

RESEARCH ARTICLE

Open Access



Carriers of human mitochondrial DNA macrohaplogroup M colonized India from southeastern Asia

Patricia Marrero¹, Khaled K. Abu-Amero², Jose M. Larruga³ and Vicente M. Cabrera^{3*}

Abstract

Background: From a mtDNA dominant perspective, the exit from Africa of modern humans to colonize Eurasia occurred once, around 60 kya, following a southern coastal route across Arabia and India to reach Australia short after. These pioneers carried with them the currently dominant Eurasian lineages M and N. Based also on mtDNA phylogenetic and phylogeographic grounds, some authors have proposed the coeval existence of a northern route across the Levant that brought mtDNA macrohaplogroup N to Australia. To contrast both hypothesis, here we reanalyzed the phylogeography and respective ages of mtDNA haplogroups belonging to macrohaplogroup M in different regions of Eurasia and Australasia.

Results: The macrohaplogroup M has a historical implantation in West Eurasia, including the Arabian Peninsula. Founder ages of M lineages in India are significantly younger than those in East Asia, Southeast Asia and Near Oceania. Moreover, there is a significant positive correlation between the age of the M haplogroups and its longitudinal geographical distribution. These results point to a colonization of the Indian subcontinent by modern humans carrying M lineages from the east instead the west side.

Conclusions: The existence of a northern route, previously proposed for the mtDNA macrohaplogroup N, is confirmed here for the macrohaplogroup M. Both mtDNA macrolineages seem to have differentiated in South East Asia from ancestral L3 lineages. Taking this genetic evidence and those reported by other disciplines we have constructed a new and more conciliatory model to explain the history of modern humans out of Africa.

Keywords: Human evolution, Mitochondrial DNA, Out of Africa

Background

From a genetic perspective built mainly on mtDNA data, the recent African origin of modern humans [1, 2] and their spread throughout Eurasia and Oceania replacing all archaic humans dwelling there, has held a dominant position in the scientific community. The recent paleogenetic discoveries of limited introgression in the genome of non-African modern humans, of genetic material from archaic, Neanderthal [3, 4] and Denisovan [5–7] hominins has been solved adding a modest archaic assimilation note to the replacement statement [8]. In the East Asia region however, the alternative hypothesis of a continuous regional evolution of modern humans from

archaic populations is supported by the slow evolution of its Paleolithic archaeological record [9] and the irrefutable presence of early and fully modern humans in China at least since 80 kya [10–12]. Moreover, recently it has been detected ancient gene flow from early modern humans into Eastern Neanderthals from the Altai Mountains in Siberia at roughly 100 kya [13]. These data contrast with the phylogenetic hypothesis of a sole and fast dispersal of modern humans out of Africa around 60 kya following a southern route [14–17]. In principle, it could be adduced, as it was in the case of the early human remains from Skhul and Qafzeh in the Levant [18], that the presence in China and Siberia of modern humans at that time was the result of a genetically unsuccessful exit from Africa. However, the fossil record shows a clinal variation along a latitudinal gradient, with decreasing ages from China to Southeast Asia [19–22]

* Correspondence: vicente.vca811@gmail.com

³Departamento de Genética, Facultad de Biología, Universidad de La Laguna, La Laguna, Tenerife, Spain

Full list of author information is available at the end of the article



ending in Australia [23]. This gradient is in the opposite direction to the expected by the southern dispersal route. Clearly, the fossil record in East Asia would be more compatible with a model proposing an earlier exit from Africa of modern humans that arrived to China following a northern route, around 100 kya. Indeed, this northern route model was evidenced from the relative relationships obtained for worldwide human populations using classical genetic markers [24, 25] and by the archaeological record [26]. Based on the phylogeography of mtDNA macrohaplogroup N, the existence of a northern route from the Levant that colonized Asia and carried modern humans to Australia was also inferred long ago [27]. However, this idea was ignored or considered a simplistic interpretation [28]. On the contrary, since the beginning, the coastal southern route hypothesis has only received occasional criticism from the genetics field [29], and discrepancies with other disciplines were mainly based on the age of exit from Africa of modern humans [30]. However, subsequent research from the fields of genetics, archaeology and paleoanthropology [31], have given additional support to the early northern route alternative. At this respect, a recent whole-genome analysis evaluating the presence of ancient Eurasian components in Egyptians and Ethiopians pointed to Egypt and Sinai as the more likely gateway in the exodus of modern humans out of Africa [32]. Furthermore, after a thoroughly revision of the evidence in support of a northern route signaled by mtDNA macrohaplogroup N [31], we realized that the phylogeny and phylogeography of mtDNA macrohaplogroup M fit better to a northern route accompanied by N than a southern coastal route as was previously suggested [27]. In fact, M in the Arabian Peninsula seems to have a recent historical implantation as in all western Eurasia. Moreover, the founder age of M in India is younger than in eastern Asia and Near Oceania and so, southern Asia might better be perceived as a receiver more than an emissary of M lineages. Recently, the unexpected detection of M lineages in Late Pleistocene European hunter-gatherers [33] has been explained as result of a back migration from the East, possibly mirroring the arrival to Africa of the haplogroup M1 in Paleolithic times [34–36], although a more ambitious interpretation has been formulated by others [37]. In this study, we propose a more conciliatory model to explain the history of *Homo sapiens* in Eurasia under the premise of an early exit from Africa following a sole northern route across the Levant to colonize the Old World.

Methods

Sampling information

A total of 206 samples from unrelated Saudi healthy donors belonging to mtDNA macrohaplogroup M were analyzed in this study, 163 of them previously published

[31, 38]. To fully characterize these M lineages, 17 complete mtDNA genomes from Saudi samples were sequenced. In addition, 7 unpublished complete mtDNA genomes from preceding studies were included [35, 39]. Only individuals with known maternal ancestors for at least three generations were considered in this study. Moreover, 4107 published complete or nearly complete mtDNA genomes belonging to macrohaplogroup M of Eurasian and Oceania origin were included in the analysis. To accurately establish the geographical ranges of the relatively rare M haplogroups, 73,215 partial sequences from the literature were screened as detailed in [31]. The procedure of human population sampling adhered to the tenets of the Declaration of Helsinki and written consent was recorded from all participants prior to taking part in the study. The study underwent formal review and was approved by the College of Medicine Ethical Committee of the King Saud University (proposal N° 09-659) and by the Ethics Committee for Human Research at the University of La Laguna (proposal NR157).

MtDNA sequencing

Total DNA was isolated from buccal or blood samples using the POREGENE DNA isolation kit from Genra Systems (Minneapolis, USA). The I and II hypervariable regions of mtDNA from 43 new Saudi Arabian samples were amplified and sequenced for both complementary strands as detailed in [40]. When necessary for unequivocal assortment into specific M subclades, the 206 Saudi M samples were additionally analyzed for haplogroup diagnostic SNPs using partial sequencing of the mtDNA fragments including those SNPs, or typed by SNaPshot multiplex reactions [41]. PCR conditions and sequencing of mtDNA genome were as previously published [27]. Successfully amplified products were sequenced for both complementary strands using the DYEnamic™ETDye terminator kit (Amersham Biosciences). Samples run on MegaBACE™ 1000 (Amersham Biosciences) according to the manufacturer's protocol. The 24 new complete mtDNA sequences have been deposited in GenBank with the accession numbers KR074233 to KR074256 (Additional file 1: Figure S1).

Previous published data compilation

Sequences belonging to specific M haplogroups were obtained from public databases such as NCBI, MITOMAP, the-1000 Genomes Project and the literature. We searched for mtDNA lineages directly using diagnostic SNPs (<http://www.mitomap.org/foswiki/bin/view/MITOMAP/WebHome>), or by submitting short fragments including those diagnostic SNPs to a BLAST search (<http://blast.st-va.ncbi.nlm.nih.gov/Blast.cgi>). Haplotypes extracted from the literature were transformed into sequences using the HaploSearch program [42]. Sequences

were manually aligned and compared to the rCRS [43] with BioEdit Sequence Alignment program [44]. Haplogroup assignment was performed by hand, screening for diagnostic positions or motifs at both hypervariable and coding regions whenever possible. Sequence alignment and haplogroup assignment was carried out twice by two independent researchers and any discrepancy resolved according to the PhyloTree database [45].

Phylogenetic analysis

Phylogenetic trees were constructed by means of the Network v4.6.1.2 program using the Reduced Median and the Median Joining algorithms in sequent order [46]. Resting reticulations were manually resolved attending to the relative mutation rate of the positions involved. Haplogroup branches were named following the nomenclature suggested by the PhyloTree database [45]. Coalescence ages were estimated by the statistics rho [47] and sigma [48], and the calibration rate proposed in [49]. Differences in coalescence ages were calculated by two-tailed t-tests. It was considered that the mean and standard error estimated for haplogroup ages from different samples and methods were normally distributed.

Global phylogeographic analysis

In this study, we deal with the earliest periods of the out of Africa spread. Given that subsequent demographic events probably eroded those early movements, spatial geographical distributions of haplogroups based on their contemporary frequencies or diversities were omitted. The presence/absence of basal M lineages in different areas was used to establish the present day geographical range for each haplogroup. Haplogroups with extensive geographic overlapping were considered as belonging to the same sub-continental region. AMOVA, CLUSTER and PCA analysis were performed to evaluate the level of geographical structure of the M haplogroups. For AMOVA we used the GenAlEx6.5 software, k-means clustering was obtained with XLSTAT statistical software, and PCA was performed using the Excel add-in Multibase package (Numerical Dynamics, Japan). The possible association between the haplogroup and fossil ages with their respective longitudinal and latitudinal geographical positions was tested by Pearson correlation analyses using the XLSTAT statistical software. For this purpose we took the intersection point between the two segments joining the largest latitudinal and longitudinal areas of each haplogroup as its geographic center. Geographic coordinates were obtained by Google Earth software (<https://earth.google.com>).

Geographic subdivision of India and regional haplogroup assignment

Attending only to geographic criteria, India was roughly subdivided into four different sampling areas: Northwest,

including Kashmir, Himachal Pradesh, Punjab, Haryana, Uttarakhand, Rajasthan, Uttar Pradesh, Gujarat, and Madhya Pradesh states; Southwest, including Maharashtra, Karnataka and Kerala; Northeast, including Bihar, Sikkim, Arunachal Pradesh, Assam, Nagaland, Meghalaya, Tripura, Jharkhand, West Bengal, Chhattisgarh and Orissa; and Southeast, represented by Andhra Pradesh, Tamil Nadu and Sri Lanka. We are very skeptical of the possibility that the actual genetic structure of India is the result of its original colonization, so the ethnic or linguistic affiliation of the samples were not considered but only its geographic origin. For the same reason, present day frequency and diversity of the haplogroups were not used but their geographic ranges and radiation ages. The criteria followed to assign haplogroups to different regions within India were a consistent detection in an area (at least 90 % of the samples) and absence or limited presence in the alternative areas (equal or less than 10 % of the samples). We considered widespread those haplogroups consistently found in all the Indian areas and also found in some of the surrounding areas as Pakistan or Iran at the west, Tibet or Nepal at the north, and Bangladesh or Myanmar at the east.

Results and discussion

Haplogroup M in western Eurasia with emphasis on Saudi Arabia

The lack of ancient and autochthonous mtDNA M lineages in western Eurasia is, at least, surprising if the colonization of the Old World by modern humans is thought to have begun through that region. Indeed, the presence of haplogroup M1 lineages in Mediterranean Europe and the Middle East has been explained as the result of secondary spreads from northern Africa where this haplogroup had a Paleolithic implantation [34–36]. Likewise, eastern Asian M lineages belonging to C, D, G and Z haplogroups mainly in Finno-Ugric-speaking populations of north and eastern Europe, seem to be the footprints of successive westward migration waves of Asiatic nomads occurred from Mesolithic period to historic times [50–58]. South Asian influences on the west have been also evidenced by the presence of Indian M4, M49 and M61 lineages in Mesopotamian remains [59]. In addition, ancestral mtDNA links between European Romani groups and northwest India populations were proved by the sharing of M5a1, M18, M25 and M35b lineages [60–64].

The mtDNA haplogroup M profile in Saudi Arabia represents about 7 % of the maternal gene pool [38, 65]. Of the 206 M haplotypes sampled (Additional file 2: Table S1), 53 % belong to the northern African M1 haplogroup, being both eastern African M1a and northern African M1b branches well represented (Additional file 2: Table S1 and Additional file 1: Figure S1). This fact contrasts with the sole presence of eastern M1a representatives in

Yemen [66]. Therefore, it is most probably that M1b lineages reached Saudi Arabia from the Levant. M lineages with indubitable Indian origin accounted for 39 % of the Saudi M pool, whereas the remaining 8 % would have a southeastern Asian source. The geographic origin of the Indian contribution seems not to be biased since 53 % of the lineages may be assigned to eastern Indian regions and 47 % to western [67–69]. However, it deserves mention that the two M lineages of the Andaman aborigines [15, 70, 71] are present in the Saudi samples. The isolate Ar2461 (Additional file 2: Table S1) has the diagnostic mutations of the Andaman branch M31a1 in the regulatory (249d, 16311) and coding (3975, 3999) regions. On the other hand, the complete mtDNA sequence of Ar1076 (Additional file 1: Figure S1) belongs to the M32c Andamanese branch, matching it with another complete sequence from Madagascar [72]. It must be stressed that this branch has been steadily found in all mtDNA reports on Madagascar [73–75]. Although these haplogroups were taken as evidence that Andamanese indigenous represent the descendants of the first out of Africa dispersal of modern humans [15, 70], more recent studies support a late Paleolithic colonization of the Andaman Islands [76, 77]. In fact, different branches of M31 have been found in the northeastern India, Nepal and Myanmar [71, 77–80], while M32c has been also detected in Indonesia [81] and Malaysia [82, 83]. Another interesting link involving India, Saudi Arabia and the Mauritius Island is the case of the haplogroup M81. It was first detected as a sole sequence in a LHON patient from India [84]. The Saudi Ar567 sample shares 215, 4254, 6620, 13590, 16129 and 16311 substitutions with this Indian sequence and, in addition it shares substitutions 151, 6170, 7954 and 16263 with a complete sequence from the Mauritius Island [85]. At first sight, the Mauritius sequence differs from that of Saudi Arabia by four different mutations occurring in a short segment of 11 bp, from 5742 to 5752. However, a closer inspection reveals they may represent two different interpretations: the C to G transversion at position 5743 in the Ma12 reading corresponds to the C deletion at the same position in the Ar567 lecture, and the G to A transition at the position 5746 in Ma12 corresponds to the A insertion at the position 5752 in Ar567. Therefore, both sequences can be only distinguished by the transition 12522, present in the Saudi sample and absent in Ma12 (Additional file 1: Figure S1). Curiously, these affinities between samples from Saudi Arabia and those from Indian Ocean islands can be extended to Saudi Q1a1 in Ar196 (Additional file 2: Table S1), which has only exact matches with MA405 sample from Madagascar [75]. Moreover, different lineages belonging to the Indian branch (M42b) in Saudi Arabia [31] and Mauritius [85] deeply link with the Australian M42a branch [86]. Other lineages in the Saudi mtDNA pool, such as M20, E1a1a1 and M7c1 point to

specific arrivals from southeastern Asia. The fact that all these lineages represent isolates in Saudi Arabia closely related to those found in their original areas, strongly supports the idea that they were incorporated to the Saudi pool in historical times as result of the Islamic expansion and by the import of workers from India and southeastern Asia by the European colonizers in the past and by the own Saudi nowadays [38]. In conclusion, Saudi Arabia lacks of ancient and autochthonous M haplogroups likewise the rest of western Eurasia.

About the origin of the North African haplogroup M1

The existence of haplogroup M lineages in Africa was first detected in Ethiopian populations by RFLP analysis [87]. Although an Asian influence was contemplated to explain the presence of this M component on the maternal Ethiopian pool, the dearth of M lineages in the Levant and its abundance in south Asia gave strength to the hypothesis that haplogroup M1 in Ethiopia was a genetic indicator of the southern route out of Africa. In addition, it was pointed out that probably this was the only successful early dispersal [88]. However, the limited geographic range and genetic diversity of M in Africa compared to India was used as an argument against this hypothesis [27, 34, 35, 67, 78, 89], instead proposing M1 as a signal of back-flow to Africa from the Indian subcontinent. However, after extensive phylogenetic and phylogeographic analyses for this marker [34–36, 67, 90], the supposed India to Africa connection was not found.

The detection in southeast Asia of new lineages that share with M1 the 14110 substitution [90, 91], gave rise to the definition of a new macrohaplogroup named M1'20'51 by PhyloTree.org Build 16 [44]. However, this substitution is not an invariable position (Additional file 2: Table S2) and, therefore, its sharing by common ancestry is not warranted [36]. We realized that, in addition, haplogroups M1 and M20 share transition 16129 which, although highly variable, would add support to this basal unification as the most parsimonious phylogenetic reconstruction (Additional file 1: Figure S1). However, a recent study reported a new mtDNA haplogroup from Myanmar, named M84, that also roots at the basal node represented by transition 14110 [92]. This haplogroup shares with M20 the transition 16272 which is more conservative than 16129 [48], therefore, weakening any specific relationship between haplogroups M1 and M20 beyond its common basal node. With the exception of M84, that seems to be limited to Myanmar, India and Southern China populations [92], the phylogeography of these haplogroups is extend and complex (Additional file 2: Table S3). M1 is found from Portugal and Senegal in the west to the Caucasus, Pakistan and Tibet at the east and, from Guinea-Bissau and Tanzania in the south to Russia at the north [34–36, 93–98] but, its highest diversity is found in Ethiopia and the Maghreb. The isolates detected at the

borders are lineages derived from the M1a branch in Russia and Tanzania and the M1b branch in Guinea Bissau and Tibet. The geographic ranges of M20 and M51 largely overlap showing a broad common area in the southeastern Asia including Myanmar and Malaysia at the west, Philippines and Hainan at the east and Tibet at the north-west (Additional file 2: Table S3). In addition, the western border of M20 further stretches to Bangladesh [99] and even to Assam in India [100] and the eastern border of M51 to Fujian and Taiwan [91]. The primary split between the ancestors of these four haplogroups probably occurred in its core area around 56 kya, coinciding with a warm climate period. Thus, it might be possible that the birth of the M1 ancestor was in southeastern Asia instead of India. Based on the scarcity and low diversity of M1 along the southern route [66], and on archaeological affinities between Levant and North Africa [34, 35], it has been suggested that the route followed by the M1 bearers to reach Africa was across the Levant not to the Bab el Mandeb strait. However, the current representatives of M1, from the Levant to the Tibet, are derived lineages of the ancestors in Africa. Until now, basal lineages of M1 have not been detected in the northern and southern hypothetical paths, so the only support for a Levantine route of M1 is the fact that all the other Eurasian lineages that returned to Africa in secondary backflows, such as N1, X1 or U6, had their most probable origins in the central or western Eurasia instead of India.

Geographical structure of the macrohaplogroup M genealogy

At global level, the mtDNA variation is phylogeographically structured [101]. For macrohaplogroup M, the regions of South Asia, East Asia and Southeast Asia have their characteristic sets of haplogroups with only minor overlapping. The same occurs in Melanesia and Australia. It is of paramount importance point out that these sets of haplogroups only share diagnostic mutations defining the basic M* node. This picture is interpreted as the result of secondary expansions from several geographically isolated centers which were reached by carriers of basic M* lineages during the primary earlier migrations. Congruently, the AMOVA analysis of 176 populations covering the main regions of Asia, Melanesia and Australia (Additional file 2: Table S4), shows that 85 % of the variance was found within populations and 15 % among the major regions ($p < 0.0001$). Furthermore, when populations were successively partitioned into k-clusters in order to minimize the within-cluster variance, the best partition was obtained for $k = 5$ (Table 1). The major regional differences explained 90.55 % of the variance. At this level, three clusters grouped together populations only belonging to Australia, Melanesia and South Asia respectively, a fourth cluster joined all Central Asian populations and the

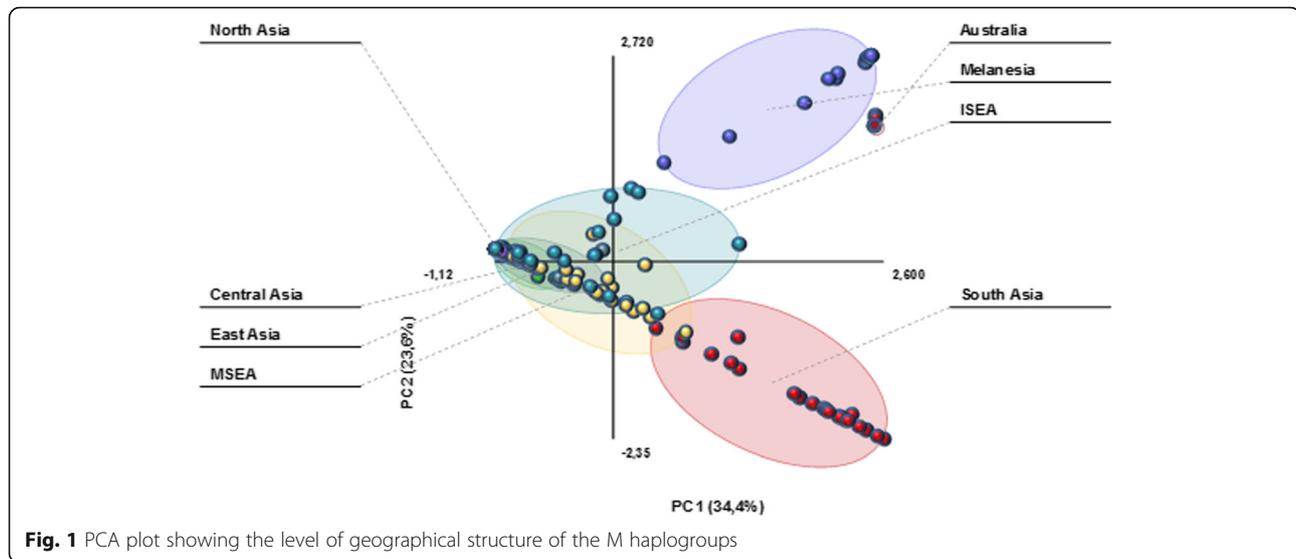
Table 1 AMOVA and k-mean Clustering results

Statistic	Variance (%)	
	Within populations	Between regions
AMOVA	85.00	15.00
k-2 Clustering	54.27	45.73
k-3 Clustering	33.00	67.00
k-4 Clustering	21.94	78.06
k-5 Clustering	9.45	90.55

majority of the East and North Asian populations together with a few Mainland (4) and Island (8) southeast Asian populations. Finally, the fifth cluster comprised the majority of the Mainland and Island southeast Asian populations and a few East (2) and South (3) Asian populations. These results are graphically visualized in the PCA plot (Fig. 1) where the first (34.4 %) and second (23.6 %) components accounted for 58 % of the variability. South Asia, Melanesia and Australia are nearly disjoint areas whereas the rest show important overlapping. As these regional genealogies can be transformed in coalescence ages, the relative role of each sub-continental area in the primitive human migrations can be approached.

The role of India in the origin and expansion of the macrohaplogroup M

The macrohaplogroup M in South Asia is characterized by a great diversity, deep coalescence age, and autochthonous nature of its lineages. These characteristics have been used for support the in-situ origin of the Indian M lineages and its rapid dispersal eastwards to colonize southeastern Asia following a southern coastal route [14, 15, 68]. However, there are several arguments against this hypothesis. Firstly, given that geographical barriers seem not to be stronger at the western border than at the eastern border of the subcontinent [102], it should be expected that the macrohaplogroup M would simultaneously radiated to both areas with India as its first center of expansion. However, while at the east its expansion crossed the Ganges and quickly reached Australia, at the west it seems to have found an insurmountable barrier at the Indus bank since M frequencies suddenly dropped from 65.5 ± 4.3 % in India to 5.3 ± 1.0 % in Iran ($p < 0.0001$) [67, 103–105]. Furthermore, unlike southeastern Asia and Australia, autochthonous M lineages have not been detected at the west of South Asia. It might be adduced that later massive expansions of western mtDNA lineages replaced the M lineages in those regions, but these western lineages have not been detected in India. Instead, it has been pointed out that the most of the extant mtDNA boundaries in South and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans [67]. In fact, only a small fraction of the specific western Asia



mtDNA lineages found in Indian populations can be ascribed to a relatively recent admixture [106].

Secondly, whenever it has been compared, the founder age of M in India is significantly ($p = 0.002$) younger than those in eastern and southeastern Asia and Australo-melanesian centers (Table 2). It has been argued that this could be due to an uneven distribution and sampling of M lineages in India [78, 90]. However, when the same weight is given to every lineage independently of its respective abundance (Table 1), the founder age slightly diminish in India and raises in East Asia compared to their respective weighted founder ages. Recently, it has been admitted that the initial radiation of macrohaplogroup M could have occurred in eastern India following the southern route [17].

Table 2 Founder ages (kya) for macrohaplogroup M in South and East Asia

South Asia	East Asia	SoutheastAsia	Oceania ^a	References
40.2 (20.8; 60.9)	55.2 (26.0; 86.7)			This study
38.4 (19.1; 58.9) ^b	56.4 (27.1; 88.1) ^b			This study
	51.6 (40.4; 63.1)	60.7 (47.4; 74.4)	68.3 (53.5; 83.6)	This study
			72.1 ± 8.0	[142]
44.6 ± 3.3	69.3 ± 5.4	55.7 ± 7.4	73.0 ± 7.9	[89]
		58.9 ± 13.6		[80]
66.0 ± 9.0 ^c	69.0 ± 7.0 ^c			[67]
36.0 ± 3.0 ^d	46.0 ± 5.0 ^d			[67]
49.4 (39.0; 60.2)	60.6 (47.3; 74.3)			[48]

^aIncluding Australia as in [143]

^bUnweighted rho

^cBased on coding region only

^dBased on synonymous positions only

In fact, the possibility that the first radiation of M were in East Asia instead of India was proposed long ago [27]. We think that it would be a hard task to detect today the original focus of the macrohaplogroup M expansion into India if it really occurred. Indeed, the mean radiation age (30.1 ± 9.3 ky) of the M haplogroups, that could be considered of widespread implantation in India, was significantly older (Two-tailed p value = 0.0147) than the mean radiation age (22.2 ± 9.6 ky) of those with more localized ranges (Additional file 2: Table S3). However, comparisons between mean radiation ages of Indian M haplogroups with eastern vs western ($p = 0.275$) and northern vs southern ($p = 0.580$) geographic ranges were not statistically significant. Furthermore, it must be taken into account the possibility that some M haplogroups with dominant implantation in northeastern India, such as M49, are secondary radiations from Myanmar [107, 108]. It is evident that the founder age of macrohaplogroup M increases eastwards from South Asia. Not only is the founder age of M in East Asia significantly greater than in South Asia but, the former is younger than the one estimated for southeast Asia despite of the probability value ($p = 0.0998$) did not reach significance probably because the estimation for the area calculated by [90] was previous to the detection of M clades with very deep ages in southeast Asia [91, 92, 109]. In turn, the founder age for M in Oceania was significantly older than the ones estimated for East Asia ($p = 0.004$) and southeast Asia ($p = 0.032$). Furthermore, there is a significant positive correlation ($R = 0.554$; $p < 0.0001$) between the age of the M haplogroups and its relative longitudinal position (Additional file 2: Table S4). This decreasing age gradient westwards is in accordance with the hypothesis that carriers of macrohaplogroup M lineages colonized India from the East instead the West. It could be argued that South Asia was eastward colonized very early by the southern

coastal expansion of modern humans out of Africa but those primitive mtDNA lineages were extinguished by genetic drift, and the subcontinent was recolonized latter by eastern groups left on the way to Australia. This argument is however in contradiction, first, with the suggestion based on past population size prediction, that the most of humanity lived in southern Asia approximately between 45 and 20 kya [110], and second, with the significantly older ($p = 0.004$) mean founder age of macrohaplogroup R in South Asia (62.5 ± 3.5 ky) than its M counterpart (45.7 ± 7.8 ky). These data are in agreement with a colonization of South Asia by two independent waves of settlers, already proposed for Indian populations according to the phylogeny and phylogeography of M and U haplogroups [106]. Curiously, this idea was abandoned in favor of a single southern route simultaneously carrying the three Eurasian mtDNA lineages (M, N and R). Now, if the born of macrohaplogroup M at some place between southeast Asia and near Oceania is accepted, the putative connection between Australia and India based on the nearly simultaneous radiation of M42a and M42b branches respectively [86] could be explained by the expansion from a nearly equidistant center of radiation, and not as a directional colonization from India to Australia following a southern route. At this respect, it seems pertinent to cite that the complexity of M42 in Australia could be greater than the currently known [111]. In addition, the tentative junction of M42 with the specific southeast Asian M74 lineage [45] based on a shared transition at 8251 gives additional support to this hypothesis.

Basal M superhaplogroups involving South Asia

As the most parsimonious topology, the existence of several M superhaplogroups, embracing different lineages based on one or a few moderately recurrent mutations have been recognized in PhyloTree.org Build 16 [44]. Following this trend we have, provisionally, unified haplogroup M11 and the recently defined haplogroup M82 [92] under macrohaplogroup M11'82 because both share the very conservative transition 8108 (Additional file 2: Table S2). Attending to their phylogeography these superhaplogroups usually link very wide geographic areas (Additional file 2: Table S2). In accordance with our suggestion that the carriers of Macrohaplogroup M colonized South Asia later than southeastern Asia and Oceania and that this mtDNA gene flow had an eastern origin we have observed that the mean radiation age of those superhaplogroups involving South Asia (39.70 ± 3.24 ky) is significantly younger ($p = 0.003$) than the one relating East and Southeast Asia (55.60 ± 2.94 ky). In this comparison we considered the South Asia-Southeast Asia-Australian link deduced from superhaplogroup M42'74 as specifically involving India. However, we considered superhaplogroup M62'68 as indicative of an East Asia-Southeast Asia link because M62 has

been found consistently in Tibet [112, 113] but only sporadically in northeast India [67].

The perspective from other genetic markers

Based on the absence of autochthonous N lineages in India and their deep age in southeastern Asia, we recently reasserted [31] the hypothesis that mtDNA macrohaplogroup N reached Australia following a northern route [27]. This time, we found additional support from Paleogenetics, which has demonstrated the introgression of DNA from Neanderthals [3, 4] and Denisovans [7, 114] of most probably northern geographical ranges, into the genome of modern humans [115]. For the specific case of India, we proposed that the N and R lineages arrived as secondary waves from the north following the Indus and Ganges-Brahmaputra banks. In our opinion, India was first colonized by modern humans from two external geographical centers situated at the northwest and northeast borders of this subcontinent [31]. The same picture was depicted, long ago, by other authors also based on mtDNA variation in India [106], although suggesting an eastward southern route for macrohaplogroup M in the same way as proposed by us [27]. Due to the lack of any mtDNA genetic evidence in support of an early migration along the southern route, we now propose an early exit of modern humans from Africa by the Levant and a unique northern route up to the Altai Mountains and then, obliged by harsh weather, down to southern China and beyond [31]. Strong evidence, supporting the primitive colonization of India by modern humans in two waves, has come from wide genome analysis. After analyzing 132 samples from 15 Indian states for more than 500,000 autosomal SNPs, the existence of two ancient populations genetically divergent was detected in India [116]. These are the ancestral northern Indians close to middle Eastern, central Asians and Europeans, and the ancestral southern Indians, distinct from the northern component and East Asians as they are from each other. Actually, the southern component is the most prevalent population in Andaman. Of particular interest is the fact that when populations of southeast Asia and Near Oceania were incorporated to these broad genome analyses, the Andaman Islanders showed a closer affinity with southeast rather than South Asian populations [117, 118]. It is evident that our mtDNA interpretation and the autosomal results give a very similar picture. In addition, it has been documented recently more Denisovan ancestry in South Asia than is expected based on existing models of history [119] but, again, the decreasing gradient of Denisovan ancestry detected from Oceanian populations to southeast and southern Asian populations easily fits in our model of an early expansion of macrohaplogroup M from the East to colonize the

Indian subcontinent. Furthermore, independent support for the existence of an early center of primitive modern humans in southeastern Asia originating very early expansions has recently come from improved phylogenetic resolution of the Y-chromosome K-M526 haplogroup. It has been detected a rapid diversification process of this haplogroup in southeast Asia-Oceania with subsequent westward expansions of the ancestors of R and Q haplogroups that make up the majority of paternal lineages in central Asia and Europe [120].

The fossil evidence

There is strong contradiction between the paleontological and the genetic interpretations about the origin of modern human in East Asia. New and reliable chronometric dating techniques applied to morphologically classified human remains have demonstrated the presence of early and fully modern humans in southern China at least since 80 kya which is in accordance to a regional continuity, or in situ evolution, of modern humans in East Asia [11, 121]. On the other hand, no apparent genetic contribution from earlier hominids was detected in the maternal [39, 91] and paternal [122–124] genetic pools of extant East Asian populations which has been taken as in support of a recent replacement of archaic humans by modern African incomers in East Asia. Although recently, genome analysis have detected introgression of Neanderthal [4] and Denisovan [7] DNA to the extant [125] and the ancient [126] genomes of modern humans in East Asia, this genetic contribution can be explained as a limited assimilation episode. Ancient DNA analysis of a morphologically early modern human from Tianyuan cave in northeast China [126] and a modern human from western Siberia of 45 ky old [127] evidenced that both were bearers of B and U mtDNA lineages respectively. These lineages are branches of haplogroup R which, in its turn, derives from macrohaplogroup N indicating thus that people in North Asia carried already fully modern mtDNA lineages around 45 kya, which is a hint of a northward secondary expansion of modern humans at that time. It seems pertinent to mention here, that in this study we found a weak but significant negative correlation ($r = -0.254$; $p = 0.029$) between the age of M haplogroups and their latitudinal geographical centers. On the other hand, dates of the East Asian fossil record (Additional file 2: Table S4) also showed a significant positive correlation ($r = 0.772$; $p = 0.0008$) with their respective southward latitudes from China. This is in agreement with our proposition that the out of Africa of modern humans occurred across the Levant and that the Skhul and Qafzeh fossil remains in Israel could be the first landmarks of that successful exit [31]. The subsequent evolution of those early modern humans in Asia might reconcile the replacement and continuity models into an inclusive synthesis. On the mtDNA side,

the main problem for this reconciliation would be an strict adherence to the chronological upper bound marked for the African exit by the coalescent age of macrohaplogroup L3 [17]. However, we think that mtDNA dating methods are still not reliable in absolute terms, mainly because we need accurate independent calibration for deep nodes and, in addition to selection, to take into account the effects of demographic parameters on the temporal variation of the substitution rate.

Turning to the probable existence of a primitive center of modern human expansion in southeast Asia, as proposed here on the basis of mtDNA haplogroup M phylogeography, the existence of a positive westward longitudinal gradient ($r = 0.551$) of the fossil record dates in southeastern Asia, with the older ages in Philippines and the youngest in Sri-Lanka (Additional file 2: Table S4), deserves mention. However, the correlation this time was not statistically significant ($p = 0.199$), mainly due to the lack of Paleolithic fossil evidence in Myanmar and India. Certainly, this counter-clock trend has been perceived by other authors as an argument against the recent straight forward out of Africa dispersal of *Homo sapiens* model [102].

The archaeological evidence

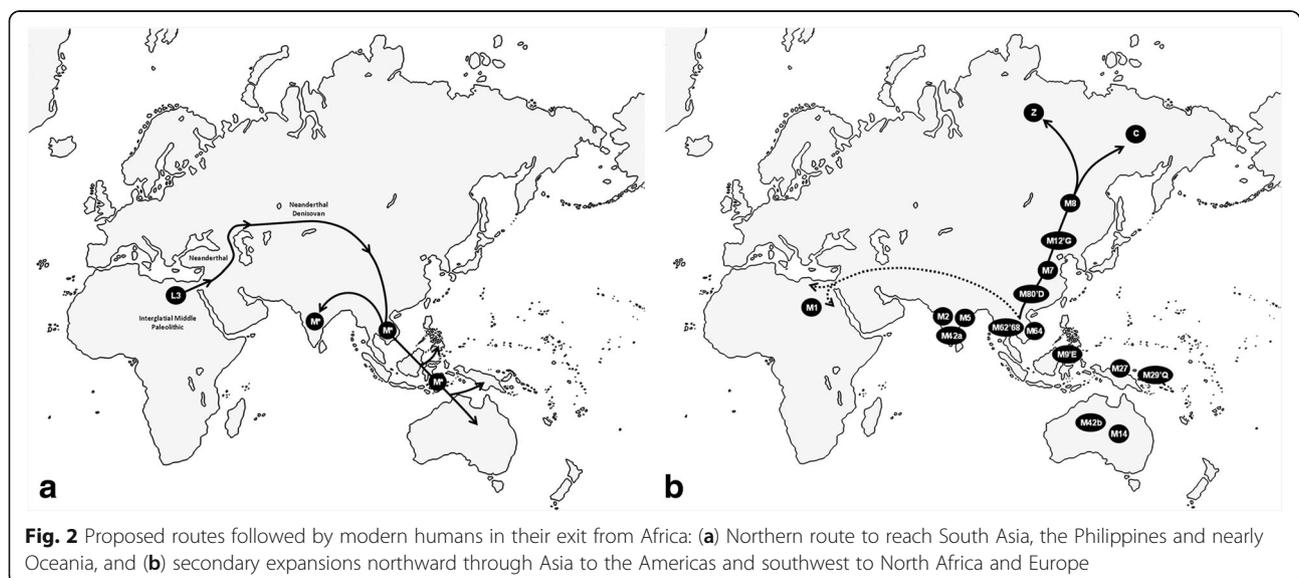
The archaeological record in East Asia also seems to be in support of the regional continuity model. The persistence and dominance of simple core-flake assemblages throughout the Paleolithic in this area sharply contrasts with the Upper Paleolithic technical and cultural innovations in western Eurasia [128]. Different lithic assemblages with potential resemblances in Africa and in southeastern Asia are crucial to interpret the southern dispersal route of modern humans across India. Blade-microblade and core-flake technologies are both present in the Indian subcontinent. They are usually contemporary and, in some cases, core-flake industries, such as the Soanian, even post-date the former [129]. Microlithics have been dated around 35-30 kya in southern India and Sri Lanka [130], although recently, it has been established that the microblade technology presents continuity in central India since 45 kya [131]. Some authors have proposed that the Indian microlithics had an African original source [17]. This model has problems to explain the absence of microblades eastwards of the subcontinent around that time, and the chronological gap between the oldest microlithic dates in India and the arrival of modern humans to Australia. Other authors consider microlithics in India as local innovations [130], however, the absence of an earlier blade technology in the Indian Paleolithic record makes an indigenous development unlikely [131]. Finally, from the age of the microblade technology in India, other authors have deduced that modern humans skirted the Indian subcontinent in the first dispersal out of Africa taking a northerly route through the middle East, central Asia, and southeastern Asia across

southern China [131]. For these authors, modern humans did not actually enter India until the time marked by the glacial climate of MIS 4. On the other hand, some core and flake industries in India have been considered as a link between those present in sub-Saharan Africa, southeast Asia and Australia [132–134]. Core sites in India with ages around 77 ky would be compatible with an early dispersal of modern humans from Africa. This model confronts problems with the later mtDNA molecular clock boundary proposed by some geneticists [16] and the lack of contemporary fossil record to confirm that this primitive technology was manufactured by modern humans. Finally, some Indian Late Pleistocene core-flake types as the Soanian seem to be related to similar industries in southeastern Asia, such as those of the Hoabinhian [135]. We suggest that microlithic technology could show the arrival of mtDNA macrohaplogroup R to India, following the north-west passage, as a branch of a global secondary dispersal of modern humans from some core area in west/central Asia that also affected Europe [26] and that later, reached North China across Siberia and Mongolia [136]. On the other hand, macrohaplogroup M came to India from southeast Asia following the northeast passage and carrying with them a simple core-flake technology.

Conclusions

A new and integrative model, explaining the time and routes followed by modern humans in their exit from Africa, is proposed in this study (Fig. 2). First, we think that the exit from Africa followed a northern route across the Levant, and that the fossils of early modern humans at Skhul and Qafzeh could be signals of this successful dispersal. These first modern humans carried undifferentiated mtDNA L3 lineages and brought primitive core-flake

technology to Eurasia [26]. The dates estimated for Skhul and Qafzeh remains (Additional file 2: Table S4) are slightly out of the range calculated for the age of mtDNA macrohaplogroup L3 (78.3, 95 % CI: 62.4; 94.9 kya) based on ancient mtDNA genomes [137]. However, they are in accordance with the presence of early modern humans in China around 100 kya, and with their subsequent presence in southeastern Asia about 70 kya. If we add the fact, also based on ancient genomes, that mtDNA lineages in northern Asia already belonged to derived B and U haplogroups around 45 kya [126, 127], we opine that the geneticists should resynchronize the mtDNA molecular clock with the Levant and East Asia fossil records instead of consider them as result of unsuccessful migrations. Second, those early modern humans went further northwards, some at least to the Altai Mountains, and in the way they occasionally mixed with other hominids as Neanderthal and Denisovans. Harsh climatic conditions dispersed them southwards erasing the mtDNA genetic footprints of this pioneer northern phase [31]. Third, the small surviving groups already carried basic N and M lineages. One of them, with only maternal N lineages, spread southwards to present-day southern China and probably, across the Sunda shelf reached Australia and the Philippines [31]. Fourth, other dispersed groups were the bearers of other N branches, including macrohaplogroup R, that enter India from the north, carrying with them the blade-microblade technology detected in this subcontinent. This technology also spread with other N and R branches to northern and western Eurasia, reaching Europe, the Levant and even northern Africa. Fifth, short after, another southeastern group carrying undifferentiated M lineages radiated from a core area, most probably localized in southeast Asia (including southern China),



reaching India westwards and near Oceania eastwards. These M haplogroup bearers brought with them at least one of the primitive core-flake technologies present in India that, therefore, had to have a southeastern Asian origin. Sixth, in subsequent mild climatic windows, demographic growth dispersed macrohaplogroup M and N northwards, most probably from overlapping areas that in time colonized northern Asia and the New World.

It seems that under archaeological grounds there are data in support of a southern route exit from Africa and across the Arabian Peninsula and the Indian subcontinent [30, 130, 132, 133, 138–141], but the lack of coetaneous fossil record leaves the hominid association to these stone tools unresolved. Even if they were modern humans, they did not leave any trace in the mtDNA gene pool of the extant populations of Arabia or India. However, there is no necessity to invoke the existence of a southern route to interpret the landscape depicted by the mtDNA phylogeny and phylogeography.

Additional files

Additional file 1: Figure S1. Phylogenetic tree of M complete sequences from this study. (XLSX 66 kb)

Additional file 2: Table S1. Haplogroup M in Saudi Arabia. In bold haplotypes from this study. Table S2. Super-haplogroup Ages and Geographic links. Table S3. Haplogroup M geographic ranges and ages in kiloyears (kya). Table S4. Populations used in the AMOVA and k-means clustering analyses. Table S5. Mitochondrial DNA M haplogroup ages and coordinates for their respective geographic centers used in the correlation analysis. Table S6. Modern human oldest fossil dating in different regions of Asia and oldest archaeological dating at the eastern side of the Wallace Line. (XLSX 91 kb)

Acknowledgements

We are grateful to Dr. Ana M. González for her contribution and helpful comments to this work. P. Marrero was supported by a postdoctoral grant from the Spanish Education Ministry (EX2009-950).

Availability of data and material

The sequence sets supporting the results of this article are available in the GenBank repository (Accession numbers: KR074233-KR074256). Additional file 2: Table S1 and Additional file 1: Figure S1. References for the published sequences used in this study are listed in supplementary. Additional file 2: Table S4 and can be directly retrieved from GenBank. All results obtained from our statistical analyses are presented in the tables and figures of this article and in the additional files.

Authors' contributions

VMC: Conceived and designed the study, analyzed the data and wrote the manuscript. PM: Edited and submitted mtDNA sequences, and contributed to the data analysis. KKA: Carried out the sequencing of the Arabian samples. JML: Contributed to the collection of sequence data and their analysis. All authors read and approved the final manuscript.

Authors' information

PM and KKA equally contributed to the study. VMC is actually retired.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The procedure of human population sampling adhered to the tenets of the Declaration of Helsinki. Written consent was recorded from all participants prior to taking part in the study. The study underwent formal review and was approved by the College of Medicine Ethical Committee of the King Saud University (proposal N° 09-659) and by the Ethics Committee for Human Research at the University of La Laguna (proposal NR157).

Author details

¹School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, Norfolk, UK. ²Glaucoma Research Chair, Department of ophthalmology, College of Medicine, King Saud University, Riyadh, Saudi Arabia. ³Departamento de Genética, Facultad de Biología, Universidad de La Laguna, La Laguna, Tenerife, Spain.

Received: 27 April 2016 Accepted: 28 October 2016

Published online: 10 November 2016

References

- Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature*. 1987;325:31–6.
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. *Science*. 1991; 253:1503–7.
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A Draft Sequence of the Neandertal Genome. *Science*. 2010;328:710–22.
- Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. 2014;505:43–9.
- Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, et al. A complete mtDNA genome of an early modern human from Kostenki, Russia. *Curr Biol CB*. 2010;20:231–6.
- Reich D, Patterson N, Kircher M, Delfin F, Nandineni MR, Pugach I, et al. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am J Hum Genet*. 2011;89:516–28.
- Meyer M, Kircher M, Gansauge M-T, Li H, Racimo F, Mallick S, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 2012;338:222–6.
- Stringer C. Why we are not all multiregionalists now. *Trends Ecol Evol*. 2014; 29:248–51.
- Gao X. Paleolithic Cultures in China: Uniqueness and Divergence. *Curr Anthropol*. 2013;54:S358–70.
- Liu W, Jin C-Z, Zhang Y-Q, Cai Y-J, Xing S, Wu X-J, et al. Human remains from Zhirendong, South China, and modern human emergence in East Asia. *Proc Natl Acad Sci U S A*. 2010;107:19201–6.
- Liu W, Martínón-Torres M, Cai Y, Xing S, Tong H, Pei S, et al. The earliest unequivocally modern humans in southern China. *Nature*. 2015;526:696–9.
- Bae CJ, Wang W, Zhao J, Huang S, Tian F, Shen G. Modern human teeth from Late Pleistocene Luna Cave (Guangxi, China). *Quat Int*. 2014;354:169–83.
- Kuhlwil M, Gronau I, Hubisz MJ, de Filippo C, Prado-Martinez J, Kircher M, et al. Ancient gene flow from early modern humans into Eastern Neanderthals. *Nature*. 2016;530:429–33.
- Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, et al. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*. 2005;308:1034–6.
- Thangaraj K, Chaubey G, Kivisild T, Reddy AG, Singh VK, Rasalkar AA, et al. Reconstructing the origin of Andaman Islanders. *Science*. 2005;308:996.
- Fernandes V, Alshamali F, Alves M, Costa MD, Pereira JB, Silva NM, et al. The Arabian cradle: mitochondrial relicts of the first steps along the southern route out of Africa. *Am J Hum Genet*. 2012;90:347–55.
- Mellars P, Gori KC, Carr M, Soares PA, Richards MB. Genetic and archaeological perspectives on the initial modern human colonization of southern Asia. *Proc Natl Acad Sci U S A*. 2013;110:10699–704.
- Bar-Yosef O. Site formation processes from a Levantine viewpoint. *Form. Process. Archaeol. Context. Goldberg, P., Nash, D.T., Petraglia, M.D. Madison: Prehistory Press; 1993. p. 13–32.*
- Détroit F, Dizon E, Falguères C, Hameau S, Ronquillo W, Sémah F. Upper Pleistocene Homo sapiens from the Tabon cave (Palawan, The Philippines): description and dating of new discoveries. *Comptes Rendus Palevol*. 2004;3:705–12.
- Barker G, Barton H, Bird M, Daly P, Datan I, Dykes A, et al. The “human revolution” in lowland tropical Southeast Asia: the antiquity and behavior of

- anatomically modern humans at Niah Cave (Sarawak, Borneo). *J Hum Evol.* 2007;52:243–61.
21. Mijares AS, Détroit F, Piper P, Grün R, Bellwood P, Aubert M, et al. New evidence for a 67,000-year-old human presence at Callao Cave, Luzon, Philippines. *J Hum Evol.* 2010;59:123–32.
 22. Demeter F, Shackelford LL, Bacon A-M, Düringer P, Westaway K, Sayavongkhamdy T, et al. Anatomically modern human in Southeast Asia (Laos) by 46 ka. *Proc Natl Acad Sci U S A.* 2012;109:14375–80.
 23. Bowler JM, Johnston H, Olley JM, Prescott JR, Roberts RG, Shawcross W, et al. New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature.* 2003;421:837–40.
 24. Cavalli-Sforza LL, Piazza A, Menozzi P, Mountain J. Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proc Natl Acad Sci U S A.* 1998;85:6002–6.
 25. Nei M, Roychoudhury AK. Evolutionary relationships of human populations on a global scale. *Mol Biol Evol.* 1993;10:927–43.
 26. Marks AE. Comments after four decades of research on the Middle to Upper Paleolithic transition. *Mitteilungen Ges Für Urgesch.* 2005;14:81–6.
 27. Maca-Meyer N, González AM, Larruga JM, Flores C, Cabrera VM. Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet.* 2001;2:13.
 28. Palanichamy MG, Sun C, Agrawal S, Bandelt H-J, Kong Q-P, Khan F, et al. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Am J Hum Genet.* 2004;75:966–78.
 29. Cordaux R, Stoneking M. South Asia, the Andamanese, and the genetic evidence for an “early” human dispersal out of Africa. *Am J Hum Genet.* 2003;72:1586–1590–1593.
 30. Petraglia MD, Alsharekh A, Breeze P, Clarkson C, Crassard R, Drake NA, et al. Hominin dispersal into the Nefud Desert and Middle palaeolithic settlement along the Jubbah Palaeolake, Northern Arabia. *PLoS One.* 2012;7:e49840.
 31. Fregel R, Cabrera V, Larruga JM, Abu-Amero KK, González AM. Carriers of Mitochondrial DNA Macrohaplogroup N Lineages Reached Australia around 50,000 Years Ago following a Northern Asian Route. *PLoS One.* 2015;10:e0129839.
 32. Pagani L, Schiffels S, Gurdasani D, Danecek P, Scally A, Chen Y, et al. Tracing the route of modern humans out of Africa by using 225 human genome sequences from Ethiopians and Egyptians. *Am J Hum Genet.* 2015;96:986–91.
 33. Posth C, Renaud G, Mittnik A, Drucker DG, Rougier H, Cupillard C, et al. Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr Biol CB.* 2016;26:827–33.
 34. Olivieri A, Achilli A, Pala M, Battaglia V, Fornarino S, Al-Zahery N, et al. The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science.* 2006;314:1767–70.
 35. González AM, Larruga JM, Abu-Amero KK, Shi Y, Pestano J, Cabrera VM. Mitochondrial lineage M1 traces an early human backflow to Africa. *BMC Genomics.* 2007;8:223.
 36. Pennarun E, Kivisild T, Metspalu E, Metspalu M, Reisberg T, Moisan J-P, et al. Divorcing the Late Upper Palaeolithic demographic histories of mtDNA haplogroups M1 and U6 in Africa. *BMC Evol Biol.* 2012;12:234.
 37. Richards MB, Soares P, Torroni A. Palaeogenomics: Mitogenomes and migrations in Europe’s past. *Curr Biol.* 2016;26:R243–6.
 38. Abu-Amero KK, Larruga JM, Cabrera VM, González AM. Mitochondrial DNA structure in the Arabian Peninsula. *BMC Evol Biol.* 2008;8:45.
 39. Tanaka M, Cabrera VM, González AM, Larruga JM, Takeyasu T, Fuku N, et al. Mitochondrial genome variation in eastern Asia and the peopling of Japan. *Genome Res.* 2004;14:1832–50.
 40. Bekada A, Fregel R, Cabrera VM, Larruga JM, Pestano J, Benhamamouch S, et al. Introducing the Algerian mitochondrial DNA and Y-chromosome profiles into the North African landscape. *PLoS One.* 2013;8:e56775.
 41. Quintáns B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A. Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Forensic Sci Int.* 2004;140:251–7.
 42. Fregel R, Delgado S. HaploSearch: a tool for haplotype-sequence two-way transformation. *Mitochondrion.* 2011;11:366–7.
 43. Andrews RM, Kubacka I, Chinney PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet.* 1999;23:147.
 44. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 1999;41:95–8.
 45. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat.* 2009;30:E386–394.
 46. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 1999;16:37–48.
 47. Forster P, Harding R, Torroni A, Bandelt HJ. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet.* 1996;59:935–45.
 48. Saillard J, Forster P, Lynnerup N, Bandelt HJ, Nørby S. mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. *Am J Hum Genet.* 2000;67:718–26.
 49. Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, et al. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet.* 2009;84:740–59.
 50. Lahermo P, Laitinen V, Sistonen P, Béres J, Karcagi V, Savontaus ML. MtDNA polymorphism in the Hungarians: comparison to three other Finno-Ugric-speaking populations. *Hereditas.* 2000;132:35–42.
 51. Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogväli E-L, et al. The western and eastern roots of the Saami—the story of genetic “outliers” told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet.* 2004;74:661–82.
 52. Ingman M, Gyllenstein U. A recent genetic link between Sami and the Volga-Ural region of Russia. *Eur J Hum Genet EJHG.* 2007;15:115–20.
 53. Nadasi E, Gyurus P, Czákó M, Bene J, Kosztolányi S, Fazekas S, et al. Comparison of mtDNA haplogroups in Hungarians with four other European populations: a small incidence of descents with Asian origin. *Acta Biol Hung.* 2007;58:245–56.
 54. Tömöry G, Csányi B, Bogácsi-Szabó E, Kalmár T, Czibula A, Csoos A, et al. Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations. *Am J Phys Anthropol.* 2007;134:354–68.
 55. Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, Savontaus M-L, et al. Migration waves to the Baltic Sea region. *Ann Hum Genet.* 2008;72:337–48.
 56. Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Rogalla U, Perkova M, et al. Origin and post-glacial dispersal of mitochondrial DNA haplogroups C and D in northern Asia. *PLoS One.* 2010;5:e15214.
 57. Derenko M, Malyarchuk B, Denisova G, Perkova M, Rogalla U, Grzybowski T, et al. Complete mitochondrial DNA analysis of eastern Eurasian haplogroups rarely found in populations of northern Asia and eastern Europe. *PLoS One.* 2012;7:e32179.
 58. Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilo A, Koshel S, et al. Ancient DNA reveals prehistoric gene-flow from Siberia in the complex human population history of North East Europe. *PLoS Genet.* 2013;9:e1003296.
 59. Witas HW, Tomczyk J, Jędrzychowska-Darńska K, Chaubey G, Płoszaj T. mtDNA from the early Bronze Age to the Roman period suggests a genetic link between the Indian subcontinent and Mesopotamian cradle of civilization. *PLoS One.* 2013;8:e73682.
 60. Gresham D, Morar B, Underhill PA, Passarino G, Lin AA, Wise C, et al. Origins and divergence of the Roma (gypsies). *Am J Hum Genet.* 2001;69:1314–31.
 61. Kalaydjieva L, Calafell F, Jobling MA, Angelicheva D, de Knijff P, Rosser ZH, et al. Patterns of inter- and intra-group genetic diversity in the Vlach Roma as revealed by Y chromosome and mitochondrial DNA lineages. *Eur J Hum Genet EJHG.* 2001;9:97–104.
 62. Malyarchuk BA, Perkova MA, Derenko MV, Vanecek T, Lazur J, Gómolcák P. Mitochondrial DNA variability in Slovaks, with application to the Roma origin. *Ann Hum Genet.* 2008;72:228–40.
 63. Mendizabal I, Valente C, Gusmão A, Alves C, Gomes V, Goios A, et al. Reconstructing the Indian origin and dispersal of the European Roma: a maternal genetic perspective. *PLoS One.* 2011;6:e15988.
 64. Gómez-Carballa A, Pardo-Seco J, Fachal L, Vega A, Cebej M, Martínón-Torres N, et al. Indian signatures in the westernmost edge of the European Romani diaspora: new insight from mitogenomes. *PLoS One.* 2013;8:e75397.
 65. Abu-Amero KK, González AM, Larruga JM, Bosley TM, Cabrera VM. Eurasian and African mitochondrial DNA influences in the Saudi Arabian population. *BMC Evol Biol.* 2007;7:32.
 66. Vyas DN, Kitchen A, Miró-Herrans AT, Pearson LN, Al-Meerri A, Mulligan CJ. Bayesian analyses of Yemeni mitochondrial genomes suggest multiple migration events with Africa and Western Eurasia. *Am J Phys Anthropol.* 2016;159:382–93.
 67. Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, et al. Most of the extant mtDNA boundaries in south and southwest Asia were

- likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet.* 2004;5:26.
68. Chandrasekar A, Kumar S, Sreenath J, Sarkar BN, Urade BP, Mallick S, et al. Updating phylogeny of mitochondrial DNA macrohaplogroup m in India: dispersal of modern human in South Asian corridor. *PLoS One.* 2009;4:e7447.
 69. Maji S, Krithika S, Vasulu TS. Phylogeographic distribution of mitochondrial DNA macrohaplogroup M in India. *J Genet.* 2009;88:127–39.
 70. Endicott P, Gilbert MTP, Stringer C, Lalueza-Fox C, Willerslev E, Hansen AJ, et al. The genetic origins of the Andaman Islanders. *Am J Hum Genet.* 2003;72:178–84.
 71. Barik SS, Sahani R, Prasad BVR, Endicott P, Metspalu M, Sarkar BN, et al. Detailed mtDNA genotypes permit a reassessment of the settlement and population structure of the Andaman Islands. *Am J Phys Anthropol.* 2008;136:19–27.
 72. Dubut V, Cartault F, Payet C, Thionville M-D, Murail P. Complete mitochondrial sequences for haplogroups M23 and M46: insights into the Asian ancestry of the Malagasy population. *Hum Biol.* 2009;81:495–500.
 73. Hurler ME, Sykes BC, Jobling MA, Forster P. The dual origin of the Malagasy in Island Southeast Asia and East Africa: evidence from maternal and paternal lineages. *Am J Hum Genet.* 2005;76:894–901.
 74. Tofanelli S, Bertocchini S, Castrì L, Luiselli D, Calafell F, Donati G, et al. On the origins and admixture of Malagasy: new evidence from high-resolution analyses of paternal and maternal lineages. *Mol Biol Evol.* 2009;26:2109–24.
 75. Capredon M, Brucato N, Tonasso L, Choessel-Cadamuro V, Ricaut F-X, Razafindrazaka H, et al. Tracing Arab-Islamic inheritance in Madagascar: study of the Y-chromosome and mitochondrial DNA in the Antemoro. *PLoS One.* 2013;8:e80932.
 76. Palanichamy MG, Agrawal S, Yao Y-G, Kong Q-P, Sun C, Khan F, et al. Comment on “Reconstructing the origin of Andaman islanders.”. *Science.* 2006;311:470. author reply 470.
 77. Wang H-W, Mitra B, Chaudhuri TK, Palanichamy MG, Kong Q-P, Zhang Y-P. Mitochondrial DNA evidence supports northeast Indian origin of the aboriginal Andamanese in the Late Paleolithic. *J Genet Genomics.* 2011;38:117–22. Yi Chuan Xue Bao.
 78. Thangaraj K, Chaubey G, Singh VK, Vanniarajan A, Thanseem I, Reddy AG, et al. In situ origin of deep rooting lineages of mitochondrial Macrohaplogroup “M” in India. *BMC Genomics.* 2006;7:151.
 79. Reddy BM, Langstieh BT, Kumar V, Nagaraja T, Reddy ANS, Meka A, et al. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia. *PLoS One.* 2007;2:e1141.
 80. Fornarino S, Pala M, Battaglia V, Maranta R, Achilli A, Modiano G, et al. Mitochondrial and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation. *BMC Evol Biol.* 2009;9:154.
 81. Hill C, Soares P, Mormina M, Macaulay V, Clarke D, Blumbach PB, et al. A mitochondrial stratigraphy for island southeast Asia. *Am J Hum Genet.* 2007;80:29–43.
 82. Nur Haslindawaty AR, Panneerchelvan S, Edinur HA, Norazmi MN, Zafarina Z. Sequence polymorphisms of mtDNA HV1, HV2, and HV3 regions in the Malay population of Peninsular Malaysia. *Int J Legal Med.* 2010;124:415–26.
 83. Maruyama S, Nohira-Koike C, Minaguchi K, Nambiar P. MtDNA control region sequence polymorphisms and phylogenetic analysis of Malay population living in or around Kuala Lumpur in Malaysia. *Int J Legal Med.* 2010;124:165–70.
 84. Khan NA, Govindaraj P, Soumitra N, Srilekha S, Ambika S, Vanniarajan A, et al. Haplogroup heterogeneity of LHON patients carrying the m.14484 T > C mutation in India. *Invest. Ophthalmol Vis Sci.* 2013;54:3999–4005.
 85. Fregel R, Seetah K, Betancor E, Suárez NM, Caval D, Caval S, et al. Multiple ethnic origins of mitochondrial DNA lineages for the population of Mauritius. *PLoS One.* 2014;9:e93294.
 86. Kumar S, Ravuri RR, Koneru P, Urade BP, Sarkar BN, Chandrasekar A, et al. Reconstructing Indian-Australian phylogenetic link. *BMC Evol Biol.* 2009;9:173.
 87. Passarino G, Semino O, Quintana-Murci L, Excoffier L, Hammer M, Santachiara-Benerecetti AS. Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *Am J Hum Genet.* 1998;62:420–34.
 88. Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. Genetic evidence of an early exit of Homo sapiens from Africa through eastern Africa. *Nat Genet.* 1999;23:437–41.
 89. Roychoudhury S, Roy S, Basu A, Banerjee R, Vishwanathan H, Usha Rani MV, et al. Genomic structures and population histories of linguistically distinct tribal groups of India. *Hum Genet.* 2001;109:339–50.
 90. Sun C, Kong Q-P, Palanichamy MG, Agrawal S, Bandelt H-J, Yao Y-G, et al. The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes. *Mol Biol Evol.* 2006;23:683–90.
 91. Kong Q-P, Sun C, Wang H-W, Zhao M, Wang W-Z, Zhong L, et al. Large-scale mtDNA screening reveals a surprising matrilineal complexity in east Asia and its implications to the peopling of the region. *Mol Biol Evol.* 2011;28:513–22.
 92. Peng M-S, He J-D, Liu H-X, Zhang Y-P. Tracing the legacy of the early Hainan Islanders—a perspective from mitochondrial DNA. *BMC Evol Biol.* 2011;11:46.
 93. Kivisild T, Reidla M, Metspalu E, Rosa A, Brehm A, Pennarun E, et al. Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. *Am J Hum Genet.* 2004;75:752–70.
 94. Gonder MK, Mortensen HM, Reed FA, de Sousa A, Tishkoff SA. Whole-mtDNA genome sequence analysis of ancient African lineages. *Mol Biol Evol.* 2007;24:757–68.
 95. Zhao M, Kong Q-P, Wang H-W, Peng M-S, Xie X-D, Wang W-Z, et al. Mitochondrial genome evidence reveals successful Late Paleolithic settlement on the Tibetan Plateau. *Proc Natl Acad Sci U S A.* 2009;106:21230–5.
 96. Malyarchuk B, Derenko M, Denisova G, Kravtsova O. Mitogenomic diversity in Tatars from the Volga-Ural region of Russia. *Mol Biol Evol.* 2010;27:2220–6.
 97. Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, Metspalu E, et al. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol.* 2012;29:359–65.
 98. Siddiqi MH, Akhtar T, Rakha A, Abbas G, Ali A, Haider N, et al. Genetic characterization of the Makrani people of Pakistan from mitochondrial DNA control-region data. *Leg Med Tokyo Jpn.* 2015;17:134–9.
 99. Gazi NN, Tamang R, Singh VK, Ferdous A, Pathak AK, Singh M, et al. Genetic structure of Tibeto-Burman populations of Bangladesh: evaluating the gene flow along the sides of Bay-of-Bengal. *PLoS One.* 2013;8:e75064.
 100. Cordaux R, Saha N, Bentley GR, Aunger R, Sirajuddin SM, Stoneking M. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. *Eur J Hum Genet EJHG.* 2003;11:253–64.
 101. Underhill PA, Kivisild T. Use of Y-chromosome and mitochondrial DNA population structure in tracing human migrations. *Au Rev G.* 2007;41:539–64.
 102. Boivin N, Fuller DQ, Dennell R, Allaby R, Petraglia MD. Human dispersal across diverse environments of Asia during the Upper Pleistocene. *Quat Int.* 2013;300:32–47.
 103. Ashrafian-Bonab M, Lawson Handley LJ, Balloux F. Is urbanization scrambling the genetic structure of human populations? A case study. *Heredity.* 2007;98:151–6.
 104. Farjadian S, Sazzini M, Tofanelli S, Castrì L, Taglioli L, Pettener D, et al. Discordant patterns of mtDNA and ethno-linguistic variation in 14 Iranian Ethnic groups. *Hum Hered.* 2011;72:73–84.
 105. Derenko M, Malyarchuk B, Bahmanimehr A, Denisova G, Perkova M, Farjadian S, et al. Complete mitochondrial DNA diversity in Iranians. *PLoS One.* 2013;8:e80673.
 106. Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, et al. Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr Biol CB.* 1999;9:1331–4.
 107. Li Y-C, Wang H-W, Tian J-Y, Liu L-N, Yang L-Q, Zhu C-L, et al. Ancient inland human dispersals from Myanmar into interior East Asia since the Late Pleistocene. *Sci Rep.* 2015;5:9473.
 108. Summerer M, Horst J, Erhart G, Weißensteiner H, Schönherr S, Pacher D, et al. Large-scale mitochondrial DNA analysis in Southeast Asia reveals evolutionary effects of cultural isolation in the multi-ethnic population of Myanmar. *BMC Evol Biol.* 2014;14:17.
 109. Zhang X, Qi X, Yang Z, Serey B, Sovannary T, Bunnath L, et al. Analysis of mitochondrial genome diversity identifies new and ancient maternal lineages in Cambodian aborigines. *Nat Commun.* 2013;4:2599.
 110. Atkinson QD, Gray RD, Drummond AJ. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Mol Biol Evol.* 2008;25:468–74.
 111. Ballantyne KN, van Oven M, Ralf A, Stoneking M, Mitchell RJ, van Oorschot RA, Kayser M. MtDNA SNP multiplexes for efficient inference of matrilineal genetic ancestry within Oceania. *Forensic Sci Int Genet.* 2012;6:425–36.
 112. Qin Z, Yang Y, Kang L, Yan S, Cho K, Cai X, et al. A mitochondrial revelation of early human migrations to the Tibetan Plateau before and after the last glacial maximum. *Am J Phys Anthropol.* 2010;143:555–69.

113. Qi X, Cui C, Peng Y, Zhang X, Yang Z, Zhong H, et al. Genetic evidence of paleolithic colonization and neolithic expansion of modern humans on the tibetan plateau. *Mol Biol Evol.* 2013;30:1761–78.
114. Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature.* 2010;468:1053–60.
115. Sawyer S, Renaud G, Viola B, Hublin J-J, Gansauge M-T, Shunkov MV, et al. Nuclear and mitochondrial DNA sequences from two Denisovan individuals. *Proc Natl Acad Sci U S A.* 2015;112:15696–700.
116. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. *Nature.* 2009;461:489–94.
117. Chaubey G, Endicott P. The Andaman Islanders in a regional genetic context: reexamining the evidence for an early peopling of the archipelago from South Asia. *Hum Biol.* 2013;85:153–72.
118. Aghakhanian F, Yunus Y, Naidu R, Jinam T, Manica A, Hoh BP, et al. Unravelling the genetic history of Negritos and indigenous populations of Southeast Asia. *Genome Biol Evol.* 2015;7:1206–15.
119. Sankararaman S, Mallick S, Patterson N, Reich D. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Curr Biol.* 2016;26:1–7.
120. Karafet TM, Mendez FL, Sudoyo H, Lansing JS, Hammer MF. Improved phylogenetic resolution and rapid diversification of Y-chromosome haplogroup K-M526 in Southeast Asia. *Eur J Hum Genet.* 2015;23:369–73.
121. Smith FH, Spencer F. The origins of modern humans. New York: Liss; 1984.
122. Ke Y. African Origin of Modern Humans in East Asia: A Tale of 12,000 Y Chromosomes. *Science.* 2001;292:1151–3.
123. Shi H, Dong Y-L, Wen B, Xiao C-J, Underhill PA, Shen P-D, et al. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. *Am J Hum Genet.* 2005;77:408–19.
124. Wang C-C, Li H. Inferring human history in East Asia from Y chromosomes. *Investig Genet.* 2013;4:11.
125. Skoglund P, Jakobsson M. Archaic human ancestry in East Asia. *Proc Natl Acad Sci U S A.* 2011;108:18301–6.
126. Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl Acad Sci U S A.* 2013;110:2223–7.
127. Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature.* 2014;514:445–9.
128. Bar-Yosef O. The Upper Paleolithic Revolution. *Annu Rev Anthropol.* 2002;31:363–93.
129. The MS, Palaeolithic L. A review of recent findings. *Mishra 2008 Low. Palaeolithic Rev. Recent Find. Man Environ.* 2008;33:14–29.
130. Dennell R, Petraglia MD. The dispersal of *Homo sapiens* across southern Asia: how early, how often, how complex? *Quat Sci Rev.* 2012;47:15–22.
131. Mishra S, Chauhan N, Singhvi AK. Continuity of microblade technology in the Indian Subcontinent since 45 ka: implications for the dispersal of modern humans. *PLoS One.* 2013;8:e69280.
132. Petraglia M, Korisettar R, Boivin N, Clarkson C, Ditchfield P, Jones S, et al. Middle Paleolithic assemblages from the Indian subcontinent before and after the Toba super-eruption. *Science.* 2007;317:114–6.
133. Clarkson C, Petraglia M, Korisettar R, Haslam M, Boivin N, Crowther A, et al. The oldest and longest enduring microlithic sequence in India: 35 000 years of modern human occupation and change at the Jwalapuram Locality 9 rockshelter. *Antiquity.* 2009;83:326–48.
134. Haslam M, Clarkson C, Petraglia MD, Korisettar R, Bora J, Boivin N, et al. Indian lithic technology prior to the 74,000 BP Toba super-eruption: searching for an early modern human signature. *Up. Palaeolithic Revolut. Glob. Perspect. Pap. Honour Sir Paul Mellars.* Boyle, K.V., Gamble, C., Bar-Yosef, O. Cambridge: MacDonald Institute of Archaeology; 2010. p. 73–84.
135. Gaillard C, Singh M, Malassé AD. Late Pleistocene to Early Holocene Lithic industries in the southern fringes of the Himalaya. *Quat Int.* 2011;229:112–22.
136. Qu T, Bar-Yosef O, Wang Y, Wu X. The Chinese Upper Paleolithic: Geography, Chronology, and Techno-typology. *J Archaeol Res.* 2013;21:1–73.
137. Fu Q, Mitnik A, Johnson PLF, Bos K, Lari M, Bollongino R, et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol CB.* 2013;23:553–9.
138. Rose JI. New Light on Human Prehistory in the Arabo-Persian Gulf Oasis. *Curr Anthropol.* 2010;51:849–83.
139. Armitage SJ, Jasim SA, Marks AE, Parker AG, Usik VI, Uerpmann H-P. The southern route “out of Africa”: evidence for an early expansion of modern humans into Arabia. *Science.* 2011;331:453–6.
140. Scerri EML, Groucutt HS, Jennings RP, Petraglia MD. Unexpected technological heterogeneity in northern Arabia indicates complex Late Pleistocene demography at the gateway to Asia. *J Hum Evol.* 2014;75:125–42.
141. Groucutt HS, Scerri EML, Lewis L, Clark-Balzan L, Blinkhorn J, Jennings RP, et al. Stone tool assemblages and models for the dispersal of *Homo sapiens* out of Africa. *Quat Int.* 2015;382:8–30.
142. Merriwether DA, Hodgson JA, Friedlaender FR, Allaby R, Cerchio S, Koki G, et al. Ancient mitochondrial M haplogroups identified in the Southwest Pacific. *Proc Natl Acad Sci U S A.* 2005;102:13034–9.
143. Kayser M. The human genetic history of Oceania: near and remote views of dispersal. *Curr Biol.* 2010;20:194–201.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

