Hum Genet (2000) 107 : 630–641
DOI 10.1007/s004390000426

ORIGINAL INVESTIGATION

Almut Nebel · Dvora Filon · Deborah A. Weiss
Michael Weale · Marina Faerman · Ariella Oppenheim
Mark G. Thomas

High-resolution Y chromosome haplotypes of Israeli and Palestinian Arabs reveal geographic substructure and substantial overlap with haplotypes of Jews

Received: 12 July 2000 / Accepted: 10 October 2000 / Published online: 21 November 2000

Abstract High-resolution Y chromosome haplotype analysis was performed in 143 paternally unrelated Israeli and Palestinian Moslem Arabs (I&P Arabs) by screening for 11 binary polymorphisms and six microsatellite loci. Two frequent haplotypes were found among the 83 detected: the modal haplotype of the I&P Arabs (~14%) was spread throughout the region, while its one-step microsatellite neighbor, the modal haplotype of the Galilee sample (~8%), was mainly restricted to the north. Geographic substructuring within the Arabs was observed in the highlands of Samaria and Judea. Y chromosome variation in the I&P Arabs was compared to that of Ashkenazi and Sephardic Jews, and to that of North Welsh individuals. At the haplogroup level, defined by the binary polymorphisms only, the Y chromosome distribution in Arabs and Jews was similar but not identical. At the haplotype level, determined by both binary and microsatellite markers, a more detailed pattern was observed. Single-step microsatellite networks of Arab and Jewish haplotypes revealed a common pool for a large portion of Y chromosomes, suggesting a relatively recent common ancestry. The two modal haplotypes in the I&P Arabs were closely related to the most frequent haplotype of Jews (the Cohen modal haplotype). However, the I&P Arab clade that includes the two Arab modal haplotypes (and makes up 32% of Arab chromosomes) is found at only very low frequency among Jews, reflecting divergence and/or admixture from other populations.

URLs for computer programs in this article are as follows:
Arlequin, http://anthropologie.unige.ch/arlequin;
BATWING, http://www.maths.abdn.ac.uk/~ijw;
Microsat, http://lotka.stanford.edu/microsat.html;
PHYLIP, http://evolution.genetics.washington.edu/phylip.html;
TreeView, http://taxonomy.zoology.gla.ac.uk/rod/rod.html

A. Nebel · D. Filon · A. Oppenheim (✉)
Department of Hematology, Hebrew University,
Hadassah Medical School and Hadassah University Hospital,
Jerusalem 91120, Israel
e-mail: Ariella@md.huji.ac.il,
Tel.: +972-2-6776753, Fax: +972-2-6423067

D.A. Weiss
Department of Anthropology, University of California,
Davis, USA

M. Weale · M.G. Thomas
Department of Biology, University College London, London, UK

M. Faerman
Laboratory of Biological Anthropology and Ancient DNA,
Hebrew University, Hadassah School of Dental Medicine,
Jerusalem, Israel

Introduction

The analysis of variable loci in the non-recombining part of the Y chromosome, which contains a record of many past mutational events, facilitates the tracing of paternal lineages. Over the last few years, an increasing number of informative Y chromosome polymorphisms, which can be typed following PCR, have become available (Jobling and Tyler-Smith 1995; Underhill et al. 1997; Hammer and Zegura 1997). Furthermore, extensive data have been assembled on the geographic distribution of different Y-specific markers among various contemporary populations (de Knijff et al. 1997; Karafet et al. 1997; Kayser et al. 1997; Skorecki et al. 1997; Hammer et al. 1997, 1998). Compound Y chromosome haplotypes comprising both binary and microsatellite polymorphisms have proved to be especially powerful tools for the investigation of population substructure (Thomas et al. 1998) and, possibly, of relationships between groups that have become obscured through considerable admixture (Thomas et al. 2000). Some high frequency modal haplotypes (signature haplotypes) may be representative of particular communities, for instance, the Cohen modal haplotype (CMH) appears to be associated with the paternally inherited Jewish priesthood (Thomas et al. 1998). Binary markers represent rare, in most cases probably unique, event polymorphisms in human evolution and, thus, allow identification of deep splits in the Y chromosome genealogy (Jobling and Tyler-Smith 1995). In contrast, microsatellite loci have a much faster mutation rate (Kayser et al. 2000) and

reflect more recent genealogical events (Roewer et al. 1996; Bosch et al. 1999).

Historically, the origin of the Arab population residing in Israel and the Palestinian Authority Area (I&P Arabs) is complex and diverse. Located at the crossroads of three continents, the Southern Levant has, throughout history, attracted many waves of immigrants and conquerors alike. Permanent human settlement in the region dates back to the Natufian Period (~12,500–10,000 BP). Based on anthropological evidence, it has been suggested that the Natufians and their descendants formed a 'core' population that can be traced to recent times, but was mixed with incoming groups (Arensburg 1973). According to historical records, major demographic events took place in the Israelite Period and during the Jewish Kingdom Period (1200–586 BC): the Assyrian and Babylonian invasions were followed by the deportation of locals and the settlement of foreign peoples (Bachi 1974). The Roman Judean Wars (66–135 AD) culminated in the destruction of the Second Temple and led to the annihilation or exile of a large portion of the Jewish population (Anderson 1995). By the fifth century AD, the majority of non-Jews and Jews had become Christians by conversion (Bachi 1974). The first millennium AD was marked by the immigration of Arab tribes, reaching its climax with the Moslem conquest from the Arabian Peninsula (633–640 AD). This was followed by a slow process of Islamization of the local population, both of Christians and Jews (Shaban 1971; Mc Graw Donner 1981). Additional minor demographic changes might have been caused by subsequent invasions of the Seljuks, Crusaders, Mongols, Mamelukes and Ottoman Turks. Recent gene-flow from various geographic origins is reflected, for example, in the heterogenous spectrum of β-globin mutations among Israeli Arabs (Filon et al. 1994).

Israeli and Palestinian Arabs share a similar linguistic and geographic background with Jews. Their genetic relationship has been the focus of several investigations. Comparative studies of classical markers (Bonné-Tamir et al. 1979), mitochondrial DNA (mtDNA) restriction haplotypes (Bonné-Tamir et al. 1986; Ritte et al. 1993), HLA (Bishara et al. 1997) and disease-related mutations (Filon et al. 1994; Peretz et al. 1997) revealed substantial genetic affinities between the two populations, yet also significant differences. Based on two Y chromosome RFLP markers it was suggested that Ashkenazi and Sephardic Jews are more closely related to Arabs from Lebanon than to Czechoslovakians (Santachiara-Benerecetti et al. 1993). Moreover, a recent survey of 18 binary Y-specific polymorphisms showed that Y chromosome haplotypes of Middle Eastern non-Jewish populations are almost indistinguishable from those of Jews (Hammer et al. 2000).

In the present study, we performed an in-depth evaluation of the Y chromosome affinity between Arabs and Jews by analysis of high-resolution haplotypes comprising both binary and microsatellite polymorphisms. This approach also allowed us to characterize in detail the I&P Arab population with regard to their genetic composition, lineage diversity, geographic substructure and possible origins.

Subjects and methods

Subjects

Y chromosomes of 143 Moslem Arabs from Israel and the Palestinian Authority Area (I&P Arabs), unrelated at the paternal great-grandfather level, were analyzed. Druze and Bedouins were excluded from this study. Most of the subjects come from rural areas where their families have lived for generations. DNA was obtained anonymously from thalassemia patients or their relatives (carriers as well as non-carriers). For the purpose of the present study, we consider this sample to be representative of the general Moslem Arab population in Israel and the Palestinian Authority Area because: (1) thalassemia is unlinked to the Y chromosome, is spread throughout the country and is not restricted to specific regions; (2) the spectra and frequencies of thalassemia mutations in any subset of males classified here according to their Y chromosomes (i.e., haplogroups) was found to be similar to each other and to the spectrum and frequency of thalassemia mutations in the general population, as previously reported (Filon et al. 1994); (3) a significantly different proportion of YAP^+ chromosomes was observed in one of the regional Arab subpopulations, the Highlands people (see Results). We used this finding to obtain independent evidence that the entire sample was not biased. We randomly collected 24 additional DNA specimens from the Highlands population. No statistical difference was detected between this control group and the Highlands sample used in the study, neither at the haplogroup level (population differentiation test: $P>0.9$) nor at the haplotype level (AMOVA: Φ_{ST} value = 0.003; $P>0.27$).

Data on Y chromosomes of Ashkenazi and Sephardic Jews (collected in Israel, Canada and the UK) were the same as previously reported (Thomas et al. 1998). Male Jews are traditionally divided into Israelites, Levites and Cohanim. In this study, only Israelites were considered. They comprise approximately 90% of the Jews and are, therefore, most representative of the Jewish population. During the diaspora, Sephardic Jews lived mostly in Arab countries (in North Africa and the Middle East) or in countries under strong Arab influence (in Spain and other regions in southern Europe), while Ashkenazim mainly resided in northern and central Europe. To ascertain that any genetic similarity between Arabs and Jews seen in this study is not due to recent admixture of Sephardim with Arabs, Ashkenazi and Sephardic Jews were treated separately.

The North Welsh, a representative European Caucasian population (Darke et al. 1998; Sellers et al. 1999), were included in the analysis because they do not have a known history of admixture with Jewish communities. The data obtained from the Welsh therefore allowed us to examine possible gene flow from Europeans to Ashkenazi Jews during the diaspora. The group of Welsh analyzed included 94 males unrelated at the paternal grandfather level. The samples were collected from villages around Llangefni, a town in North Wales.

The study was approved by the Hebrew University Committee for Ethics in Research.

Typing of Y chromosome DNA polymorphisms

The DNA samples, prepared from peripheral blood by standard protocols, were typed for 17 different Y chromosome DNA polymorphisms in three multiplex PCRs as described (Thomas et al. 1999): UEP1-PCR: YAP (*DYS287*; Hammer 1994), 92r7 (Mathias et al. 1994), SRY4064 (Whitfield et al. 1995), SRY+465 (Shinka and Nakahori, personal communication), sY81 (*DYS271*; Seielstad et al. 1994), Tat (Zerjal et al. 1997); UEP2-PCR: M9, M13, M17, M20 (Underhill et al. 1997), SRY10831 (Whitfield et al. 1995); MS-PCR: *DYS19*, *DYS388*, *DYS390*, *DYS391*, *DYS392*, *DYS393* (Jobling and Tyler-Smith 1995).

Table 1 Y chromosome haplogroup distribution in Arabs, Jews and Welsh (*P* YAP insert present, *N* no YAP insert, *A* adenine, *C* cytosine, *G* guanine, *T* thymine)

Population	I&P Arabs		Ashkenazi Jews[a]		Sephardic Jews[a]		North Welsh[b]	
Haplogroup	*n*	Frequency	*n*	Frequency	*n*	Frequency	*n*	Frequency
1 NCGCAT[c]	100	0.699	42	0.618	32	0.628	7	0.074
2 NTGCAT[c]	14	0.098	12	0.176	12	0.235	84	0.894
3 PCACAT[c]	29	0.203	14	0.206	7	0.137	3	0.032
Total	143	1.000	68	1.000	51	1.000	94	1.000

[a]Data on Ashkenazi and Sephardic Israelites as reported (Thomas et al. 1998)
[b]Data on Welsh are available upon request from one of the authors (M.G.T.)
[c]Haplogroups are defined by the allele status at six binary markers in the following order: YAP, 92r7, SRY4064, SRY+465, sY81, Tat

Table 2 Y chromosome haplogroup distribution in four I&P Arab subpopulations (*P* YAP insert present, *N* no YAP insert, *A* adenine, *C* cytosine, *G* guanine, *T* thymine, *G+* guanine present, *G-* guanine deleted)

Population		I&P Arabs (total)		North		Lowlands		Gaza		Highlands	
Haplogroup		*n*	Frequency	*n*	Frequency	*n*	Frequency	*n*	Frequency	*n*	Frequency
1	**NCGCAT[a]**	**100**	**0.699**	**39**	**0.722**	**23**	**0.793**	**15**	**0.714**	**23**	**0.589**
1A	CGG+AG[b]	88	0.615	35	0.648	20	0.690	12	0.571	21	0.538
1B	GGG+AG[b]	10	0.070	4	0.074	3	0.103	1	0.048	2	0.051
1C	CCG+AA[b]	2	0.014	–	–	–	–	2	0.095	–	–
2	**NTGCAT[a]**	**14**	**0.098**	**8**	**0.148**	**4**	**0.138**	**1**	**0.048**	**1**	**0.026**
2A	GGG+AG[b]	12	0.084	6	0.111	4	0.138	1	0.048	1	0.026
2B	GGG-AA[b]	2	0.014	2	0.037	–	–	–	–	–	-
3	**PCACAT[a] CGG+AG[b]**	**29**	**0.203**	**7**	**0.130**	**2**	**0.069**	**5**	**0.238**	**15**	**0.385**
Total		143	1.000	54	1.000	29	1.000	21	1.000	39	1.000

[a]Haplogroups are defined by the allele status at six binary markers in the following order: YAP, 92r7, SRY4064, SRY+465, sY81, Tat
[b]Haplogroups are additionally defined by the allele status at five binary markers in the following order: M9, M13, M17, M20, SRY10831

Since the Y chromosomes of Jews had not been typed for the five markers in UEP2-PCR, the Arab chromosomes were, for the purpose of Arab-Jewish comparisons, classified into three haplogroups based on the allelic state of the six binary markers in UEP1-PCR. For intra-population analysis, the Arab chromosomes were classified into six haplogroups based on all 11 binary markers. Haplogroups were defined using the nomenclature presented in Tables 1 and 2.

Statistical and genealogical analyses

Genetic identity (*I*), haplotype diversity (*h*) and its sampling variance were calculated as described by Nei (1987), using unbiased estimates. Interpopulation comparisons were made using the exact test of population differentiation (Raymond and Rousset 1995) and the analysis of molecular variance (AMOVA; Excoffier et al. 1992) included in the software package Arlequin (Version 1.1; Schneider et al. 1997). Φ_{ST} values in AMOVA were calculated using the sum of squared allele size differences (R_{ST}) as a measure of microsatellite haplotype distance (Michalakis and Excoffier 1996). Microsat (Version 1.5d; Minch 1997) was used to compute the genetic distance measure average square distance (ASD; Goldstein et al. 1995). UPGMA and unrooted neighbor joining (NJ) haplotype trees based on ASD were drawn with the phylogeny inference package, PHYLIP (Version 3.5c; Felsenstein 1995) and displayed using TreeView (Version 1.5). Relationships of haplotypes were also visualized by constructing networks of microsatellite haplotypes for each haplogroup (Cooper et al. 1996). In these networks, all haplotypes are linked to one or more other haplotypes without inferring unobserved intermediate states. This method excluded all those haplotypes (44 of the 140 different haplotypes found in Arabs and Jews) that were at least two microsatellite mutational steps removed from their nearest link in the network. We note that, with the exception of three, all unlinked haplotypes were singletons. Tests for differences in genetic distance of populations *X* and *Y* from a third population *Z* were carried out by bootstrapping a relevant statistic *D* and deriving a 95% confidence interval for the difference D_{XZ}-D_{YZ}.

Estimate of population expansion date

In order to estimate dates for the start of population growth amongst the I&P Arabs, we carried out full-likelihood Bayesian analysis using the BATWING (*B*ayesian *A*nalysis of *T*rees *W*ith *I*nternal *N*ode *G*eneration) program (Wilson and Balding 1998). The extended version of the program used here assumes an unbounded single stepwise mutation model (Moran 1975) for the microsatellite loci and a coalescent process under a model of exponential population growth from an initially constant-size population. Binary polymorphisms were used only to condition the space of permissible trees, but otherwise did not contribute to the likelihood. Informative priors were based on previous data or reasonable ranges. The initial effective population size was given a prior distribution of Gamma(4,0.001), which has 2.5%, 50% and 97.5% quantiles of 1090, 3672 and 8767. The population growth rate (*r*) per generation was given a Gamma(1.5,100) prior (2.5%, 50%, 97.5% quantiles = 0.0011, 0.012, 0.047). The microsatellite mutation rate per generation, μ, was given a Gamma(12,5862) prior

(2.5%, 50%, 97.5% quantiles = 0.0011, 0.0020, 0.0034). This prior is based on data from four published studies (Heyer et al. 1997; Bianchi et al. 1998; Kayser et al. 1997; Kayser et al. 2000), restricted to the same microsatellite loci as those used here. The 12 observed mutational events out of 5862 meioses were combined with an (improper) uniform pre-prior. Generation time was set at 25 years.

Results

Genetic composition of the I&P Arab population

The six binary markers YAP, 92r7, SRY4064, SRY+465, sY81 and Tat defined three haplogroups among the I&P Arabs (Table 1). Haplogroups 1 and 2 lack the YAP insertion, while haplogroup 3 is YAP^+/sY81-A. The additional five polymorphisms, M9, M13, M17, M20 and SRY10831 further divided haplogroups 1 and 2 into three and two groups, respectively (Table 2). The two sets of binary markers together with the six microsatellites defined 83 different compound haplotypes, as listed in the appendix. No instance of homoplasy of microsatellite haplotypes across haplogroups was observed. The large number of singletons (73.5%) contributed to the high haplotype diversity (h) of 0.971±0.008 for all the Arab samples. Haplotype 25 was the most common (~14%) and was, therefore, called the modal haplotype of the Israeli and Palestinian Arabs. Haplotype 21, one of its single-step neighbors, was observed almost exclusively in individuals from the Lower Galilee area, where it was the modal haplotype. It made up 18.5% of the samples from the Galilee. Together with their one-step neighbors, both modal haplotypes comprise 31% of the sample.

Using BATWING analysis, the median estimate for the start of exponential population growth from an initially constant-size population was 7780 years BP (95% credibility interval = 2780–22856 years BP). The proportion of the posterior distribution for population expansion dates that is post the onset of the Neolithic Period in the Southern Levant (10,300 BP; Bar-Yosef 1995) is 69.2%. The median estimate for the population growth rate during the expansion period was 0.76% per generation (95% credibility interval = 0.26%–2.24% per generation).

Genetic affiliation of Arabs with Jews

For the Jews, data on only six binary polymorphisms and six microsatellites were available (Thomas et al. 1998). Therefore, a comparison of the Arab and Jewish samples was performed using 12-marker haplotypes.

The population differentiation test based on haplogroup frequencies (Table 1) showed a statistically significant difference between Arabs and Sephardic Jews (P<0.05) due to a higher frequency in Sephardic Jews of haplogroup 2, but no significant differences were found between Arabs and Ashkenazi Jews nor between the two Jewish communities. However, all three were distinguished from the Welsh (P<0.0005). Table 1 shows that the major proportion of Y chromosomes of Arabs and Jews belonged to haplogroup 1, while most of the Welsh chromosomes were of haplogroup 2. The frequency of haplogroup 2 chromosomes in Arabs (~10%) was significantly lower than that in Sephardic Jews (~24%, P<0.018), and lower, but not significantly so, than that in Ashkenazi Jews (~18%, P>0.05).

Table 3 presents the extent of I as pairwise comparisons between populations based on the frequencies of binary plus microsatellite haplotypes. Using bootstrap tests on I values, we found that both Sephardic and Ashkenazi Jews were significantly closer to I&P Arabs than Arabs were to Welsh (P<0.001 in both cases), and no significant difference was found in the genetic identity of Arabs to one Jewish group compared to the other (P=0.816).

Sephardic Jews were less distant to Arabs than to Welsh. However, Ashkenazi Jews and Welsh were closer to one another than Ashkenazim and Arabs. Examination of the data revealed that this finding is due to the fact that two Ashkenazi individuals carried the most common Welsh haplotype (~27%). This haplotype has also been observed at high frequencies in other European countries, including England (~15%), Frysland (~13%) and Norway (~6%) (manuscript in preparation). We note that 8 of the 12 chromosomes shared between Jews (both Ashkenazim and Sephardim) and Welsh belonged to haplogroup 2, which was found in the Welsh at a frequency of ~89% (Table 1). In contrast, most chromosome sharing between Arabs and Jews involved haplogroup 1. Neither of the two Arab modal haplotypes were detected in Jews. However, three Arab individuals from different regions carried the CMH (haplotype 27). None of these three modal haplotypes were seen in the Welsh sample.

The genetic differences among the four populations were assessed by AMOVA, including chromosomes from

Table 3 Genetic identity (I) between populations

Population	Sephardic Jews	Ashkenazi Jews	North Welsh
I&P Arabs	0.144	0.134	0.018
Sephardic Jews	–	0.706	0.093
Ashkenazi Jews	–	–	0.174

Table 4 Y chromosome variation between populations estimated by AMOVA (Φ_{ST})[a]

Population	I&P Arabs		Ashkenazi Jews		Sephardic Jews	
	Φ_{ST}[b]	P[c]	Φ_{ST}[b]	P[c]	Φ_{ST}[b]	P[c]
Ashkenazi Jews	0.067	<0.001	–	–	–	–
Sephardic Jews	0.055	0.002	0.001	0.330	–	–
North Welsh	0.391	<0.001	0.290	<0.001	0.376	<0.001

[a]AMOVA was performed on chromosomes from all three haplogroups
[b]Φ_{ST}: ratio of variance component due to differences among populations over the total variance
[c]P the probability of observing a more extreme Φ_{ST} after 1000 randomizations than those generated in the analysis

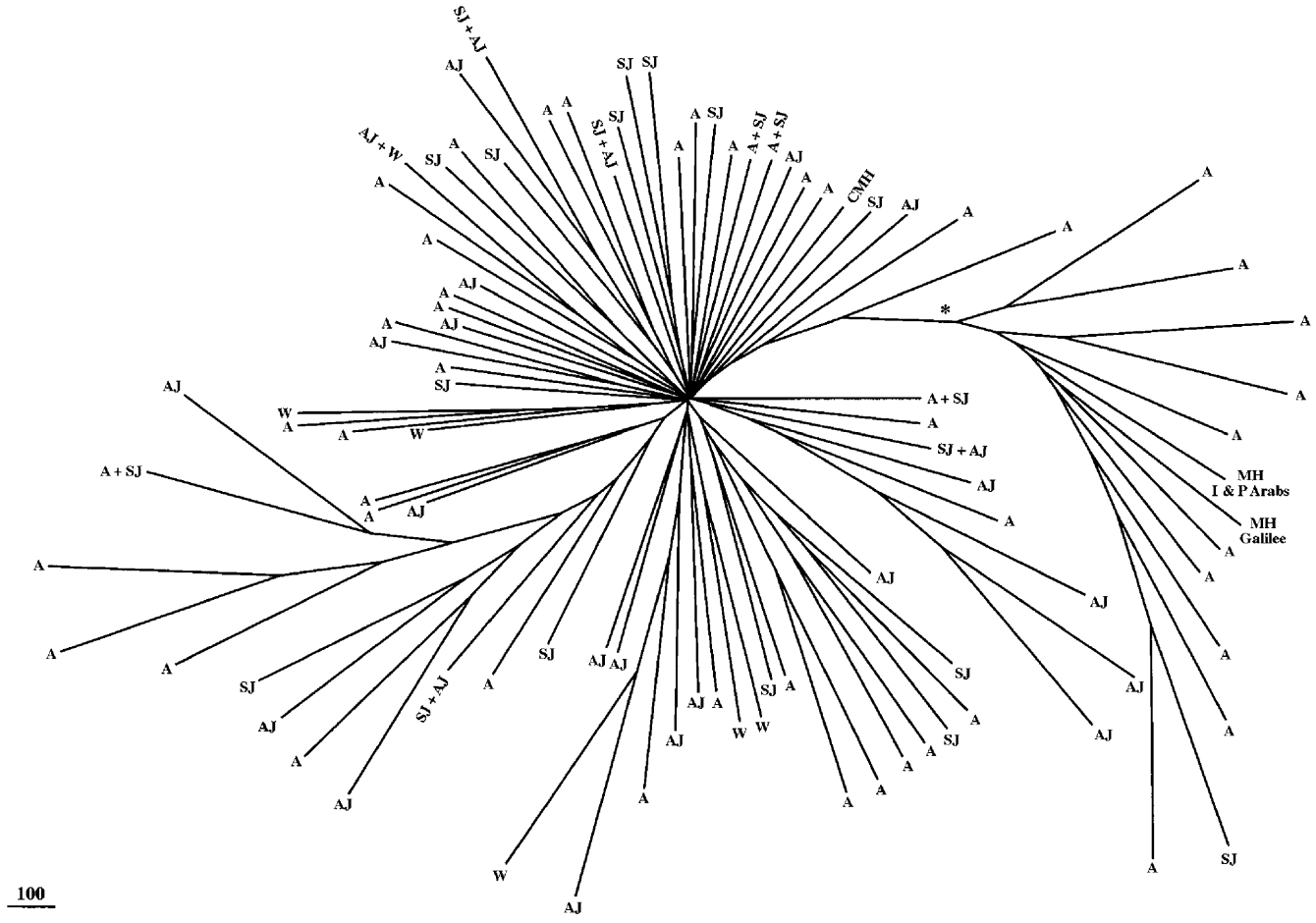

Fig. 1 Unrooted neighbor joining tree based on ASD was drawn for haplogroup 1 haplotypes of Arabs, Jews and Welsh. Bootstrap analysis on 500 resamplings showed support for the Arab clade of about 50% (marked with an *asterisk*). The numbers of the 12 Arab haplotypes that define the Arab clade are: 7–9, 18–26 (as listed in the appendix). (*A* I&P Arab, *AJ* Ashkenazi Jew, *SJ* Sephardic Jew, *W* Welsh, *MH I&P Arabs* modal haplotype of the I&P Arabs, *MH Galilee* modal haplotype of the Galilee sample, *CMH* Cohen modal haplotype)

all three haplogroups (Table 4). The Φ_{ST} values (the proportion of total variance attributable to inter-population differences) between Arabs and both Jewish communities were significant, yet low relative to the values obtained in the analysis of the Welsh outgroup with the three other populations. The Φ_{ST} values calculated for Ashkenazi and Sephardic Jews revealed no significant difference.

High repeat numbers for *DYS388* (≥15) were common in Arabs and both Jewish populations. Europeans have been shown to exhibit mainly *DYS388* alleles with short repeat lengths (Kayser et al. 1997; Kittles et al. 1998, 1999). High repeat numbers for *DYS388* (≥15) have, so far, been found at high frequency only in populations originating in the Middle East (Thomas et al. 2000) and are restricted to haplogroup 1. Compared to 54% of the Ashkenazi and 53% of the Sephardic Jews, 72% of Arabs with a haplogroup 1 chromosome carried a *DYS388* repeat number ≥15. In the Welsh, only one of the seven individuals belonging to haplogroup 1 had a high repeat number for *DYS388*.

Genealogical trees of Arab, Jewish and Welsh haplotypes were constructed separately for each haplogroup, using both UPGMA and NJ on ASD distances. The NJ tree for haplogroup 1 is presented as an example in Fig. 1. The haplotypes of Arabs, Jews and Welsh were intermingled and no population-specific clustering was observed, with the exception of a single branch in haplogroup 1. This branch or clade, which was identical in both UPGMA and NJ trees, comprised almost exclusively Arab haplotypes, including the modal haplotypes of the I&P Arabs and the Galilee sample (Fig. 1). 32% of the 143 Arab chromosomes belonged to this "I&P Arab clade" that contained only one non-Arab chromosome, that of a Sephardic Jew. Bootstrap analysis on 500 resamplings showed support for the Arab clade of about 50% (NJ) and 10% (UPGMA), respectively.

Evidence for an Arab clade can also be seen in the single-step network of haplogroup 1 chromosomes in which the haplotypes were linked to each other without inferring unobserved intermediate states (Fig. 2). The network excluded all those haplotypes that could not be linked and therefore comprised fewer haplotypes than the corresponding trees. The bottom part of the network with the two Arab modal haplotypes and their one-step neighbors

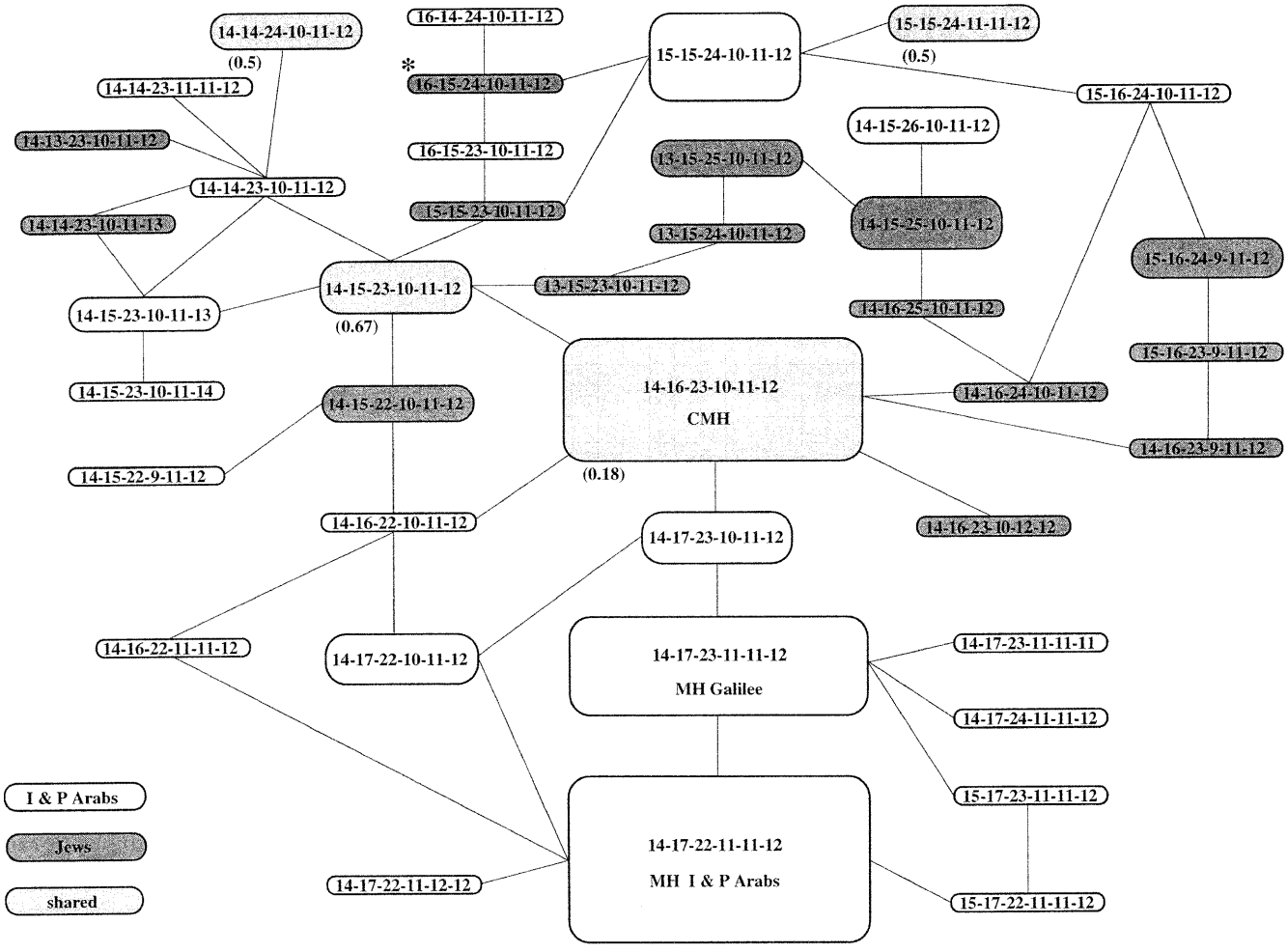

Fig. 2 Network of haplogroup 1 Y chromosome haplotypes of Arabs and Jews. Alleles for the microsatellite loci are listed in the order *DYS19*, *DYS388*, *DYS390*, *DYS391*, *DYS392*, *DYS393*. *Lines* represent single microsatellite mutation steps. The network was drawn manually by starting with the most common haplotype and sequentially adding adjacent ones. All possible adjacent relationships were indicated by *connecting lines*. The arrangement of the haplotypes within the network is arbitrary. The length of the lines connecting haplotypes as well as the distance between unlinked haplotypes do not reflect phylogenetic proximity. *Colour* indicates the population in which the haplotype is present. The *box area* surrounding each haplotype is proportional to its frequency in both populations. *Numbers in parenthesis* under shared haplotypes denote the proportion of Arab chromosomes. The haplotype marked with an *asterisk* was also found in a single Welsh individual. (*MH I&P Arabs* modal haplotype of the I&P Arabs, *MH Galilee* modal haplotype of the Galilee sample, *CMH* Cohen modal haplotype)

contains 9 of the 13 haplotypes of the Arab clade. The other Arab clade haplotypes, including the one of the Sephardic Jew, are not present as they were at least two microsatellite mutational steps removed from their nearest link in the network. Networks were also constructed separately for the other two haplogroups (not shown). Apart from the Arab clade, Jewish, Arab and shared Y chromosomes were, to a substantial extent, intermingled throughout the networks of all three haplogroups.

Is there a geographic substructure in the Arab population?

Most Arab families in the rural regions have lived in the same local area for generations. Although Israel and the Palestinian Authority Area are relatively small, it was of interest to investigate whether there was evidence of any geographic structure within the Arab population. Four regions were designated as previously described (Filon et al. 1994): the North comprises the mountainous region of the Upper and the Lower Galilee; the fertile plain along the Mediterranean Sea is divided into the central Lowlands and the Gaza area; the Highlands to the east of the coastal plain are a mountain range that covers Samaria and Judea (including Jerusalem).

The analysis of the four Arab subpopulations was based on the 17-marker haplotypes (Table 2). At the haplogroup level, the Y chromosome distribution was similar for all populations, except for the Highlands people. As shown by the population differentiation test, they were distinguished from those in the Lowlands ($P<0.01$) and the North ($P<0.05$), but not from those in the Gaza area. The Highlands people are also the only population who differed significantly from the Ashkenazi ($P<0.05$) and Sephardic Jews ($P<0.001$). If haplogroup 3 (YAP^+) chromosomes are excluded from the analysis, the only signifi-

cant difference that remains is between the Highlands and Sephardic Jews, $P<0.05$. Thus, the higher frequency of haplogroup 3 is the principal distinguishing feature of the Highlands population. In a genealogical tree including all YAP^+ haplotypes, no association between haplotypes and a particular subpopulation was observed (not shown). All four Arab subpopulations were distinguished from the Welsh ($P<0.0005$).

The close relationship of the four Arab subpopulations is reflected in the high proportion of shared chromosomes. Each had between 36% (the Highlands) and 52% (the Gaza area) of the Y chromosomes in common with the other Arab subpopulations. The modal haplotype of the I&P Arabs was detected in all four subpopulations.

Discussion

Geographic origin of haplogroups

The largest proportion (~62%) of Arab Y chromosomes belonged to haplogroup 1A, the most geographically widespread haplotype today (Underhill et al. 1997; Hammer et al. 1998). Haplogroup 1C, represented by two Arab chromosomes in this study, has so far only been reported in Sudan (Underhill et al. 1997). Haplogroup 3 (chromosomes with the SRY4064-A polymorphism on a YAP^+ background), estimated to have arisen about 30,000 years BP in Africa (Hammer et al. 1998), constitutes about 20% of the overall Arab sample.

The most recent YAP^+ haplotype, with the A to G transition at sY81, has been estimated to have originated in sub-Saharan Africa between 10,000 and 20,000 years BP (Hammer et al. 1998). Outside its area of origin, the presence of this haplogroup is thought to indicate admixture with sub-Saharan Africans. It has been reported, for example, in Egypt (2%) and the Arabian Peninsula (5%; Hammer et al. 1998). Other sub-Saharan markers have been documented in the I&P Arabs, including mtDNA RFLP defined haplotypes, blood groups and sickle cell anemia (Sandler et al. 1979; Bonné-Tamir et al. 1986; Rund et al. 1990; Ritte et al. 1993). It is, therefore, intriguing that this YAP^+ haplotype with the A to G transition at sY81 was not observed in the I&P Arabs sampled in this study.

Haplogroups 1B and 2, which comprise about 17% of the examined Arab Y chromosomes, are likely to be of Eurasian origin (Underhill et al. 1997; Hurles et al. 1998). In Europe, the 92r7 T-allele, which distinguishes haplogroup 2 from haplogroup 1, shows a decreasing north-south gradient, ranging from 89% in Welsh (this study) to 48% in north Italians and 20% in Greeks (Mitchell et al. 1997). It would seem that the Arabs, with a haplogroup 2 frequency of 10%, represent a population in the southern part of the geographic cline.

Of the polymorphic variants particularly associated with Asian populations (SRY4064-G/YAP^+, SRY+465-T, Tat-C, M20-G and M17-G^-) only the last was detected here, in two individuals (haplogroup 2B). To date, M17-G^- has been observed at high frequency in central and western Asia and in India/Pakistan (Underhill et al. 1997).

Our median estimate for the start of exponential population growth falls within the Neolithic Period. Although this was unsurprising, given that population growth is expected to follow the development of farming, the credibility interval is wide, with only 69.2% of the posterior distribution falling on or after the onset of the Neolithic Period in the Southern Levant. This reflects, to some extent, the limited information contained in a sample of only 143 chromosomes. However, recent admixture events may also contribute to this wide estimate.

I&P Arab population structure

The similarity of the four Arab subpopulations is attested by the large extent of Y chromosome and haplotype sharing among the people residing in the various regions. This homogeneity is not surprising given that the studied area is relatively small and that there are no natural geographical barriers. However, the haplotype distribution also reflects a certain degree of regional isolation. This is seen in the Highlands people, who differ from two of the three other subpopulations by a significantly higher frequency of YAP^+ chromosomes. The genetic distinctiveness of the Highlands people has also been noted with regard to mutations in the β-globin gene. The predominant β-thalassemia mutation in the region, IVS 1–6, reaches 40%, which is fourfold higher than in any other part of the country, where another mutation, IVS 1–110, predominates (Filon et al. 1994). The locally restricted modal haplotype of the Galilee sample also suggests possible geographic structuring, although this is not statistically established. A possible explanation for the restricted presence of this haplotype could be a recent, localized population expansion in the Galilee. The observed regional differences in the distribution of Y chromosome haplotypes may be the result of drift and/or founder effect and may subsequently have been enhanced by traditional endogamous marriage practices, such as intrafamilial marriages, male polygamy and patrilocality.

Genetic affinities of Arabs to Jews

Arabs are more closely related to Jews than they are to the Welsh, indicating a more recent common ancestry. At the haplogroup level, the Y chromosome distribution was similar in both Arabs and Jews, although a significant difference was found between Arabs and Sephardic Jews in haplogroup 2 frequencies. This finding does not necessarily conflict with the results of Hammer et al. (2000), who found no significant differences between Jews and non-Jewish Middle-Eastern populations, as the difference reported here was only just significant ($P<0.05$) and Hammer et al. used a lower significance level ($\alpha=0.03$).

The incorporation of microsatellites in the analysis revealed a more complex pattern of genetic affinities and

differences among the populations. Arabs and Jews had approximately 18% of their chromosomes in common. In addition, they were characterized by closely related modal haplotypes, each displaying a high repeat number *DYS388* allele typical of Middle Eastern populations (Thomas et al. 2000). Furthermore, with the exception of the Arab clade, the haplotypes of the two populations appeared interspersed throughout the NJ and UPGMA trees of all three haplogroups, reflecting a close genealogical relationship given the high observed mutation rate of Y chromosome microsatellites (Kayser et al. 2000). In the haplotype networks, Jewish chromosomes provided connecting links that were missing in the Arab-only networks and vice versa. A substantial portion of Arab (82%) and Jewish Y chromosomes (70%) comprised the mixed chromosome pools of each of the three haplogroups (as inferred from the networks). Our findings corroborate previous studies that suggested a common origin for Jewish and non-Jewish populations living in the Middle East (Santachiara-Benerecetti et al. 1993; Peretz et al. 1997; Hammer et al. 2000).

However, the present study, using high-resolution haplotypes, also revealed statistically significant differences between Arabs and Jews. Both populations were characterized by distinct modal haplotypes that were infrequent in the other population. Notably, about one third of the Arab individuals carried I&P Arab clade haplotypes that were observed once in the Sephardic Jews studied here and in a single Ashkenazi Cohen out of 306 male Jews tested (Thomas et al. 1998). Although Arabs and Jews showed a high frequency of haplotypes with a *DYS388* repeat number ≥15, the distribution of the alleles was different in each population. *DYS388* allele 17 was found almost exclusively in Arabs, while allele 16 was common among Jews. In addition, both Sephardic and Ashkenazi Jews had higher frequencies of haplogroup 2 chromosomes than Arabs. These differences presumably reflect divergence over time due to genetic drift and/or gene-flow from other populations. Both Ashkenazi and Sephardic Jews are likely to have experienced admixture with various populations during the 2000 years of diaspora. The higher frequency of haplogroup 2 chromosomes in Jews may reflect the influx of foreign lineages. Two Ashkenazi Jews in our sample were found to share the most common haplotype of the Welsh, a frequent and widespread haplotype in northern Europe. The presence of European Y chromosomes in Ashkenazi Jews has also been reported previously based on two RFLP markers (Santachiara-Benerecetti et al. 1993).

The occurrence of less than 1% of I&P Arab clade chromosomes in the Ashkenazi and Sephardic samples is noteworthy since they shared many other haplotypes with Arabs. The low haplotype diversity of the Arab clade chromosomes, as seen in the network (Fig. 2), suggests that they descended from a relatively recent common ancestor. Arab clade chromosomes could have been present in the common ancestral population of Arabs and Jews, and drifted to high frequencies in one of the subgroups following population isolation. The event leading to this isolation might have been the acceptance of the monotheistic Jewish religion by a subset of the population, or geographic separation due to the expulsion of Jews after the destruction of the Second Temple in AD 79. Alternatively, the Arab clade could have been introduced through gene-flow, perhaps by the immigration of Arab tribes in the first millennium AD. In this regard, it is of interest that Arab clade chromosomes were observed in 8 out of 49 Moslem Arabs (16%) from the Hadramaut in Yemen (Thomas et al. 2000). Further studies are needed to clarify the origin of the Arab clade.

According to historical records part, or perhaps the majority, of the Moslem Arabs in this country descended from local inhabitants, mainly Christians and Jews, who had converted after the Islamic conquest in the seventh century AD (Shaban 1971; Mc Graw Donner 1981). These local inhabitants, in turn, were descendants of the core population that had lived in the area for several centuries, some even since prehistorical times (Gil 1992). On the other hand, the ancestors of the great majority of present-day Jews lived outside this region for almost two millennia. Thus, our findings are in good agreement with historical evidence and suggest genetic continuity in both populations despite their long separation and the wide geographic dispersal of Jews.

Acknowledgements We wish to thank Prof. Patricia Smith for stimulating discussions and encouragement and Prof. Hagai Ben-Shamai for providing the historical insight. We are grateful to Dr. Moien Kanaan and Mahmoud Abd El-Latif for supplying control DNA samples. This work was partially supported by a research grant from the Israeli Ministry of Science, Culture and Sport, and by funding to M.G.T. from the Nuffield Foundation (NUF-NAL).

Appendix

Distribution of Y chromosome haplotypes in four I&P Arab subpopulations

Haplotype No.	Allele status at						Arab subpopulations[a]				
	DYS19	*DYS388*	*DYS390*	*DYS391*	*DYS392*	*DYS393*	N	L	G	H	Total
Haplogroup 1A											
1	16	15	23	10	11	12	1				1
2	16	14	24	10	11	12		1			1
3	16	13	24	11	11	13		1			1
4	16	13	23	10	12	14				1	1
5	16	12	23	10	11	12		1			1
6	16	12	22	10	11	14	1		1		2
7	15	18	23	12	11	13	1				1
8	15	17	23	11	11	12	1				1
9	15	17	22	11	11	12				1	1
10	15	16	24	10	11	12			1		1
11	15	15	24	11	11	12				1	1
12	15	15	24	10	11	12		3	1	1	5
13	15	15	22	10	8	12	1				1
14	15	13	25	10	11	12		1			1
15	15	13	23	10	12	12	1				1
16	15	12	22	11	11	14		2			2
17	15	10	23	10	12	13	2				2
18	14	17	24	11	11	12		1			1
19	14	17	23	13	11	12	1				1
20	14	17	23	11	11	11		1			1
21	14	17	23	11	11	12	10			2	12
22	14	17	23	10	11	12	2			1	3
23	14	17	22	12	11	13		1			1
24	14	17	22	11	12	12			1		1
25	14	17	22	11	11	12	5	4	6	5	20
26	14	17	22	10	11	12		1		2	3
27	14	16	23	10	11	12	1	1	1		3
28	14	16	23	9	11	16				1	1
29	14	16	22	11	11	12	1				1
30	14	16	22	10	11	12	1				1
31	14	15	26	10	11	12	2				2
32	14	15	23	10	11	14				1	1
33	14	15	23	10	11	13				2	2
34	14	15	23	10	11	12	1		1		2
35	14	15	22	11	11	13	2				2
36	14	15	22	9	11	12		1			1
37	14	14	25	11	11	12		1			1
38	14	14	24	10	11	12				1	1
39	14	14	23	11	11	12				1	1
40	14	14	23	10	11	12	1				1
41	14	12	22	11	11	14				1	1
Haplogroup 1B											
42	15	12	24	10	14	13			1		1
43	15	12	23	11	13	15		1			1
44	15	12	23	10	13	13	1				1
45	15	12	23	9	13	13				1	1
46	14	12	23	10	14	13		1			1
47	14	12	23	10	14	12	1				1
48	14	12	22	10	14	11				1	1
49	14	12	22	10	13	13	1				1
50	13	12	24	10	13	13	1	1			2

Haplotype No.	Allele status at						Arab subpopulations[a]				
	DYS19	*DYS388*	*DYS390*	*DYS391*	*DYS392*	*DYS393*	N	L	G	H	Total
Haplogroup 1C											
51	15	11	22	10	11	13			1		1
52	14	15	23	9	11	12			1		1
Haplogroup 2A											
53	15	12	24	11	13	13				1	1
54	15	12	24	10	13	13	1	1			2
55	15	12	24	10	13	12	1				1
56	14	12	25	11	13	12	1				1
57	14	12	25	10	13	13			1		1
58	14	12	24	11	14	12	1				1
59	14	12	24	11	13	12	1	2			3
60	14	12	23	10	10	12		1			1
61	13	12	24	10	14	13	1				1
Haplogroup 2B											
62	16	12	25	11	11	13	1				1
63	15	12	25	11	11	12	1				1
Haplogroup 3											
64	16	12	24	10	11	13				1	1
65	15	13	24	11	11	13		1			1
66	15	12	24	11	11	13				1	1
67	15	12	24	10	11	14			1		1
68	15	12	24	10	11	13	1			1	2
69	14	13	24	10	11	12			1		1
70	14	12	25	10	11	14	1		1		2
71	14	12	25	10	11	13				1	1
72	14	12	25	9	11	13				1	1
73	14	12	24	10	11	13	1			1	2
74	13	12	26	10	11	14				4	4
75	13	12	25	10	12	13			1		1
76	13	12	25	10	11	14				1	1
77	13	12	25	10	11	13	1	1		1	3
78	13	12	24	11	12	13	1				1
79	13	12	24	10	11	12	2				2
80	13	12	24	10	10	13			1		1
81	13	12	24	9	11	13				1	1
82	13	12	23	10	11	13				1	1
83	11	12	23	10	12	13				1	1
Total							54	29	21	39	143
h[b]							0.958	0.973	0.929	0.974	0.971
v[c]							0.018	0.018	0.051	0.014	0.008

[a]Arab subpopulations: *N* North, *L* Lowlands, *G* Gaza, *H* Highlands
[b]*h* Y chromosome haplotype diversity
[c]*v* sampling variance

References

Anderson J (1995) The impact of Rome on the periphery: the case of Palestina – Roman Period (63 BCE-324 CE). In: Levy T (ed) The archaeology of society in the Holy Land. Leicester University Press, London, pp 446–468

Arensburg B (1973) The people in the land of Israel from the Epipaleolithic to present times. Doctoral Dissertation, University of Tel Aviv, Tel Aviv

Bachi R (1974) The population of Israel. The Hebrew University, Jerusalem

Bar-Yosef O (1995) Earliest food producers – pre-pottery Neolithic. In: Levy T (ed) The archaeology of society in the Holy Land. Leicester University Press, London, pp 190–204

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

Bishara A, Brautbar C, Nagler A (1997) HLA and graft-versus-host disease: a population-based study of HLA phenotypes of Jewish and Arab bone marrow transplanted patients in Israel. Leuk Res 21:111–118

Bonné-Tamir B, Asbel S, Kennet R (1979) Genetic markers: benign and normal traits of Ashkenazi Jews. In: Goodman R, Motulsky A (eds) Genetic diseases among Ashkenazi Jews. Raven Press, New York, pp 54–76

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

Bosch E, Calafell F, Santos F, Pérez-Lezaun A, Comas D, Benchemsi N, Tyler-Smith C, et al (1999) Variation in short tandem repeats is deeply structured by genetic background on the Y chromosome. Am J Hum Genet 65:1623–1638

Cooper G, Amos W, Hoffman D, Rubinsztein D (1996) Network analysis of human Y microsatellite haplotypes. Hum Mol Genet 11:1759–1766

Darke C, Guttridge M, Thompson J, McNamara S, Street J, Thomas M (1998) HLA class I and II gene and haplotype frequencies in blood donors from Wales. Exp Clin Immunogenet 15:69–83

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

Felsenstein J (1995) PHYLIP (phylogeny inference package), version 3.5c. University of Washington, Seattle, USA

Filon D, Oron V, Krichevski S, Shaag A, Shaag Y, Warren T, Goldfarb A, et al (1994) Diversity of β-globin mutations in Israeli ethnic groups reflect recent historic events. Am J Hum Genet 54:836–843

Gil M (1992) A history of Palestine. Cambridge University Press, Cambridge, pp 643–1099

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

Hammer MF (1994) A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. Mol Biol Evol 11:749–761

Hammer M, Zegura S (1997) The role of the Y chromosome in human evolutionary studies. Evol Anthropol 4:116–134

Hammer M, Spurdle A, Karafet T, Bonner M, Wood E, Novelletto A, Malaspina P, et al (1997) The geographic distribution of human Y chromosome variation. Genetics 145:787–805

Hammer M, Karafet T, Rasanayagam A, Wood E, Altheide T, Jenkins T, Griffiths R, et al (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol Evol 15:427–441

Hammer M, Redd A, Wood E, Bonner M, Jarjanazi H, Karafet T, Santachiara-Benerecetti A, et al (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, Knijff P de (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803

Hurles M, Irven C, Nicholson J, Taylor P, Santos F, Loughlin J, Jobling M, et al (1998) European Y chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. Am J Hum Genet 63:1793–1806

Jobling M, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. Trends Genet 11:449–456

Karafet T, Zegura S, Vuturo-Brady J, Posukh O, Osipova L, Wiebe V, Romero F, et al (1997) Y chromosome markers and trans-Bering Strait dispersals. Am J Phys Anthropol 102:301–314

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Krüger C, et al (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

Kittles R, Perola M, Peltonen L, Bergen A, Aragon R, Virkkunen M, Linnoila M, Goldman D, Long J (1998) Dual origins of Finns revealed by Y chromosome haplotype variation. Am J Hum Genet 62:1171–1179

Kittles R, Bergen A, Urbanek M, Virkkunen M, Linnoila M, Goldman D, Long J (1999) Autosomal, mitochondrial and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. Am J Phys Anthropol 108:381–399

Knijff P de, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110: 134–140

Mathias N, Bayés M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. Hum Mol Genet 3:115–123

Mc Graw Donner F (1981) The early Islamic conquests. Princeton University Press, Princeton

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142:1061–1064

Minch E (1997) Microsat, the microsatellite distance program, version 1.5d. Stanford University, USA

Mitchell R, Earl L, Fricke B (1997) Y-chromosome specific alleles and haplotypes in European and Asian populations: linkage disequilibrium and geographic diversity. Am J Phys Anthropol 104:167–176

Moran PAP (1975) Wandering distributions and the electrophoretic profile. Theor Popul Biol 8:318–330

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

Peretz H, Mulai A, Usher S, Zivelin A, Segal A, Weisman Z, Mittelman M, et al (1997) The two common mutations causing factor XI deficiency in Jews stem from distinct founders: one of ancient Middle Eastern origin and another of more recent European origin. Blood 7:2654–2659

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

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

Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, Knijff P de (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. Hum Mol Genet 5:1029–1033

Rund D, Kornhendler N, Shalev O, Oppenheim A (1990) The origin of sickle cell alleles in Israel. Hum Genet 85:521–524

Sandler S, Kravitz C, Sharon R, Hermani D, Ezekiel E, Cohen T (1979) The Duffy blood group system in Israel: Jews and Arabs. Vox Sang 37:41–46

Santachiara-Benerecetti A, 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

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) Arlequin ver. 1.1: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland

Seielstad MT, Herbert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, Cavalli-Sforza LL (1994) Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. Hum Mol Genet 3:2159–2161

Sellers J, Thompson J, Guttridge MG, Darke C (1999) Human platelet antigens: typing by PCR using sequence-specific primers and their distribution in blood donors resident in Wales. Eur J Immunogenet 26:393–397

Shaban MA (1971) Islamic history AD 600–750: a new interpretation. Cambridge University Press, Cambridge.

Skorecki K, Selig S, Blazer S, Bradman R, Bradman N, Warburton P, Ismajlowicz M, et al (1997) Y chromosomes of Jewish priests. Nature 385:32

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

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

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

Underhill P, Jin L, Lin A, Mehdi S, Jenkins T, Vollrath D, Davis R, et al (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Res 7:996–1005

Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. Nature 378:379–380

Wilson IJ, Balding DJ (1998) Genealogical inference from microsatellite data. Genetics 150:499–510

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, et al (1997) Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. Am J Hum Genet 60:1174–1183