

Title:

Counterbalancing the time-dependence effect on the human mitochondrial DNA molecular clock

Author:

Vicente M. Cabrera

Affiliation:

Departamento de Genética, Universidad de La Laguna, E-38271 La Laguna, Tenerife, Spain

Corresponding author:

Vicente M. Cabrera. E-mail: vicente.vca811@gmail.com

Abstract:

We propose a new method for estimating the coalescent age of phylogenetically related sequences that takes into account the observed time dependency of molecular rate estimates. Applying this method to human mitochondrial DNA data we have obtained significantly older ages for the main events of human evolution than in previous estimates. These ages are in close agreement with the most recent archaeological and paleontological records.

Main text:

During the last three decades, the application of mitochondrial DNA (mtDNA) variation to studies of human evolution has dominated the scientific scenario. Recently, this small molecule is being substituted by the analysis of whole genomes. However, before turning the page, it would be convenient to solve the patent contradictions between the mtDNA molecular clock time estimations and those proposed by paleontological and archaeological data. A key achievement of early mtDNA analyses was the dating and origin of the most recent common female ancestor of all living women around 200 kya in Africa¹. However, in flagrant contradiction, hominin fossils and associated Middle Stone Age artifacts from Jebel Irhoud in Morocco have been recently aged at 315 ± 34 kya^{2,3}. These older dates have been confirmed in a study of ancient African genomes that estimated modern human divergence at 350 to 260 kya⁴. Another controversial milestone of the mtDNA molecular clock was the dating, based on the coalescence age of macrohaplogroup L3, of the dispersal of modern humans out of Africa 50 to 70 kya⁵. This is against the presence of early modern human remains in the Levant at the Skhul and Qafzeh caves dated at about 80-120 kya⁶, the presence of middle stone age industries in the Arabian peninsula with similar dates around 80-130 kya⁷⁻⁹, the recent discovery in southern China of unequivocally modern human teeth dated to 80-120 kya¹⁰, or the recently reported detection of an ancient gene flow from early modern humans into the ancestors of eastern Neanderthals more than 100 kya¹¹. Furthermore, the most recent finding of a *Homo sapiens* maxilla found at Misliya Cave, Israel, dated to 177-194 kya¹², could significantly anticipate the exit of homo sapiens from Africa. Curiously, these dates within the MIS-7 last interstadial, are in agreement with the age proposed for an ancient African hominin introgression into the European Neanderthals¹³. In addition, OSL dating of stratigraphic

undisturbed basal stone tool assemblages in Madjedbebe¹⁴ placed the human colonization of Australia around 65 kya with minor age uncertainties of only $\pm 3-4$ kyr. The above date is significantly older than the 43-47 kya coalescence age estimated recently from Australian aboriginal mitogenomes¹⁵.

The efficiency of the molecular clock¹⁶ is based on the reliability of several implicit assumptions as a) the correctness of the mutation rate (μ) of the gene under study; b) constancy with time of the rate of molecular evolution ($\bar{\mu}$) and c) rate homogeneity among the different lineages involved in the phylogeny. To accomplish the first point, in this study we have used the full-length mtDNA germ-line mutation rate of 1.3×10^{-8} mutations per site per year (assuming a generation time of 20 years) and its derived rate scalar of one mutation every 4651 years estimated by others¹⁷. This mtDNA mutation rate is lower than estimates in most pedigree studies which the authors explain because, at analyzing two tissues, they could discard somatic heteroplasmies. In this respect, it has to be mentioned that second-generation massive sequencing has made possible the direct calculation of the germ-line human genomic mutation rate which resulted in half of the phylogenetic mutation rate. Thus, doubling the estimated divergence dates of Africans and proposing that crucial events in human evolution have occurred earlier than suggested previously¹⁸.

It is well known that rates of molecular evolution are not constant neither at interspecific nor intraspecific levels¹⁹. In general, they decline with increasing divergence time, but the rate of decay differs among taxa. This time-dependent pattern has also been observed for human mtDNA in coding as well as non-coding regions^{20,21}. Purifying selection on deleterious mutations and mutation saturation have been suggested as the main forces responsible for this time rate decay²⁰. However, the unrealistic large effective population sizes required to explain the long-term persistence of significantly deleterious mutations cast doubts on whether purifying selection alone can explain the observed rate acceleration²². It has also been found unlikely that the apparent decline in rates over time is due to mutational saturation²⁰. Congruently, correcting for the effects of purifying selection and saturation has only slightly modified the mtDNA evolutionary mutation rate, providing molecular times still in apparent contradiction with archaeological and paleontological ages^{20,23}. Demographic processes such as serial bottlenecks and expansions have also been proposed to explain the differences in rate estimation over time²¹. It seems evident that some adjustment should be implemented to correct the time dependency of the molecular clock. In this paper, we propose a practical approach to counteract the time dependency effect on the molecular rate estimates.

Since early molecular analyses, it was observed that rates of homologous nuclear DNA sequence evolution differ between taxonomic groups²⁴ which was also extensible to mtDNA²⁵. Later, significant differences in rates of molecular evolution between mtDNA human lineages were also detected at haplogroup level²⁶⁻³⁰. Different relaxed molecular-clock methods have been implemented to incorporate rate variation among lineages^{31,32}. However, the application of these methods to the human mtDNA has yield age estimates for the main milestones of human evolution in agreement with previous molecular estimates^{33,34}. In this paper, when distributing mutations of lineages with significant rate differences within coalescent periods, we have used a simple proportionality criterion. We allowed a window

from 0 to 5 mutations between sequences as it has been demonstrated that, under a Poisson distribution, even in an extended period of 10,000 years, there could be lineages still carrying the same mutations that their common ancestor, and lineages that have accumulated five new mutations with probabilities higher than 0.05 percent³⁵.

In a population, the fixation time of a mutation (forward) or the coalescence time to the most recent common ancestor (backward), is usually calculated by the estimator $\bar{\tau} = 4N_e\mu$ (being N_e the effective population size). For the haploid mtDNA genome, $\bar{\tau}$ equals to $N_{ef}\mu$ (being N_{ef} the female effective population size). Although under strict neutral theory parameters, the rate of substitution is equated to the mutation rate³⁶, the same author warned us that a clear distinction exists between mutation rate (μ) and mutation substitution ($\bar{\tau}$). The former refers to the change of genetic material at the individual level, and the latter refers to that at the population level³⁷. Thus, only when N_e is constant across generations, keeping small or large sizes, $\bar{\tau}$ equals to μ . This holds because with large constant sizes the number of new mutations incorporated into the population ($N_e\mu$) increases but, on the same path, the probability of fixation ($1/N_e$) decreases. Contrarily, with small constant sizes, the number of new mutations decreases but the probability of fixation of any of them increases at a similar level. It is widely admitted that N fluctuated largely during the human history and that global exponential population growth is happening since recent times. In this paper, we have taken into account the changes in population size that occurred backward in time and its influence on the rate of gene substitution. When the population size fluctuates across generations, the probability of fixation of a mtDNA neutral variant (q) is no longer the $1/N$ constant. It will depend on the difference in population size of the next generation (N_1) with respect to the initial size (N_0):

$$q = \frac{N_1}{N_0} \times \frac{1}{N_0}$$

For example, if N_1 is twice the size of N_0 , q equals $2/N_0$ and, on the contrary, if N_1 is half the size of N_0 , q equals $1/2N_0$. As a consequence, the rate of substitution for neutral mutations in a population with fluctuating size depends on the change in size between generations:

$$\bar{\tau} = \frac{N_1}{N_0} \times \mu$$

As human populations have been growing exponentially for several centuries, we should counterbalance this effect from the present-day generation (N_n) going backward in time by inverting the fraction between consecutive generations (N_{n-1}/N_n). Notice that this dependence might explain the differences in rate estimation over time observed empirically (Henn et al. 2009). We will take into consideration this important relationship for the calculation of $\bar{\tau}$.

Several statistics, based on DNA polymorphism, exist to estimate the parameter $\bar{\tau}$. One, S , the number of segregating sites per nucleotide in a sample of sequences³⁸ is strongly dependent on the sample size. A second, π , is defined as the average number of nucleotide differences per site in a sample of sequences³⁹. These two estimators were used to implement a statistical method for testing the neutral mutation hypothesis⁴⁰. A third statistic, rho (ρ), is referred to as the mean number of nucleotide differences of a sample of sequences compared to their common ancestral type⁴¹. This last statistic is calculated from a rooted phylogenetic tree relating the sampled sequences. Although the accuracy of molecular dating with the rho statistic has been questioned because it shows downward biased data estimations, large

asymmetric variances and strong dependency of demographic factors⁴², it is still the most used method to measure intraspecific mtDNA divergence events in humans. In this paper, we propose the use of a modified rho (ρ_m) that significantly improves the molecular date estimation of key events in the human history based on mtDNA genome data. In order to make explicit our modifications to the classical rho, we have depicted, in figure 1, a real genealogy constructed from five lineages (a), and an idealized star-like phylogeny of the same five lineages (b) supposing a population exponential growth short after a severe bottleneck⁴³. The number of lineages sampled is represented by n_i ; t_i are the time periods defined by progressive coalescent events from the tips to the MRCA root; i represents the number of independent lineages left after successive coalescences, μ_i is the number of mutations accumulated during each coalescent period, and m_i the number of mutations accumulated along each lineage. Mutations in the star-like phylogeny are distributed into periods following the pattern found in the real phylogeny. As the accumulation of mutations along each lineage is an individual process driven by the mutation rate, μ , and distributed as independent Poisson processes, ρ , the average number of mutations per lineage, has the same value irrespective of the topology. However, as a consequence of the fact that the lineages in the real tree are not independent, because of their shared genealogy, the variance decay is much slower ($1/\log n$) than in the independent star-like tree ($1/n$)⁴⁴. As a consequence of this, for the rho calculation, mutations within lineages in the star-like phylogeny are counted only once. On the contrary, in the real phylogeny, only mutations occurring at the tips are counted only once while mutations in subsequent coalescent periods going backward to the MRCA node are counted as many times as the number of the period to which they belong. For this reason, mutations in the older periods are overrepresented in the rho calculation. Because of lineage independence, star-like phylogenies are statistically optimal to calculate ρ and π estimators. Under this topology, as mutations along lineages are counted from the root to the tips in ρ and from tip to tip in π pairwise comparisons, the value of π doubles that of ρ . Thus, to correct for dependence in the real tree, we propose a modified rho statistic (ρ_m) that is the summation of classical rho statistics calculated for each coalescent period in the tree:

$$\rho_m = \sum_{i=2}^n \rho_i$$

This compound Poisson distribution is also Poisson distributed, therefore mean and variance are equal and, the standard deviation is the square root of this variance. Another important difference between the real and star-like genealogies is that in the first the number of lineages decreases as one stepwise function across coalescent periods while in the second the number of lineages is constant until the root is reached. Equating the number of lineages in the sample to the effective population size in the population, we should take into account this backward real decrease in N_e to improve the estimation of the MRCA age. In an ideal coalescent model we should have $i-1$ population decreasing sizes, but in real phylogenies, in addition to bifurcations, there are also multifurcations and lineages with long internal segments without any branching event. Even so, we applied to each rho in consecutive periods going backward the reverse proportion used to counteract the time dependence effect on the evolutive rate. That is, multiplying in each $i-1$ period the mutation rate μ by $(i-1)/i$ and leaving the μ rate as calculated from germ-line estimations for the most recent period, comprising the tips of all the lineages sampled. With this method, we have obtained a time-dependent scaled mutation

rate, $\bar{\mu}$, that gave human mtDNA intraspecific ages congruent with the archaeological and paleontological calibrated nodes representing key events in the human history.

$$\bar{\mu} = \frac{1}{\mu} \left(\rho_n + \sum_{i=2}^{n-1} \frac{i+1}{i} \rho_i \right)$$

Performing calculations on the empirical genealogy (Figure 1a), we have obtained an age of $22,276.8 \pm 149$ years using the standard ρ and age of $30,396 \pm 174$ years, 1.36 times greater, when using the time-dependent $\bar{\mu}$ estimator proposed here (Table 1).

To apply this statistic to real mtDNA human data we need a rooted tree relating the sampled sequences. Using coalescent methodology we could obtain a probabilistic tree. However, in the case of human mtDNA, we have a very contrasted phylogenetic tree⁴⁵, constructed using the Network program⁴⁶. In it, mutations are placed hierarchically from the tips to the root and multiple hits, identified by network reticulations, have been resolved attending to the relative mutation rate of the positions involved²³. Thus, following this standard, we constructed an African mtDNA genome-based phylogenetic tree, using 86 previously published complete mtDNA sequences, in which all the main African haplogroups are represented (Figure S1). Likewise, using 142 already published complete mtDNA sequences, we constructed a second phylogenetic tree for the Australasian specific haplogroup P (Figure S2). Finally, to test a more recent human colonization, we constructed a third tree including 48 already published complete mtDNA sequences belonging to the Americas specific haplogroup B2 (Figure S3). Using these trees, we applied our proposed time-dependent based estimator to calculate coalescence ages of several essential nodes of the human history (Tables S1 to S11). We found a TMRCA for all the extant human African mtDNAs of $315,801 \pm 17,827$ years (Table 2 and S2) which is highly compatible with the recent archaeological and paleontological estimations of modern human origin around 315,000 ya^{2,3}. It has been proposed an early out-of-Africa of modern humans recently, carrying haplogroup L3 precursor lineages, in a favorable time window around 125,000 ya⁴⁷, that is in harmony with the age calculated here for the L3'4 split in Africa of $165,610 \pm 12,869$ ya as a lower bound (Tables 2 and S3). Furthermore, this age frame is compatible with the presence of modern humans in the Levant⁶ and in China¹⁰ around 100,000 ya. In the same paper, a return to Africa of basal L3 lineages over 75kya was also suggested. Again, the age for the L3 African expansion calculated with the method reported here of $112,829 \pm 10,622$ ya makes this suggestion feasible (Tables 2 and S4). Furthermore, under this new temporal window the great morphological variability of the Skhul/Qafzeh remains and their corresponding wide range of ages (120-80 kya), could easily fit into the whole molecular period proposed elsewhere⁴⁷, beginning with the out-of-Africa of early modern humans and finishing with their early return to the same Continent, carrying basic L3 lineages (125-75 kya). However, notice that a return to Africa from the Arabian peninsula would also be supported by the dates estimated from the archaeological record of the region. On the other hand, our TMRCA for the Austrasia haplogroup P ($103,267 \pm 10,332$ ya) is also in agreement with an early presence of modern humans in Asia (Tables 2 and S6, S7, S8, S9). Thus, it represents a lower bound for the colonization of Philippines⁴⁸, Sumatra⁴⁹ and Australia¹⁴ around 65,000 to 73,000 ya. It has to be mentioned that from the genome sequencing of an Aboriginal Australian⁵⁰, it was deduced that Aboriginal Australians are descendants of human dispersal into Eastern Asia that occurred about 62-75 kya. Finally, the age of human expansion into the American Continent, deduced from the haplogroup B2

phylogeny, is around 37,000 ya (Table 2, and table S11). This age is in support of a pre-Clovis occupation of the New World, well before the last glacial maximum. As the calculated ages have wide statistical confidence intervals (Table 2), different models could be adjusted into their frames. For example, we might suppose an earlier out of Africa matching the Misliya maxilla dated to 177-194 kya, then, the Skhul and Qafzeh remains, dated around 80-130 kya might signal the return to Africa of the carriers of the mtDNA basal haplogroup L3 lineages instead of the out of Africa of early anatomically modern humans as proposed here. In the same way, the controversial presence of *H. sapiens* in Java⁵¹ and Sulawesi⁵² as early as 120 kya would fit into the window age of the Australasian mtDNA haplogroup P (Table 2). Sure, future archaeological discoveries and more precise fossil dating will outline the most appropriate model.

REFERENCES

1. Cann, R.L., Stoneking, M., and Wilson, A.C. Mitochondrial DNA and human evolution. *Nature* 325, 31–36.
2. Hublin, J.-J., Ben-Ncer, A., Bailey, S.E., Freidline, S.E., Neubauer, S., Skinner, M.M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., et al. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature* 546, 289–292.
3. Richter, D., Grün, R., Joannes-Boyau, R., Steele, T.E., Amani, F., Rué, M., Fernandes, P., Raynal, J.-P., Geraads, D., Ben-Ncer, A., et al. (2017). The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. *Nature* 546, 293–296.
4. Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R., Vicente, M., Steyn, M., Soodyall, H., et al. (2017). Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science* 358, 652–655.
5. Soares, P., Alshamali, F., Pereira, J.B., Fernandes, V., Silva, N.M., Afonso, C., Costa, M.D., Musilová, E., Macaulay, V., Richards, M.B., et al. (2012). The Expansion of mtDNA Haplogroup L3 within and out of Africa. *Mol. Biol. Evol.* 29, 915–927.
6. Grün, R., and Stringer, C.B. (1991). Electron spin resonance dating and the evolution of modern humans. *Archaeometry* 33, 153–199.
7. Armitage, S.J., Jasim, S.A., Marks, A.E., Parker, A.G., Usik, V.I., and Uerpmann, H.-P. (2011). The southern route “out of Africa”: evidence for an early expansion of modern humans into Arabia. *Science* 331, 453–456.
8. Rose, J.I., Usik, V.I., Marks, A.E., Hilbert, Y.H., Galletti, C.S., Parton, A., Geiling, J.M., Cerný, V., Morley, M.W., and Roberts, R.G. (2011). The Nubian complex of Dhofar, Oman: an African middle stone age industry in southern Arabia. *PLoS One* 6, e28239.

9. Groucutt, H.S., Scerri, E.M., Lewis, L., Clark-Balzan, L., Blinkhorn, J., Jennings, R.P., Parton, A., and Petraglia, M.D. (2015). Stone tool assemblages and models for the dispersal of *Homo sapiens* out of Africa. *Quaternary International* 382, 8–30.
10. Liu, W., Martínón-Torres, M., Cai, Y., Xing, S., Tong, H., Pei, S., Sier, M.J., Wu, X., Edwards, R.L., Cheng, H., et al. (2015). The earliest unequivocally modern humans in southern China. *Nature*.
11. Kuhlwilm, M., Gronau, I., Hubisz, M.J., de Filippo, C., Prado-Martinez, J., Kircher, M., Fu, Q., Burbano, H.A., Lalueza-Fox, C., de La Rasilla, M., et al. (2016). Ancient gene flow from early modern humans into Eastern Neanderthals. *Nature* 530, 429–433.
12. Hershkovitz, I., Weber, G.W., Quam, R., Duval, M., Grün, R., Kinsley, L., Ayalon, A., Bar-Matthews, M., Valladas, H., Mercier, N., et al. (2018). The earliest modern humans outside Africa. *Science* 359, 456–459.
13. Posth, C., Wißing, C., Kitagawa, K., Paganí, L., van Holstein, L., Racimo, F., Wehrberger, K., Conard, N.J., Kind, C.J., Bocherens, H., et al. (2017). Deeply divergent archaic mitochondrial genome provides lower time boundary for African gene flow into Neanderthals. *Nature Communications* 8, 16046.
14. Clarkson, C., Jacobs, Z., Marwick, B., Fullagar, R., Wallis, L., Smith, M., Roberts, R.G., Hayes, E., Lowe, K., Carah, X., et al. (2017). Human occupation of northern Australia by 65,000 years ago. *Nature* 547, 306–310.
15. Tobler, R., Rohrlach, A., Soubrier, J., Bover, P., Llamas, B., Tuke, J., Bean, N., Abdullah-Highfold, A., Agius, S., O’Donoghue, A., et al. (2017). Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* 544, 180–184.
16. Zuckerkandl, E., and Pauling, L. (1965). Evolutionary divergence and convergence in proteins. *Evolving Genes and Proteins* 97, 97–166.
17. Rebolledo-Jaramillo, B., Su, M.S.-W., Stoler, N., McElhoe, J.A., Dickins, B., Blankenberg, D., Korneliussen, T.S., Chiaromonte, F., Nielsen, R., Holland, M.M., et al. (2014). Maternal age effect and severe germ-line bottleneck in the inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences* 111, 15474–15479.
18. Scally, A., and Durbin, R. (2012). Revising the human mutation rate: implications for understanding human evolution. *Nature Reviews Genetics* 13, 745–753.
19. Molak, M., and Ho, S.Y. (2015). Prolonged decay of molecular rate estimates for metazoan mitochondrial DNA. *PeerJ* 3, e821.
20. Ho, S.Y., Phillips, M.J., Cooper, A., and Drummond, A.J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* 22, 1561–1568.
21. Henn, B.M., Gignoux, C.R., Feldman, M.W., and Mountain, J.L. (2009). Characterizing the time dependency of human mitochondrial DNA mutation rate

- estimates. *Molecular Biology and Evolution* 26, 217–230.
22. Woodhams, M. (2006). Can deleterious mutations explain the time dependency of molecular rate estimates? *Mol. Biol. Evol.* 23, 2271–2273.
23. Soares, P., Ermini, L., Thomason, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V., and Richards, M.B. (2009). Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* 84, 740–759.
24. Britten, R.J. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231, 1393–1399.
25. HASEGAWA, M., and KISHINO, H. (1989). Heterogeneity of tempo and mode of mitochondrial DNA evolution among mammalian orders. *The Japanese Journal of Genetics* 64, 243–258.
26. Torroni, A., Rengo, C., Guida, V., Cruciani, F., Sellitto, D., Coppa, A., Calderon, F.L., Simionati, B., Valle, G., Richards, M., et al. (2001). Do the four clades of the mtDNA haplogroup L2 evolve at different rates? *The American Journal of Human Genetics* 69, 1348–1356.
27. Maca-Meyer, N., González, A.M., Pestano, J., Flores, C., Larruga, J.M., and Cabrera, V.M. (2003). Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. *BMC Genetics* 4, 1.
28. Howell, N., Elson, J.L., Turnbull, D., and Herrnstadt, C. (2004). African haplogroup L mtDNA sequences show violations of clock-like evolution. *Molecular Biology and Evolution* 21, 1843–1854.
29. Merriwether, D.A., Hodgson, J.A., Friedlaender, F.R., Allaby, R., Cerchio, S., Koki, G., and Friedlaender, J.S. (2005). Ancient mitochondrial M haplogroups identified in the Southwest Pacific. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13034–13039.
30. Pierron, D., Chang, I., Arachiche, A., Heiske, M., Thomas, O., Borlin, M., Pennarun, E., Murail, P., Thoraval, D., Rocher, C., et al. (2011). Mutation rate switch inside Eurasian mitochondrial haplogroups: impact of selection and consequences for dating settlement in Europe. *PLoS One* 6, e21543.
31. Duchêne, S., Lanfear, R., and Ho, S.Y. (2014). The impact of calibration and clock-model choice on molecular estimates of divergence times. *Molecular Phylogenetics and Evolution* 78, 277–289.
32. Battistuzzi, F.U., Filipowski, A., Hedges, S.B., and Kumar, S. (2010). Performance of relaxed-clock methods in estimating evolutionary divergence times and their credibility intervals. *Mol. Biol. Evol.* 27, 1289–1300.
33. Fu, Q., Mittnik, A., Johnson, P.L., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., et al. (2013). A revised timescale for human

evolution based on ancient mitochondrial genomes. *Current Biology* 23, 553–559.

34. Rieux, A., Eriksson, A., Li, M., Sobkowiak, B., Weinert, L.A., Warmuth, V., Ruiz-Linares, A., Manica, A., and Balloux, F. (2014). Improved calibration of the human mitochondrial clock using ancient genomes. *Molecular Biology and Evolution* 31, 2780–2792.

35. Larruga, J.M., Marrero, P., Abu-Amero, K.K., Golubenko, M.V., and Cabrera, V.M. (2017). Carriers of mitochondrial DNA macrohaplogroup R colonized Eurasia and Australasia from a southeast Asia core area. *BMC Evolutionary Biology* 17, 115.

36. Kimura, M., and others (1968). Evolutionary rate at the molecular level. *Nature* 217, 624–626.

37. Kimura, M. (1991). The neutral theory of molecular evolution: a review of recent evidence. *The Japanese Journal of Genetics* 66, 367–386.

38. Watterson, G. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7, 256–276.

39. Nei, M., and Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A.* 76, 5269–5273.

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

41. Forster, P., Harding, R., Torroni, A., and Bandelt, H.J. (1996). Origin and evolution of Native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.* 59, 935–945.

42. Cox, M.P. (2008). Accuracy of molecular dating with the rho statistic: deviations from coalescent expectations under a range of demographic models. *Human Biology* 80, 335–357.

43. Slatkin, M., and Hudson, R.R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562.

44. Donnelly, P., and Tavaré, S. (1995). Coalescents and genealogical structure under neutrality. *Annu. Rev. Genet.* 29, 401–421.

45. Van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* 30, E386–94.

46. Bandelt, H.-J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16, 37–48.

47. Cabrera, V.M., Marrero, P., Abu-Amero, K.K., and Larruga, J.M. (2018). Carriers of mitochondrial DNA macrohaplogroup L3 basal lineages migrated back to Africa from Asia around 70,000 years ago. *BMC Evol. Biol.* 18, 98.

48. Mijares, A.S., Détroit, F., Piper, P., Grün, R., Bellwood, P., Aubert, M., Champion, G., Cuevas, N., De Leon, A., and Dizon, E. (2010). New evidence for a 67,000-year-old human presence at Callao Cave, Luzon, Philippines. *Journal of Human Evolution* 59, 123–132.
49. Westaway, K.E., Louys, J., Awe, R.D., Morwood, M.J., Price, G.J., Zhao, J., Aubert, M., Joannes-Boyau, R., Smith, T., Skinner, M.M., et al. (2017). An early modern human presence in Sumatra 73,000-63,000 years ago. *Nature* 548, 322.
50. Rasmussen, M., Guo, X., Wang, Y., Lohmueller, K.E., Rasmussen, S., Albrechtsen, A., Skotte, L., Lindgreen, S., Metspalu, M., Jombart, T., et al. (2011). An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334, 94–98.
51. Westaway, K.E., Morwood, M.J., Roberts, R.G., Rokus, A.D., Zhao, J., Storm, P., Aziz, F., van den Bergh, G., Hadi, P., Jatmiko, et al. (2007). Age and biostratigraphic significance of the Punung Rainforest Fauna, East Java, Indonesia, and implications for Pongo and Homo. *J. Hum. Evol.* 53, 709–717.
52. Van den Bergh, G.D., Li, B., Brumm, A., Grün, R., Yurnaldi, D., Moore, M.W., Kurniawan, I., Setiawan, R., Aziz, F., Roberts, R.G., et al. (2016). Earliest hominin occupation of Sulawesi, Indonesia. *Nature* 529, 208.

Figure 1 legend: A) Empirical coalescent tree of five lineages (a to e), with four coalescent periods (t) and mutations along branches (numbers). B) Ideal star-like tree of the same five lineages.

Table 1: Coalescence age for Figure 1 Tree using the compound rho

Period	Lineages	Mutations	Rho	$i/i + 1$	μ	$1/\mu$	Years
2	2	3	1.50	0.67	1.44×10^{-4}	6944	10416
3	3	6	2.00	0.75	1.61×10^{-4}	6211	12422
4	4	2	0.50	0.80	1.72×10^{-4}	5814	2907
5	5	5	1.00	1.00	2.15×10^{-4}	4651	4651

Coalescence age for Figure 1 tree: $30,396 \pm 5.513$

Table 2: Coalescence age estimates for several human mtDNA evolutive events

Haplogroup split	Evolutive event	Mean age in years	95% Coefficient interval
L0/L1'2'5'6'4'3	Most recent African common ancestor	317,814 ya	(352,755-282,873 ya)
L3'4	Out of Africa	165,610 ya	(190,833-140,387 ya)
L3	Return to Africa of L3	112,829 ya	(133,648-92,010 ya)
P	Reaching the Pacific	106,752 ya	(127,003-86,501 ya)
P	Reaching Australia	108,034 ya	(128,406-87,662 ya)
P	Reaching Philippines	111,545 ya	(132,244-90,846 ya)
P	Reaching New Guinea	112,070 ya	(132,818-91,322 ya)
B2	Expansion Americas	37,701 ya	(49,735-25,667 ya)

Figure 1

