Am. J. Hum. Genet. 72:1058–1064, 2003

Report

Extensive Female-Mediated Gene Flow from Sub-Saharan Africa into Near Eastern Arab Populations

Martin Richards,[1] Chiara Rengo,[2,3] Fulvio Cruciani,[2] Fiona Gratrix,[4] James F. Wilson,[5] Rosaria Scozzari,[2] Vincent Macaulay,[6] and Antonio Torroni[7]

[1]Department of Chemical and Biological Sciences, University of Huddersfield, Huddersfield, United Kingdom; [2]Dipartimento di Genetica e Biologia Molecolare, Università di Roma "La Sapienza," and [3]Istituto di Medicina Legale, Università Cattolica del Sacro Cuore, Rome; [4]Department of Biological Sciences, Imperial College, and [5]Department of Biology, University College London, London; [6]Department of Statistics, University of Oxford, Oxford, United Kingdom; and [7]Dipartimento di Genetica e Microbiologia, Università di Pavia, Pavia, Italy

We have analyzed and compared mitochondrial DNA variation of populations from the Near East and Africa and found a very high frequency of African lineages present in the Yemen Hadramawt: more than a third were of clear sub-Saharan origin. Other Arab populations carried ~10% lineages of sub-Saharan origin, whereas non-Arab Near Eastern populations, by contrast, carried few or no such lineages, suggesting that gene flow has been preferentially into Arab populations. Several lines of evidence suggest that most of this gene flow probably occurred within the past ~2,500 years. In contrast, there is little evidence for male-mediated gene flow from sub-Saharan Africa in Y-chromosome haplotypes in Arab populations, including the Hadramawt. Taken together, these results are consistent with substantial migration from eastern Africa into Arabia, at least in part as a result of the Arab slave trade, and mainly female assimilation into the Arabian population as a result of miscegenation and manumission.

Contacts between eastern Africa and Arabia have existed since the establishment of obsidian exchange networks as early as the 7th millennium B.C. They were stimulated by the growth of the Egyptian state from the 4th millennium onward, with possible settlements of people from Arabia in the Horn of Africa as early as the 3rd and 2nd millennia B.C. The Afro-Arabian Tihama cultural complex, for which an African origin seems most likely, arose in the mid-2nd millennium. In addition to the coastal site of Adulis in Eritrea and sites farther inland in Eritrea, Ethiopia, and Sudan, it is represented on the Saudi coastal plains and the western and southern coasts of Yemen. Other traditions appear to have spread in the opposite direction (Fattovich 1997).

Southern Arabs gained control of the Red Sea trade routes in the 12th century B.C., and the first kingdom, Saba, arose in Yemen in ~800 B.C. As a result, Eritrea and Ethiopia were gradually incorporated into the area of Arab influence. These contacts crystallized in the formation of the Ethio-Sabean state of Daamat on the Tigrean plateau in ~600 B.C. The archaeological evidence suggests that this is likely to have been the result of small-scale colonization by several Arabian groups (including Sabeans) and acculturation of the indigenous population (Fattovich 1997). This was succeeded, following the decline of Saba and Daamat and several centuries of isolation, by the kingdom of Aksum in ~A.D. 100, which formed as part of the Roman exchange network extending from Egypt as far as India, trading via the Red Sea port of Adulis. Aksum survived for 800 years and occupied southern Arabia for part of this period. Utilitarian Aksumite pottery has been found in large quantities in deposits from the 5th and 6th centuries in the Yemen Hadramawt, suggesting that there may have been substantial immigration during that period. From the 7th century onward, with the rise of Islam, Muslim communities became established along the coast of Eritrea and Somalia, spreading inland from the late 1st millennium onward (Fattovich 1997). Concomitantly, the Arab slave trade, operating from pre-Islamic times but at its height between A.D. 650 and 1900, carried millions of men and women from the Nile Valley, the Horn of Africa,

Received December 2, 2002; accepted for publication January 21, 2003; electronically published March 10, 2003.

Address for correspondence and reprints: Dr. Martin Richards, Department of Chemical and Biological Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, United Kingdom. E-mail: m.b.richards@hud.ac.uk

0002-9297/2003/7204-0030$15.00

and the eastern African coast across the Red Sea to Arabia and many more millions across the Sahara (Segal 2001).

To evaluate the extent and nature of gene flow between Africa and the Near East, we have examined mtDNA variation in numerous populations from the two geographic areas and compared it with that observed in the Y chromosome. These genetic systems are uniparentally transmitted and trace out the female and male genealogies. Moreover, their sequence variation has been subdivided into a number of clades, termed "haplogroups," each of which has originated at one time and place and been subsequently dispersed by the movement of people. Thus, haplogroups tend to be geographically restricted and are particularly informative for detecting gene flow between continents. The haplogroup profiles of Africans and Eurasians are sufficiently distinct to allow us to address the question of historical gene flow between eastern Africa and Arabia and whether the same patterns are seen in both female and male lines of descent.

We studied the variation in the first hypervariable segment (HVS-I) of the mtDNA control region (encompassing at minimum nps 16090–16365 but as much as nps 16017–16497 in many cases; numbering as per Anderson et al. [1981]) in 533 subjects from Arabic-speaking communities and 960 from non-Arabic–speaking communities in the Near East, including 344 Near Eastern Jews (table 1). (An additional set of 38 Yemeni Jews [Richards et al. 2000] was also included but was kept distinct, since some members may be from the same sample as those reported by Thomas et al. [2002].) We compared these with 74 Ethiopians and 46 Ethiopian Jews, 246 additional East Africans, 447 West Africans, 132 Central Africans, and 98 Khoisan and 416 Mozambican Bantu speakers from southern Africa (table 1). Haplogroup affiliation was defined primarily on the basis of the observed HVS-I motif, with some additional testing of diagnostic RFLP markers to ensure that misassignments were avoided. These were: haplogroups H (−7025 *Alu*I), HV (−14766 *Mse*I), V (−4577 *Nla*III), U (+12308 *Hin*fI), K (−9052 *Hae*II), JT (+4216 *Nla*III), T (+15606 *Alu*I), F1 (−12406 *Hin*cII), R (+12703 *Mbo*II), M (+10397 *Alu*I, +10394 *Dde*I), D (−5176 *Alu*I), N (+10871 *Mnl*I), X (+14465 *Acc*I), W (+8249 *Ava*II, –8994 *Hae*III), A (+663 *Hae*III), L3b (+10084 *Taq*I), L3d (−8616 *Mbo*I), L3e (+2349 *Mbo*I), and paragroup L1, along with haplogroup L2 (+3592 *Hpa*I).

We also analyzed Y-chromosome variation in 323 subjects from Arabic-speaking communities and 1,525 from non-Arabic–speaking communities in the Near East. For comparison, we also included 214 Ethiopians, 530 West Africans, 36 Central Africans, 53 southern African Bantu speakers, and 129 southern African Khoisan speakers (table 2).

Human mtDNA and Y-chromosome variation was interpreted within the following phylogeographic framework. Human mtDNA falls into clades L1a–f, L2, L3b, L3d, and L3e, and a paraphyletic cluster L3*, which characterize sub-Saharan Africans (Chen et al. 1995; Watson et al. 1997; Alves-Silva et al. 2000; Salas et al. 2002); and two further clades within L3, namely N and M, which are largely pan-Eurasian and East/South Asian, respectively (Torroni et al. 1994; Kivisild et al. 1999;

Table 1

Percentages of mtDNA Haplogroups L1–L3A, U6, M1, and (pre-HV)1 in Near Eastern and African Populations

Group and Population/Region	*N*	Haplogroup Frequencies (%)			
		L1–L3A	U6	M1	(pre-HV)1
Arab Near East:					
Iraqis[a]	116	9	1	1	4
Bedouin[b]	29	10	7	7	17
Palestinians[a]	117	15	1	2	3
Syrians[a]	69	9	4	0	6
Jordanians[a]	146	14	0	2	0
Yemen Hadramawt[b]	56	34	0	4	7
Non-Arab Near East:					
Georgians[c]	105	0	0	0	2
Armenians[a]	192	0	0	0	1
Azerbaijanis[a]	48	0	0	0	0
Turks[a]	218	1	0	0	1
Kurds[a]	53	4	0	0	2
Near Eastern Jews:					
Iraqi Jews[c]	56	0	0	7	0
Iranian Jews[c]	75	0	0	1	1
Ashkenazi Jews[c]	78	0	3	0	3
Georgian Jews[c]	70	0	0	0	0
Yemen Jews[c]	65	11	0	0	17
Yemen Jews[a]	38	5	0	0	26
East Africa:					
Ethiopians[c]	74	55	3	10	8
Ethiopian Jews[c]	46	52	0	15	15
Somalians[d]	27	70	0	11	11
Nubians[e]	80	58	0	10	9
Southern Sudanese[e]	76	92	0	4	0
Kenyans[d]	63	95	3	2	0
West Africa:					
Overall[d,f,g,h]	447	89	3	0	0
Central Africa:					
Mbuti and Biaka[f]	37	100	0	0	0
Equatorial Guineans[i]	95	99	1	0	0
Southern Africa:					
Kung and Khwe[f,j]	98	100	0	0	0
Mozambicans[k,l]	416	100	0	0	0

[a] Richards et al. (2000).
[b] Di Rienzo and Wilson (1991).
[c] Thomas et al. (2002).
[d] Watson et al. (1997).
[e] Krings et al. (1999).
[f] Vigilant et al. (1991).
[g] Graven et al. (1995).
[h] Rando et al. (1998).
[i] Mateu et al. (1997).
[j] Chen et al. (2000).
[k] Pereira et al. (2001).
[l] Salas et al. (2002).

Table 2

Percentages of Y-Chromosome Haplogroups A, E, and E3a in Near Eastern and African Populations

Group and Population/Region	*N*	Haplogroup Frequencies (%)		
		A[a]	E	E3a
Arab Near East:				
Yemen Hadramawt[b]	49	NA	10	4
Yemen Sena[b]	27	NA	0	0
Bedouin[c]	32	0	19	0
Palestinians[c]	143	1	20	0
Syrians[d]	72	1	23	1
Non-Arab Near East:				
Georgians[e]	68	0	1	0
Armenians[e]	734	0	5	0
Kurds[c]	95	0	7	0
Azeris[e]	29	0	7	0
Turks[e]	173	0	10	0
Greeks[e]	132	0	27	0
Kurdish Jews[c]	99	0	12	0
Sephardi Jews[c]	78	0	19	0
Ashkenazi Jews[c]	79	0	23	0
Africa:				
Ethiopians[f,g]	214	13	68	1
West Africans[f,g,h]	530	1	81	69
Central Africans[f]	36	0	67	58
Southern African Bantu[f]	53	6	81	62
Southern African Khoisan[f,h]	129	33	54	36

Note.—Y-chromosome nomenclature is that established by the Y Chromosome Consortium (2002).

[a] NA = not available.
[b] Thomas et al. (2000).
[c] Nebel et al. (2001).
[d] Wilson et al. (2001).
[e] Weale et al. (2001).
[f] Underhill et al. (2000).
[g] Semino et al. (2002).
[h] Cruciani et al. (2002).

Quintana-Murci et al. 1999; Richards and Macaulay 2000, 2001; Richards et al. 2000). For convenience, we refer to African L3 lineages (which do not form a clade) as "L3A" (Rando et al. 1998) and to the entire set of sub-Saharan haplogroups as "L1–L3A." Approximately 85% of Near Eastern lineages fall into a set of haplogroups within haplogroup N that are characteristic of west Eurasians: H, (pre-HV)1, HV1, U1–U7, K, J, T, I, W, X, and N1b, most of which probably originated in the Near Eastern/Caucasus region (Torroni et al. 1996; Macaulay et al. 1999; Richards et al. 2000). Among these is the African haplogroup U6, which is of ancient Eurasian ancestry (since haplogroup U probably evolved in the Near East), but diverged in North Africa over the past 20,000–40,000 years and can therefore be regarded as characteristic of indigenous North Africans (Rando et al. 1998; Macaulay et al. 1999). Its presence in the Near East, therefore, implies gene flow from North Africa.

Two haplogroups, (pre-HV)1 and M1, have a distribution that spans the Red Sea. Haplogroup (pre-HV)1 occurs widely throughout the Near East, reaching highest frequency in Arabia, but is also common in Ethiopia and Somalia (Watson et al. 1997; Richards et al. 2000). Given its close phylogenetic relationship with other Eurasian clusters, this haplogroup probably originated in the Near East and spread later into eastern Africa. Haplogroup M1, however, has been assigned an African origin, despite the facts that (i) all other basal branches of haplogroup M are restricted to South Asia, East Asia, and Australasia, and (ii) the diversity of M in Asia is greater than in Africa (Quintana-Murci et al. 1999). It is restricted to the Near East and north and eastern Africa, concentrated in Somalia and Ethiopia (Watson et al. 1997). It is therefore unclear whether any particular M1 sequence type in the Near East arrived recently from Africa; an Asian origin with back-migration to Africa is possible.

Our estimates of sub-Saharan African ancestry in the Near East are, therefore, based on haplogroup L1–L3A lineages, but we show also the distribution of these other clusters present both in the Near East and East Africa.

Our Y-chromosome nomenclature follows the phylogenetic scheme established by the Y Chromosome Consortium (2002). The principal diagnostic haplogroups for sub-Saharan Africa are the ancient haplogroups A (defined by the derived form of the M91 marker) and B (derived at M60 and M181) and a derived form of haplogroup E, E3a (derived at M2), which occurs in Bantu-speaking populations at high frequency. The remainder of haplogroup E (with the YAP^+ M1 marker and derived at M96) is also present throughout Africa and in the Near East (Hammer et al. 1998, 2001; Scozzari et al. 1999, 2001; Rosser et al. 2000; Semino et al. 2000; Thomas et al. 2000; Underhill et al. 2000, 2001; Cruciani et al. 2002). It is unfortunate that haplogroup B was not distinguished from some Eurasian haplogroups before the introduction of high-resolution binary markers (Underhill et al. 2000) and cannot, therefore, be used to identify gene flow between eastern Africa and Eurasia in most published data sets. However, haplogroup B occurs only at low frequencies in eastern Africa and has not been detected outside Africa in global surveys (Underhill et al. 2000, 2001; Hammer et al. 2001), so this is unlikely to lead to serious bias. Indeed, Y-chromosome variation is often characterized by very steep frequency gradients, which, in some respects, may make phylogeographic inferences more straightforward than for other markers.

The distribution of mtDNA haplogroups in the Near East and Africa is shown in table 1. Haplogroups L1–L3A have so far been found at highest frequencies in Central Africans and southern African Khoisan speakers, where they comprise ~100% of extant lineages. For example, the sampled Pygmies include only L1 and L2 lineages.

L1–L3A make up 89% of mtDNAs in West Africa, >90% in southern East Africa, ~70% in Somalia, and ~55% in Ethiopia.

The reason for the lower frequency of haplogroups L1–L3A in Ethiopia is the presence both of haplogroups (pre-HV)1 and M1 (at high frequencies) and of the west Eurasian haplogroups T, J, U, and HV, which are indicative of substantial gene flow from the Near East. West Eurasian mtDNAs are elsewhere very rare in sub-Saharan Africa, the main previously discovered examples having entered Nubia from Egypt (Krings et al. 1999) and into the western Sahara from northwest Africa (Rando et al. 1998). In the case of Ethiopia, the west Eurasian types mostly match types in Arabia, with only a couple of exceptions of rare derived types not previously seen elsewhere. Haplogroups (pre-HV)1 and M1 are found primarily both in eastern Africa and the Near East. In Ethiopia, by contrast with the other west Eurasian types, instances of both (pre-HV)1 and M1 types tended to be unique types or to match others found only in eastern Africa. These patterns—as well as the rather higher frequencies of (pre-HV)1 and M1 in Ethiopia—suggest that the (pre-HV)1 and M1 lineages in Ethiopia may be the result of fairly ancient interactions between East Africa and Eurasia, whereas the other west Eurasian types may be the result of more recent historical gene flow. Our estimate of recent Near Eastern mtDNA input (22%) amounts to considerably more than the 5% of Passarino et al. (1998). Contra Passarino et al. (1998), this does not seem to differ from the recent Y-chromosome input from the Near East signaled by Y-chromosome haplogroup J, which they estimate at ~25% but which has been subsequently shown to vary substantially in different Ethiopian populations (Semino et al. 2002).

Haplogroups L1–L3A in the Near East reach their highest frequency in the Yemen Hadramawt (~35%). Other Arab populations—Palestinians, Jordanians, Syrians, Iraqis, and Bedouin—have ~10%–15% of lineages of sub-Saharan African origin. These types are rarely shared between different Arab populations. By contrast, non-Arab Near Eastern populations—Turks, Kurds, Armenians, Azeris, and Georgians—have few or no such lineages, suggesting that gene flow from Africa has been specifically into Arab populations. For comparison, southern European mtDNAs include only ~2% of these lineages, and northern Europeans <1% (Richards et al. 2000). The only European region to stand out is Iberia, where ~4% of mtDNAs belong to these clusters, probably a trace of the medieval Moorish conquests (Côrte-Real et al. 1996; Richards et al. 2000). There is also evidence from one sample (Semino et al. 1989) that in parts of Sicily, which was held by the Arabs between 825–1091 A.D., haplogroups L1 and L2 amount to ~4%.

We compared the frequency of haplogroups L1–L3A in Jewish communities from the Near East with that in non-Jewish communities residing historically in the same area (table 1). Near Eastern Jewish groups almost entirely lack haplogroups L1–L3A (as, indeed, do Ashkenazi Jews [Thomas et al. 2002]). The only exception is in Jews from Yemen, but, even here, these lineages amount only to a quarter of their frequency in the non-Jewish sample from the Hadramawt. (L1 and L2 types are completely absent.)

These observations enable us to put a tentative time frame on the widespread appearance of sub-Saharan lineages in the Near East. The eastern Jewish communities probably emerged no earlier than 586 B.C., at the time of the destruction of the First Temple (Mourant et al. 1978). The mtDNAs in these communities show evidence of strong founder effects involving distinct sequence types in each case, in some cases matching types found locally. Therefore, it seems likely that substantial conversion of indigenous women took place at the time of the founding of these communities, followed by subsequent isolation (Thomas et al. 2002). In the case of the Yemeni Jews, this scenario is supported by blood-group evidence and the historical record. Indeed, in the case of the Yemeni Jews, it is likely that the main conversions took place between 70 A.D. and 132 A.D. (Mourant et al. 1978).

This being the case, the lack of haplogroups L1–L3A in these communities implies that these haplogroups were probably absent at the time at which the Jewish communities were founded. This suggests that much of the African gene flow into the neighboring Arab populations most likely postdates the 5th century B.C., in the case of the northern Arab populations, and perhaps even the 2nd century A.D., in the case of the Yemen.

It is conceivable that haplogroups L1–L3A have been lost from the Jewish communities as a result of genetic drift, although the independent loss of both L1 and L2 from all Jewish groups seems unlikely. However, several other lines of evidence also support recent introgression. More than half of the Yemen L1–L3A lineages occur at the tips of the mtDNA tree (cf. Salas et al. 2002), indicating that they have been generated by mutation relatively recently. Furthermore, a majority of the L1–L3A lineages in the Hadramawt—such as members of L2a, L2d, L3b, and L3d—trace back ultimately to West Africa, so that it is likely that they were delivered to East Africa by the Bantu dispersals. Supporting this suggestion, all of the L2a types in the Hadramawt occur at elevated frequency in the Bantu speakers of Mozambique (Pereira et al. 2001; Salas et al. 2002). Moreover, the chief L1a type in the Hadramawt also occurs at elevated frequency in Bantu speakers and is implicated in the Bantu dispersals, albeit having been picked up in East Africa en route (Salas et al. 2002). Bantu speakers are thought to have become first established to the east of the Great Lakes region somewhat <2,000 years ago (Phillipson 1993). Assuming that the sub-Saharan African input into Ara-

bia is indeed directly from east Africa (rather than including a component from west or southeastern Africa), as is most likely on historical and geographical grounds (Segal 2001), this again limits the main spread into Arabia to within the last ~2,000 years.

In addition to lineages of west Eurasian and sub-Saharan African origin, the Yemen Hadramawt sample also includes several haplogroup M lineages, some originating in south Asia and some in southeast Asia. These connections may testify to female gene flow along the trading networks established eastward from the Red Sea, along the Indian Ocean, and as far as southeast Asia.

Ethiopian Y-chromosome variation has been characterized in some detail (Passarino et al. 1998; Underhill et al. 2000; Semino et al. 2002). These data show that, as in their mtDNA, Ethiopians differ from other sub-Saharan African populations in their haplogroup profiles. They show ~68% haplogroup E (including only ~1% the derived haplogroup E3, which characterizes sub-Saharan Africans), ~13% haplogroup A, ~5% the African haplogroup B, and ~3% the Eurasian haplogroup K*. The remainder is mainly haplogroup J, which occurs at high frequency in most Near Eastern populations.

By contrast, throughout the Near East, haplogroup A is virtually absent, for example, in Bedouin, Palestinians, and Syrians, as well as in Turks, Kurds, Armenians, Azeris, and Georgians and several Jewish groups (table 2). Haplogroup E is present in both Arab and Jewish populations throughout the Near East, as well as at high frequencies throughout most of Africa (Scozzari et al. 1999, 2001; Underhill et al. 2000; Cruciani et al. 2002). However, its distribution in the Near East suggests an ancient presence in the region, rather than indicating recent gene flow: it is present not only in Near Eastern Arab populations, but also in several groups of Jews (12%–23%) and Turks (~10%), declining to <5% as one moves toward the Caucasus and Europe. Further supporting the suggestion of an ancient presence rather than recent gene flow from East Africa, haplogroup E occurs at only ~10% in the Yemen Hadramawt, substantially lower than most other Arab and Jewish groups in the Near East. Yet this is precisely the region in which female-mediated gene flow from Eastern Africa reaches its highest levels. Only ~4% of the Hadramawt sample is the derived sub-Saharan African form, E3a, which indicates recent gene flow from Africa. This subclade is virtually absent from all other Near Eastern populations sampled. Moreover, haplogroup E is entirely absent from a second Yemen sample from Sena (Thomas et al. 2000).

Taken together, these results indicate that historical male-mediated gene flow from Ethiopia to Yemen has been low, in striking contrast to the results from mtDNA. This seems likely to be true for other parts of the Near East, too, again in contrast to the pattern from mtDNA.

The mtDNAs of sub-Saharan origin that are present in Near Eastern populations show a striking phylogeographic pattern. They are virtually restricted to Arab populations, occurring at ~35% in the Yemen Hadramawt and ~10%–15% in other Arab populations. They are absent or almost so from Turks, Armenians, Azeris, Georgians, and Near Eastern Jews and present at lower levels in Turkish Kurds. In Europe, they are detected at appreciable levels only in regions which experienced Arab rule during the medieval period. This pattern suggests that most female gene flow from sub-Saharan Africa into the Near East probably took place relatively recently, within the last ~2,500 years. Y-chromosome data indicate that recent male gene flow was substantially less. This appears to be the case even for the Yemen, where more than a third of mtDNAs derive from Africa.

In summary, these results are consistent with mainly female migration from eastern Africa into Arab communities within the last few thousand years. There have been many opportunities for such migrations between eastern Africa and southern Arabia during this period. However, the most likely explanation for the presence of predominantly female lineages of African origin in other parts of the Arab world is that these may trace back to women brought from Africa as part of the Arab slave trade, assimilated into the Arabian population as a result of miscegenation and manumission. Indeed, unlike the situation in the Americas, there are no substantial communities of African descent in the Near East today. This is thought to be because relatively few men—mainly employed in manual labor and military service or castrated and employed as eunuchs—left descendants. Women, by contrast, were imported specifically for the sexual gratification of élite males and for their reproductive potential. The practice of manumission meant that their offspring were born free. Female slaves were, therefore, readily integrated into Islamic society (Lewis 1992; Segal 2001).

A number of recent studies have compared Y-chromosome and mtDNA variation and drawn conclusions about sex-specific migration (Underhill et al. 2001; Salas et al. 2002). Some of these have been associated with patrilocality versus matrilocality (Oota et al. 2001) and others with ethnic-specific long-range dispersal patterns (Thomas et al. 2002). By contrast, this study indicates the long-term effects of a particular socioeconomic system, based on slavery, on the gene pool of an entire region.

Acknowledgments

We thank Mike Weale, Mark Thomas, and Hans-Jürgen Bandelt, for critical advice, and David Goldstein and Neil Bradman, for support. This research also received support from the Italian Ministry of the University, Progetti Ricerca Interesse Nazionale 2002 (to A.T. and R.S.), Progetto MIUR-CNR Genomica Funzionale (to A.T.), Fondo d'Ateneo per la

Ricerca 2002 dell'Università di Pavia (to A.T.), the Istituto Pasteur Fondazione Cenci Bolognetti, Università di Roma "La Sapienza" (to R.S.), Grandi Progetti di Ateneo, Università di Roma "La Sapienza" (to R.S.), and a Research Career Development Fellowship from the Wellcome Trust (to V.M.).

References

Alves-Silva J, da Silva Santos M, Guimarães PEM, Ferreira ACS, Bandelt H-J, Pena SDJ, Prado VF (2000) The ancestry of Brazilian mtDNA lineages. Am J Hum Genet 67:444–461

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Chen Y-S, Olckers A, Schurr TG, Kogelnik AM, Huoponen K, Wallace DC (2000) mtDNA variation in the South African Kung and Khwe—and their genetic relationships to other African populations. Am J Hum Genet 66:1362–1383

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

Côrte-Real HBSM, Macaulay VA, Richards MB, Hariti G, Issad MS, Cambon-Thomsen A, Papiha S, Bertranpetit J, Sykes BC (1996) Genetic diversity in the Iberian peninsula determined from mitochondrial sequence analysis. Ann Hum Genet 60:331–350

Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scozzari R, Underhill PA (2002) A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am J Hum Genet 70: 1197–1214

Di Rienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci USA 88:1597–1601

Fattovich R (1997) The Near East and eastern Africa: their interaction. In: Vogel JO (ed) Encyclopedia of precolonial Africa. AltaMira Press, Walnut Creek, pp 479–484

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

Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol Evol 15:427–441

Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL (2001) Hierarchical patterns of global human Y-chromosome diversity. Mol Biol Evol 18:1189–1203

Kivisild T, Kaldma K, Metspalu E, Parik J, Papiha S, Villems R (1999) The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the Old World. In: Papiha S, Deka R, Chakraborty R (eds) Genomic diversity: applications in human population genetics. Plenum, New York, pp 135–152

Krings M, Salem AH, Bauer K, Geisert H, Malek AK, Chaix L, Simon C, Welsby D, Di Rienzo A, Utermann G, Sajantila A, Pääbo S, Stoneking M (1999) mtDNA analysis of Nile River valley populations: a genetic corridor or a barrier to migration? Am J Hum Genet 64:1166–1176

Lewis B (1992) Race and slavery in the Middle East: an historical enquiry. Oxford University Press, Oxford

Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonné-Tamir B, Sykes B, Torroni A (1999) The emerging tree of west Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am J Hum Genet 64: 232–249

Mateu E, Comas D, Calafell F, Pérez-Lezaun A, Abade A, Bertranpetit J (1997) A tale of two islands: population history and mitochondrial sequence variation of Bioko and São Tomé, Gulf of Guinea. Ann Hum Genet 61:507–518

Mourant AE, Kopec AC, Domaniewska-Sobczak K (1978) The genetics of the Jews. Clarendon Press, Oxford

Nebel A, Filon D, Brinkman B, Majumder PP, Faerman M, Oppenheim A (2001) The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. Am J Hum Genet 69:1095–1112

Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M (2001) Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29:20–21

Passarino G, Semino O, Quintana-Murci L, Excoffier L, Hammer M, Santachiara-Benerecetti AS (1998) Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. Am J Hum Genet 62:420–434

Pereira L, Macaulay V, Torroni A, Scozzari R, Prata MJ, Amorim A (2001) Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade. Ann Hum Genet 65:439–458

Phillipson DW (1993) African archaeology. Cambridge University Press, Cambridge

Quintana-Murci L, Semino O, Bandelt H-J, Passarino G, McElreavey K, Santachiara-Benerecetti AS (1999) Genetic evidence for an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. Nat Genet 23:437–441

Rando JC, Pinto F, González AM, Hernández M, Larruga JM, Cabrera VM, Bandelt H-J (1998) Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European, near-Eastern, and sub-Saharan populations. Ann Hum Genet 62:531–550

Richards M, Macaulay V (2000) Genetic data and the colonization of Europe: genealogies and founders. In: Renfrew C, Boyle K (eds) Archaeogenetics: DNA and the population prehistory of Europe, McDonald Institute for Archaeological Research, Cambridge, pp 139–151

——— (2001) The mitochondrial gene tree comes of age. Am J Hum Genet 68:1315–1320

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, et al (2000) Tracing European founder lineages in the Near Eastern mitochondrial gene pool. Am J Hum Genet 67:1251–1276

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D,

Amorim A, Amos W, et al (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543

Salas A, Richards M, De la Fe T, Lareu M-V, Sobrino B, Sánchez-Diz P, Macaulay V, Carracedo A (2002) The making of the African mtDNA landscape. Am J Hum Genet, 71:1082–1111

Scozzari R, Cruciani F, Pangrazio A, Santolamazza P, Vona G, Moral P, Latini V, Varesi L, Memmi MM, Romano V, De Leo G, Gennarelli M, Jaruzelska J, Villems R, Parik J, Macaulay V, Torroni A (2001) Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. Hum Immunol 62:871–884

Scozzari R, Cruciani F, Santolamazza P, Malaspina P, Torroni A, Sellitto D, Arredi B, Destro-Bisol G, De Stefano G, Rickards O, Matinez-Lebarga C, Modiano D, Biondi G, Moral P, Olckers A, Wallace DC, Novelletto A (1999) Combined use of biallelic and microsatellite Y-chromosome polymorphisms to infer affinities among African populations. Am J Hum Genet 65:829–846

Segal R (2001) Islam's black slaves. Atlantic, London

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. Science 290: 1155–1159

Semino O, Santachiara-Benerecetti AS, Falaschi F, Cavalli-Sforza LL, Underhill PA (2002) Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. Am J Hum Genet 70:265–268

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

Thomas MG, Parfitt T, Weiss DA, Skorecki K, Wilson JF, le Roux M, Bradman N, Goldstein DB (2000) Y chromosomes traveling south: the Cohen modal haplotype and the origins of the Lemba—the "black Jews of southern Africa." Am J Hum Genet 66:674–686

Thomas MG, Weale ME, Jones AL, Richards M, Smith A, Redhead N, Torroni A, Scozzari R, Gratrix F, Tarakegn A, Wilson JF, Capelli C, Bradman N, Goldstein DB (2002) Founding mothers of Jewish communities: geographically separated Jewish groups were independently founded by very few female ancestors. Am J Hum Genet 70:1411–1420

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus M-L, Wallace DC (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang JG, Droma T, Wallace DC (1994) Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93: 189–199

Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA, Oefner PJ, Cavalli-Sforza LL (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet 65:43–62

Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonné-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ (2000) Y chromosome sequence variation and the history of human populations. Nat Genet 26:358–361

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Watson E, Forster P, Richards M, Bandelt H-J (1997) Mitochondrial footprints of human expansions in Africa. Am J Hum Genet 61:691–704

Weale ME, Yepiskoposyan L, Jager RF, Hovhannisyan N, Khudoyan A, Burbage-Hall O, Bradman N, Thomas MG (2001) Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group. Hum Genet 109:659–674

Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB (2001) Genetic evidence for different male and female roles during cultural transitions in the British Isles. Proc Natl Acad Sci USA 98:5078–5083

Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–348