Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa

Lluís Quintana-Murci[1,5], Ornella Semino[1], Hans-J. Bandelt[2], Giuseppe Passarino[1,3], Ken McElreavey[4] & A. Silvana Santachiara-Benerecetti[1]

The out-of-Africa scenario[1] has hitherto provided little evidence for the precise route by which modern humans left Africa. Two major routes of dispersal have been hypothesized: one through North Africa into the Levant, documented by fossil remains[2], and one through Ethiopia along South Asia, for which little, if any, evidence exists[3]. Mitochondrial DNA (mtDNA) can be used to trace maternal ancestry. The geographic distribution and variation of mtDNAs can be highly informative in defining potential range expansions and migration routes in the distant past. The mitochondrial haplogroup M, first regarded as an ancient marker of East-Asian origin[4,5], has been found at high frequency in India[6] and Ethiopia[7], raising the question of its origin. (A haplogroup is a group of haplotypes that share some sequence variations.) Its variation and geographical distribution suggest that Asian haplogroup M separated from eastern-African haplogroup M more than 50,000 years ago. Two other variants (489C and 10873C) also support a single origin of haplogroup M in Africa. These findings, together with the virtual absence of haplogroup M in the Levant and its high frequency in the South-Arabian peninsula, render M the first genetic indicator for the hypothesized exit route from Africa through eastern Africa/western India. This was possibly the only successful early dispersal event of modern humans out of Africa.

The presence (~20%) of the 'Asian' mtDNA haplogroup M, defined by the 10400 C→T transition, is unique in eastern Africa[7]. As in Asia, the haplogroup M in Ethiopia belongs to a superhaplogroup (L3), which encompasses almost all Eurasian and a considerable number of African mtDNAs (refs 8,9). We previously advanced[7] three hypotheses to explain the presence of haplogroup M in eastern Africa: (i) it has evolved through an independent mutation; (ii) it has been acquired through exchange with Asians; or (iii) it was present in the ancient Ethiopians and subsequently spread toward Asia. To better understand the origin of eastern African M and its relationship with its Asian counterpart, we have characterized, through high-resolution restriction analysis, 25 eastern-African M molecules and, through sequencing of the mtDNA hypervariable segment I (HVS-I), the African and 27 Indian M samples.

The two groups are considerably different (Table 1). The eastern-African group, named M1, is characterized within M by a consensus HVS-I motif defined by four transitions at nt 16,129, 16,189, 16,249 and 16,311. This HVS-I signature motif is not found either in our or other[10] Indian samples, whereas it characterizes most of the M types sporadically observed in the Mediterranean area[11,12] and 7% of Nile Valley sequences[13]. The phylogenetic relationships encompassing both RFLP and HVS-I variations in eastern Africans and Indians are shown (Figs 1 and 2, respectively). The eastern-African data yield a non-star-like phylogeny with four tentative subgroups (I–IV) of closely related sequences. By contrast, the Indian M mtDNAs form

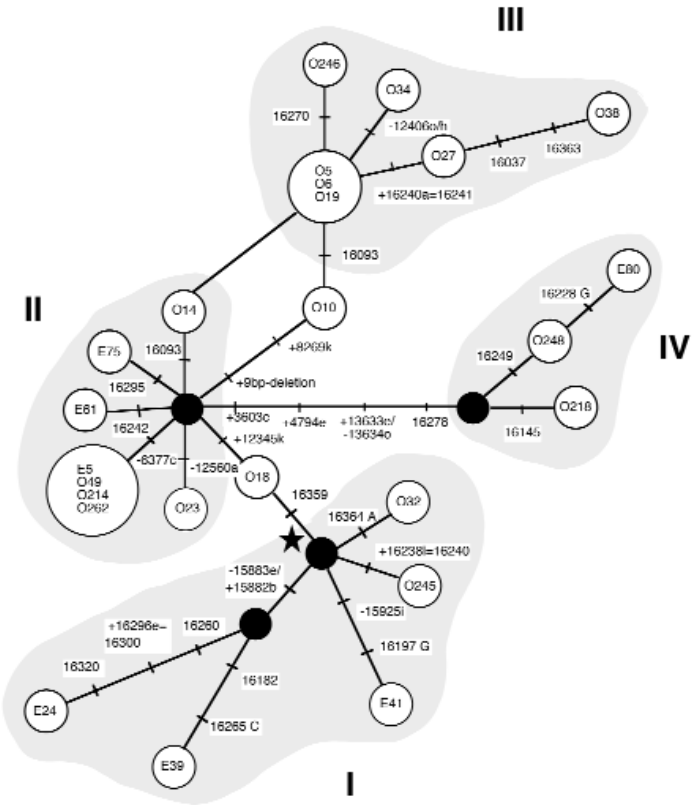

Fig. 1 The full median network[24] of RFLP and HVS-I variation in the eastern-African clade M1. Numbers along links indicate transitions or transversions (when suffixed by A, G or C), or RFLP site changes (when suffixed by a small letter, see Table 1), so that the derived states are indicated. Length polymorphisms in the C-run (nt 16,184–16,193) are disregarded. The order of mutations on single links is arbitrary. The node labelled with a star represents the coalescence type of this eastern-African clade M1 and differs from CRS at nt 16,129, 16,189, 16,223, 16,249 and 16,311, as well as at sites 10,394c, 10,397a and 16,517e. The network is partitioned into four starlike trees (I, II, III, IV) harbouring closely related types; O18 and O10 may be intermediate between I and II, and II and III, respectively. The *Rsa*I site at nt 12,345 (separating II, III, IV from I) has never been observed outside M1.

[1]*Dipartimento di Genetica e Microbiologia, University of Pavia, Italy.* [2]*Mathematisches Seminar, Universität Hamburg, Germany.* [3]*Dipartimento di Biologia Cellulare, University of Calabria, Italy.* [4]*Unité d'Immunogénétique Humaine, Institut Pasteur, Paris, France.* [5]*Present address: Unité d'Immunogénétique Humaine, Institut Pasteur, Paris, France. Correspondence should be addressed to A.S.S.-B. (e-mail: santa@ipvgen.unipv.it).*

Table 1 • RFLP and HVS-I mtDNA haplotypes of eastern-African and Indian M samples

Sample[a]	RFLP haplotype[b]	HVS-I haplotype[c]
eastern African:		
E5	+16517e; +12345k; –6377c	**129 189 223 249 311** 359
E24	+16517e; –15883e/+15882b; +16296c	**129 189 223 249** 260 300 **311** 320
E39	+16517e; –15883e/+15882b	**129** 182 **189 223 249** 265C **311**
E41	+16517e; –15925i	**129 189** 197G **223 249 311**
E61	+16517e; +12345k	**129 189 223** 242 **249 311** 359
E75	+16517e; +12345k	**129 189 223 249** 295 **311** 359
E80	+16517e; +12345k; +3603c; +4794e; +13633e/–13634o	**129 189 223** 228G 278 **311** 359
O5	+16517e; +12345k; region V 9-bp del; +8269k	093 **129 189 223 249 311** 359
O6	+16517e; +12345k; region V 9-bp del; +8269k	093 **129 189 223 249 311** 359
O10	+16517e; +12345k; region V 9-bp del; +8269k	**129 189 223 249 311** 359
O14	+16517e; +12345k	093 **129 189 223 249 311** 359
O18	+16517e	**129 189 223 249 311** 359
O19	+16517e; +12345k; region V 9-bp del; +8269k	093 **129 189 223 249 311** 359
O23	+16517e; +12345k; –12560a	**129 189 223 249 311** 359
O27	+16517e; +12345k; region V 9-bp del; +8269k; +16240a	093 **129 189 223** 241 **249 311** 359
O32	+16517e	**129 189 223 249 311** 364A
O34	+16517e; +12345k; region V 9-bp del; +8269k; –12406o/h	093 **129 189 223 249 311** 359
O38	+16517e; +12345k; region V 9-bp del; +8269k; +16240a	037 093 **129 189 223** 241 **249 311** 359 363
O49	+16517e; +12345k; –6377c	**129 189 223 249 311** 359
O214	+16517e; +12345k; –6377c	**129 189 223 249 311** 359
O218	+16517e; +12345k; +3603c; +4794e; +13633e/–13634o	**129** 145 **189 223 249** 278 **311** 359
O245	+16517e; +16238l	**129 189 223** 240 **249 311**
O246	+16517e; +12345k; region V 9-bp del; +8269k	093 **129 189 223 249** 270 **311** 359
O248	+16517e; +12345k; +3603c; +4794e; +13633e/–13634o	**129 189 223** 278 **311** 359
O262	+16517e; +12345k; –6377c	**129 189 223 249 311** 359
Indian[d]:		
PU191	+16517e;	129 223 291
PU193	+16517e; +16125k	126 150 223
PU202	+16517e; +626e; +3388c/3391a; +8148e; +8249b/–8250e; +9156e; +13332e	189 223 278 352 362
PU214	–5176a	189 223 299 316
PU224	+16517e; –6904j; –8074a; –8309c; –9470c	129 223
PU242	+16517e; +16125k	126 223
PU249	n.c.	212T 214 223
PU256	+16517e; +11163j	093 166del 223
I60	+16517e; +16125k; –12212a	102 126 223 301 344
I64	+16517e; –10364e; –10631c; +14440k	172 189 223 256 294
I76	+16517e; –12560a; –15925i; –16303k	223 304
I872	+16517e; –9052n/–9053f; –16310k	189 223 311 360 363A
I874	n.c.	189 223
I889	+16517e; –12560a; –15925i; –16303k; –7859j; –8882c; +9134g; +9524a	223 231 304
I904	+16517e; +6856l; +16296c	129 189 192 223 300 362
I906	+9678a; –11362a	189 223 262 360 363A 366+C
I914	+16517e; +10431j	223 234
I916	+16517e; –10364e; –10631c; +14440k	129 223
AP22	+16517e; –3534c/–3537a; +12507b	213 223 231 356 362
AP24	+16517e; +16125k; +12185e	126 223
AP62	+16517e; –6000c; –16329k	189 193 223 327 330 360
AP63	+16517e; –6000c; –16329k	189 193 223 330
AP72	+16517e; +6618e; –7859j; +16238l; –16310k	145 223 240 261 311
AP75	+16517e; –3534c/–3537a; +12507b	223 231 356 362
AP76	+16517e; +245l; –15954j	223 318T 354
AP77	+16517e; –3534c/–3537a; +12507b; 15754c	213 223 231 356 362
AP84	+16517e;	179 223 294 357

The eastern-African HVS-I signature motif is shown in bold; n.c., not classified. [a]E, Ethiopians (Amhara); O, Oromo from Kenya/Ethiopia; PU, Punjab; I, Indians from New Delhi and Nepal; AP, Indians from Andhra Pradesh. [b]Restriction sites differences in addition to +10394c and +10397a relative to the Anderson sequence (CRS). a, *Alu*I; b, *Ava*II; c, *Dde*I; e, *Hae*III; f, *Hha*I; g, *Hin*fI; h, *Hpa*I; i, *Msp*I; j, *Mbo*I; k, *Rsa*I; l, *Taq*I; m, *Bam*HI; n, *Hae*II; o, *Hin*cII. Sites separated by a slash indicate simultaneous changes for different enzymes. [c]Nucleotides (minus 16,000). Mutations relative to the CRS are transitions unless specified by the base change. Nucleotides at 16,355 and 16,360 for PU214, at 16,376 for O19, and at 16,378 for I872 could not be determined unambiguously. Length polymorphisms in the C-run are not indicated. Del, deletion; +C, C insertion. [d]The Indian RFLP haplotypes are from Passarino *et al.*[6]

a star-like phylogeny centred at the root of M. The African and Indian M phylogenies are highly divergent. They share single transitions at one of the hypervariable M1 motif positions, and both lack the recurrent 15925 *Msp*I site. These are probably coincidental events.

We calculated the coalescence time for haplogroup M mtDNAs in eastern Africa and India using the statistic ρ (ref. 14) by RFLP and HVS-I. The age of eastern-African haplogroup M, calculated using RFLP data (48,000±15,000), is compatible with that of Indian haplogroup M (56,000±7,000; Table 2), whereas a lower age is obtained using HVS-I data. In fact, the root of the eastern-African M1 clade is separated from the root of M by four transitions (Fig. 3), which, even if hypervariable, needed some time to be accumulated (a few thousand or a few ten-thousand years). A somewhat more recent coalescence of haplogroup M in eastern Africa (that is, in the area of the Oromos) than in Asia can be explained by coalescent theory: a small localized population cannot maintain the same level of diversity as a population that expanded and spread over an entire continent. Thus, it is plausible that haplogroup M already existed in eastern Africa approximately 60,000 years ago. The four subgroups (I–IV) of M1 (Fig. 1) may have similar ages of approximately 10,000–20,000 years. These values may correspond to local expansions of these subclades followed by

dispersals, leading to the sporadic presence of M along the Mediterranean.

As for the three hypotheses advanced on the origin of haplogroup M, the T→C transitions at nt 489 and 10,873 are informative. In fact, they appear potentially correlated with the 10400C→T transition (defining haplogroup M) according to 16 full mtDNA sequences from Eurasia (mainly Japan; refs 15,16). Here, the 11 haplotypes bearing 10400T also bore 489C and 10873C. Whereas 489C was observed in Europe[17], where it seems to be confined to haplogroup J, no information was available on the geographic distribution of 10873C. We found 489C (Table 3) in all Indian and eastern-African haplogroup M mtDNAs analysed, but not in the non-M haplogroup controls, including 20 Africans representing all African main lineages (6 L1, 4 L2, 10 L3) and 11 Asians. These findings, and the lack of positive evidence (given the RFLP status) that the 10400 C→T transition defining M has happened more than once, suggest that it has a single common origin, but do not resolve its geographic origin. Analysis of position 10873 (the *Mnl*I RFLP) revealed that all the M molecules (eastern African, Asian and those sporadically found in our population surveys) were 10873C (Table 3). As for the non-M mtDNAs, the ancient L1 and the L2 African-specific lineages[5], as well as most L3 African mtDNAs, also carry 10873C. Conversely, all non-M mtDNAs of non-African origin analysed so far carry 10873T. These data indicate that the transition 10400 C→T, which defines haplogroup M, arose on an African background characterized by the ancestral state 10873C, which is also present in four primate (common and pygmy chimps, gorilla and orangutan) mtDNA sequences[18]. The phylogenetic position of the eastern African M1 clade within the human mtDNA phylogeny is shown (Fig. 3). The network, which has been constructed disregarding the transitions at 489 and 10,873, indicates that the eastern-African M1 clusters with the Asian M subgroups. The location of the 489 and 10,873 transitions is predicted by our analysis (Table 3). Thus, in the context of the out-of-Africa model[1], the following scenario can be envisaged: haplogroup M originated in eastern Africa approximately 60,000 years ago and was carried toward Asia. This agrees with the proposed date of an out-of-Africa expansion approximately 65,000 years ago[10]. After its arrival in Asia, the haplogroup M founder group went through a demographic and geographic expansion. The remaining M haplogroup in eastern Africa did not spread, but remained localized up to approximately 10,000–20,000 years ago, after which it started to expand.

On the basis of the fossil record in the Middle East (Israel, Skhul/Qafzeh; ref. 2), it has been proposed that modern humans expanded from Africa into the Levant approximately 100,000 years ago. The presence of modern humans in the Middle East would reflect a brief episodic occupation not followed by further expansion, and only an exodus approximately 60,000 years ago would have led to the peopling of Eurasia[2,3,19]. The estimated age of M falls into this period. If the virtual absence of this marker in the Levant is not due to subsequent migrations that erased its traces, its presence at rather high frequencies (~16%) in the South-Arabian peninsula (A. Torroni, pers. comm.) makes it the first genetic indicator of the migration route of *Homo sapiens sapiens* through eastern Africa along the coast toward Southeast Asia, Australia and the Pacific Islands, and of its first expansion into eventually all of Asia. Y-chromosome and chromosome-21 markers[20] also support the African dispersal scenario, and in particular, other nuclear and mtDNA markers indicate eastern Africa as the origin of both African and Eurasian expansions[9,21].

Table 2 • Estimated ages for haplogroup M in eastern Africa and India

Population	Sample size	Age (y)*	
		HVS-I	RFLP
eastern Africa	25	36,000±11,000 (ρ=1.76±0.57)	48,000±15,000 (ρ =1.96±0.62)
India	25	53,000±7,000 (ρ = 2.64±0.32)	56,000±7,000 (ρ =2.32±0.30)
Tibet[4]	36	–	45,000±6,000 (ρ =1.86±0.23)
Mongolia[30]	45	67,000±5,000	– (ρ=3.33±0.27)

*Estimates accompanied by lower bound on standard deviation.

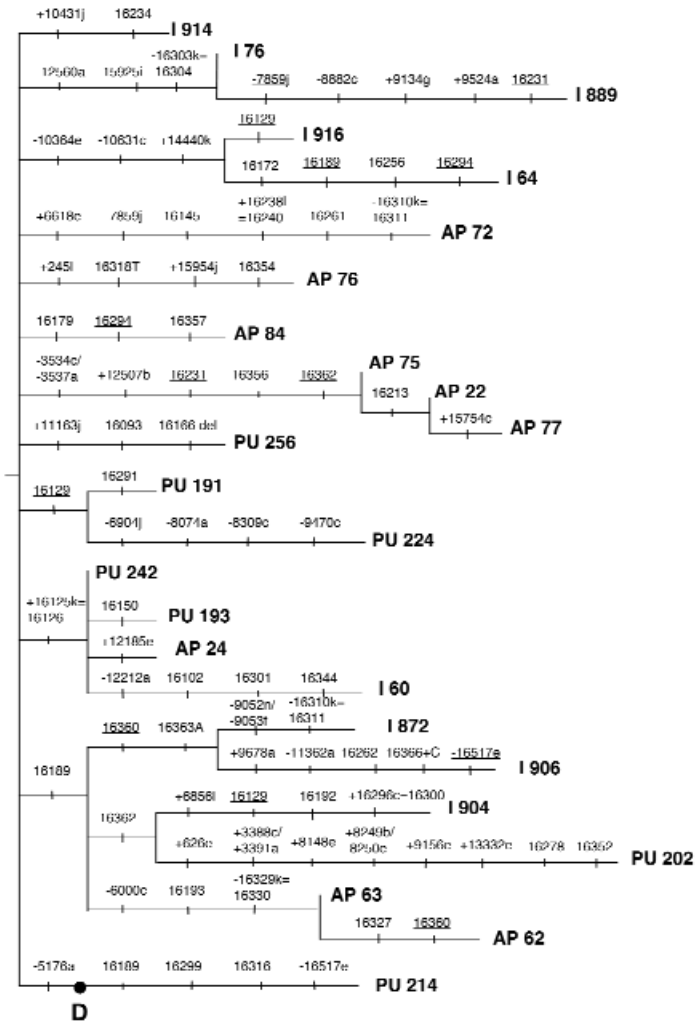

Fig. 2 A most parsimonious tree for haplogroup M in India. This tree (included in the reduced median network; data not shown) would not postulate any back mutation; HVS-I data are from Table 1 (excluding I 874 and PU 249) and RFLP data are from Passarino *et al.*[6]. Recurrent mutations are underlined. Length polymorphisms in the C-run are not displayed.

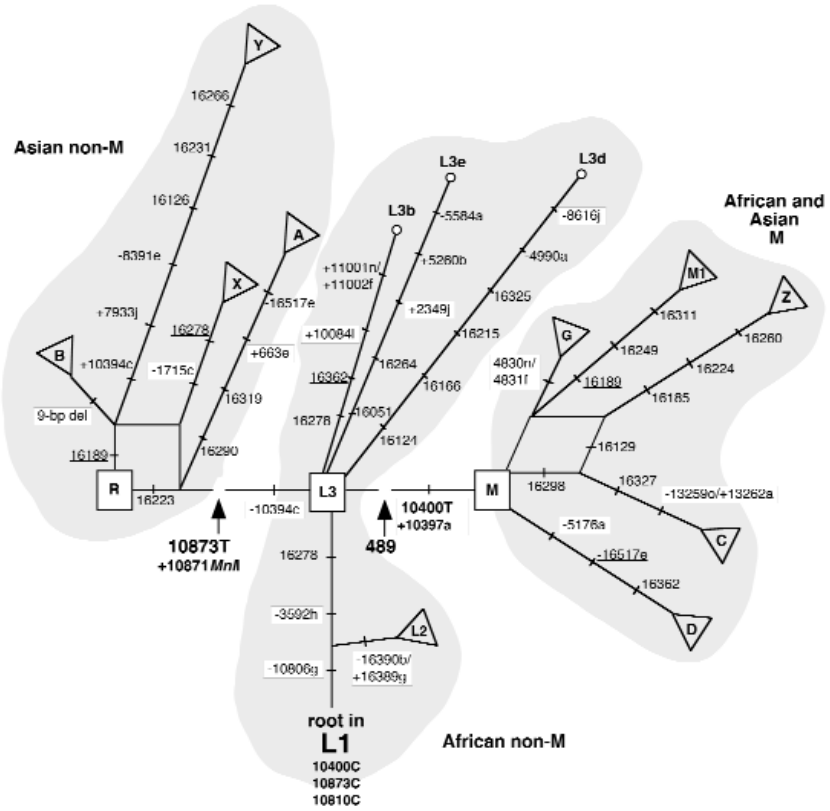

Fig. 3 Phylogenetic position of the eastern-African M1 clade within the human mtDNA phylogeny. The network, rooted in the African haplogroup L1, is composed of most parsimonious trees of L3 based on ancestral combined RFLP and HVS-I types (disregarding the additional positions 489 and 10,873; of (i) the nested super-haplogroups R, M, L3 (indicated by squares); (ii) the clade M1; (iii) the East-Asian/Native-American haplogroups A, B, C, D, G, X, Y, Z; and (iv) the African haplogroup L2 (all indicated by triangles) as inferred from previous data[4,12,25–27]. Three haplogroups (L3b, L3d, L3e) within L3 are represented by individual combined types (small circles), deduced from combined data[5,28,29]. The Indian tree/network is not shown here, but it can directly be attached to the M ancestor, except for a potential D lineage (PU214), with only spurious sharing of singular HVS-I sites such as 16129, 16189 and so on. The phylogenetic location of the mutations at nt 489 and 10,873 (arrow) was predicted by our analysis. The seemingly shared mutation at nt 16,129 (by G, Z and M1) is very likely an accidental parallelism. The ancestral states 10400C, 10810C and 10873C are fixed in L1 (as analysed so far) and are present in the ape sequences[18]. The full M1 network of Fig. 1 may be attached without creating additional reticulation. R includes most European mtDNA haplogroups[12]. Links are labelled as in Figs 1 and 2. The order of mutations on single links is arbitrary. Diagnostic mutations are in boxes.

Methods

High-resolution restriction analysis. We performed high-resolution restriction analysis as described[4].

Sequences of the mitochondrial HVS-I region. We amplified the HVS-I region as described[22]. We determined nucleotide sequences directly from gel-purified PCR products using an ABI 370A automatic sequencer and the ABI Prism dideoxyterminator system (Perkin-Elmer).

Analysis of the mtDNA 489T→C and 10873T→C transitions. We revealed the 489T→C transition by sequencing as described for HVS-I. We amplified the mtDNA fragment from nt 394 to nt 605 with the following

Table 3 • Distribution of the mtDNA 489 and 10873 variants

	Population		Haplogroup M[a]		Haplogroup non-M[a]					
			C	T		C			T	
			L3	L3	L1	L2	L3	L1	L2	L3
489 nt										
	African	Ethiopian	10					4	2	8
	n=30	Senegalese						2	2	2
	Asian	Indian	7							6
	n=20	Tharu	2							5
Total	50		19					6	4	21
10,873 nt										
	African	Ethiopian	22		4	2	4			7
	n=114	Senegalese			17	10	32			4
		Pygmy			6	6				
	Asian	Indian	14							9
	n=58	Tharu	5							6
		Chinese	14							10
	Other	Hungarian	1							8
	n=74	Georgian	1							6
		Turkish	1							7
		Lebanese								3
		Jewish								2
		Greek	1							4
		Iraqi[b]	2		4	4	2			13
		Italian[b]	1			2				12
Total	246		62		31	24	38			91

[a]The mtDNA sample is partitioned into L1 (+10,806 *Hin*fI), L2 (+16,390 *Hin*fI/–16,389; *Ava*II, 16390A) and L3 (–3,592 *Hpa*I). [b]Note the presence of African lineages in these populations.

primers: 489For, 5′–CAAATTTTATCTTTTGGCGGT–3′; 489Rev, 5′–ATTGCTTTGAGGAGGTAAGC–3′. We used the following cycling conditions: 94 °C for 2 min, and then 30 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s. We tested the 10873T→C transition by digestion with *Mnl*I (New England Biolabs) of the amplified mtDNA appropriate fragment followed by agarose gel electrophoresis[4].

Statistic analyses. We calculated the coalescence times of eastern-African and the Asian subgroups of M via ρ (ref. 14). We used the statistic ρ as an unbiased estimator of the age of the respective ancestral root type, given calibrations of mutational rates with respect to HVS-I and RFLPs. We adopted the calibration that $\rho_{HVS\text{-}I}=1$ corresponds to 20,180 y (ref. 14) and $\rho_{RFLP}=1$ corresponds to 24,420 y (ref. 23). We considered only transitions (and no transversions) and ignored length polymorphisms in the C-run (nt 16,184–16,193) in the case of HVS-I, and for the RFLPs, we disregarded major indel events such as the region V 9-bp deletion and the hypervariable 16517 *Hae*III changes.

We approximated the variance of ρ from below by ρ/n (where n is the sample size) in the case of star-like phylogenies. For the clade M1, we obtained a more realistic lower bound by splitting ρ into two summands, ρ′ and ρ′′, so that ρ′ accounts for the averaged distance within the four subgroups (I, II, III, IV) to their centres and the intermediate types O18 and O10, whereas ρ′′ is the average of the (frequency-) weighted distances from those four plus two types to the root of M1. We then took 4 as the artificial sample size for the second component, so that $s=\sqrt{[(\rho'/25)+(\rho''/4)]}$ served as a lower bound for the unknown standard deviation of ρ.

Acknowledgements

We thank F. Falaschi for the Kenya Oromo part of the samples; A. Torroni for information on the haplogroup M frequency in the South-Arabian peninsula; D. Toniolo for the use of the ABI sequencer; R. Ricotti for help in DNA sequencing; and V. Macaulay and S. Nørby for drawing our attention to the Ozawa et al.[15] and Ozawa[16] papers. This research was supported by the Ministry of University (PRIN ex 40%; to A.S.S.-B.). L.Q.-M. is a Marie Curie Fellow of the EC (ERBFMBICT-961830).

Received 23 April; accepted 25 October 1999.

1. Lewin, R. Africa: cradle of modern humans. *Science* **237**, 1292–1295 (1987).
2. Stringer, C.B., Grün, R., Schwarcz, H.P. & Goldberg, P. ESR dates for the hominid burial site of Skhul in Israel. *Nature* **338**, 756–758 (1989).
3. Cavalli-Sforza, L.L., Menozzi, P. & Piazza, A. *The History and Geography of Human Genes* (Princeton University Press, Princeton,1994).
4. Torroni, A. *et al.* Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.* **53**, 563–590 (1993).
5. Chen, Y.S., Torroni, A., Excoffier, L., Santachiara-Benerecetti, A.S. & Wallace, D.C. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am. J. Hum. Genet.* **57**, 133–149 (1995).
6. Passarino, G., Semino, O., Bernini, L.F. & Santachiara-Benerecetti, A.S. Pre-Caucasoid and Caucasoid genetic features of Indian population revealed by mtDNA polymorphisms. *Am. J. Hum. Genet.* **59**, 927–934 (1996).
7. Passarino, G. *et al.* Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *Am. J. Hum. Genet.* **62**, 420–434 (1998).
8. Excoffier, L. & Langaney, A. Origin and differentiation of human mitochondrial DNA. *Am. J. Hum. Genet.* **44**, 73–85 (1989).
9. Watson, E., Forster, P., Richards, M. & Bandelt, H.-J. Mitochondrial footprints of human expansions in Africa. *Am. J. Hum. Genet.* **61**, 691–704 (1997).
10. Mountain, J.L. *et al.* Demographic history of India and mtDNA-sequence diversity. *Am. J. Hum. Genet.* **56**, 979–992 (1995).
11. Côrte-Real, H.B.S.M. *et al.* Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Ann. Hum. Genet.* **60**, 331–350 (1996).
12. Macaulay, V. *et al.* The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* **64**, 232–249 (1999).
13. Krings, M. *et al.* mtDNA analysis of Nile River Valley populations: a genetic corridor or a barrier to migration? *Am. J. Hum. Genet.* **64**, 1166–1176 (1999).
14. Forster, P., Harding, R., Torroni, A. & Bandelt, H.-J. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.* **59**, 935–945 (1996).
15. Ozawa, T. *et al.* Patients with idiopathic cardiomyopathy belong to the same mitochondrial DNA gene family of Parkinson's disease and mitochondrial encephalomyopathy. *Biochem. Biophys. Res. Commun.* **177**, 518–525 (1991).
16. Ozawa, T. Mechanism of somatic mitochondrial DNAs mutations associated with age and diseases. *Biochim. Biophys. Acta* **1271**, 177–189 (1995).
17. Lutz, S., Weisser, H.J., Heizmann, J. & Pollak, S. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. *Int. J. Legal. Med.* **111**, 67–77 (1998); erratum: **112**, 145–150 (1999).
18. Horai, S., Hayasaka, K., Kondo, R., Tsugane, K. & Takahata, N. Recent African origin of modern human revealed by complete sequences of hominoid mitochondrial DNAs. *Proc. Natl Acad. Sci. USA* **92**, 532–536 (1995).
19. Lahr, M.M. & Foley, R.A. Multiple dispersals and modern human origins. *Evol. Anthropol.* **3**, 48–60 (1994).
20. Jin, L. *et al.* Distribution of haplotypes from a chromosome 21 region distinguishes multiple prehistoric human migrations. *Proc. Natl Acad. Sci. USA* **96**, 3796–3800 (1999).
21. Tishkoff, S.A. *et al.* Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* **271**, 1380–1387 (1996).
22. Vigilant, L.A., Wilson, A.C. & Harpending, H. African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–1507 (1991).
23. Torroni, A. *et al.* mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* **62**, 1137–1152 (1998).
24. Bandelt, H.-J., Forster, P., Sykes, B.C. & Richards, M.B. Mitochondrial portraits of human populations using median networks. *Genetics* **141**, 743–753 (1995).
25. Brown, M.D. *et al.* mtDNA haplogroup X: ancient link between Europe/Western Asia and North America? *Am. J. Hum. Genet.* **63**, 1852–1861 (1998).
26. Starikovskaya, Y.B., Sukernik, R.I., Schurr, T.G., Kogelnik, A.M. & Wallace, D.C. mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of ancient Beringia and the peopling of the New World. *Am. J. Hum. Genet.* **63**, 1473–1491 (1998).
27. Schurr, T.G., Sukernik, R.I., Starikovskaya, Y.B. & Wallace, D.C. Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. *Am. J. Phys. Anthropol.* **108**, 1–39 (1999).
28. Graven, L. *et al.* Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. *Mol. Biol. Evol.* **12**, 334–345 (1995).
29. Rando, J.C. *et al.* Mitochondrial DNA analysis of Northwest African populations reveals genetic exchanges with European, Near-Eastern, and sub-Sharan populations. *Ann. Hum. Genet.* **62**, 531–550 (1998).
30. Kolman, C.J., Sambuughin, N. & Bermingham, E. Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. *Genetics* **142**, 1321–1334 (1996).