Am. J. Hum. Genet. 70:1411–1420, 2002

Founding Mothers of Jewish Communities: Geographically Separated Jewish Groups Were Independently Founded by Very Few Female Ancestors

Mark G. Thomas,[1,*] Michael E. Weale,[1,*] Abigail L. Jones,[1] Martin Richards,[3] Alice Smith,[2] Nicola Redhead,[2] Antonio Torroni,[5,6] Rosaria Scozzari,[6] Fiona Gratrix,[2] Ayele Tarekegn,[1] James F. Wilson,[2] Cristian Capelli,[2] Neil Bradman,[1] and David B. Goldstein[2]

[1]The Centre for Genetic Anthropology, Departments of Biology and Anthropology, and [2]Department of Biology, University College London, London; [3]Department of Chemical and Biological Sciences, University of Huddersfield, Huddersfield, United Kingdom; [4]Bruce Rappaport Faculty of Medicine and Research Institute, Technion and Rambam Medical Center, Haifa, Israel; [5]Dipartimento di Genetica e Microbiologia, Università di Pavia, Pavia, Italy; and [6]Dipartimento di Genetica e Biologia Molecolare, "La Sapienza" di Roma, Rome

We have analyzed the maternally inherited mitochondrial DNA from each of nine geographically separated Jewish groups, eight non-Jewish host populations, and an Israeli Arab/Palestinian population, and we have compared the differences found in Jews and non-Jews with those found using Y-chromosome data that were obtained, in most cases, from the same population samples. The results suggest that most Jewish communities were founded by relatively few women, that the founding process was independent in different geographic areas, and that subsequent genetic input from surrounding populations was limited on the female side. In sharp contrast to this, the paternally inherited Y chromosome shows diversity similar to that of neighboring populations and shows no evidence of founder effects. These sex-specific differences demonstrate an important role for culture in shaping patterns of genetic variation and are likely to have significant epidemiological implications for studies involving these populations. We illustrate this by presenting data from a panel of X-chromosome microsatellites, which indicates that, in the case of the Georgian Jews, the female-specific founder event appears to have resulted in elevated levels of linkage disequilibrium.

Introduction

Before the Second World War (1939–1945) and the founding of the modern state of Israel (1948), there were many long-standing separate Jewish communities in Europe, North Africa, and Asia. All of them claimed an origin in one or another dispersal from Israel and Judea. However, the origins of small minority communities founded before the 16th century are rarely well documented. For some Jews (e.g., the Babylonian Jews and modern Iraqi Jews), evidence exists of ancient Jewish communities in the same locations as in present times, but gaps often exist in the records of intervening centuries (Rejwan 1985, p. 143). In no case is there clear evidence of unbroken genetic continuity from early dispersal events to the present (de Lange 1984, p. 15; Encyclopaedia Judaica 1972).

A further difficulty in reconstructing past events is that the extent of travel in earlier times between Jewish communities, once they had become established, is not well understood. On the one hand, separation was sufficient for religious customs to vary from place to place (Reif 1993). On the other hand, intercommunity travel for social, religious, and trading purposes could be extensive (Beinart 1992). At the end of the 15th century, many Jewish communities in North Africa and western Asia were augmented as a consequence of the expulsions from Spain in 1492 (Ben-Sasson 1976).

How Jewish identity was determined in antiquity is also unclear. Conversion to Judaism was not uncommon in the pre-Christian Roman Empire and, in the 1st millennium A.D., the ruling classes of more than one polity adopted Judaism as the state religion (e.g., Himyar and Khazaria). Although tribal affiliation and priestly status were determined by paternal descent in ancient Israel, since Talmudic times (circa 200 B.C. to 500 A.D.), Jewish status has been defined, in the absence of conversion, by maternal descent (Encyclopaedia Judaica 1972).

Studies using both classical and molecular markers have provided evidence both for the common genetic origin of Jewish communities and for admixture between Jewish communities and their geographical neighbors (Mourant et al 1978; Livshits et al. 1991; Bonné-Tamir

Received November 2, 2001; accepted for publication March 6, 2002; electronically published April 30, 2002.

Address for correspondence and reprints: Dr. David Goldstein, Department of Biology, University College London, Gower Street, London WC1E 6BT, United Kingdom

* The first two authors contributed equally to this work.

0002-9297/2002/7006-0004$15.00

et al. 1992). Studies of the paternally inherited Y chromosome have indicated that Ashkenazic Jews, Sephardic Jews, and Lebanese all have high frequencies of haplogroup J (as defined by the Y Chromosome Consortium 2002), in contrast to the populations of central Europe (Ritte et al. 1993*a*; Santachiara-Benerecetti et al. 1993). Thomas et al. (1998) reported that a particular microsatellite haplotype (the "Cohen modal haplotype," now known to be within haplogroup J) occurs at a high frequency in priests of Jewish groups that had been geographically separated for a long period. This led them to suggest that it could be a genetic signature of the ancient Hebrew population. Subsequent analyses have confirmed a substantial sharing of Y-chromosome haplotypes among different Jewish communities and also between Jewish and non-Jewish Near Eastern populations (Hammer et al. 2000; Nebel et al. 2000).

Previous low-resolution RFLP studies of the maternally inherited mtDNA of Jews, using five or six restriction enzymes, have also revealed patterns interpreted both in terms of common origin and local admixture. Ritte et al. (1993*b*) found that genetic distances among seven Jewish communities and Israeli Arabs were comparable to those found among five globally dispersed populations, with Ethiopian Jews appearing more as an outgroup than Israeli Arabs. Tikochinski et al. (1991) and Ritte et al. (1992) found that genetic diversity within Jewish populations was generally lower than in populations with a geographically extensive distribution, such as whites, Asians, Australians, and Africans, but was greater than that found in geographically restricted populations such as New Guineans, a pattern they attributed to an unusually polymorphic ancestral Jewish population, a high rate of growth in Jewish populations, or introgression events from neighboring populations. Ritte et al. (1993*a*) compared mtDNA and Y-chromosome haplotypes in six Jewish communities and found consistently lower genetic diversity in the mtDNA than in the Y-chromosome haplotypes, although differences in mutational processes between these two marker systems make it very difficult to ascribe such differences with certainty to demographic effects.

Here we directly compare the paternal and maternal genetic histories of Jewish populations through paired analyses of Y chromosome and mtDNA diversity in Jewish populations and relevant host populations among whom the Jewish populations have been living for many generations. The use of paired comparisons (Jewish vs. non-Jewish) overcomes the difficulties that arise when making direct comparisons between mtDNA and Y-chromosome data (c.f. Ritte et al. 1993*a*). Given that the mutational processes are the same in both Jewish and non-Jewish populations within a genetic system, any consistent differences that are found must arise as the result of different population processes. The comparison of maternal and paternal histories—in the present case, mostly by use of the same population samples—is of particular interest, because of the long-standing religious definition of Jewish identity referred to above. If this system of matrilineal inheritance of Jewish identity has been strictly followed, we could expect it to be reflected in systematic differences in the pattern of mtDNA and Y-chromosome genetic variation within and among Jewish populations.

Subjects and Methods

Population Samples

Definitions for the communities included in this study were as follows. Ashkenazic Jews are Jews practicing the Ashkenazic rite associated with the communities of northern Europe. The origin of this community is the subject of historical dispute. Bene Israel (Indian Jews) are members of the present day Bene Israel community of the Bombay region; local tradition claims that the Bene Israel are descendants of refugees from the persecutions of Antiochus Epiphanus (175–163 B.C.). Beta Israel (Ethiopian Jews) are Jews who were born—or whose parents or grandparents were born—in Ethiopia; local tradition claims descent from noblemen who accompanied Menelik on his journey from Israel/Judea at the beginning of the 1st millennium B.C. Bukharan Jews are members of the present day Bukharan Jewish community in Uzbekistan; local tradition claims that the community originates from the expulsion from Israel in the 8th century B.C., and Jews from Persia may have joined the Bukharan Jewish community in the 3d to 7th century A.D. Georgian Jews are those associated with communities of the former Soviet Republic of Georgia; the community's origin is obscure, but there is evidence of a Jewish community in Tbilisi around the 4th century A.D. Iranian Jews are Jews who were born—or whose parents or grandparents were born—in Iran; the origin of the community is obscure, but it is thought that it may be an offshoot of the Iraqi community. Iraqi Jews are Jews who were born—or whose parents or grandparents were born—in Iraq; Jewish tradition places the origin of this community in the exile following the destruction of the first temple in 586 B.C. Moroccan Jews are Jews who were born—or whose parents or grandparents were born—in Morocco; the origin of Moroccan Jewry is obscure, but Jews have been present in the region at least since late antiquity. Yemenite Jews are Jews who were born—or whose parents or grandparents were born—in Yemen; their origin is obscure, but the ancient Jewish kingdom of Himyar was located in present day Yemen. Germans were used as a control group for the Ashkenazic Jews, as putative descendants of communities living in the areas that were settled early

by the Ashkenazim. The Bukharan Jews were collected in Uzbekistan, and the Bene Israel were collected in India and Israel. The other Jewish groups were collected in Israel. For the neighboring, non-Jewish host populations, samples were collected from self-identifying members of the various populations in the countries and regions indicated by their labels. Israeli Arabs/Palestinians were also included to represent an extant population living in the region to which Jews trace their origin. Individuals, so far as is known, are unrelated to the grandparent level. To obtain appropriate control subjects, Indians were sampled from the Bombay region, and Ethiopians were sampled among Amharic speakers. All sample donors were adult males, and appropriate informed consent was obtained in each case.

Y-chromosome markers typed include UEP markers 92R7, M9, M13, M17, M20, SRY 465, SRY 4064, SRY 10831, sY81, and YAP, and microsatellite markers DYS19, DYS388, DYS390, DYS391, DYS392, and DYS393. Typing protocols were as described by Thomas et al. (1999). The data for Israeli Arabs/Palestinians are as published elsewhere (Nebel et al. 2000).

Sequencing of mtDNA was performed using ABI BigDye Terminator Cycle Sequencing kits, and reactions were run on an ABI 3100 genetic analyzer using POP6 polymer and capillaries. Hypervariable segment 1 (HVS-1) was amplified in 10-μl reaction volumes containing 0.35 μM of primers conL1-mod (TAA ACT AAT ACA CCA GTC TTG TAA ACC) and conH1-mod (CCC TGA AGT AGG AAC CAG ATG), 0.13 units *Taq* polymerase (HT Biotech), 9.3 nM *Taq*Start monoclonal antibody (Clontech), 200 μM dNTPs and a buffer supplied with the polymerase. Cycling parameters were as follows: preincubation for 5 min at 95°C, followed by 37 cycles of 1 min at 93°C, 1 min at 55°C, and 1 min at 72°C; products were purified using MicroCLEAN (Microzone Ltd.). Forward-strand sequencing was performed in 15-μl reaction volumes using 1 μl BigDye termination mix, 5 μl Better Buffer (Microzone), and 0.16 μM primer conL2 (CAC CAT TAG CAC CCA AAG CT). Reverse-strand sequencing was performed on a subset of the samples, including all of those with the polycytosine tract at nucleotide positions 16182–16193, as above, except the reverse primer conH3 (CGG AGC GAG GAG AGT AGC) was used. We sequenced a minimum of 332 bp (average 450 bp) from the mtDNA HVS-1 in samples from each of the nine geographically isolated Jewish groups and from six of the eight host populations. The sequencing-error rate was assessed by repeat sequencing on 49 randomly selected samples. One difference was found (a reading error) providing an estimated per-sequence error rate of 0.02 and a per-nucleotide error rate of 5×10^{-5}. The new sequence data were combined with previously published mtDNA HVS-1 sequences from the following populations: Israeli Arabs/Palestinians (Di Rienzo and Wilson 1991; Richards et al. 2000), Syrians (Richards et al. 2000), and Germans (Richards et al. 1996; Hofmann et al. 1997). Together, these data all overlap between sites 16090–16365, according to the numbering scheme of Anderson et al. (1981), and this was the region used to construct mtDNA haplotypes.

To assess long-range patterns of linkage disequilibrium (LD), 34 microsatellite markers, spanning 150 Mb of the X chromosome, were typed in the Georgian Jews and in men from Britain ($n = 48$), Norway ($n = 78$), and Turkey ($n = 65$). The markers were typed as described by Wilson and Goldstein (2000). The loci are DXS7103, 1232, 8027, 8087, 8036, 984, 1067, 8009, 8061, 8085, 1204, 8092, 8099, 996, 8078, 1212, 1036, 1205, 8081, 1223, 1053, 8013, 8098, 1220, 8086, 8014, 8038, 1206, 8073, 1062, 1192, 1211, 1203, and 8068.

Statistical Analysis

Genetic diversity, h, was estimated from the unbiased formula given by Nei (1987). Pairwise genetic distance, F_{ST}, was estimated according to the formula of Reynolds et al. (1983). It should be noted that F_{ST} is used here as a convenient statistic summarizing multidimensional differences in allele frequencies, which affords comparison with previous studies (e.g., Seielstad et al. 1998). No further assumptions regarding the underlying population genetic model were applied in its interpretation, other than a general monotonic relationship between F_{ST} and genetic differences. Sampling distributions of h and F_{ST} were obtained by bootstrap resampling of individuals, according to observed haplotype frequencies in each population sample. This method of bootstrapping is designed to estimate sampling variance only, and it does not address whether differences in h or F_{ST} are due to drift or some other cause. The latter issue was, instead, addressed by looking for consistent differences between groups (i.e., between Jewish and non-Jewish–host pairs).

Tests for differences in h and F_{ST} were performed using conservative double testing under two procedures. In the first procedure, the difference in bootstrap sampling distributions was found, and the smaller area of the resulting distribution, lying to one side of zero, was doubled to yield a two-tailed P value. However, this procedure may lead to inflated type I error rates due to the bias in the bootstrap sampling distributions. This bias arises from the inherent drawback of bootstrapping, namely that sampling distributions are estimated by resampling an observed sample, rather than resampling from the original population. Because the observed sample is expected to have less genetic diversity than the original population (Nei 1987), bootstrap distributions of both h and F_{ST} are biased, and the correct distributions are not directly recoverable. We therefore also employed

Table 1

Y Chromosome Diversity within 18 Jewish and Non-Jewish Populations

Jewish Data Set			Host and Non-Jewish Data Set			
Population[a]	*n*	*h*[b] (SE)	Population[c]	*n*	*h*[b] (SE)	Difference[d]
AshJ	78	.973 (.0091)	Ger	88	.968 (.0080)	ns
MorJ	176	.980 (.0034)	Ber	60	.876 (.0372)	s (other)
IrqJ	57	.969 (.0103)	Syr	72	.979 (.0091)	ns
IrnJ	77	.891 (.0297)	…	…	…	…
GeoJ	70	.940 (.0126)	Geo	106	.975 (.0075)	s
BkhJ	36	.933 (.0249)	Ubk	40	.990 (.0091)	s
YemJ	66	.970 (.0094)	Yem	62	.977 (.0074)	ns
EthJ	51	.946 (.0173)	Eth	193	.972 (.0040)	ns
IndJ	76	.948 (.0099)	Hin	41	.978 (.0123)	ns
…	…	…	IsrA/Pal	143	.970 (.0079)	…
Mean		.950			.965	

[a] AshJ = Askenazic Jews; MorJ = Moroccan Jews; IrqJ = Iraqi Jews; IrnJ = Iranian Jews; GeoJ = Georgian Jews; BkhJ = Bukharan Jews; YemJ = Yemeni Jews; EthJ = Ethiopian Jews (Beta Israel); and IndJ = Indian Jews (Bene Israel).

[b] h = unbiased gene diversity (Nei 1987) based on unique event polymorphism+ microsatellite haplotypes; 462 haplotypes were seen in the combined data set. Standard errors (in parentheses) are based on bootstrap resampling according to observed population frequencies.

[c] Host and non-Jewish data sets are paired with Jewish data sets on same row. Ger = Germans; Ber = Berbers; Syr = Syrians; Geo = Georgians; Ubk = Uzbekistanis; Yem = Yemenis; Eth = Ethiopians; Hin = Indian Hindus; and IsrA/Pal = Israeli Arabs/Palestinians.

[d] s = the difference in h between Jewish and host data sets is significant ($P < .05$, on the basis of estimated bootstrap variances); ns = not significant; s (other) = Moroccan Jews are significantly more diverse than Berbers.

a second procedure in which a standard two-tailed z test was performed on the observed difference in unbiased h or F_{ST} estimates, and the bootstrap distribution of this difference was used only to provide an estimate of the sampling variance. Although this second procedure formally depends on assumptions of normality that are unlikely to hold, the properties of the bootstrap sampling distributions for h and F_{ST} are such that the real sampling distributions are likely to have less variance than the bootstrap ones, which therefore tends to make these z tests conservative (M.E.W., unpublished data). As a final conservative measure, only the larger of the two P values produced by the above two procedures was used to assess significance.

Results

mtDNA and Y Chromosome Diversity

Comparison of Y-chromosome and mtDNA patterns reveals a striking contrast between the maternal and paternal genetic heritage of Jewish populations. On the Y chromosome, there is no consistent pattern of lower diversity in Jewish communities when compared with their non-Jewish host populations; in two cases, diversity is significantly lower in the Jewish groups; in one case, it is higher; and, in the rest, differences are not significant (table 1). However, the pattern in the mtDNA is quite different. In each case, the Jewish community has a significantly lower mtDNA diversity than its paired host population (table 2). Indeed, every Jewish population has a lower mtDNA diversity than *any* non-Jewish population. This finding indicates that mistakes in associating particular host populations with Jewish populations would be very unlikely to affect our results. When Fisher's method for combining significance tests (Sokal and Rohlf 1995) is used, the overall significance level for the mtDNA pattern of lower Jewish diversity is $P = 10^{-15}$. When ratios of mtDNA to Y-chromosome diversity were calculated, to standardize the mtDNA results in relation to the other genetic system, the ratio for the Jewish data sets (mean 0.954) was again found, in all but one case (the Ethiopian Jews), to be less than the ratio for the non-Jewish host (mean 1.027).

Even more striking than this, however, is the high frequency of particular mtDNA haplotypes in the Jewish populations. No host population in our sample has an mtDNA modal frequency >12% (mean 7.7%) (table 3). In contrast, seven of the Jewish populations have a modal frequency >12% (mean 22.6%), and some of the Jewish groups have much higher frequencies. In particular, Moroccan Jews, the Bene Israel, and Georgian Jews have modal frequencies of 27.0%, 41.3%, and 51.4%, respectively, which are all higher than those observed in

Table 2

mtDNA Diversity within 18 Jewish and Non-Jewish Populations

Jewish Data Set			Host and Non-Jewish Data Set			
Population	*n*	*h* (SE)	Population	*n*	*h* (SE)	Difference
AshJ	78	.973 (.0069)	Ger	174	.988 (.0031)	s
MorJ	115	.917 (.0206)	Ber	60	.984 (.0093)	s
IrqJ	56	.936 (.0176)	Syr	69	.988 (.0072)	s
IrnJ	75	.969 (.0077)	…	…	…	…
GeoJ	70	.720 (.0542)	Geo	105	.984 (.0051)	s
BkhJ	33	.938 (.0220)	Ubk	37	.994 (.0080)	s
YemJ	65	.923 (.0165)	Yem	56	.988 (.0059)	s
EthJ	48	.971 (.0113)	Eth	74	.994 (.0033)	s
IndJ	75	.801 (.0423)	Hin	41	.994 (.0076)	s
…	…	…	IsrA/Pal	117	.994 (.0031)	
Mean		.906			.990	

Note.— A total of 615 haplotypes were seen in the combined data set. Standard errors (in parentheses) are based on bootstrap resampling according to observed population frequencies. Abbreviations are defined in table 1.

any of the other populations. Again, this pattern is not seen on the Y chromosome, where the modal frequencies in Jewish populations (mean 15.2%; range 7.4%–31.2%) are not significantly different from those seen in host populations (mean 13.6%; range 8.1%–33.3%, $P = .588$, using paired-sample *t* test). In most European and Near Eastern populations, the highest frequency mtDNA type is the HVS-1 Cambridge Reference Sequence (CRS). This type occurs at 16%, on average, in Europe, and at 6%, on average, in the Near East (Richards et al. 2000). This pattern is reflected in our data, in that all of the seven European and Near Eastern non-Jewish populations have the CRS as their modal haplotype. However, only two of the nine Jewish populations have the CRS as their modal haplotype, while, among the other seven, each has a different modal haplotype. Thus, among the nine Jewish groups there are eight different mtDNA types that are modal with an unusually high frequency. Apart from the CRS, none of the other Jewish modal haplotypes are represented in the Israeli Arab/Palestinian data set, in contrast to the similarities between Ashkenazic Jews, Sephardic Jews, Israeli Arabs/Palestinian, and Lebanese populations reported for the Y chromosome (Santachiara-Benerecetti et al. 1993; Nebel et al. 2000).

mtDNA and Y Chromosome Differentiation

We compared mean F_{ST} values for all Jewish populations among Jewish pairwise comparisons (36 pairs), all Jewish-host comparisons (8 pairs), and all host-host comparisons (28 pairs) (table 4). We found no significant differences among these categories when Y-chromosome data were used. Using mtDNA variation, however, we find that the Jewish-Jewish comparisons have significantly greater F_{ST} values than the host-host comparisons do, apparently because of the presence of different high-frequency haplotypes within most of the Jewish populations.

Not only does mtDNA show greater differentiation among the Jewish populations than among the hosts, but it shows significantly greater differentiation among the Jewish populations than does the Y chromosome (table 4). This means the lack of evidence for founder effects on the Y chromosome cannot be attributed to the obliteration of a founder signal due to a faster mutation rate, because, under this scenario, the Y chromosomes would be more, not less, differentiated among populations.

These results therefore suggest that an extreme founder effect has occurred in the maternal but not paternal genetic histories of most Jewish populations. Greater geographic structuring of the mtDNA than the Y chromosome is an unusual pattern (Seielstad 1998). To assess whether this is specific to the Jewish populations, we also compared mtDNA and Y-chromosome structuring among the host populations. Among the latter populations we found the more usual pattern of greater Y-chromosome differentiation (table 4). This demonstrates that the unusual pattern observed among the Jewish populations is not associated with the geographic areas from which they derive but rather with their unique demographic histories.

Origins of Modal mtDNA Types in Jewish Communities

Unfortunately, in many cases, it is not possible to infer the geographic origin of the founding mtDNAs within the different Jewish groups with any confidence. It would appear that the founder effects on the maternal side have been so severe that mtDNA frequencies in the Jewish populations are very different from those found in any non-Jewish population. The non-CRS modal haplotypes

Table 3

Frequently Encountered mtDNA Haplotypes

HSV-I Sequence (Sites 16090–16365)[a]	Frequency among[b]																	
	AshJ (78)	MorJ (115)	IrqJ (56)	IrnJ (75)	GeoJ (70)	BkhJ (33)	YemJ (65)	EthJ (48)	IndJ (75)	Ger (174)	Ber (60)	Syr (69)	Geo (105)	Ubk (37)	Yem (56)	Eth (74)	Hin (41)	IsrA (117)
10% frequency in ⩾1 population:																		
CRS	9.0	27.0	1.8	8.0	2.9	6.1	0	0	8.0	6.9	11.7	10.1	9.5	5.4	7.1	0		6.8
343	2.6	.9	17.9	0	0	0	0	0	0	1.1	1.7	0	0	0	0	0	0	0
126, 145, 222, 261	0	0	12.5	2.7	0	0	9.2	0	0	0	0	0	0	0	0	0	0	1.7
93, 126, 261	0	0	10.7	1.3	0	0	0	0	0	0	0	1.4	0	0	0	0	0	0
355	0	0	0	0	51.4	0	0	0	0	0	0	2.9	1.0	0	0	0	0	0
145	0	.9	0	0	10.0	0	0	0	0	0	0	0	0	0	0	0	0	0
126, 193, 278	0	0	0	1.3	0	15.2	0	0	0	.6	1.7	0	0	0	0	0	0	0
184, 265T	2.6	0	0	0	0	15.2	0	0	0	0	0	0	0	0	0	0	0	0
129, 223	2.6	0	1.8	1.3	0	12.1	0	0	0	1.1	0	0	0	0	0	0	2.4	0
274	1.3	0	0	0	0	0	20.0	0	0	.6	0	0	0	0	0	0	0	0
126, 304, 362	0	0	0	0	0	0	12.3	0	0	0	0	0	0	0	1.8	0	0	0
126, 305T, 362	0	0	0	0	0	0	0	12.5	0	0	0	0	0	0	0	1.4	0	0
166, 223, 311	0	0	0	0	0	0	0	0	41.3	0	0	0	0	0	0	0	0	0
223, 304	0	0	0	0	0	0	0	0	10.7	0	0	0	0	0	0	0	7.3	0
Modal or joint modal, <10% frequency:																		
183C, 189, 249	0	0	0	9.3	0	0	1.5	0	0	0	0	0	0	0	0	0	0	0
311	0	3.5	0	0	0	0	0	0	0	4.0	1.7	1.4	2.9	5.4	0	1.4		.9
223, 227, 278, 362	0	0	0	0	0	0	0	0	0	0	0	0	0	5.4	0	0	0	0
111, 129, 223, 257A, 261	0	0	0	0	0	0	0	0	0	0	0	0	0	5.4	0	0	0	0
278, 362	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.1	0	0
126, 355, 362	0	0	0	0	0	0	1.5	0	0	0	0	1.4	1.0	0	1.8	4.1	0	.9
126, 187, 189, 223, 264, 270, 278, 289, 293, 311	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.1	0	0

[a] HVS-1 nomenclature is described by Macaulay et al. (1999), except that 16,000 has been subtracted from all HVS-1 sequence mutation sites. CRS = matches Cambridge Reference Sequence (Anderson et al. 1981).

[b] Abbreviations are defined in footnote to table 1. Numbers in parentheses are sample sizes. Frequencies ⩾10% are underlined.

Table 4

Mean F_{ST} Values in mtDNA and Y Chromosome Data

Comparison (No. of Pairs)[a]	Mean Pairwise F_{ST} (SE) for mtDNA	Mean Pairwise F_{ST} (SE) for Y	Difference
J-J (36)	.090 (.0094)	.040 (.0053)	s
J-H (8)	.047 (.0056)	.032 (.0036)	s
H-H (28)	.005 (.0020)	.032 (.0054)	s
Difference	s	ns	

Note.— s = all relevant comparisons of mean are significant ($P < .05$); ns = not significant.

[a] J-J = comparison among the nine Jewish data sets; J-H = comparison of Jewish and host data sets; and H-H = comparison among the eight host data sets.

in the Jewish populations are generally rare in the non-Jewish populations. The CRS, on the other hand, is too ubiquitous to allow it to be pinpointed to anything other than a general Eurasian origin.

For example, the most extreme founder effect is seen in the Georgian Jews, of whom 51% possess the same haplotype. The Georgian Jewish modal type is matched by a single individual in the Georgian sample. However, a search of the mtDNA database (Richards et al. 2000) shows that it also occurs in Syria (2/69 individuals) and Iraq (1/116). One directly derived type is present in two Georgians, but derived types are also found in the North Caucasus (2/208 individuals), Turkey (1/218), Armenia (1/191), and Sicily (1/90). For the Georgian modal haplotype, there is therefore no clear indication of provenance, although an indigenous origin is certainly possible, given the data.

In two cases, however, comparison with the published data does provide some indication of the possible geographic origins of the modal types. The modal type in the Bene Israel is a one-step mutational neighbor of a haplotype present in the Indian sample, as well as being a one-step neighbor of a type previously identified in India (Kivisild et al. 1999*a*, 1999*b*). Similarly, the commonest type in the Ethiopian Jewish sample is also present in the non-Jewish Ethiopian sample and occurs in the worldwide mtDNA database only in Somalia (Watson et al. 1997). Other high-frequency haplotypes in the Ethiopian Jewish sample are also found almost entirely in Africa (data not shown). The lack of an indication of a Middle Eastern origin for these haplotypes, on the basis of the Richards database, makes local recruitment a more reasonable explanation in these two cases.

Discussion

The greatly reduced mtDNA diversity in the Jewish populations in comparison with the host populations, together with the wide range of different modal haplotypes found in different communities, indicates female-specific founding events in the Jewish populations. Although we cannot be certain whether this occurred immediately after the establishment of the communities or over a longer period of time, a simple explanation for the exceptional pattern of mtDNA variation across Jewish populations is that each of the different Jewish communities is composed of descendants of a small group of maternal founders. After the establishment of these communities, inward gene flow from the host populations must have been very limited.

The differences among the Jewish populations in mtDNA haplogroup frequencies indicates that the Jewish groups formed independently around (at least) eight small, distinct nuclei of women. The severity of these demographic events was sufficiently great to drive an unusual pattern of geographic variation among the Jewish populations. Although it has been commonly found that Y-chromosome variation shows greater geographic structure than the mtDNA (e.g., Seielstad et al 1998; Wilson et al., 2001), this pattern is reversed in the Jewish populations, which show greater differentiation for the mtDNA than for the Y chromosome. Jewish populations therefore appear to represent an example in which cultural practice—in this case, female-defined ethnicity—has had a pronounced effect on patterns of genetic variation. An analogous example has recently been reported by Oota et al. (2001), who showed that matrilocal communities in northern Thailand have patterns of mtDNA and Y chromosome variation distinct from nearby patrilocal communities.

The pattern in Ashkenazic Jews is of particular interest. Despite the common opinion that this population has undergone a strong founder event, it has a modal haplotype with a frequency similar to that of its host

Table 5

Comparison of X Chromosome LD among Georgian Jews and non-Jewish British Men

Additional Condition[a]	No. of Marker Pairs with a Distance (in mB) of ⩽.5	⩽1	⩽2	⩽5	⩽10
Overall No.	9	13	30	50	87
Georgian Jews:					
$P \leqslant .05$	8	10	16	20	27
$P \leqslant .01$	2	4	8	10	12
$P \leqslant .001$	2	2	4	5	5
British men:					
$P \leqslant .05$	3	3	4	4	4
$P \leqslant .01$	1	1	1	1	1
$P \leqslant .001$	1	1	1	1	1

[a] Overall No. = number of pairwise combinations among 34 X-chromosome microsatellite markers that satisfy the distance condition. An additional condition was applied according to the statistical signficance of association within each population. *P* values were calculated from an extension of Fisher's exact test.

population (9.0% vs. 6.9%), providing little evidence of a strong founder event on the female side. The possibility remains, however, that present-day Ashkenazic Jews may represent a mosaic group that is descended on the maternal side from several independent founding events.

Kruglyak (1999) showed that human populations are not expected to maintain LD over long genetic distances, unless they have experienced sharp founder effects or bottlenecks in the recent past. Reich et al. (2001) have recently demonstrated relatively long-range LD in northern European populations, possibly as a result of a strong bottleneck associated with either the Last Glacial Maximum or the initial Out-of-Africa movement into Eurasia. Our results show that, in some populations, such effects may be specific to one or other of the sexes. In the more extreme cases, these founder effects may have been sufficiently severe to affect the epidemiological properties of the populations in question (Wright et al 1999). To illustrate this point, we present (table 5) a comparison of levels of LD (as measured by the number of marker pairs in significant association) in Georgian Jews (the Jewish population showing the greatest reduction in mtDNA diversity) and a British sample, using a panel of 34 microsatellite markers spanning 150 Mb of the X chromosome. Even when compared with a northern European sample, in which high levels of long-range LD might be expected (Reich et al. 2001), the Georgian Jewish sample shows consistently elevated levels of long-range LD. The same pattern is found when the Georgian Jews are compared with Norwegian and with Turkish samples (data not shown). These results demonstrate that demographic events restricted to only one of the sexes can be of considerable epidemiological significance. The LD in the Georgian Jews also indicates that some Eurasian populations may be of particular use in population-based gene mapping but that it is hard to know which ones without detailed genetic analysis.

Acknowledgments

We wish to express our gratitude to the individuals who very kindly provided samples for this study and to Telethon-Italy (grant E.0890) and the Istituto Pasteur Fondazione Cenci Bolognetti, for their kind support, to Michael Stumpf, for sample collection, and to Karl Skorecki, for valuable discussion. D.B.G. holds a Royal Society/Wolfson Research Merit Award.

References

Alves-Silva J, da Silva Santos M, Guimarães PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Ferreira Prado V (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

Beinart H (1992) Atlas of medieval Jewish history. Simon & Schuster, New York

Ben-Sasson HH (1976) The Middle Ages. In: Ben-Sasson HH (ed) A history of the Jewish people. Harvard University Press, Cambridge, MA, pp 570–573

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 38:341–352

Bonné-Tamir B, Zoossmann-Diskin A, Ticher A (1992) Genetic diversity among Jews reexamined: preliminary analyses at the DNA level. In: Bonné-Tamir B, Adam A (eds) Genetic diversity among Jews: diseases and markers at the DNA level. Oxford University Press, Oxford, pp 80–94

Chen YS, Huoponen K, Torroni A, Gericke GS, Olckers A, Wallace DC (1995*a*) Mitochondrial DNA variation in African Bushman. Am J Hum Genet Suppl 57:A161

Chen YS, 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

de Lange N (1984) Atlas of the Jewish world. Phaidon Press, Oxford

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

Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bosch E, Martínez-Arias R, Clarimon J, Facchini F, Fiori G, Luiselli D, Pettener D, Bertranpetit J (1998) Trading genes along the Silk Road: mtDNA sequences and the origin of central Asian populations. Am J Hum Genet 63:1824–1838

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

Encyclopaedia Judaica (1972). Keter Publishing, Jerusalem

Hammer MF, Redd AJ, Wood ET, Bonner MR, Jarjanazi H, Karafet T, Santachiara-Benerecetti S, Oppenheim A, Jobling MA, Jenkins T, Bonné-Tamir B (2000) Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. Proc Natl Acad Sci USA 97:6769–6774

Hofmann S, Jaksch M, Bezold R, Mertens S, Aholt S, Paprotta A, Gerbitz KD (1997) Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlations with D loop variants and association with disease. Hum Mol Genet 6: 1835–1846

Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, Parik J, Watkins WS, Dixon ME, Papiha SS, Mastana SS, Mir MR, Ferak V, Villems R (1999*a*) Deep common ancestry of Indian and western Eurasian mtDNA lineages. Curr Biol 9:1331–1334

Kivisild T, Kaldma K, Metspalu E, Parik J, Papiha S, Villems R (1999*b*) 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

Kolman CJ, Sambuughin N, Bermingham E (1996) Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 142:1321–1334

Krings M, Salem AH, Bauer K, Geisert H, Malek A, 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

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139–144

Livshits G, Sokal RR, Kobyliansky E (1991) Genetic affinities of Jewish populations. Am J Hum Genet 49:131–146

Lucotte G, Smets P (1999) Origins of Falasha Jews studied by haplotypes of the Y chromosome. Hum Biol 71:989–993

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, Perez-Lezaun A, Abade A, Bertranpetit J (1997) A tale of two islands: population history and mitochondrial sequence variation of Bioko and Sao Tome, 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, Weiss DA, Weale M, Faerman M, Oppenheimer A, Thomas MG (2000) High-resolution Y chromosome haplotypes of Israeli and Palestinian Arabs reveal geographic substructure and substantial overlap with haplotypes of Jews. Hum Genet 107:630–641

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

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

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, Cabrera VM, Larruga JM, Hernandez M, Gonzalez AM, Pinto F, Bandelt H-J (1999) Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. Ann Hum Genet 63:413–428

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

Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES (2001) Linkage disequilibrium in the human genome. Nature 411:199–204

Reif SC (1993) Judaism and Hebrew prayer. Cambridge University Press, Cambridge

Rejwan N (1985) The Jews of Iraq. Westview Press, Boulder, CO

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779

Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt H-J, Sykes B (1996) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203

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

Ritte U, Neufeld E, Bonné-Tamir B (1992) Types of mitochondrial DNA among Jews. In: Bonné-Tamir B, Adam A (eds) Genetic diversity among Jews: diseases and markers at the DNA level. Oxford University Press, Oxford, pp 51–59

Ritte U, Neufeld E, Broit M, Shavit D, Motro U (1993*a*) The differences among Jewish communities—maternal and paternal contributions. J Mol Evol 37:435–440

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

Santachiara-Benerecetti AS, Semino O, Passarino G, Torroni A, Brdicka R, Fellous M, Modiano G (1993) The common, near-Eastern origin of Ashkenazi and Sephardi Jews supported by Y-chromosome similarity. Ann Hum Genet 57: 55–64

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

Sokal RR, Rohlf FJ (1995) Biometry, 3d ed. WH Freeman, New York

Thomas MG, Bradman N, Finn HM (1999) High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome. Hum Genet 105:577–581

Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Goldstein DB (1998) Origins of Old Testament priests. Nature 394:138–140

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

Torroni A, Bandelt H-J, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savantaus M-L, Bonné-Tamir B, Scozzari R (1998) mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62:1137–1152

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 popu-

lation and its adaptation to high altitude. Am J Phys Anthropol 93:189–199

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

Wilson JF, Goldstein DB (2000) Consistent long-range linkage disequilibrium generated by admixture in a Bantu-Semitic hybrid population. Am J Hum Genet 67:926–935

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 98:5079–5083

Wright AF, Carothers AD, Pirastu M (1999) Population choice in mapping genes for complex diseases. Nat Genet 23:397–404

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