuals, which was comparable to the effective population sizes of other endangered mammals (*30*, *31*). This size (1280) might have been overestimated because of hidden population structure (*32*). Naturally occurring population size fluctuations might have further increased the extinction risk for our ancestors during the bottleneck. The bottleneck could also have increased the inbreeding level of our ancestors, thus contributing to the 65.85% loss in present-day human genetic diversity.

The ancient population size reduction that occurred around 930 kyr BP was likely driven by climatic changes during the Early to Middle Pleistocene transition (*33*, *34*). During this transition period known as the "0.9 Ma event" (Ma, million years ago), glaciations were changed from predominantly short-term to long-term events with more extreme thermic intensity, especially at the peak of glaciation. This event resulted in a decrease in marine surface temperature to the lowest that occurred during the entire transition period (*33*), with an inferred long period of drought and extensive wildlife turnover in Africa and Eurasia (*35*).

The existence of the ancient severe bottleneck could explain the extreme scarcity of the available hominin fossil record in Africa and Eurasia between 950 and 650 kyr BP (Fig. 5 and fig. S24). In Africa, only a few fossil specimens dated in this time period have been found, including the cranial fragments from Gombore in Ethiopia and the fossil samples from Tighenif in Algeria (*36*, *37*). Although the taxonomic statuses of these fossils are still not clear, they have features resembling those of later fossils attributed to *Homo heidelbergensis*. They are different from the coeval *Homo antecessor* from a paleoanthropological site in Spain (Atapuerca, Gran Dolina), and some scholars considered *H. antecessor* as a possible alternative for the last common ancestor (LCA) (*38*). During the same chronological interval, the East Asian fossil record contains specimens identified as *Homo erectus* (*39*). It does not appear that East Asian *H. erectus* is connected to the ancient severe bottleneck because it is unlikely to have contributed to the lineage leading to modern humans (*38*). In addition, coincident with this bottleneck, two ancestral chromosomes are believed to have fused to form chromosome 2 in humans around 900 to 740 kyr BP (*40*, *41*). Therefore, the ancient severe bottleneck possibly marks a speciation event leading to the emergence of the LCA shared by Denisovans, Neanderthals, and modern humans, whose divergence has been dated to about 765 to 550 kyr BP (*38*, *42*, *43*).

A rapid population recovery was detected in all 10 African populations with a 20-fold increase in size around 813 kyr BP. Control of fire could be part of the explanation for this population expansion, which is shown by the early archaeological evidence found in Israel dated about 790 kyr BP (*44*). Other factors, such as climatic changes (*33*, *34*), might also be a driving force for this rapid population recovery.

The ancient severe bottleneck was not detected in previous SFS-based analyses (*6*, *10*, *12*, *14*). This failure might be due to the use of predefined demographic models. In this study, we found that the likelihood must be accurately calculated to detect the severe bottleneck (fig. S1). The use of other methods such as Stairway Plot—which may not have sufficient resolution power for estimation of ancient population size history—is another possible reason for the failure (*6*).

Our study revealed that an extremely small human population lasted for about 117 kyr around 930 to 813 kyr BP. Many questions remain unanswered, such as where these individuals lived, how they overcame the catastrophic climate changes, and how the ancient population remained so small for so long. Further studies are warranted to investigate these matters to obtain a more detailed picture of human evolution during the Early to Middle Pleistocene transition.

REFERENCES AND NOTES

1. J. Galway-Witham, C. Stringer, *Science* **360**, 1296–1298 (2018).
2. C. M. Schlebusch *et al.*, *Science* **358**, 652–655 (2017).
3. A. Meneganzin, T. Pievani, G. Manzi, *Evol. Anthropol.* **31**, 199–212 (2022).
4. M. Stoneking, J. Krause, *Nat. Rev. Genet.* **12**, 603–614 (2011).
5. S. Ramachandran *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15942–15947 (2005).
6. X. Liu, Y. X. Fu, *Nat. Genet.* **47**, 555–559 (2015).
7. H. Li, R. Durbin, *Nature* **475**, 493–496 (2011).

8. J. Terhorst, J. A. Kamm, Y. S. Song, *Nat. Genet.* **49**, 303–309 (2017).
9. H. Li, W. Stephan, *PLOS Genet.* **2**, e166 (2006).
10. L. Excoffier, I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, M. Foll, *PLOS Genet.* **9**, e1003905 (2013).
11. R. C. Griffiths, S. Tavaré, *Math. Comput. Model.* **23**, 141–158 (1996).
12. R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, *PLOS Genet.* **5**, e1000695 (2009).
13. Y. X. Fu, *Theor. Popul. Biol.* **145**, 95–108 (2022).
14. J. Jouganous, W. Long, A. P. Ragsdale, S. Gravel, *Genetics* **206**, 1549–1567 (2017).
15. Y. X. Fu, *Theor. Popul. Biol.* **48**, 172–197 (1995).
16. A. Polanski, M. Kimmel, *Genetics* **165**, 427–436 (2003).
17. D. Zivković, T. Wiehe, *Genetics* **180**, 341–357 (2008).
18. A. Polanski, A. Bobrowski, M. Kimmel, *Theor. Popul. Biol.* **63**, 33–40 (2003).
19. A. Bhaskar, Y. X. R. Wang, Y. S. Song, *Genome Res.* **25**, 268–279 (2015).
20. A. Scally, R. Durbin, *Nat. Rev. Genet.* **13**, 745–753 (2012).
21. A. Kong *et al*., *Nature* **488**, 471–475 (2012).
22. A. Bhaskar, Y. S. Song, *Ann. Stat.* **42**, 2469–2493 (2014).
23. The 1000 Genomes Project Consortium, *Nature* **526**, 68–74 (2015).
24. A. Bergström *et al*., *Science* **367**, eaay5012 (2020).
25. J. C. Fay, C. I. Wu, *Genetics* **155**, 1405–1413 (2000).
26. H. C. Harpending *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1961–1967 (1998).
27. L. Speidel, M. Forest, S. Shi, S. R. Myers, *Nat. Genet.* **51**, 1321–1329 (2019).
28. E. M. L. Scerri *et al*., *Trends Ecol. Evol.* **33**, 582–594 (2018).
29. Z. Hao, P. Du, Y. H. Pan, H. Li, *Hum. Genet.* **141**, 273–281 (2022).
30. H. Li *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14079–14084 (2016).
31. Y. X. Chen *et al*., *Zool. Res.* **43**, 523–527 (2022).
32. J. Wakeley, *Genetics* **153**, 1863–1871 (1999).
33. M. J. Head, P. L. Gibbard, in *Early-Middle Pleistocene Transitions: The Land–Ocean Evidence*, M. J. Head, P. L. Gibbard, Eds. (Geological Society of London, 2005), pp. 1–18.
34. P. U. Clark *et al*., *Quat. Sci. Rev.* **25**, 3150–3184 (2006).
35. M. J. Head, B. Pillans, S. A. Farquhar, *Episodes* **31**, 255–259 (2008).
36. A. Profico, F. Di Vincenzo, L. Gagliardi, M. Piperno, G. Manzi, *J. Anthropol. Sci.* **94**, 41–63 (2016).
37. A. Mounier, F. Marchal, S. Condemi, *J. Hum. Evol.* **56**, 219–246 (2009).
38. A. Bergström, C. Stringer, M. Hajdinjak, E. M. L. Scerri, P. Skoglund, *Nature* **590**, 229–237 (2021).
39. J.-J. Bahain *et al*., *Anthropologie* **121**, 215–233 (2017).
40. B. Poszewiecka, K. Gogolewski, P. Stankiewicz, A. Gambin, *BMC Genomics* **23**, 616 (2022).
41. T. R. Dreszer, G. D. Wall, D. Haussler, K. S. Pollard, *Genome Res.* **17**, 1420–1430 (2007).
42. K. Prüfer *et al*., *Nature* **505**, 43–49 (2014).
43. D. Reich *et al*., *Nature* **468**, 1053–1060 (2010).
44. N. Goren-Inbar *et al*., *Science* **304**, 725–727 (2004).
45. H. Li, Genomic inference of a severe human bottleneck during the Early to Middle Pleistocene transition, version 3, Mendeley Data, 2023; https://doi.org/10.17632/xmf5r8nzrn.3.
46. W. Hu *et al*., FitCoal v1.1., Zenodo, 2023; https://doi.org/10.5281/zenodo.7857456.

ACKNOWLEDGMENTS

We thank D. Živković for sharing his codes to calculate the expected branch lengths, X. Liu for providing simulation results, and Bio-Med Big Data Center at Shanghai Institute of Nutrition and Health for providing a computing facility. **Funding:** National Natural Science Foundation of China grants 32270674, 91131010, and 91731304 (H.L.); National Natural Science Foundation of China grant 31100273 (Y.H.P.); National Natural Science Foundation of China grants 91631304 and 32130011 (Y.X.F.); National Natural Science Foundation of China grant 82171801 (Z.H.); Strategic Priority Research Program of the Chinese Academy of Sciences grant XDPB17 (H.L.); National Key Research and Development Project grant 2022YFF1203202 (H.L.); National Institutes of Health grant R01HG009524 (Y.X.F.); Education Bureau of Jinan and Shandong First Medical University grant JNSX2021046 (Z.H.); Key Laboratory of Brain Functional Genomics at East China Normal University grant 20SKBFGK2 (Y.H.P. and H.L.); Shanghai Institute of Nutrition and Health grant JBGSRWBD-SINH-2021-10 (H.L.); China Postdoctoral Science Foundation grant 2022M711978 (Z.H.); Shandong Provincial Natural Science Foundation grant ZR2022QC062 (Z.H.); and Shandong Provincial Postdoctoral Innovation Talent Support Program grant SDBX2022012 (Z.H.). **Author contributions:** Conceptualization: W.H., Z.H., F.D.V., G.M., Y.X.F., Y.H.P., and H.L.; Methodology: W.H., Z.H., Y.H.P., and H.L.; Software: W.H., Z.H., and H.L.; Investigation: W.H., Z.H., P.D., F.D.V., G.M., J.C., Y.X.F., Y.H.P., and H.L.; Visualization: W.H., Z.H., P.D., J.C., F.D.V., and G.M.; Funding acquisition: Z.H., Y.X.F., Y.H.P., and H.L.; Supervision: G.M., Y.H.P., and H.L.; Writing: W.H., Z.H., P.D., F.D.V., G.M., Y.X.F., Y.H.P., and H.L. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Raw data are deposited at Mendeley (*45*), and FitCoal is archived at Zenodo (*46*). The download links for 1000GP and HGDP-CEPH data are available in table S2.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq7487
Materials and Methods
Supplementary Text
Figs. S1 to S60
Tables S1 to S23
References (*47*–*96*)
MDAR Reproducibility Checklist

Submitted 28 April 2022; resubmitted 7 February 2023
Accepted 11 July 2023
10.1126/science.abq7487

Genomic inference of a severe human bottleneck during the Early to Middle Pleistocene transition

Wangjie Hu, Ziqian Hao, Pengyuan Du, Fabio Di Vincenzo, Giorgio Manzi, Jialong Cui, Yun-Xin Fu, Yi-Hsuan Pan, and Haipeng Li

Science, **381** (6661), .
DOI: 10.1126/science.abq7487

Editor's summary

Today, there are more than 8 billion human beings on the planet. We dominate Earth's landscapes, and our activities are driving large numbers of other species to extinction. Had a researcher looked at the world sometime between 800,000 and 900,000 years ago, however, the picture would have been quite different. Hu *et al*. used a newly developed coalescent model to predict past human population sizes from more than 3000 present-day human genomes (see the Perspective by Ashton and Stringer). The model detected a reduction in the population size of our ancestors from about 100,000 to about 1000 individuals, which persisted for about 100,000 years. The decline appears to have coincided with both major climate change and subsequent speciation events. —Sacha Vignieri