Multiple Origins of Ashkenazi Levites: Y Chromosome Evidence for Both Near Eastern and European Ancestries

Doron M. Behar,[1,*] Mark G. Thomas,[2,*] Karl Skorecki,[1] Michael F. Hammer,[4] Ekaterina Bulygina,[2] Dror Rosengarten,[1] Abigail L. Jones,[2] Karen Held,[2] Vivian Moses,[2] David Goldstein,[3] Neil Bradman,[2] and Michael E. Weale[2]

[1]Bruce Rappaport Faculty of Medicine and Research Institute, Technion and Rambam Medical Center, Haifa, Israel; [2]The Centre for Genetic Anthropology and [3]The Centre for Population Genetics and Human Health, Department of Biology, University College London, London; and [4]Division of Biotechnology, University of Arizona, Tucson, AZ

Previous Y chromosome studies have shown that the Cohanim, a paternally inherited Jewish priestly caste, predominantly share a recent common ancestry irrespective of the geographically defined post-Diaspora community to which they belong, a finding consistent with common Jewish origins in the Near East. In contrast, the Levites, another paternally inherited Jewish caste, display evidence for multiple recent origins, with Ashkenazi Levites having a high frequency of a distinctive, non–Near Eastern haplogroup. Here, we show that the Ashkenazi Levite microsatellite haplotypes within this haplogroup are extremely tightly clustered, with an inferred common ancestor within the past 2,000 years. Comparisons with other Jewish and non-Jewish groups suggest that a founding event, probably involving one or very few European men occurring at a time close to the initial formation and settlement of the Ashkenazi community, is the most likely explanation for the presence of this distinctive haplogroup found today in >50% of Ashkenazi Levites.

Introduction

Jewish identity, since at least Talmudic times (~100 B.C.E.–500 C.E.), has been acquired either by maternal descent from a Jewish woman or by rabbinically authorized conversion. Only in recent years have some strands of the Jewish religion accepted paternal descent as a qualifying criterion. Within the Jewish community, however, membership in the three male castes (Cohen, Levi, and Israelite) is determined by paternal descent. Cohanim (plural of "Cohen," the Hebrew word for priest) are, in Biblical tradition, the descendants of Aaron the brother of Moses; Levites are, in that tradition, considered to be those male descendants of Levi, the third son of the patriarch Jacob and paternal ancestor of Aaron, who are not Cohanim. The Cohanim have both rights and duties in religious law, as well as being subject to restrictions that do not apply to the other castes. They are, for example, called first to the reading of the Torah in synagogue and forbidden entry into a cemetery. Levites have some rights similar to those

of Cohanim (for example, exemption from payment of a special tax on the birth of a first-born male) but are not subject to the particular restrictions placed on the Cohanim. Strict adherence to the qualifying rules would mean that the male descendants of men who were not Jews at birth could be Israelites but not Cohanim or Levites (Encyclopaedia Judaica 1972). It is estimated that Cohanim and Levites each comprise ~4% of the Jewish people (Bradman et al. 1999).

In addition to classification by caste, Jews, on the basis of their ancestry and religious practice, can be assigned to one or other of a few long-standing, geographically separated Jewish communities, the most numerous of which are the Ashkenazi and Sephardi groupings (Reif 1993). The term "Ashkenaz" describes a relatively compact area of Jewish settlement in northwestern Europe, including northeastern France and northern Germany, where Jewish settlement is documented dating back to at least the 6th century C.E. From the 10th century, Ashkenazi Jews spoke a common language (Yiddish), written with Hebrew characters but borrowing its lexicon mostly from German. By the 16th century, Jews speaking this language and following the Ashkenazi religious rite and cultural tradition populated communities extending from the Loire in the west to the Dnieper in the east and from Rome in the south to the Danish border in the north. During the past 500 years, there has been rapid population growth, culminating in an estimated population size of ~8 million Ashkenazi Jews just prior to the outbreak of World War

II. There is uncertainty concerning the relative contributions to Ashkenazi Jewry of, on one hand, western versus eastern immigration of Jews and, on the other hand, internally generated population growth versus conversion to Judaism. In particular, it has been suggested that subjects of the Khazar Empire (located to the northeast of the Black Sea), who had adopted Judaism in the last quarter of the first millennium C.E., were an important constituent of the nascent Ashkenazi community (Encyclopaedia Judaica 1972).

The term "Sephardi" originally described Jews descended from the communities that existed in Spain prior to the expulsion in 1492 C.E. However, current usage applies this designation to all descendants of the communities of North Africa and the Near East who follow the Sephardi rite of worship and cultural traditions. It is thought that, prior to the middle of the 20th century, gene flow between the Ashkenazi and non-Ashkenazi groups was relatively restricted.

The purported different modes of transmission of Levite and Cohen versus Israelite status provide a priori expectations about patterns of genetic variation on the paternally inherited nonrecombining region of the Y chromosome (NRY). In particular, because of recent shared ancestry, Cohanim and Levites would be expected to display lower gene diversity of NRY haplotypes than would Israelites. In addition, the distribution of haplotype frequencies, in the absence of drift, should be similar (*a*) in Cohanim and (*b*) in Levites across the Ashkenazi-Sephardi division, given that this division occurred after the founding of these two groups.

In fact, previous studies have indeed shown that the NRYs of Ashkenazi and Sephardi Cohanim are genetically more closely related to each other than they are to the NRYs of Israelites or non-Jews (Skorecki et al. 1997; Thomas et al. 1998). This pattern arises primarily from differences in the frequency of a particular NRY haplotype (the Cohen Modal Haplotype [CMH], defined by six rapidly mutating microsatellites [Thomas et al. 1998]), and a cluster of closely related haplotypes within a single haplogroup (defined by slowly mutating unique event polymorphisms [UEPs]). Chromosomes belonging to this haplotype and its related cluster were found at high frequency among Cohanim but at a much lower frequency among Israelites. Furthermore, the pattern of diversity within the cluster was found to be consistent with descent from a common ancestor who lived between 2,100 and 3,900 years ago. The CMH is also found, at lower frequency, in non-Jewish populations in the Near East, which would be consistent with its origin in this geographic region. However, the same studies found high frequencies of multiple haplogroups in the Levites, indicating that no single recent origin could be inferred for the majority of this group, despite an oral tradition of a patrilineal descent similar to that

of the Cohanim. Moreover, a cluster of closely related NRY haplotypes was identified within a distinctive deep-rooting NRY clade that was found at much higher frequency among Ashkenazi Levites than in either Sephardi Levites or any other Jewish group. However, the reasons for this difference in the Ashkenazi Levites were not explored.

Given the importance of the paternally defined Levite caste in Jewish history and tradition, the multiple theories of the ethnogenesis of the Ashkenazi Jewish community, and a suggestion that Yiddish is a relexified Slavic tongue (Wexler 1993), we undertook a detailed investigation of the paternal genetic history of Ashkenazi Levites and compared the results with matching data from neighboring populations among which the Ashkenazi community lived during its formation and subsequent demographic expansion. By analyzing NRY haplotypes, we have revealed (*a*) a plausible historical explanation for the multiple paternal histories of the Levite caste, (*b*) the probable period during which a European introgression into the caste took place, and (*c*) the likely demographic scale of the event.

Subjects and Methods

Sampling

We analyzed NRY variation in a set of 988 unrelated males from Ashkenazi Jews, Sephardi Jews, and four non-Jewish European populations. Populations were categorized using a combination of geographic, religious, and ethnohistorical criteria and individual affiliation within a category was determined according to self-designation by the participant providing the sample. Three of the four non-Jewish groups (88 Germans, 112 Sorbians, and 306 Belarusians) were chosen because of their geographic location relative to the ancestral European communities of the Ashkenazi Jews, and also because of prior knowledge of the geographic areas known to have high frequencies of the NRY haplogroup within which the distinctive Ashkenazi Levite high-frequency cluster was previously reported. This haplogroup is designated "R1a1" according to the Y Chromosome Consortium (2002). An additional non-Jewish group consisting of 83 samples previously collected in Norway (Weale et al. 2002) was chosen, to represent a geographic location that excluded Jewish entry until the middle of the 19th century and is known to be outside the area in which Ashkenazi communities originated.

The Jewish samples comprised 236 Ashkenazi Jews (AJ) and 163 Sephardi Jews, who were further divided into 100 Ashkenazi Israelites (AI), 76 Ashkenazi Cohanim (AC), 60 Ashkenazi Levites (AL), 63 Sephardi Israelites (SI), 69 Sephardi Cohanim (SC), and 31 Sephardi Levites (SL). The 60 AL samples were collected

from Ashkenazi males who identified themselves as Levites, with a paternal ancestry from one of the following nine Ashkenazi Jewish communities; Austria-Hungary (10), Belarus (4), France (6), Germany (10), Lithuania (8), Netherlands (5), Poland (7), Romania (4), and Russia (6). Current political borders (including the current borders of Austria and Hungary) were used to define geographic origin. When the Ashkenazi Levite sample was split into western (France, Netherlands, and Germany) versus eastern (all others) communities, no significant differences were found in haplogroup or haplotype frequencies (using the exact test of Raymond and Rousset [1995]), but we note that the sample size is too small to have power to detect anything other than very large differences in frequency. Information regarding paternal lineage was collected from each of the participants. Paternal population affiliation was determined by the oldest male-line ancestor that the participants identified, which extended back to at least the grandparental level. The participants were not related through their paternal lineages. The ethics review committees of the participating institutions approved the sampling protocols.

Genotyping

Haplogroups.—In all population samples except the Ashkenazi Levites and Germans, we typed a set of 12 biallelic markers including one *Alu* insertion and 11 SNPs to investigate the data reported herein. The 12 markers were: 92R7, M9, M13, M17, M20, SRY_{+465}, SRY_{4064}, SRY_{10831}, sY81, Tat, YAP, and p12f2. PCR protocols for detection of these polymorphisms have been reported elsewhere (Thomas et al. 1999; Rosser et al. 2000). These 12 markers give rise to the following 10 observed haplogroups, labeled according to the Y Chromosome Consortium (2002) nomenclature: Y*(xBR,A3b2), BR*(xDE,JR), E*(xE3a), J, K*(xL,N3,O2b,P), L, N3, P*(xR1a), R1a*, and R1a1. We note that one of these is a very rare haplogroup (R1a*) that has previously been observed only in two Armenian men (Weale et al. 2001). We now report the same haplogroup in one additional Belarusian man. The Ashkenazi Levites were typed for a set of 25 biallelic markers including P9, M216, M217, YAP, P2, M35, P14, M201, P15, M52, P19, 12f2b, M172, M9, M20, LLY22g, Tat, M175, P27, P36, M207, M173, SRY_{10831}, M17, and P25. PCR protocols for detection of these polymorphisms have been reported elsewhere (Hammer et al. 2001; Underhill et al. 2001). This set of biallelic markers allowed the designation of the Ashkenazi Levite chromosomes into the same 10 observed haplogroups. The exact labeling of the Ashkenazi Levite haplogroups, by use of all 25 biallelic markers, is reported in table A (online only). The German data are reported elsewhere (Capelli et al. 2003) and were not typed for

M13, M20, SRY_{10831}, and SRY_{+465} as these markers distinguish haplogroups (taking other typed markers into account) that are extremely rare in Europe and the Near East (Underhill et al. 2000). In place of YAP, SRY_{4064}, and sY81, the German sample was typed for M35, as intermediate haplogroups (derived for YAP, ancestral for M35) are, again, very rare in this part of the world.

Haplotypes.—We genotyped all samples for six STRs, DYS19 (equivalent to DYS394), DYS388, DYS390, DYS391, DYS392, and DYS393, according to protocols previously reported by Thomas et al. (1999). In the case of the Ashkenazi Levite samples, we also typed six additional STRs (DYS385a, DYS385b, DYS389I, DYS389II, DYS426, and DYS439), to allow more accurate estimation of the time to most recent common ancestor (TMRCA). PCR primers and conditions for the typing of the additional six STRs are as described by Redd et al. (2002), multiplexes I and II. PCR products were analyzed on ABI310 and ABI3100 DNA Analyzers, and fragment lengths were converted to repeat numbers by the use of allelic ladders. We define DYS389CD as equivalent to DYS389I, and we define DYS389AB as equivalent to DYS389II minus DYS389I (Rolf et al. 1998).

Statistical Analysis

Estimation of Time to Most Recent Common Ancestor.—If the root haplotype is known, central estimates can be obtained of the TMRCA that are independent of population demography and shape of the genealogy and are dependent only on the mutation model assumed (Stumpf and Goldstein 2001). We obtained central estimates of the TMRCA under two models of microsatellite mutation. Under the simple stepwise mutation model (S-SMM), mutation rate is independent of microsatellite repeat number and mutations are equally likely to result in an increase or decrease in length. It can be shown that $\hat{T} = \bar{\Delta}/\bar{\mu}$ is an unbiased estimator of the TMRCA (in generations), where $\bar{\Delta}$ is the average squared difference in repeat length between each sampled NRY and the root haplotype, averaged over loci, and $\bar{\mu}$ is the S-SMM mutation rate (per generation) averaged over loci (Goldstein et al. 1995*a*, 1995*b*; Slatkin 1995). Locus-specific estimates of mutation rates are currently hampered by lack of data, so we chose to estimate a common mutation rate by combining data on all tri- and tetra-nucleotide NRY microsatellites for which direct data (including mean repeat sizes) are available. Combining the pedigree-based data of Bianchi et al. (1998), Forster et al. (1998), Heyer et al. (1997), and Kayser et al. (2000) leads to an estimate of $\bar{\mu} = 18/9347 = 0.00193$. A more recent pedigree-based study estimated a higher mutation rate of 0.0042 for the same 12 STRs as were used for the Ashkenazi Levites

(Bonne-Tamir et al. 2003), as did a study based on mutational analysis of single sperm cells (Holtkemper et al. 2001). In contrast, some other studies based on dating deep-rooting Y chromosome clades have suggested lower "evolutionary" mutation rates (Caglia et al. 1997; Forster et al. 2000). These considerations introduce additional uncertainty in the true value of $\bar{\mu}$ for our loci (Zhivotovsky et al. 2001). We allowed for some uncertainty in $\bar{\mu}$ when calculating confidence limits for the TMRCA, and we note that the length-dependent model described below may also accommodate some locus-specific variation in mutation rate (Carvalho-Silva et al. 1999). Two microsatellite loci typed in our study, DYS385a and DYS385b, are homologous, so that reported repeat lengths cannot be assigned unambiguously to this locus. This problem is overcome by adding the two repeat lengths together and treating this as a single "locus" with a mutation rate of 2μ.

The second microsatellite mutation model we used was a linear length dependent stepwise mutation model (LLD-SMM). Under this model, motivated by a proportional slippage mutation mechanism, the mutation rate is determined by the linear equation $\mu = a + bL$, where L is the current repeat size. Length-dependency has been proposed as one possible explanation for large inter-locus variability in Y chromosome microsatellites (Forster et al. 2000). It can be shown that $\hat{T} = \bar{\Delta}/\bar{\mu}^0$ is an unbiased estimator of the TMRCA, where $\bar{\Delta}$ is as defined above and $\bar{\mu}^0$ is the mutation rate for the root haplotype, averaged over loci (Calabrese et al. 2001). Again, due to limited data we assume common values for a and b over all loci, and use the estimates reported by Stumpf and Goldstein (2001), anchored by assuming $\mu = 0.00193$ when $L = 15.77$ (from the direct data, described above, of Bianchi et al. [1998], Forster et al. [1998], Heyer et al. [1997], and Kayser et al. [2000]). This results in estimates of $a = -4.03 \times 10^{-3}$ and $b = 3.78 \times 10^{-4}$. The two homologous loci DYS385a and DYS385b were accommodated by adding their repeat lengths together into a single value of L and treating them as a single "locus" with $\mu = 2a + bL$. If negative values of μ were produced, these were treated as zero.

Unlike the central estimates, CIs on the TMRCA are dependent on population demography, regardless of the mutation model used. These were found by Monte Carlo simulation, using a coalescent-with-exponential-growth model scaled in units of the TMRCA, incorporating uncertainty in mutation rate, and implemented in the program Ytime.

Estimation of genetic diversity and distance. — Genetic diversity was measured using Nei's unbiased h statistic (Nei 1987). Tests for population differentiation were performed using the exact test of Raymond and Rousset (1995). Two common measures of genetic distance are F_{ST} (Reynolds et al. 1983) and $1-I$, where I is Nei's

Genetic Identity. We found a poor correlation between F_{ST} and $1-I$ for the population samples analyzed in this study ($r = 0.383$ among the 10 population data sets using haplotype frequencies and using all 45 pairwise combinations). We propose that $1-I$ is a better measure of genetic distance (in the sense of a metric summarizing useful aspects of interpopulation differences in gene frequency), for comparisons among populations where recent admixture is a likely demographic model. This is because if two different populations A and B admix to form a population C, the expected immediate I value between C and either A or B is dependent only on the admixture proportion between A and B, the I value between A and B, and the relative gene diversities (via the ratio $(1 - h_A)/(1 - h_B)$), whereas the expected F_{ST} value is dependent not only on these factors but also on the absolute diversities of A and B. This results in an unwanted dependency of F_{ST} on the absolute gene diversities of A and B (F_{ST} is larger when the gene diversities are smaller, because F_{ST} is a measure of between-population diversity divided by the sum of between- and within-population diversities). This unwanted dependency is avoided by the use of $1-I$. We note that the same arguments apply to Nei's Genetic Distance, $D = -\log(I)$, as to $1-I$, since there is a monotonic relationship between the two. We found the same conclusions when we analyzed our data using D as we did using $1-I$.

Visualization of genetic-distance patterns. — We visualized patterns of genetic differentiation using Principal Coordinates Analysis, performed on I to form the similarity matrix. Values along the main diagonal of the similarity matrix, representing the similarity of each population sample to itself, were calculated from the estimated genetic distance between two copies of the same population sample (for I values, the resulting self-similarity values simplify to $n/(n - 1)$, where n is the sample size).

Results

The haplogroup frequencies found in the six Jewish and four non-Jewish data sets are shown in table 1. The table confirms the presence of R1a1 as the modal haplogroup in the Ashkenazi Levite Jews (52% of chromosomes). This haplogroup is found at similarly high frequencies in the two Slavonic-speaking populations (Sorbians and Belarusians), but at a maximum frequency of only 5.8% among the other five Jewish data sets (mean frequency 3.2%). Pairwise genetic similarities (Nei's genetic identity I) and P values for population differentiation at the haplogroup level are shown in table 2, and figure 1 displays the principal coordinates plot derived from the genetic similarity values. Figure 1 reflects the unusual distribution of Ashkenazi Levite haplogroup frequencies.

Table 1

Haplogroup Frequencies across Six Jewish and Four Non-Jewish Sample Sets

	FREQUENCY IN SAMPLE SET[a]									
HAPLOGROUP	AC (n = 76)	SC (n = 69)	AL (n = 60)	SL (n = 31)	AI (n = 100)	SI (n = 63)	Ger (n = 88)	Nor (n = 83)	Sorb (n = 112)	Bel (n = 306)
Y*(xBR,A3b2)	…	…	…	…	…	.0159	…	…	…	…
BR*(xDE,JR)	.0263	.0290	.0333	.1935	.1400	.0794	.3295	.4458	.1786	.2582
E*(xE3a)	.0395	.0435	.2000	.0968	.2200	.1905	.0341	.0120	.0625	.0458
J	.8684	.7536	.1000	.3226	.3700	.3651	.0341	.0241	.0268	.0327
K*(xL,N3,O2b,P)	.0263	.0290	.0333	.2258	.0800	.1270	…	…	…	.0033
L	…	…	…	.0323	…	…	…	…	…	…
N3	…	…	…	…	…	…	.0114	.0361	.0089	.1046
P*(xR1a)	.0263	.0870	.1167	.0968	.1500	.2063	.4659	.2651	.0893	.0425
R1a*	…	…	…	…	…	…	…	…	…	.0033
R1a1	.0132	.0580	.5167	.0323	.0400	.0159	.1250	.2169	.6339	.5098

[a] Population sample labels: AC = Ashkenazi Cohen; SC = Sephardi Cohen; AL = Ashkenazi Levite; SL = Sephardi Levite; AI = Askenazi Israelite; SI = Sephardi Israelite; Ger = German; Sorb = Sorb; Bel = Belarusian; Nor = Norwegian.

The other Jewish data sets separate according to whether they are Cohanim or non-Cohanim, and, within these two groups, there are no significant differences between Ashkenazi and Sephardi data sets (all tests for population differentiation performed using the exact test of Raymond and Rousset (1995). The genetic similarities among these five Jewish data sets range from 0.79 to 1.0. In contrast, the Ashkenazi Levites cluster more with the Slavonic data sets than they do with the other Jewish data sets. The genetic similarities with the other Jewish data sets range from 0.22 to 0.47, whereas the *I* values with the Sorbian and Belarusian data sets are 0.95 and 0.88, respectively. When a bootstrap test is used, the *I* value for Ashkenazi Levites with Sorbians and Belarusians is, in both cases, significantly higher than the *I* value for Ashkenazi Levites with Sephardi Israelites, the most similar Jewish data set to the Ashkenazi Levites (*P* = .004 for Sorbians and .008 for Belarusians).

At the haplotype level (combining UEP information with the six microsatellite loci), significant differences in haplotype frequency emerge between all data sets (table 3), but the same underlying patterns of similarity remain (fig. 2). Full haplotype information for all data sets is available in supplemental tables A and B (online only). The genetic similarities between the Ashkenazi Levites and the other Jewish data sets range from 0.04 to 0.18, whereas the *I* values with the Sorbian and Belarusian data sets are 0.52 and 0.49 respectively. Using a bootstrap test, the *I* value for Ashkenazi Levites with Sorbians and Belarusians is again significantly higher in both cases than the corresponding *I* value for Ashkenazi Levites with their most similar Jewish group, the Ashkenazi Israelites (*P* = .006 for Sorbians and .004 for Belarusians).

Within haplogroup R1a1, the microsatellite haplotypes found in the AL data set are tightly clustered around a modal haplotype (16-12-25-10-11-13) that comprises 74% of Ashkenazi Levites within this haplogroup, and 38% of Ashkenazi Levites overall (table 4). This modal haplotype is evenly distributed across the geographically defined communities from which the Ashkenazi Levite sample was taken (see the "Subjects and Methods" section and table A), so that clues to its origin could not be found from these data. The very high frequency of this modal haplotype makes the genetic diversity of Ashkenazi Levite NRYs within the R1a1 haplogroup much lower than in the non-Jewish comparative data sets in which R1a1 is found at high frequency (table 5). Bootstrap tests indicate that, within R1a1, *h* is significantly lower in AL than in the Sorbian, Belarusian, and Norwegian data sets (*P* < .001 in each case). The *P* value for the comparison with the German data set is of borderline significance (*P* = .064), but the number of R1a1 chromosomes within this data set is small (*n* = 11), resulting in a larger sampling variance. The additional six microsatellites genotyped for Ashkenazi Levites only reduce the modal haplotype frequency within the R1a1 haplogroup to 58%, also confirming the high degree of haplotype homogeneity (table 6). As is explained in the "Discussion" section, this arrangement of R1a1 haplotypes within the Ashkenazi Levite data set is consistent with common descent from a recent ancestor and is unlikely to have resulted from a large number of founding lineages. By assuming that the most recent common ancestor (MRCA) possessed the modal haplotype and by using a male intergeneration time of 25 years, we estimated a mean TMRCA of 663 years before present under the Simple Stepwise Mutation Model and a mean time of 1,000 years before present under the Linear Length-Dependent Stepwise Mutation Model (see the "Subjects and Methods" section). These dates coincide with the historically estimated time when compact settlements of Jews in northwest Europe began and the important religious

communities of Mainz and Worms were established (Encyclopaedia Judaica 1972), but they also come with very wide CIs. If we assume a genealogy generated under no population growth, the 95% CIs are 0–2,425 years before present under the simple stepwise mutation model, and 0–3,672 years before present under the linear length-dependent stepwise mutation model. However, the Ashkenazi Jewish population has undergone rapid population growth in the past 1,000 years, estimated at a 17.5% growth rate per generation in the Ashkenazi Jewish population over this period (Della-Pergola 2001). The starlike pattern for the R1a1 haplotypes for AL is consistent with this pattern of population expansion for the Ashkenazi Jewish population in general. By use of this estimate of 17.5% for growth and under the assumption of 4.5 million male Ashkenazi Jews in 1939, a 1:5 ratio of effective population size to census population size due to reproductive variance and the fact that Ashkenazi Levites make up 4% of Ashkenazi Jews, the 95% CI is reduced substantially to 244–1,570 years before present (BP) under the simple stepwise mutation model, and 375–2,248 years BP under the linear length-dependent stepwise mutation model. It should be noted that our estimations of the TMRCA rely on pedigree mutation rates. It has been argued that "evolutionary" mutational rates may occur at a lower rate (Caglia et al. 1997; Forster et al. 2000). If so, the TMRCA is underestimated by our method. On the other hand, if our assumption of a single founder is false, then the TMRCA of this clade would overestimate the founding event. An alternative method for dating the founding event, again under the assumption of a single founder, derives from consideration of the fact that it is easier to affect a large frequency change in the Ashkenazi Levite male gene pool when it is small than when it is large. We allowed for a possible initial reproductive advantage to the founder—for example,
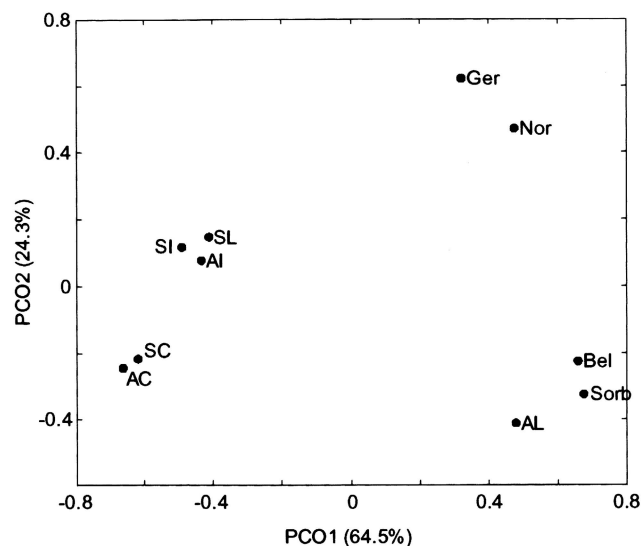
Figure 1   Principal coordinates plot of the genetic identity values (haplogroup level) shown in table 2. Axis labels indicate the percentage explained by the first two principal axes.

due to elevated wealth or status—such that he contributed, at most, 100 grandsons to the Ashkenazi Levite population (i.e., 20 grandsons to the effective population size, under the assumption of a 1:5 ratio, as above), after which time the differential reproductive fitness was lost. As above, we assumed a general Ashkenazi Levite growth rate of 17.5% per generation and an Ashkenazi Levite effective population size of 36,000 in 1939. Performing coalescent simulations with exponential growth, we found that the 20 grandsons needed to be added at least 35 generations ago to provide a chance of >5% that 31 of 60 modern Ashkenazi Levites sampled would descend from this founding group. Under

Table 2

Pairwise Comparisons at Haplogroup Level among Six Jewish and Four Non-Jewish Sample Sets

	AC	SC	AL	SL	AI	SI	Ger	Nor	Sorb	Bel
AC	...	.407	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SC	1.002	...	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
AL	.216	.293	...	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SL	.793	.821	.381	...	.111	.200	<.001	<.001	<.001	<.001
AI	.821	.856	.468	.977	...	.548	<.001	<.001	<.001	<.001
SI	.819	.860	.425	.983	1.021	...	<.001	<.001	<.001	<.001
Ger	.106	.192	.426	.507	.510	.523	...	.048	<.001	<.001
Nor	.089	.159	.515	.531	.461	.398	.927	...	<.001	<.001
Sorb	.071	.144	.953	.273	.276	.205	.475	.656	...	.002
Bel	.089	.152	.875	.342	.307	.212	.510	.745	.975	...

NOTE.—Lower left triangle = Nei's Genetic Identity, I; upper right triangle = P values for population differentiation (exact test of Raymond and Rousset 1995), uncorrected for multiple pairwise comparisons. Population sample labels are the same as in table 1.

the assumption of a male intergeneration time of 25 years, this implies an original founding event that occurred at least 990 years BP (925 years before 1939). Taken together, these separate time estimates provide a timeframe that coincides with the initial formation and early settlement of the Ashkenazi Jewish population. Finally, we used the same growth model to investigate how many males could have moved into the Ashkenazi Levite population at the same time as the founder and yet have at least a 5% chance of leaving no descendents in a sample of 60 Ashkenazi Levites today. Because we allow for a potentially very small effective population size, it is possible for an introgression to comprise a sizeable proportion of the Ashkenazi Levite NRY gene pool at the time and yet correspond to only a relatively small number of individuals. If the event took place 37 generations before 1939 (990 years BP), such "silent" introgressors could have made up as much as 19% of the Ashkenazi Levite gene pool, equivalent to an addition of 50 men to the existing Ashkenazi Levite population. The corresponding figures for 45 generations before 1939 (1,190 years BP) and 55 generations before 1939 (1,440 years BP) are 53% (33 men) and 94% (10 men).

Discussion

Our study confirms the previously reported finding of a caste-specific high-frequency haplogroup within the Ashkenazi Levites (Thomas et al. 1998). The presence of this haplogroup, R1a1, within Ashkenazi Levites is striking for several reasons. Firstly, it is found at high frequency in the Ashkenazi Levites but not in Sephardi Levites or any other Jewish grouping examined so far. This means that in paternal ancestry Ashkenazi and Sephardi Levites are genetically dissimilar, unlike Ashkenazi and Sephardi Cohanim (fig. 1). The Ashkenazi and
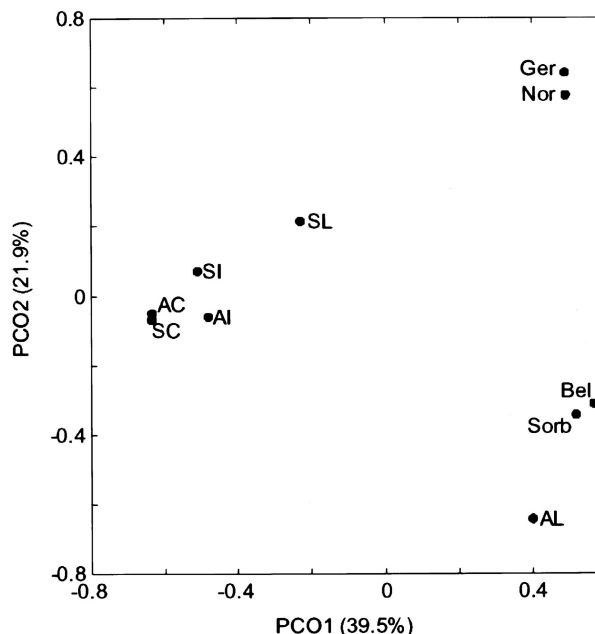
Figure 2 Principal coordinates plot of the genetic identity values (haplotype level) shown in table 4. Axis labels indicate the percentage explained by the first two principal axes.

Sephardi Israelites are also relatively similar to each other, which is consistent with the previous reports of shared overall paternal Near Eastern ancestries for these populations (Hammer et al. 2000; Nebel et al. 2000). Secondly, the microsatellite haplotypes within this haplogroup form a tight cluster within the Ashkenazi Levites, indicative of very recent origin from a single common ancestor. Thirdly, the haplogroup is extremely rare in other Jewish groups and in non-Jewish groups of Near Eastern origin, but is found at high frequency in populations of eastern European origin. This contrasts with

Table 3

Pairwise Comparisons at Haplotype Level among Six Jewish and Four Non-Jewish Sample Sets

	AC	SC	AL	SL	AI	SI	Ger	Nor	Sorb	Bel
AC	...	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SC	.945	...	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
AL	.060	.081	...	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SL	.349	.335	.040	...	.002	.028	<.001	<.001	<.001	<.001
AI	.723	.765	.181	.421	...	.004	<.001	<.001	<.001	<.001
SI	.711	.766	.126	.491	.807	...	<.001	<.001	<.001	<.001
Ger	.019	.028	.028	.198	.152	.179	...	.036	<.001	<.001
Nor	.067	.075	.101	.138	.131	.215	.799	...	<.001	<.001
Sorb	.087	.096	.525	.104	.191	.160	.295	.334	...	<.001
Bel	.028	.076	.490	.056	.201	.081	.332	.334	.613	...

NOTE.—Lower left triangle = Nei's Genetic Identity, *I*; upper right triangle = *P* values for population differentiation (exact test of Raymond and Rousset 1995), uncorrected for multiple pairwise comparisons. Population sample labels are the same as in table 1.

Table 4

Microsatellite Haplotype Frequencies within Haplogroup R1a1 across Six Jewish and Four Non-Jewish Sample Sets

MICROSATELLITE HAPLOTYPE[a]	FREQUENCY IN SAMPLE SET									
	AC	SC	AL	SL	AI	SI	Ger	Nor	Sorb	Bel
14 12 25 10 11 13	...	...	...	...	...	...	...	...	...	.0033
15 12 24 10 11 12	...	...	...	...	...	...	...	...	...	.0033
15 12 24 10 11 13	...	...	.0333	...	...	...	...	...	...	.0098
15 12 24 11 11 13	...	...	...	...	...	...	...	.0120	.0179	.0098
15 12 25 10 11 13	.0132	...	.0333	.0323	.0100	...	.0114	.0482	.0536	.0294
15 12 25 10 12 13	...	...	...	...	...	...	...	...	...	.0065
15 12 25 11 11 13	...	...	...	...	...	...	...	.0241	.0089	.0458
15 12 25 11 11 14	...	...	...	...	...	...	.0114	.0120	...	...
15 12 26 10 11 13	...	...	...	...	...	...	...	...	...	.0065
15 12 26 11 11 13	...	...	...	...	...	...	...	.0241	...	.0065
15 13 25 10 13 13	...	...	...	...	...	...	...	...	...	.0033
15 13 25 11 11 13	...	...	...	...	...	...	...	...	.0089	...
15 14 25 11 11 13	...	...	...	...	...	...	...	.0120	...	...
16 11 24 11 11 14	...	...	...	...	...	...	...	...	...	.0033
16 12 24 10 11 12	...	...	.0167	...	...	...	...	...	...	...
16 12 24 10 11 13	...	...	.0167	...	...	.0159	...	.0120	.0893	.0131
16 12 24 10 11 14	...	...	...	...	...	...	...	...	...	.0065
16 12 24 10 12 13	...	...	...	...	...	...	...	...	...	.0033
16 12 24 11 11 13	...	.0145	...	...	...	...	...	...	...	.0229
16 12 25 10 11 12	...	...	...	...	...	...	.0114	...	...	...
16 12 25 10 11 13	...	.0145	.3833	...	.0100	...	...	.0120	.1071	.0850
16 12 25 11 11 12	...	...	...	...	...	...	...	...	...	.0065
16 12 25 11 11 13	...	.0290	...	...	.0200	...	.0568	.0361	.0179	.1209
16 12 25 11 11 14	...	...	...	...	...	...	...	...	...	.0033
16 12 25 11 12 13	...	...	...	...	...	...	...	...	...	.0065
16 12 25 12 11 13	...	...	...	...	...	...	...	...	...	.0033
16 12 26 10 11 13	...	...	...	...	...	...	...	...	.0089	...
16 12 26 11 11 13	...	...	...	...	...	...	...	...	.0089	.0131
16 13 25 10 11 13	...	...	...	...	...	...	...	...	...	.0163
16 13 25 10 11 15	...	...	...	...	...	...	...	...	...	.0033
16 13 25 11 11 14	...	...	...	...	...	...	...	...	...	.0033
17 10 24 10 11 13	...	...	...	...	...	...	...	.0120	...	...
17 12 23 11 11 14	...	...	...	...	...	...	...	.0120	...	...
17 12 24 10 11 13	...	...	.0167	...	...	...	.0114	...	.0268	.0033
17 12 24 10 11 14	...	...	...	...	...	...	...	...	.0089	...
17 12 24 11 11 13	...	...	...	...	...	...	...	...	.0089	...
17 12 25 10 11 12	...	...	...	...	...	...	...	...	.0089	...
17 12 25 10 11 13	...	...	.0167	...	...	...	.0114	...	.1786	.0425
17 12 25 10 11 14	...	...	...	...	...	...	...	...	...	.0033
17 12 25 10 12 13	...	...	...	...	...	...	...	...	.0089	...
17 12 25 11 11 13	...	...	...	...	...	...	...	...	.0357	.0131
17 12 26 10 11 13	...	...	...	...	...	...	.0114	...	.0268	.0033
17 12 26 10 12 13	...	...	...	...	...	...	...	...	...	.0033
17 13 25 10 11 13	...	...	...	...	...	...	...	...	...	.0033
17 13 25 11 11 13	...	...	...	...	...	...	...	...	...	.0033
17 13 26 11 11 13	...	...	...	...	...	...	...	...	.0089	...

NOTE.—Population sample labels are the same as in table 1.

[a] Composed of the repeat sizes of microsatellite loci, in the following order: DYS19, DYS388, DYS390, DYS391, DYS392, DYS393.

the Cohen Modal Haplotype, for example, which belongs to a haplogroup that is more likely to be of Near Eastern origin. Finally, the haplogroup represents ~50% of all Ashkenazi Levite NRYs, which indicates that, as a group, the Ashkenazi Levites have heterogeneous or-igins in comparison with the Cohanim, who are domi-nated by a single haplogroup and whose origins are con-sistent with an event antedating the Diaspora.

The greatly elevated frequency of haplogroup R1a1 only within the Ashkenazi Levites suggests an event spe-

Table 5

Sample Size and Microsatellaite Diversity within Haplogroup R1a1

	VALUE IN SAMPLE SET									
VARIABLE	AC	SC	AL	SL	AI	SI	Ger	Nor	Sorb	Bel
No. of Subjects in R1a1 Haplogroup	1	4	31	1	4	1	11	18	71	156
Haplotype Diversity[a]	...	...	.4516	...	...	...	.8182	.9281	.86721	.896
SE for Haplotype Diversity	...	...	.1097	...	...	...	.1191	.0401	.0245	.0142

[a] By use of the unbiased gene diversity formula (Nei 1987), for cases where $n > 10$.

cific to the Ashkenazi Levites. This is borne out by the fact that the haplogroup is virtually absent from the Sephardi Levites, which would indicate an event that occurred after the separate formation of the Ashkenazi and Sephardi groupings, and also by the very recent inferred date for the common ancestor of the NRYs within this haplogroup in the Ashkenazi Levites, which also supports an event occurring after the Ashkenazi split from other Jewish populations.

The pattern of microsatellite haplotype diversity within haplogroup R1a1 in the Ashkenazi Levites suggests that any founding event is unlikely to have involved a large number of founding lineages contributing to today's Ashkenazi Levite NRY gene pool. If large numbers were involved, then the lack of microsatellite diversity would imply that all the founding lineages were very closely related to each other. Furthermore, under this scenario the microsatellite diversity seen today would partly result from diversity among the founding lineages, in which case the evidence would suggest a founding event even more recent than that estimated above, under the assumption of a single founding individual. Although a more recent event is not impossible, the current range of estimated dates fits very nicely with the known origins of the Ashkenazi, a time when the estimated population size was small and which would therefore make it easier for the descendents of a single founder to expand to high frequency within the population, as our simulations demonstrate. Furthermore, if there had been a large influx of founder lineages, the haplotype frequency distribution of the source population would be expected to match the haplotype frequency distribution found in Ashkenazi Levites today. Such a matching source population has yet to be identified. In the case of the two non-Jewish populations of Eastern European origin examined in this study, neither the Sorbians nor the Belarusians would be suitable candidates. This is because, in both cases, the modal haplotype within haplogroup R1a1 for the Sorbians and the Belarusians is either completely absent or found only as a singleton within the Ashkenazi Levite sample (table 4).

Although the number of founding lineages contributing to the current Ashkenazi Levite NRY gene pool appears to be small and could involve only a single

founding male, it is possible that the founding event also coincided with a larger introgression event but subsequent drift led to the loss of these other lineages from the NRY gene pool. Our simulations suggest that up to 50 men could have been involved in such an event. This result assumes a 5:1 ratio of real:effective population size, and there are processes that could make this ratio even more pronounced (see, for example, the work of Gagnon and Heyer [2001]). However, this would result in such small effective population sizes and so much resultant drift that one would not expect to find the close NRY genetic similarity observed between Ashkenazi and Sephardi Cohanim, who make up a percentage of the Jewish population very similar to the Levites'. Two further arguments against a recent large introgression of non-Jews into the Ashkenazi Levite caste are: (*a*) it would have to breach a well-regulated rabbinically controlled barrier, and (*b*) it would most likely leave some prominent trace in the historical record—which it has not.

For the reasons stated above, it is likely that the event leading to a high frequency of R1a1 NRYs within the Ashkenazi Levites involved very few, and possibly only one, founding father. A question, then, arises regarding the possible origins of the founder(s). Haplogroup R1a1 is found at very low frequency in other Jewish groups (table 1). It is possible, therefore, that this haplogroup was also present at very low frequency in the Ashkenazi Levites, before some event specific to the Ashkenazi Levites occurred that led to a founder who, by chance, had this very rare haplogroup and whose descendents became very numerous within the Ashkenazi Levites. Likewise, the haplogroup is also found at very low frequency within some non-Jewish populations of Near Eastern origin (data not shown). It is therefore also possible that a conversion event prior to the formation of the Ashkenazi grouping led to the founding of this haplogroup and its emergence at high frequency within the Ashkenazi Levites. Although it is not possible to formally refute either of these two possible explanations, it would be a remarkable coincidence that the geographic origins and demographic expansion of the Ashkenazi are within Northern and Eastern Europe and that this haplogroup is found at very high frequency within neighboring non-Jewish populations of European origin

Table 6

Microsatellite Haplotype Frequencies within Haplogroup R1a1 for 12 Microsatellites Typed in the Ashkenazi Levite Sample

6-Locus[a] and Constituent 12-Locus[b] Haplotypes	Count	Frequency in Haplogroup	Frequency in Sample
15 12 24 10 11 13:			
[15 11 14 12 13 17 24 10 11 13 12 14]	2	.0645	.0333
15 12 25 10 11 13:			
[15 11 14 12 13 17 25 10 11 13 12 14]	1	.0323	.0167
[15 11 14 12 13 16 25 10 11 13 12 14]	1	.0323	.0167
16 12 24 10 11 12:			
[16 11 14 12 13 18 24 10 11 12 12 14]	1	.0323	.0167
16 12 24 10 11 13:			
[16 11 14 12 13 17 24 10 11 13 12 14]	1	.0323	.0167
16 12 25 10 11 13:			
[16 11 14 12 13 16 25 10 11 13 12 14]	4	.1290	.0667
[16 11 14 12 13 18 25 10 11 13 12 14]	1	.0323	.0167
[16 11 14 12 13 17 25 10 11 13 12 14]	18	.5806	.3000
17 12 24 10 11 13:			
[17 11 14 12 13 17 24 10 11 13 12 14]	1	.0323	.0167
17 12 25 10 11 13:			
[17 11 14 12 13 17 25 10 11 13 12 14]	1	.0323	.0167

[a] Composed of the repeat sizes of microsatellite loci, in the following order: DYS19, DYS388, DYS390, DYS391, DYS392, DYS393.

[b] Shown in brackets and composed of the repeat sizes of microsatellite loci, in the following order: DYS19, DYS385a, DYS385b (arbitrarily given larger repeat size over DYS385a), DYS388, DYS389-I, DYS389-II, DYS390, DYS391, DYS392, DYS393, DYS426, DYS439.

but not at high frequency elsewhere. An alternative explanation, therefore, would postulate a founder(s) of non-Jewish European ancestry, whose descendents were able to assume Levite status.

If a European origin for the Ashkenazi Levite haplogroup R1a1 component is accepted as a reasonable possibility, it is of interest to speculate further on the possible timing, location, and mechanism of this event. Because the modal haplotype of haplogroup R1a1 found in the Ashkenazi Levites is found at reasonably high frequency throughout the eastern European region, it is not possible to use genetic information to pinpoint the exact origin of any putative founder from the currently available data sets. Intriguingly, the Sorbian tongue, relexified with a German vocabulary, has been proposed as the origin of Yiddish, the language of the Ashkenazim, but there has been no suggestion of an association between Ashkenazi Levites in particular and the Sorbian language. One attractive source would be the Khazarian Kingdom, whose ruling class is thought to have converted to Judaism in the 8th or 9th century (Dunlop 1967). This kingdom flourished between the years 700 C.E. and 1016 C.E. It extended from northern Georgia in the south to Bulgar on the Volga River in the north and from the Aral Sea in the east to the Dnieper River in the west—an area that falls within a region in which haplogroup R1a1 NRYs are found at high frequency (Rosser et al. 2000). Archival material also

records migration of Khazars into the Hungarian Duchy of Taskony in the 10th century. The break-up of the Khazar Empire following their defeat by invading Rus led to the flight of some Khazars to central and northern Europe. Although neither the NRY haplogroup composition of the majority of Ashkenazi Jews nor the microsatellite haplotype composition of the R1a1 haplogroup within Ashkenazi Levites is consistent with a major Khazar or other European origin, as has been speculated by some authors (Baron 1957; Dunlop 1967; Ben-Sasson 1976; Keys 1999), one cannot rule out the important contribution of a single or a few founders among contemporary Ashkenazi Levites.

Finally, it is interesting to speculate on the possible mechanism by which the descendant of a non-Jew or convert could have acquired Levite status. The fact that Ashkenazi Cohanim NRYs show no evidence for an introgression of this nature suggests a lesser degree of stringency for the assumption of Levite status than for the assumption of Cohen status. This may be because there are more rights and duties associated with the Cohen status than with that of the Levite, leading to more rigorous protection of the former. Cohanim, for example, are called upon, on special occasions, to bless the assembled congregation and are prohibited from marrying divorcees and converts, religious laws that do not apply to Levites. Indeed, Talmudic sources may possibly be interpreted to support the notion of differences

in the social, religious, and legal barriers that relate to the assumption of Cohen and Levite status. These include descriptions of the possible assumption of Levite status other than through patrilineal descent, in a Talmudic passage describing a debate regarding the potential assignment of Levite status to a man (and his descendants) whose father was a non-Jew and whose mother was the daughter of a Levite. Such differences could have provided the backdrop for the sanctioned acceptance of Levite status other than through patrilineal descent.

The comparative study of patterns of NRY variation among Ashkenazi Jews and other populations has revealed evidence for an unexpected and unusual historical event, which was not appreciated using other, more conventional historical approaches. This finding may motivate historians and social scientists to seek further information regarding the possibility of such an event and, more generally, to include information gleaned from studies of DNA variation in the repertoire of tools used to uncover historical events of interest.

Acknowledgments

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

Ytime program, http://www.ucl.ac.uk/tcga/software/

References

Baron SW (1957) A social and religious history of the Jews. Vol. III. The Jewish Publication Society of America, Philadelphia, pp 173–222

Ben-Sasson HH (1976) Jewish autonomy from the Black Death to the Reformation. In: Ben-Sasson HH (ed) A history of the Jewish people. Harvard University Press, Cambridge, MA, pp 593–611

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, Lopez-Camelo JS (1998) Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. Am J Hum Genet 63:1862–1871

Bonne-Tamir B, Korostishevsky M, Redd AJ, Pel-Or Y, Kaplan ME, Hammer MF (2003) Maternal and paternal lineages of the samaritan isolate: mutation rates and time to most recent common male ancestor. Ann Hum Genet 67:153–164

Bradman N, Thomas M, Goldstein D (1999) The genetic origins of Old Testament priests. In: Renfrew CE (ed) Population specific polymorphisms. Cambridge University Press, Cambridge, United Kingdom, pp 31–44

Caglia A, Novelletto A, Dobosz M, Malaspina P, Ciminelli BM, Pascali VL (1997) Y-chromosome STR loci in Sardinia and continental Italy reveal islander-specific haplotypes. Eur J Hum Genet 5:288–292

Calabrese PP, Durrett RT, Aquadro CF (2001) Dynamics of microsatellite divergence under stepwise mutation and proportional slippage/point mutation models. Genetics 159:839–852

Capelli C, Redhead N, Abernethy JK, Gratrix F, Wilson JF, Moen T, Hervig T, Richards M, Stumpf MPH, Underhill PA, Bradshaw P, Shaha A, Thomas MG, Bradman N, Goldstein DB (2003) A Y chromosome census of the British Isles. Curr Biol 13:979–984

Carvalho-Silva DR, Santos FR, Hutz MH, Salzano FM, Pena SD (1999) Divergent human Y-chromosome microsatellite evolution rates. J Mol Evol 49:204–214

DellaPergola S (2001) Jewish demography 2001. In: DellaPergola S, Even S (eds) Papers in Jewish demography. The Hebrew University of Jerusalem Press, Jerusalem, pp 11–33

Dunlop DM (1967) The history of the Jewish Khazars. Schocken Books, New York

Encyclopaedia Judaica (1972) Keter Publishing, Jerusalem

Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, Brinkmann B (1998) Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. Mol Biol Evol 15:1108–1114

Forster P, Rohl A, Lunnemann P, Brinkmann C, Zerjal T, Tyler-Smith C, Brinkmann B (2000) A short tandem repeat-based phylogeny for the human Y chromosome. Am J Hum Genet 67:182–196

Gagnon A, Heyer E (2001) Intergenerational correlation of effective family size in early Quebec (Canada). Am J Hum Biol 13:645–659

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995a) An evaluation of genetic distances for use with microsatellite loci. Genetics 139:463–471

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995b) Genetic absolute dating based on microsatellites and the origin of modern humans. Proc Natl Acad Sci USA 92:6723–6727

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

Hammer MF, Redd AJ, Wood ET, Bonner MR, Jarjanazi H, Karafet T, Santachiara-Benerecetti S, Oppenheim A, Jobling MA, Jenkins T, Ostrer H, Bonne-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

Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803

Holtkemper U, Rolf B, Hohoff C, Forster P, Brinkmann B (2001) Mutation rates at two human Y-chromosomal microsatellite loci using small pool PCR techniques. Hum Mol Genet 10:629–633

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Kruger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A (2000) Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 66:1580–1588

Keys D (1999) Catastrophe: an investigation into the origins of the modern world. Ballantine Books, New York

Nebel A, Filon D, Weiss DA, Weale M, Faerman M, Oppenheim 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 evolutionaty genetics. Columbia University Press, New York

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Redd AJ, Agellon AB, Kearney VA, Contreras VA, Karafet T, Park H, de Knijff P, Butler JM, Hammer MF (2002) Forensic value of 14 novel STRs on the human Y chromosome. Forensic Sci Int 130:97–111

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

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

Rolf B, Meyer E, Brinkmann B, de Knijff P (1998) Polymorphism at the tetranucleotide repeat locus DYS389 in 10 populations reveals strong geographic clustering. Eur J Hum Genet 6:583–588

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

Skorecki K, Selig S, Blazer S, Bradman R, Bradman N, Waburton PJ, Ismajlowicz M, Hammer MF (1997) Y chromosomes of Jewish priests. Nature 385:32

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Stumpf MP, Goldstein DB (2001) Genealogical and evolutionary inference with the human Y chromosome. Science 291: 1738–1742

Thomas MG, Bradman N, Flinn 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

Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, 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, Bonne-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

Weale ME, Weiss DA, Jager RF, Bradman N, Thomas MG (2002) Y chromosome evidence for Anglo-Saxon mass migration. Mol Biol Evol 19:1008–1021

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

Wexler P (1993) The Ashkenazi Jews: a Slavo-Turkic people in search of a Jewish identity. Slavica Publishers, Columbus, OH

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

Zhivotovsky LA, Goldstein DB, Feldman MW (2001) Genetic sampling error of distance $\delta\mu^2$ and variation in mutation rate among microsatellite loci. Mol Biol Evol 18:2141–2145