Am. J. Hum. Genet. 64:1166–1176, 1999

mtDNA Analysis of Nile River Valley Populations: A Genetic Corridor or a Barrier to Migration?

Matthias Krings,[1] Abd-el Halim Salem,[2] Karin Bauer,[1] Helga Geisert,[1] Adel K. Malek,[3] Louis Chaix,[4] Christian Simon,[4] Derek Welsby,[5] Anna Di Rienzo,[6] Gerd Utermann,[7] Antti Sajantila,[8] Svante Pääbo,[1] and Mark Stoneking[1, 9]

[1]Max Planck Institute for Evolutionary Anthropology, Leipzig, and Department of Zoology, University of Munich, Munich; [2]Anatomy Department, Faculty of Medicine, Suez Canal University, Ismaelia, Egypt; [3]Department of Anatomy, College of Medicine, University of Assiut, Assiut, Egypt; [4]Museé d'Histoire Naturelle, Departement d'Archezoologie, Route de Malagnou, Geneva; [5]Department of Egyptian Antiquities, The British Museum, London; [6]Department of Human Genetics, The University of Chicago, Chicago; [7]Institut für Medizinische Biologie und Genetik, Universität Innsbruck, Innsbruck; [8]Department of Forensic Medicine, University of Helsinki, Helsinki; and [9]Department of Anthropology, Pennsylvania State University, University Park

Summary

To assess the extent to which the Nile River Valley has been a corridor for human migrations between Egypt and sub-Saharan Africa, we analyzed mtDNA variation in 224 individuals from various locations along the river. Sequences of the first hypervariable segment (HV1) of the mtDNA control region and a polymorphic *Hpa*I site at position 3592 allowed us to designate each mtDNA as being of "northern" or "southern" affiliation. Proportions of northern and southern mtDNA differed significantly between Egypt, Nubia, and the southern Sudan. At slowly evolving sites within HV1, northern-mtDNA diversity was highest in Egypt and lowest in the southern Sudan, and southern-mtDNA diversity was highest in the southern Sudan and lowest in Egypt, indicating that migrations had occurred bidirectionally along the Nile River Valley. Egypt and Nubia have low and similar amounts of divergence for both mtDNA types, which is consistent with historical evidence for long-term interactions between Egypt and Nubia. Spatial autocorrelation analysis demonstrates a smooth gradient of decreasing genetic similarity of mtDNA types as geographic distance between sampling localities increases, strongly suggesting gene flow along the Nile, with no evident barriers. We conclude that these migrations probably occurred within the past few hundred to few thousand years and that the migration from north to south was either earlier or lesser in the extent of gene flow than the migration from south to north.

Received July 17, 1998; accepted for publication February 4, 1999; electronically published March 12, 1999.

Address for correspondence and reprints: Dr. Mark Stoneking, Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, D-04103 Leipzig, Germany. E-mail: stoneking@eva.mpg.de

0002-9297/99/6404-0029$02.00

Introduction

A characteristic of human populations is that they move. The earliest documented evidence of such population movements, or migrations, involves the initial expansion of our hominid ancestors out of Africa, perhaps as long as 1.8 million years ago (Swisher et al. 1994). The inference of a second, major expansion out of Africa, which involved anatomically modern humans, ~100,000 years ago receives strong support from genetic, fossil, and archaeological evidence (Lahr 1994; Lieberman 1995; Stoneking 1996; Tishkoff et al. 1996; Foley and Lahr 1997; Krings et al. 1997; Cavalli-Sforza 1998). More recently, attention has been focused on subsequent migrations, and, again, both the genetic evidence and the morphological evidence point to multiple dispersals of modern humans during the past 100,000 years (Harpending et al. 1993; Lahr and Foley 1994; Stoneking et al. 1997).

To understand the human species, then, requires knowledge of what dispersals have occurred, and, ultimately, what has caused them. However, reconstruction of dispersal events on the basis of either genetic or fossil evidence is fraught with difficulty. Seemingly simple questions—such as whence, when, and by what route the populations migrated—may become much more complex when all of the various possible sources, routes, and migration times are considered. Multiple migrations to, from, or through a particular region may also complicate matters. Limitations on sample availability, which are unavoidable in either genetic or fossil studies, can severely compromise the ability to address all of the pertinent possibilities.

To mitigate these potential complexities, we have undertaken a study of migrations along the Nile River Valley, by analyzing mtDNA variation in human populations along the Nile. The Nile River Valley offers several attractive features for this study, the primary one

being the geographic and climatic features of the Nile River. The Nile River is formed by two main tributaries, the White Nile and the Blue Nile, originating in Lake Victoria and the Ethiopian highlands, respectively, that unite in the area of Khartoum, in the southern Sudan. From there, the river flows in a generally northerly direction for >2,000 km, transversing the Sahara. Along the way, it supports human life in a strip of arable land that is generally ≤10 km wide. The Nile River Valley thus can be regarded as an extremely long and narrow oasis in which humans live essentially distributed in a one-dimensional pattern and where most of the population interactions can be expected to have taken place along the river. This situation has prevailed since climatic changes at the end of the last glaciation, some 10,000–15,000 years ago.

Another advantage of studying the Nile River Valley populations is that the long and comprehensive written record of Egyptian history, going back >5,000 years, affords the rare opportunity to compare migrations inferred from genetic evidence with documented events. Moreover, the existence of large quantities of ancient skeletal and mummified remains from the Nile River Valley offers the potential, via ancient-DNA analysis, for investigation of temporal variation in migration patterns.

Although the potential of the Nile River Valley to serve as a corridor for human migration seems obvious, some archeological evidence suggests, instead, that there was a significant and long-standing frontier zone between the northern and southern states in Lower Nubia (Alexander 1988). Moreover, the human populations distributed along the length of the Nile River Valley do exhibit cultural and linguistic differences (Grimes 1996). Egypt, at the northern end of the Nile, is one of the oldest known centralized states in the world; both ancient Egyptian and Arabic (spoken today in Egypt) are Afro-Asiatic languages. Nubia comprises the southern part of Egypt and the northern part of the Sudan and, historically, has consisted of kingdoms that were only occasionally united politically; Nubian languages constitute one branch of the Nilo-Saharan language family. The populations along the southernmost part of the Nile River in the southern Sudan are pastoral nomads with a tribal organization, and they speak languages within the Nilotic branch of the Nilo-Saharan family. Thus, this evidence suggests that there was a barrier between the northern and southern portions of the Nile River Valley and that the latter was a cul-de-sac, rather than a corridor, for human migration.

We report here our analysis of mtDNA variation in contemporary Nile River Valley populations in Egypt, Nubia, and the southern Sudan. By virtue of its strictly maternal and haploid mode of inheritance, mtDNA has already proved to be useful in the tracing of the migrations leading to the colonization of Polynesia (Melton et al. 1995; Redd et al. 1995; Sykes et al. 1995) and the New World (Wallace and Torroni 1992; Merriwether et al. 1996). In addition, comparison with the large amount of existing data on human mtDNA variation enables us to confidently assign mtDNA types in Nile River Valley populations to northern (Eurasian) or southern (sub-Saharan African) affiliation and to use this information to infer migrations. Given the cultural and linguistic differences that exist between Egyptian, Nubian, and southern Sudanese populations, the question that we address here is whether the Nile River Valley has been a corridor or a cul-de-sac for human migrations.

Subjects and Methods

Whole blood, serum, or head hairs were obtained from 224 individuals, including 68 from Egypt, 80 from Nubia, and 76 from the southern Sudan (fig. 1). Detailed information on the geographic origin and mtDNA sequence of each sample has been deposited in the HvrBase database (Burckhardt et al. 1999). Blood samples were kept in an equal volume of 100 mM Tris, 100 mM EDTA (pH 8.0), 2% SDS buffer at ambient temperature while in the field and at 4°C after arrival at the laboratory. Serum samples that were provided for this study had been kept at −20°C after separation from cells. Hairs were plucked by the donors and were kept in 70% ethanol at ambient temperature in the field and at −20°C after arrival at the laboratory.

DNA from blood and sera was extracted by means of a standard phenol/chloroform protocol (Sambrook et al. 1989) followed by centrifugal dialysis with Centricon 30 microconcentrators (Amicon). DNA from hairs was extracted as described elsewhere (Higuchi et al. 1988; Vigilant et al. 1989), except for the omission of the 1-butanol extraction step. A negative control was included in each set of extractions.

The hypervariable region 1 (HV1) of the mtDNA control region was amplified, and the sequence at positions 16024–16383 (Anderson et al. 1981) was determined, by means of primers and methods described elsewhere (Watson et al. 1996). Sequences were aligned by means of the ESEE program (Cabot and Beckenbach 1989).

A 219-bp fragment containing a polymorphic *Hpa*I restriction site at position 3592 (Denaro et al. 1981) was amplified by means of primers L03526 (5′-CAT CAC CCT CTA CAT CAC CG-3′) and H03706 (5′-ATT GTT TGG GCT ACT GCT CG-3′). A 782-bp control fragment, containing a nonpolymorphic *Hpa*I restriction site at position 5691, was amplified by means of primers L05269 (5′-TTG CCC AAA TGG GCC ATT AT-3′) and H06012 (5′-TGG CCC AGG TCG GCT CGA AT-3′). *Hpa*I-restriction analysis was done by addition of 8.0 μl of 10 × *Hpa*I buffer, 1 unit of enzyme (Gibco BRL), 45

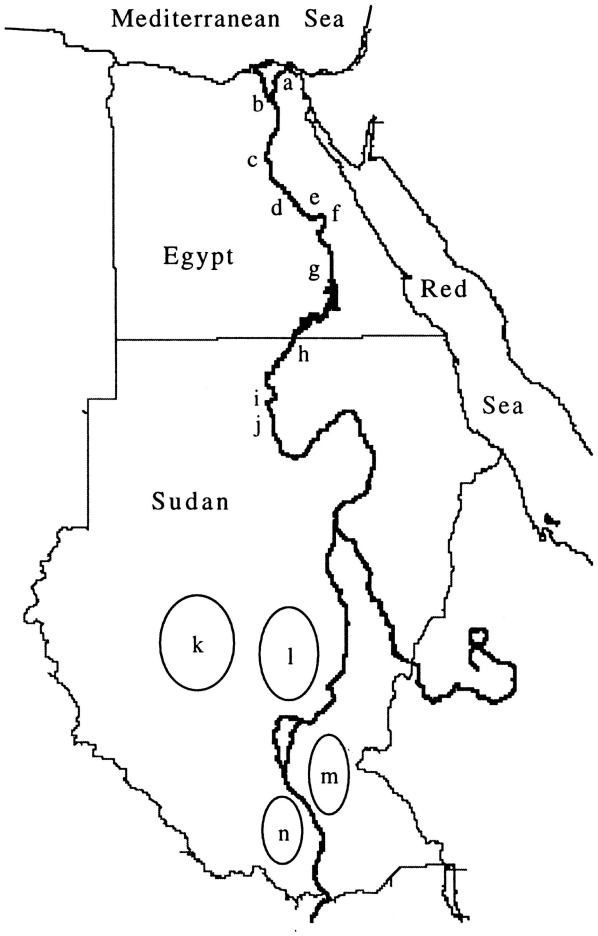

Figure 1 Map of Egypt, Nubia, and the Sudan, showing sampling localities from the Nile River Valley. The current political borders of Egypt and the Sudan are indicated by the thinner lines, and the Nile River Valley is indicated by the thicker lines. Nubia is generally considered to extend from Assuan in Egypt to Ad-Debba in the Sudan; individuals from localities g–j self-identified as Nubians, whereas individuals from localities a–f self-identified as Egyptians. The approximate territories of the Sudanese tribal groups are indicated by ellipses. Sampling localities (with sample sizes [*n*]) are designated as follows: a = Mansoura ($n = 40$); b = Monofia ($n = 1$) and Chephen ($n = 1$); c = Minia ($n = 2$); d = Assiut ($n = 19$); e = Sohag ($n = 2$); f = Kena ($n = 1$); g = Assuan ($n = 11$); h = Wadi Halfa ($n = 1$); i = Kerma ($n = 40$); j = Dongola ($n = 14$); k = Nuba ($n = 11$); l = Shilluk ($n = 8$); m = Nuer ($n = 14$); n = Dinka ($n = 43$). Two Egyptians and 14 Nubians were from unknown localities.

μl of the 219-bp PCR product containing the polymorphic *Hpa*I site, and 26 μl of the 782-bp control PCR product. Ten microliters of the digestion reaction were electrophoresed through a 2% agarose gel containing ethidium bromide and were visualized by means of UV light. Restriction digestion of the 219-bp PCR product resulted in one fragment, of 219 bp, when the *Hpa*I site was absent and in two fragments, of 131 bp and 88 bp, when the *Hpa*I site was present, whereas the 782-bp control PCR product yielded fragments of 444 bp and 338 bp.

mtDNA-type diversity, *h*, for the HV1 sequences (excluding the *Hpa*I site) was calculated according to equation 8.5 of Nei (1987), and the net mtDNA divergence between populations, d_A, was calculated according to equation 10.21 of Nei (1987) (these equations can also be found in the footnotes to tables 1 and 4, respectively). The significance of differences among populations, with respect to mtDNA-sequence distributions, was assessed by the analysis of molecular variance (AMOVA) procedure (Excoffier et al. 1992), by means of the computer program ARLEQUIN (Schneider et al. 1997). For some analyses, nucleotide sites were partitioned according to the number of mutations observed at each site, as determined elsewhere by a phylogenetic analysis of worldwide human mtDNA sequences (Hasegawa et al. 1993).

Additional insights into the pattern of mtDNA differentiation along the Nile River Valley were obtained via spatial autocorrelation analysis (Bertorelle and Barbujani 1995). We calculated the autocorrelation index for DNA analysis (AIDA), *II* (Bertorelle and Barbujani, 1995), which ranges from -1 to $+1$, where, for a given geographic-distance class, a positive value for *II* indicates genetic similarity for samples separated by that geographic distance whereas a negative value indicates dissimilarity. The statistical significance of *II* can be estimated by comparison with a null distribution generated by randomization of the assignment of haplotypes to geographic coordinates while the observed sample size at each geographic coordinate is retained. If there is no spatial structure to the distribution of haplotypes, then *II* values are not expected to deviate significantly from 0, regardless of geographic-distance class, whereas gradients (or clines) are detected by correlograms (plots of *II* values vs. geographic-distance classes) that exhibit a decline, from significant positive values at small geographic-distance classes to significant negative values at large distance classes; more-complicated patterns that reflect local barriers to migration or drift effects can also be observed (Bertorelle and Barbujani 1995).

For the Nile River Valley data, samples of ambiguous geographic origin were excluded, leaving 189 samples for the AIDA analysis. After invariable sites were deleted, each mtDNA sequence plus the *Hpa*I site was transformed into a string of binary digits, as described by Excoffier et al. (1992). The geographic coordinates for each sample were determined from a map (scale 1: 4,000,000); cities were considered a single point on the map, as were the areas of origin of the southern Sudanese samples (Dinka, Nuba, Nuer, and Shilluk). The AIDA program (Bertorelle and Barbujani 1995) was used to calculate *II* values, and the statistical significance of *II* was assessed by randomized assignment of haplotypes to geographic coordinates 250 times, as described above.

Several analyses with various parameters were performed: (1) four to nine exactly equal–size geographic-distance classes, (2) four to nine distance classes with approximately equal numbers of pairwise comparisons per class, and (3) various combinations of arbitrarily chosen distance classes. The distance class at 0 km, which measures genetic relatedness within populations, was considered an independent class in all analyses.

Results

Identification of Northern mtDNA Types versus Southern mtDNA Types

To identify the probable geographic origin of mtDNA types from the Nile River Valley, the published literature on mtDNA-sequence and restriction-site polymorphisms was scanned for polymorphisms that differ significantly, in frequency, between Eurasian (northern) and sub-Saharan African (southern) populations. Three sites (fig. 2) that exhibit significant frequency differences were selected for analysis in Nile River Valley mtDNA types; two of these sites (16223 and 16311) are in HV1 and hence were determined on the basis of sequence analysis of HV1, whereas the third is a polymorphic *Hpa*I site, at position 3592, that was typed by *Hpa*I digestion of a PCR product containing this site.

Northern mtDNA types are associated with a C at 16223, a T at 16311, and the absence of the *Hpa*I site, whereas southern mtDNA types are associated with a T at 16223, a C at 16311, and the presence of the *Hpa*I site (fig. 2). Therefore, Nile River Valley mtDNA types with the northern character at two or three of the sites were classified as originating from the north, whereas those with the southern character at two or three of the sites were classified as originating from the south. Utilizing three sites in this manner should minimize incorrect classification of mtDNA types; however, because the two sites in HV1 are subject to repeated mutations (Hasegawa et al. 1993), we were concerned that some incorrect classification might nevertheless occur. We therefore compared the entire HV1 sequences of the Nile River Valley mtDNA types versus a worldwide database of 4,079 HV1 sequences (Handt et al. 1998) and, on the basis of whether it was either identical to or differed by one substitution from database mtDNA types with exclusively northern or southern affiliations, independently classified the origin of each unique Nile River Valley mtDNA sequence. Approximately one-third of the Nile River Valley mtDNA types could be unambiguously classified on the basis of this database comparison; the results were nearly completely concordant with the classification based on the three sites, with the single discrepancy involving an Egyptian mtDNA that, on the basis of the three sites, was classified as northern but, on the basis of the database comparison, was classified as southern because it was identical to sequences found in two Songhai from Mali and two Kikuyu from Kenya (Watson et al. 1996). Because alteration of the classifi-

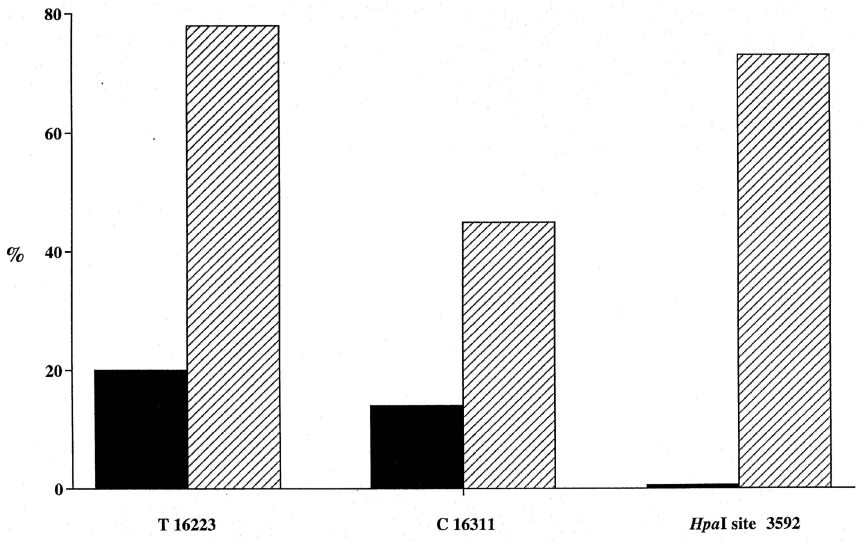

Figure 2 Frequencies of three mtDNA polymorphisms in Eurasian (*blackened bars*) and African (*hatched bars*) mtDNA types. The frequencies of T at 16223 and C at 16311 are based on 1,873 Eurasian and 624 African sequences (Handt et al. 1998). The frequency of the *Hpa*I site at 3592 is based on 1,355 Eurasian and 754 African mtDNA types (Denaro et al. 1981; Johnson et al. 1983; Bonné-Tamir et al. 1986; Brega et al. 1986; Cann et al. 1987; Harihara et al. 1988; Scozzari et al. 1988, 1994; Vilkki et al. 1988; Semino et al. 1989, 1991; Tikochinski et al. 1991; Ballinger et al. 1992; Soodyall and Jenkins 1992; Ritte et al. 1993; Chen et al. 1995; Graven et al. 1995).

Table 1

Genetic Diversity in Nile River Valley mtDNA Types

	Sample Size	No. of mtDNA Types	h[a]	MPD	Proportion of mtDNA Types (%)	
					Northern	Southern
Egypt	68	58	.99 ± .01	7.0	74.6	25.4
Nubia	80	53	.98 ± .01	8.4	45.1	54.9
Southern Sudan	76	65	.99 ± .01	8.6	19.7	80.3

[a] $h = (1 - \Sigma p_i^2)n/(n - 1)$, where n is the sample size and p_i is the frequency of the ith mtDNA type.

cation of this one sequence does not significantly change any of the results that follow, this Egyptian mtDNA was still classified as northern, in accordance with the results from use of the three sites.

A phylogenetic analysis of the Nile River Valley HV1 sequences, in comparison with various Eurasian and sub-Saharan African sequences, is also generally concordant with the north-south classification based on the three sites (data not shown). In addition, it has been proposed that the *Hpa*I site at 3592 has a single origin in sub-Saharan Africa (Chen et al. 1995), which means that, according to our scheme, all mtDNA types with this site should be classified as southern. This is indeed the case; all mtDNA types with the *Hpa*I site gain also had at least one of the other two southern markers. Overall, these additional analyses support the use of the three sites for classification of the geographic affiliations of Nile River Valley mtDNA types.

Analysis of Nile River Valley mtDNA Types

The sequence of a 360-bp segment of HV1 and the polymorphic *Hpa*I site at position 3592 were determined for 224 individuals from the Nile River Valley. Several groups of varying sample size are included within the Egyptian, Nubian, and southern Sudanese samples (fig. 1); to determine whether the pooling of groups within these larger regions is justified, the AMOVA procedure was used to assess the significance of population substructure within each region, for those localities with reasonable sample sizes. There were no significant differences, in the distribution of HV1 sequences, either between the two groups within Egypt (Mansoura and Assiut) or between four groups within the southern Sudan (Dinka, Nuer, Nuba, and Shilluk), a finding that justified the pooling of these groups within each region, but the three Nubian groups (Kerma, Assuan, and Dongola) did exhibit statistically significant heterogeneity (analysis not shown). However, further analysis showed that the heterogeneity within Nubia involved only the comparison of Assuan with Dongola and that it may reflect the small sample sizes of these two groups (12 and 14 individuals, respectively). Furthermore, the Nubian localities did not differ significantly in the frequency of northern versus southern mtDNA types ($\chi^2 = 4.38$, df 2, $P > .05$). Therefore, to keep the Nubian sample size comparable to that for Egypt and that for the southern Sudan, the Nubian groups were pooled as well, resulting in final sample sizes of 68 Egyptians, 80 Nubians, and 76 southern Sudanese.

High levels of mtDNA variation were found in all three samples (table 1). The amount of mtDNA diversity in the Egyptian, Nubian, and southern Sudanese samples was nearly indistinguishable, whereas the mean pairwise differences (MPD) were slightly lower in Egypt than in Nubia and the southern Sudan. However, the distribution of northern and southern mtDNA types differed significantly among the three samples ($\chi^2 = 53.69$, df 2, $P < .001$), with the frequency of northern types decreasing and the frequency of southern types increasing, from Egypt to Nubia to the southern Sudan (table 1). The AMOVA procedure also indicates that all three samples differ significantly from one another with respect to the distribution of HV1 sequences (data not shown). Thus, although the overall levels of mtDNA variation are similar, there are, nonetheless, significant differences in the composition of the mtDNA gene pools of the Egyptian, Nubian, and southern Sudanese samples.

Examination of diversity separately for northern and southern mtDNA types did not reveal any clear patterns (table 2). The diversity and MPD for northern mtDNA types were highest in Egypt and the southern Sudan and were slightly lower in Nubia. For southern mtDNA types, the highest diversity and MPD were observed in the southern Sudan, and the lowest were observed in Egypt. Furthermore, the AMOVA procedure indicates that the northern HV1 sequences differ significantly among Egypt, Nubia, and the southern Sudan (table 3), whereas the southern HV1 sequences do not differ significantly among Egypt, Nubia, and the southern Sudan, with the exception of the Nubia–southern Sudan comparison (table 3).

One factor that might obscure trends in mtDNA variation is the very high rate of mtDNA mutation at certain nucleotide positions (Hasegawa et al. 1993; Wakeley

Table 2

Genetic Diversity, in Nile River Valley, of Northern and Southern mtDNA Types

mtDNA Type	Sample Size	No. of mtDNA Types	*h*	MPD
Northern:				
Egypt	51	45	.99 ± .01	6.2
Nubia	34	19	.93 ± .02	4.4
Southern Sudan	11	9	.96 ± .05	4.6
Southern:				
Egypt	17	13	.95 ± .04	6.5
Nubia	46	34	.96 ± .02	9.0
Southern Sudan	65	56	.99 ± .01	9.0

NOTE.—Data are as defined in table 1.

1993); that is, a diversity reduction associated with a migration event might not be evident because mutations at rapidly evolving sites have led to a recovery of diversity (Sajantila et al. 1996). We therefore examined mtDNA diversity separately at slowly evolving and rapidly evolving subsets of nucleotide positions, using Hasegawa et al.'s (1993) classification. Twenty of the 110 polymorphic sites in the Nile River Valley mtDNA types were not observed to be polymorphic in Hasegawa's compilation and hence cannot be classified by his scheme. We chose to exclude these sites from this analysis, but including them and classifying them as slowly evolving sites does not alter the trends discussed in the following paragraph (data not shown).

Clear trends in diversity were shown when only slowly evolving sites were considered (fig. 3). The diversity in the southern Sudanese, with respect to northern mtDNA types, is dramatically reduced when only slowly evolving sites (i.e., sites that, according to Hasegawa et al. [1993], have mutated only once) are considered; the 9 different types among 11 northern mtDNA types in the southern Sudan are reduced to just 1 type when rapidly evolving sites are excluded. Even though there are only 11 southern Sudanese with northern mtDNA types, this reduction to 1 type for slowly evolving sites is still significantly less ($P = .03$) than would be expected if 11 individuals were chosen at random from the 57 Egyptians with northern mtDNA types. The diversity in northern mtDNA types in Egypt and Nubia are both reduced to a similar extent, but not as greatly as is observed for northern mtDNA types from the southern Sudan, when only slowly evolving sites are considered.

For southern mtDNA types, the Egyptian and Nubian mtDNA diversity at slowly evolving sites is reduced considerably, whereas the reduction in mtDNA diversity in the southern Sudan is not as great. Again, Egypt and Nubia exhibit comparable reductions in mtDNA diversity. For both northern and southern mtDNA types in all three samples, the mtDNA diversity based only on rapidly evolving sites (i.e., sites that, according to Hasegawa et al. [1993] have mutated three or more times) is nearly identical to that based on all sites.

These analyses show that the diversity of northern mtDNA types is highest in Egypt and lowest in the southern Sudan and that the diversity of southern mtDNA types is highest in the southern Sudan and lowest in Egypt, suggesting that gene flow has occurred in both directions along the Nile River Valley. To further investigate the Nile River Valley populations' spatial structuring with respect to mtDNA variation, we performed a spatial autocorrelation analysis (Bertorelle and Barbujani 1995). Various ways of partitioning the data with respect to geographic distance (equal geographic distances per class, equal numbers of pairwise comparisons per distance class, or arbitrary distance divisions) all yielded quite similar results, with significantly positive *II* values observed for small distance classes and significantly negative values observed for large distance classes; a representative result is shown in figure 4. The smooth decrease, from small to large geographic distance classes observed in the *II* values is indicative of a gradient or cline in the distribution of mtDNA types along the Nile River Valley.

The spatial autocorrelation analysis thus indicates that at least one migration has occurred along the Nile River Valley, although this analysis does not indicate the direction(s) of the migration(s). However, the overall trends in mtDNA diversity—in particular, the finding of (1) both northern and southern mtDNA types in all three Nile River Valley populations and (2) both the highest diversity of northern mtDNA types at the northern end of the Nile River Valley and the highest diversity of southern mtDNA types at the southern end of the Nile River Valley—suggest that migrations have occurred both from north to south and from south to north along the Nile River Valley. To obtain some idea of the relative antiquity of these migrations, we computed the average d_A for northern and southern mtDNA types separately (table 4). Two estimates of mtDNA divergence were used to convert the d_A estimates into time: a slow, phylogenetic rate of 33%/million years (myr) (Ward et al. 1991)

Table 3

AMOVA Results (*P* Values) for Nile River Valley Population Comparisons Involving Northern and Southern mtDNA Types

	Egypt	Nubia	Southern Sudan
Egypt	...	.016	.002
Nubia	.348	...	.001
Southern Sudan	.132	.012	...

NOTE.—Probability that a randomly generated Φst value will exceed the observed value (based on 1,000 replications), for comparisons involving northern mtDNA types (above the diagonal) and southern mtDNA types (below the diagonal).

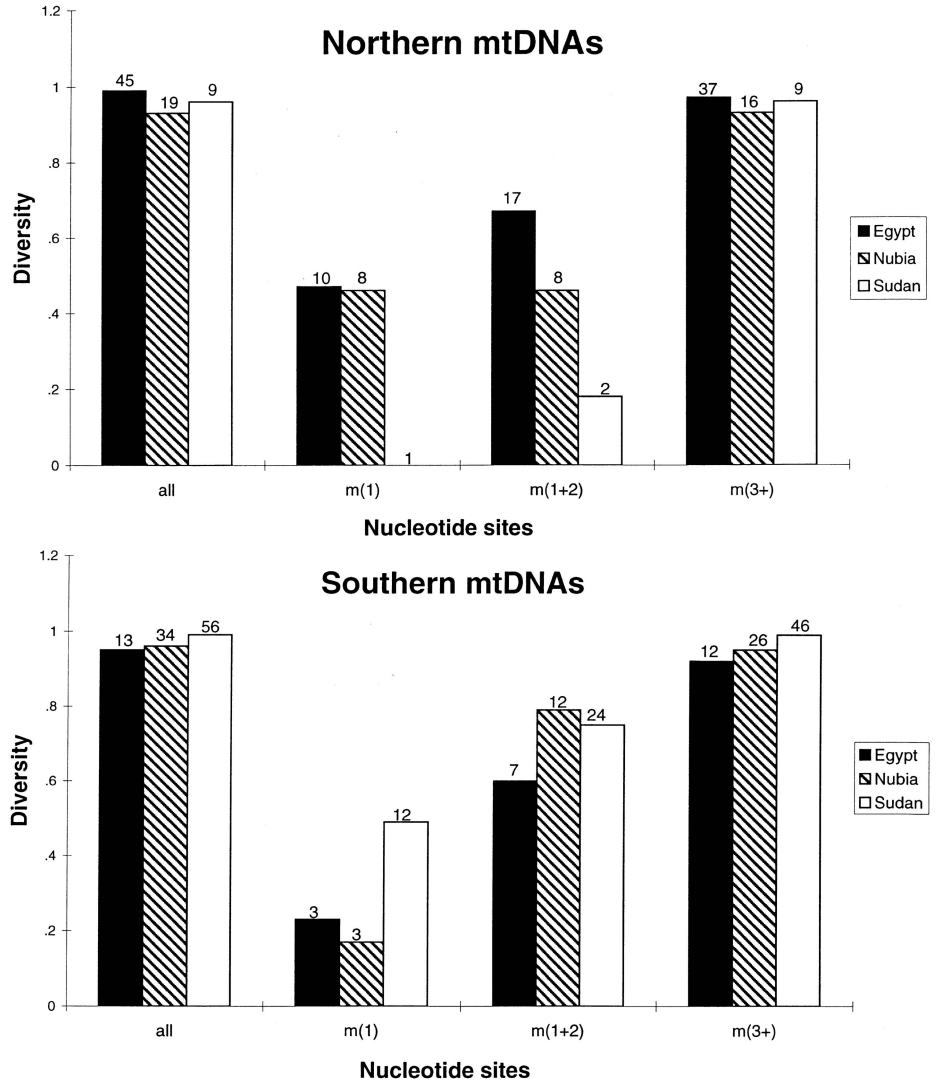

Figure 3 Diversity, in Egyptians, Nubians, and southern Sudanese, of northern mtDNA types (*top*) and southern mtDNA types (*bottom*), for various classes of nucleotide sites. The number at the top of each bar denotes the number of different mtDNA types. The term "all" refers to all nucleotide sites, whereas "m(1)," "m(1+2)," and "m(3+)" refer to sites that mutated once, once or twice, and three or more times, respectively, according to the phylogenetic analysis by Hasegawa et al. (1993); the m(1+2) class thus includes the m(1) class. Note that, for northern mtDNA types, there was just one type observed in the southern Sudan, when only the m(1) sites are considered, and hence the diversity is 0.

and a fast rate, based on family studies, of 234%/myr (Jazin et al. 1998). These results (table 4) indicate that the Nile River Valley populations separated relatively recently (within the past few hundred to few thousand years) and that the population-separation times based on northern mtDNA types are larger than the population-separation times based on southern mtDNA types. As discussed in more detail below (see Discussion), however, these times should be interpreted cautiously, because of the assumptions involved in this analysis.

Discussion

The goal of this study was to ascertain whether the Nile River Valley has been a corridor for human population movements or cultural and linguistic differences between populations distributed along the Nile River Valley have been a barrier to migration. Our results, based on mtDNA analysis of 224 individuals, indicate that migrations have indeed occurred, both from north to south and from south to north. This is supported by

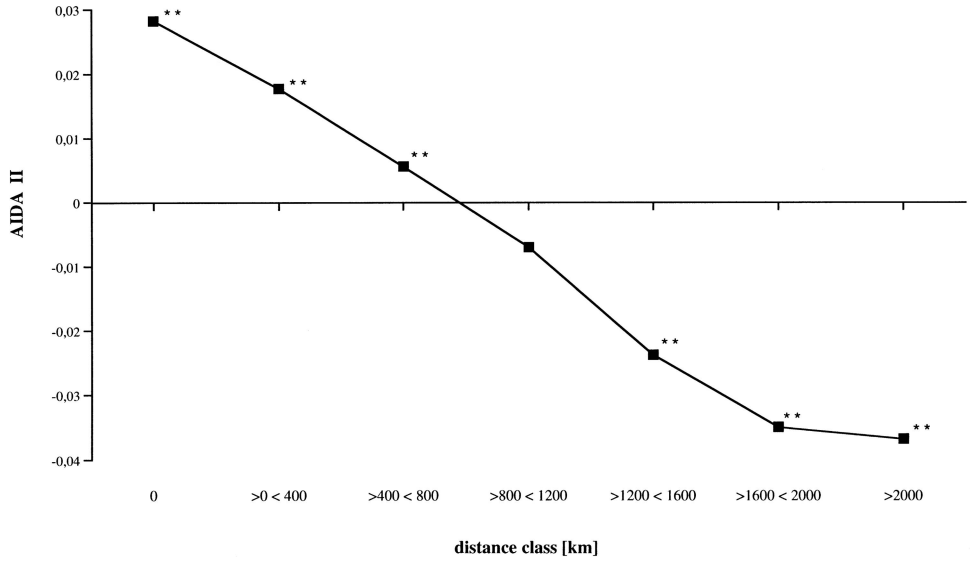

Figure 4 Correlogram of the AIDA *II* for 189 mtDNA haplotypes (HV1 sequences plus the *Hpa*I restriction site) from the Nile River Valley. The value of *II* is the *Y*-axis, whereas the *X*-axis is the width of arbitrarily chosen distance classes. The double asterisks (**) denote that the value of *II* for all distance classes except 800–1,200 km differs significantly ($P < .01$) from 0.

both the finding of northern mtDNA types at the southern end of the Nile River Valley and the finding of southern mtDNA types at the northern end of the Nile River Valley.

Additional evidence for these migrations comes from the observation that the diversity associated with northern mtDNA types decreases from north to south whereas the diversity associated with southern mtDNA types decreases from south to north. However, these diversity trends were not evident until rapidly evolving nucleotide sites were removed from the analysis. Apparently, mutations at these rapidly evolving sites have regenerated mtDNA diversity that was lost during the migrations. This is analogous to the situation in Finland, where genetic epidemiology seems to indicate that a bottleneck occurred during the founding of the Finnish population (Norio et al. 1973). A study of mtDNA and Y-chromosome diversity in Finns (Sajantila et al. 1996) found a reduction in Y-chromosome diversity relative to that in other European populations but found no comparable reduction in overall mtDNA diversity. However, mtDNA diversity at slowly evolving sites is reduced in the Finns, in agreement with the Y-chromosome results, thus supporting the contention that the Finns have experienced a bottleneck. Both the present study and that by Sajantila et al. (1996) illustrate the importance of considering the evolutionary dynamics of the mtDNA control region (Hasegawa et al. 1993; Wakeley et al. 1993) when one is attempting to draw population-history inferences from patterns of mtDNA variation.

The smooth, decreasing gradient in the AIDA *II* values that is observed in the spatial autocorrelation analysis also supports the conclusion of past migrations along the Nile River Valley. Moreover, the lack of any significant fluctuations in the correlograms suggests that there have not been any significant local barriers to migration or significant local drift effects. However, although the mtDNA results indicate that migrations have occurred in both directions along the Nile River Valley, these migrations have not been extensive enough to genetically homogenize the mtDNA gene pools of Nile River Valley populations. Significant differences exist among Egyptian, Nubian, and southern Sudanese mtDNA types, with Nubians appearing to be more similar to Egyptians than to the southern Sudanese and with Egyptians and the southern Sudanese exhibiting the greatest differences. This is perhaps to be expected, given that, during the second millenium B.C., Pharaonic Egypt colonized part or most of Nubia and that Nubian kings conquered Egypt during the 8th and 7th centuries B.C. (Shaw and Nicholson 1995). Thus, there is historical evidence documenting much direct interaction between Egypt and Nubia.

This greater similarity between Egypt and Nubia than between either of them and the southern Sudan is also apparent in the estimated divergence times among these three populations (table 4); for both southern and northern mtDNA types, the divergence time is approximately the same for Egypt and Nubia, and both exhibit approximately equivalent divergence times with respect to

Table 4

d_A and Separation-Time Estimates among Nile River Valley Populations

	Northern mtDNA Types			Southern mtDNA Types		
		Time, for Rate of (years)			Time, for Rate of (years)	
Comparison	$d_A \times 10^{-5}$	33%/myr	234%/myr	$d_A \times 10^{-5}$	33%/myr	234%/myr
Egypt-Nubia	37	1,120	160	36	1,105	155
Nubia–southern Sudan	171	5,200	735	52	1,580	225
Egypt–southern Sudan	200	6,070	855	51	1,540	215

Note.—$d_A = d_{xy} - (d_x + d_y)/2$, where d_{xy} is the observed mean pairwise sequence divergence between populations x and y and where d_x and d_y are the observed mean pairwise sequence divergence within populations x and y, respectively.

the southern Sudan. However, these estimated divergence times should not be taken literally. First, there are large uncertainties in these estimates, owing (in part) to uncertainty with regard to the rate of mtDNA evolution. We have chosen to use two mtDNA-evolution rates that should bracket the "true" value: a slow rate of 33%/myr (Ward et al. 1991), based on phylogenetic comparisons of human and chimpanzee mtDNA types; and a fast rate of 234%/myr (Jazin et al. 1998), based on family studies. Incidentally, the fast rate produces excessively recent estimates of population separation times, when one considers the geographic distances involved. This problem of too-recent times of divergence has been noted by others trying to apply the fast rates derived from family studies to phylogenetic analyses of human mtDNA variation (Jazin et al. 1998). Second, these divergence times reflect actual population-separation times only in the absence of subsequent migrations. In other words, the relatively recent migrations inferred from table 4 could, in fact, reflect much older migrations that were followed by limited amounts of genetic exchange among these populations. This limited genetic exchange would have to be of sufficient magnitude to increase the similarity of the populations, but not so great as to eliminate the statistically detectable genetic differences that still remain between these populations. Third, founder events during the migrations could artificially inflate the estimated divergence time. For example, the northern mtDNA types in the southern Sudan could reflect a small group of Egyptians that migrated very recently to the southern Sudan if, by chance, the migrant mtDNA types differed sufficiently from the overall composition of the Egyptian mtDNA to result in an apparent divergence time of 855–6,070 years. Fourth, the greater diversity observed for northern mtDNA types in Egypt and for southern mtDNA types in the Sudan could indicate gene flow from Eurasia into Egypt only, and/or gene flow from sub-Saharan Africa into the Sudan only, thus altering diversity levels in these populations. Simulation analyses might be a fruitful means by which to explore the combinations of initial population separation times, founder events, and subsequent amounts of gene flow that can be accommodated by the data.

Nonetheless, we can infer that the migration of northern mtDNA types to the south is older than the migration of southern mtDNA types to the north (or that there has been less gene flow from north to south than from south to north along the Nile River Valley) and that Egypt and Nubia have had more genetic contact than either has had with the southern Sudan. Moreover, we can tentatively infer that these migrations occurred recently enough to fall within the period of the documented historical record of human populations in the Nile River Valley. Thus, it is tempting to try to relate these migrations to specific historical events (Shaw and Nicholson 1995). For example, the migration from north to south may coincide with the Pharaonic colonization of Nubia, which occurred initially during the Middle Kingdom (12th Dynasty, 1991–1785 B.C.) and more permanently during the New Kingdom, from the reign of Thotmosis III (1490–1437 B.C.). The migration from south to north may coincide with the 25th Dynasty (730–655 B.C.), when kings from Napata in Nubia conquered Egypt. Of course, additional migrations documented during the Ptolemeic, Roman, and Arabic times are also likely to have contributed to the current distribution of mtDNA types along the Nile River Valley.

The migrations inferred to have occurred along the Nile River Valley are based on an analysis of a single genetic locus that is informative for maternal ancestry only; similar analyses of autosomal and Y-chromosomal loci, which are in progress, are needed to confirm and extend the conclusions of the present study. Indeed, the Nile River Valley may provide an excellent test of the hypothesis that female migration has been greater than male migration (Seielstad et al. 1998). Another potential way of testing the hypothesized correlations between migrations inferred from patterns of mtDNA variation and

the historical record would be via ancient-DNA analysis; that is, the temporal variation of migration patterns along the Nile River Valley could be directly assessed, if authentic ancient DNA could be retrieved from a suitable number of specimens. Indeed, a recent study did report success in obtaining DNA from the remains of 15 of 29 ancient Nubians that are ~2,000 years old (Fox 1997). These individuals were screened for the *Hpa*I site at position 3592; the frequency of this site in the ancient Nubians (26.7%) is not significantly different from that in contemporary Nubians in the present study (32.5%). Thus, it may prove feasible to use ancient-DNA analysis to investigate the temporal stability of migration patterns along the Nile River Valley. However, in a study of 132 mummies and skeletons of late pre-Dynastic and early Pharaonic age, only 2 samples yielded reproducible sequences when cloned amplification products derived from independent extracts were compared (M. Krings, unpublished data). Thus, the prospects for a comprehensive survey of ancient-DNA variation along the Nile River Valley do not look promising (Handt et al. 1996), and further technical advances in ancient DNA analysis will doubtless be required.

Acknowledgments

We thank O. Behrend for samples, G. Barbujani for the AIDA program, S. Meyer and L. Vigilant for discussions and help, and the Deutsche Forschungsgemeinschaft (support to S.P.) and the Boehringer Ingelheim Fonds (support to M.K.) for financial support.

Electronic-Database Information

The URL for data in this article is as follows:

HvrBase database, http://www.eva.mpg.de/hvrbase

References

Alexander J (1988) The Saharan divide in the Nile Valley: the evidence from Qasr Ibrim. Afr Archeol Rev 6:73–90

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457–465

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen K-H, et al (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130:139–152

Bertorelle G, Barbujani G (1995) Analysis of DNA diversity by spatial autocorrelation. Genetics 140:811–819

Bonné-Tamir B, Johnson MJ, Natali A, Wallace DC, Cavalli-Sforza LL (1986) Human mitochondrial DNA types in two Israeli populations—a comparative study at the DNA level. Am J Hum Genet 38:341–351

Brega A, Scozzari R, Maccioni L, Iodice C, Wallace DC, Bianco I, Cao A, et al (1986) Mitochondrial DNA polymorphisms in Italy. I. Population data from Sardinia and Rome. Ann Hum Genet 50:327–338

Burckhardt F, von Haeseler A, Meyer S (1999) HvrBase: compilation of mtDNA control region sequences from primates. Nucleic Acids Res 27:138–142

Cabot EL, Beckenbach AT (1989) Simultaneous editing and of multiple nucleic acid and protein sequences with ESEE. Comput Appl Biosci 5:233–234

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36

Cavalli-Sforza LL (1998) The DNA revolution in population genetics. Trends Genet 14:60–65

Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC (1995) Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J Hum Genet 57:133–149

Denaro M, Blanc H, Johnson MJ, Chen KH, Wilmsen E, Cavalli-Sforza LL, Wallace DC (1981) Ethnic variation in HpaI endonuclease cleavage patterns of human mitochondrial DNA. Proc Natl Acad Sci USA 78:5768–5772

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Foley R, Lahr MM (1997) Mode 3 technologies and the evolution of modern humans. Camb Archaeol J 7:3–36

Fox CL (1997) mtDNA analysis in ancient Nubians supports the existence of gene flow between sub-Sahara and North Africa in the Nile valley. Ann Hum Biol 24:217–227

Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, Excoffier L (1995) 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

Grimes BF (ed) (1996) Ethnologue: languages of the world. Summer Institute of Linguistics, Dallas

Handt O, Krings M, Ward RH, Pääbo S (1996) The retrieval of ancient human DNA sequences. Am J Hum Genet 59: 368–376

Handt O, Meyer S, von Haeseler A (1998) Compilation of human mtDNA control region sequences. Nucleic Acids Res 26:126–129

Harihara S, Saitou N, Hirai M, Gojobori T, Park KS, Misawa S, Ellepola SB, et al (1988) Mitochondrial DNA polymorphism among five Asian populations. Am J Hum Genet 43: 134–143

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. Curr Anthropol 34:483–496

Hasegawa M, Di Rienzo A, Kocher TD, Wilson AC (1993) Toward a more accurate time scale for the human mitochondrial DNA tree. J Mol Evol 37:347–354

Higuchi R, von Beroldingen CH, Sensabaugh GF, Erlich HA (1988) DNA typing from single hairs. Nature 332:543–546

Jazin E, Soodyall H, Jalonen P, Lindholm E, Stoneking M, Gyllensten U (1998) Mitochondrial mutation rate revisited: hot spots and polymorphism. Nat Genet 18:109–110

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL (1983) Radiation of human mitochondria DNA types analyzed by restriction endonuclease cleavage patterns. J Mol Evol 19:255–271

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neandertal DNA sequences and the origin of modern humans. Cell 90:19–30

Lahr MM (1994) The multiregional model of modern human origins: a reassessment of its morphological basis. J Hum Evol 26:23–56

Lahr MM, Foley R (1994) Multiple dispersals and modern human origins. Evol Anthropol 3:48–60

Lieberman DE (1995) Testing hypotheses about recent human evolution from skulls. Curr Anthropol 36:159–197

Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, Martinson J, Stoneking M (1995) Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. Am J Hum Genet 57:403–414

Merriwether DA, Hall WW, Vahlne A, Ferrell RE (1996) mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. Am J Hum Genet 59:204–212

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Norio R, Nevanlilnna HR, Perheentupa J (1973) Hereditary diseases in Finland. Ann Clin Res 5:109–141

Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro ASM, Stoneking M (1995) Evolutionary history of the COII/tRNA-lys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Biol Evol 12:604–615

Ritte U, Neufeld E, Prager EM, Gross M, Hakim I, Khatib A, Bonné-Tamir B (1993) Mitochondrial DNA affinity of several Jewish communities. Hum Biol 65:359–385

Sajantila A, Salem A-H, Savolainen P, Bauer K, Gierig C, Pääbo S (1996) Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc Natl Acad Sci USA 93:12035–12039

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) Arlequin: an exploratory population genetics software environment. University of Geneva, Geneva

Scozzari R, Torroni A, Semino O, Cruciani F, Spedini G, Santachiara-Benerecetti AS (1994) Genetic studies in Cameroon: mitochondrial DNA polymorphisms in Bamileke. Hum Biol 66:1–12

Scozzari R, Torroni A, Semino O, Sirugo G, Brega A, Santachiara-Benerecetti AS (1988) Genetic studies on the Senegal population. I. Mitochondrial DNA polymorphisms. Am J Hum Genet 43:534–544

Seielstad M, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. Nat Genet 20:278–280

Semino O, Torroni A, Scozzari R, Brega A, De Benedictis G, Santachiara-Benerecetti AS (1989) Mitochondrial DNA polymorphisms in Italy. III. Population data from Sicily: a possible quantitation of maternal African ancestry. Ann Hum Genet 53:193–202

Semino O, Torroni A, Scozzari R, Brega A, Santachiara-Benerecetti AS (1991) Mitochondrial DNA polymorphisms among Hindus: a comparison with the Tharu of Nepal. Ann Hum Genet 55:123–126

Shaw I, Nicholson P (1995) British museum dictionary of ancient Egypt. British Museum Press, London

Soodyall H, Jenkins T (1992) Mitochondrial DNA polymorphisms in Khoisan populations from Southern Africa. Ann Hum Genet 56:315–324

Stoneking M (1996) Mitochondrial DNA variation and human evolution. In: Jackson M, Strachan T, Dover G (eds) Mitochondrial DNA variation and human evolution. BIOS Scientific, Oxford, pp 263–281

Stoneking M, Fontius JJ, Clifford SL, Soodyall H, Arcot SS, Saha N, Jenkins T, et al (1997) *Alu* insertion polymorphisms and human evolution: evidence for a larger population size in Africa. Genome Res 7:1061–1071

Swisher CC, Curtis GH, Jacob T, Getty AG, Suprijo A, Widiasmoro (1994) Age of the earliest known hominids in Java, Indonesia. Science 263:1118–1121

Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M (1995) The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. Am J Hum Genet 57: 1463–1475

Tikochinski Y, Ritte U, Gross SR, Prager EM, Wilson AC (1991) mtDNA polymorphism in two communities of Jews. Am J Hum Genet. 48:129–136

Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, Bonné-Tamir B, et al (1996) Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science 271:1380–1387

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354

Vilkki J, Savontaus M-L, Nikoskelainen EK (1988) Human mitochondrial DNA types in Finland. Hum Genet 80: 317–321

Wakeley J (1993) Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. J Mol Evol 37:613–623

Wallace DC, Torroni A (1992) American Indian prehistory as written in the mitochondrial DNA: a review. Hum Biol 64: 403–416

Ward RH, Frazier BL, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 88:8720–8724

Watson E, Bauer K, Aman R, Weiss G, von Haeseler A, Pääbo S (1996) mtDNA sequence diversity in Africa. Am J Hum Genet 59:437–444