

Doron M. Behar · Daniel Garrigan · Matthew E. Kaplan
Zahra Mobasher · Dror Rosengarten · Tatiana M. Karafet
Lluís Quintana-Murci · Harry Ostrer · Karl Skorecki
Michael F. Hammer

Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations

Received: 31 July 2003 / Accepted: 29 November 2003 / Published online: 22 January 2004

© Springer-Verlag 2004

Abstract The molecular basis of more than 25 genetic diseases has been described in Ashkenazi Jewish populations. Most of these diseases are characterized by one or two major founder mutations that are present in the Ashkenazi population at elevated frequencies. One explanation for this preponderance of recessive diseases is accentuated genetic drift resulting from a series of dispersals to and within Europe, endogamy, and/or recent rapid population growth. However, a clear picture of the manner in which neutral genetic variation has been affected by such a demographic history has not yet emerged. We have examined a set of 32 binary markers (single nucleotide polymorphisms; SNPs) and 10 microsatellites on the non-recombining portion of the Y chromosome (NRY) to investigate the ways in which patterns of variation differ between Ashkenazi Jewish and their non-Jewish host populations in Europe. This set of SNPs defines a total of 20 NRY haplogroups in these populations, at least four of which

are likely to have been part of the ancestral Ashkenazi gene pool in the Near East, and at least three of which may have introgressed to some degree into Ashkenazi populations after their dispersal to Europe. It is striking that whereas Ashkenazi populations are genetically more diverse at both the SNP and STR level compared with their European non-Jewish counterparts, they have greatly reduced within-haplogroup STR variability, especially in those founder haplogroups that migrated from the Near East. This contrasting pattern of diversity in Ashkenazi populations is evidence for a reduction in male effective population size, possibly resulting from a series of founder events and high rates of endogamy within Europe. This reduced effective population size may explain the high incidence of founder disease mutations despite overall high levels of NRY diversity.

Electronic Supplementary Material Supplementary material is available in the online version of this article at <http://dx.doi.org/10.1007/s00439-003-1073-7>

D.M. Behar and D. Garrigan contributed equally to this work
Electronic database information: URLs for the data in this article are as follows:

ARLEQUIN, <http://lgb.unige.ch/arlequin/>

D. M. Behar · D. Rosengarten · K. Skorecki
Bruce Rappaport Faculty of Medicine and Research Institute,
Technion and Rambam Medical Center, Haifa, Israel

D. Garrigan · M. E. Kaplan · Z. Mobasher · T. M. Karafet
M. F. Hammer
Division of Biotechnology, University of Arizona,
Tucson, Arizona, USA

L. Quintana-Murci
CNRS URA1961, Institut Pasteur, Paris, France

H. Ostrer
Human Genetics Program,
New York University School of Medicine, New York, USA

K. Skorecki (✉)
Department of Nephrology and Molecular Medicine,
Technion and Rambam Medical Center,
POB 9602, 31096 Haifa, Israel
Tel.: +972-4-8542853, Fax: +972-4-8542946,
e-mail: skorecki@techunix.technion.ac.il

Introduction

The contemporary Ashkenazi Jewish population is thought to have descended from a founding population that originated in the Near East and migrated to Europe within the last two millennia (Goodman 1979a). Following the destruction of the Second Temple in 70 AD and the Roman exile, southern Italy provided an initial destination for Jewish populations dispersing from the Near East. A northbound migration from Italy, beginning in the 4th century AD and proceeding at least until the 10th century, gave rise to a nucleus of Ashkenazi Jewry in the Rhine valley (Ostrer 2001). The eastern expansion of the Ashkenazi settlement, starting after the 11th century, produced what was to become the largest Jewish population prior to World War II. Demographic data provide evidence for a dramatic expansion in size; from an estimated number of ~25,000 in 1300 AD, the Ashkenazi population had grown to more than 8.5 million by the beginning of the 19th century (Weinryb 1972).

This complicated demographic history of prolonged migration and recent growth, plus endogamy in isolated populations (Bonne-Tamir and Adams 1992; Weinryb 1972), may help us understand why there is an elevated frequency of more than 25 recessive disease alleles in Ashkenazi populations (Ostrer 2001; Risch et al. 2003). However, neither the evolutionary processes responsible for the high incidence of genetic diseases nor the time at which these processes occurred within Ashkenazi population history have been elucidated. For example, Risch et al. (1995) have proposed that founder effects resulting from the dynamics of population growth in the 16th to 19th centuries, especially in the northern Jewish Pale of Settlement (Lithuania and Belarus), explain most, if not all, of the genetic diseases observed at high frequency in the Ashkenazi population today. This hypothesis is supported by the inference of a recent age of the single founder mutation (~350 years) that causes early-onset idiopathic torsion dystonia (Risch et al. 1995). On the other hand, the much older estimated age of the factor XI type II mutation (~3,000 years), which has a high frequency in both Ashkenazi and Iraqi Jewish populations, implies that its frequency is largely independent of the recent demographic upheavals peculiar to the Ashkenazi population (Goldstein et al. 1999). This raises the possibility that the elevated frequency of some genetic diseases is the result of accentuated genetic drift in a population ancestral to the major Jewish groups. An alternative explanation for the persistence of disease alleles in the Ashkenazi population is positive selection (i.e., heterozygote advantage) in Jewish populations confronted with novel environments (Chakravarti and Chakraborty 1978; Diamond 1994; Goodman 1979b; Jorde 1992). Furthermore, previous reports of high levels of NRY diversity in the Ashkenazi Jewish

population and low levels of admixture with European non-Jewish populations (Hammer et al. 2000; Thomas et al. 2002), appear to be inconsistent with a demographic history favoring recent founder effect and population expansion as an explanation for the apparent preponderance of Ashkenazi founder disease mutations.

Because natural selection acts in a locus-specific manner and demographic processes such as gene flow and changes in population size affect the whole genome, studies of neutral genetic variation have the potential to elucidate the role of non-selective evolutionary processes influencing disease allele frequencies in Ashkenazi populations. In particular, haploid regions of the genome such as mitochondrial DNA (mtDNA) and the non-recombining portion of the Y chromosome (NRY) that are unusually sensitive to genetic drift should be useful for detecting the effects of bottlenecks in Ashkenazi populations. Currently, the question of whether mtDNA data show evidence of a bottleneck in Ashkenazi populations is unresolved. Whereas Thomas et al. (2002) have found no evidence for an Ashkenazi mtDNA bottleneck, we have detected a strong reduction in effective size of mtDNA in the Ashkenazi population (Behar et al. 2004). In studies of Jewish NRY variation published to date, many contemporary Jewish communities, including Ashkenazim, have been traced to a common Middle Eastern source population existing several thousand years ago (Hammer et al. 2000; Nebel et al. 2001; Thomas et al. 2002). However, these studies possess limited power to address questions about Ashkenazi demographic history as a result of the small sample sizes and the number of NRY markers employed.

In the present study, we have genotyped a set of 32 binary (single nucleotide polymorphism; SNP) markers and 10 Y short tandem repeats (Y-STRs) in a sample of 442

Table 1 Lineage-based and mutation-based names of the 20 AJ and NJ haplogroups/paragroups in Fig. 1

Lineage-based name	Mutation-based name ^a	Derived state at:	Ancestral state at:
E*(xE3)	E-SRY ₄₀₆₄ *	SRY ₄₀₆₄	P2
E3a	E-P1 ^b	P1	
E3b(xE3b1, E3b2)	E-M35*	M35	M78, M81
E3b1	E-M78 ^b	M78	
E3b2	E-M81 ^b	M81	
F*(xG, H, I, J, K)	F-P14*	P14	M201, P19, 12f2a, M9, M52
G*(xG2)	G-M201*	M201	P15
G2	G-P15	P15	
H	H-M52	M52	
I	I-P19	P19	
J*(xJ2)	J-12f2a*	12f2a	M172
J2	J-M172 ^b	M172	
K*(xL, N, O, P)	K-M9*	M9	LLY22g, M175, M20, P27
L	L-M20	M20	
N*(xN3)	N-LLY22g*	LLY22g	
P*(xQ,R)	P-P27*	P27	P36
Q	Q-P36	P36	
R1*	R-M173*	M173	P25
R1a1	R-M17 ^b	M17	
R1b	R-P25 ^b	P25	

^aAbbreviated without parenthetical system

^bNo downstream markers typed; hence, lineage referred to as haplogroup

Ashkenazi Y chromosomes tracing to 10 Jewish communities in western and eastern Europe. Patterns of Ashkenazi diversity resulting from this high resolution analysis are compared with those of geographically matched non-Jewish European host populations typed with the same set of markers to address the following questions: (1) what are the major paternal founding lineages of the Ashkenazi population, (2) what is the rate of admixture between Ashkenazim and European non-Jewish populations, (3) which Y chromosome lineages introgressed from host European non-Jewish populations, and (4) is there a detectable signature of an Ashkenazi Y chromosome bottleneck? These results may also help address the question of sex-specific demographic processes that may have resulted in different male and female effective population sizes (Thomas et al. 2002).

Subjects and methods

Population samples

Buccal swab samples were collected with informed consent from unrelated individuals of Ashkenazi Jewish origin according to procedures approved by the University of Arizona and Rambam Medical Center Human Subjects Committees. Each of the volunteers reported the birthplace of their father, grandfather, and in most cases, great-grandfather. The Jewish samples were subdivided into Ashkenazi (AJ) groupings based on a combination of geographic, religious, and ethno-historical criteria: western Ashkenazi Jews (WAJ) included 50 French Jews (FrJ), 39 German Jews (GeJ), and 23 Dutch Jews (DuJ); eastern Ashkenazi Jews (EAJ) included 39 Austro-Hungarian Jews (AuJ), 40 Byelorussian Jews (BeJ), 37 Lithuanian Jews (LiJ), 58 Polish Jews (PoJ), 47 Romanian Jews (RoJ), 54 Russian Jews (RuJ), and 55 Ukraine Jews (UkJ). The French Jews were collected in the Rhine Valley region to better represent the presumed location of origin of the early Jewish settlement of Ashkenaz. Of the 442 Ashkenazi Jews sampled, 25 and 23 individuals identified themselves as Cohen and Levite, respectively (for definitions of Cohen and Levite see Behar et al. 2003; Skorecki et al. 1997; Thomas et al. 1998). The European non-Jewish (NJ) samples consisted of: western non-Jews (WNJ) including 64 French (Fr), 34 Germans (Ger), and 31 Austrians (Aus); eastern non-Jews (ENJ) including 56 Hungarians (Hun), 50 Poles (Pol), 54 Romanians (Rom), and 59 Russians (Rus). A subset of the SNPs listed in Table 1 was previously typed in some European non-Jewish samples (i.e., SRY₄₀₆₄, P1, P14, P15, P19, 12f2a, M9, P27, M17, and P25; Hammer et al. 2001), and typing data for most of these SNPs were presented for Russian non-Jews in Karafet et al. (2002). The data for the remaining 10 SNPs in Table 1 and for all SNPs in the Hungarian sample are presented here for the first time.

Terminology

We have followed the terminological conventions recommended by the Y Chromosome Consortium (YCC 2002) for naming NRY lineages. Capital letters A–R identify the 18 major NRY clades or haplogroups. Lineages not defined on the basis of a derived character state represent interior nodes of the tree and are potentially paraphyletic (Hammer and Zegura 2002). Thus, the term “paragroup” (rather than haplogroup) is used to describe these lineages, and these paragroups are distinguished by the star symbol (*); the term lineage is as a synonym for haplogroup or paragroup at any level of the tree. For convenience, the term haplogroup or lineage is occasionally used as a synonym for paragroup. Lineages excluded from a haplogroup are listed in Table 1 after an initial “x” symbol within parentheses after the haplogroup name for the offi-

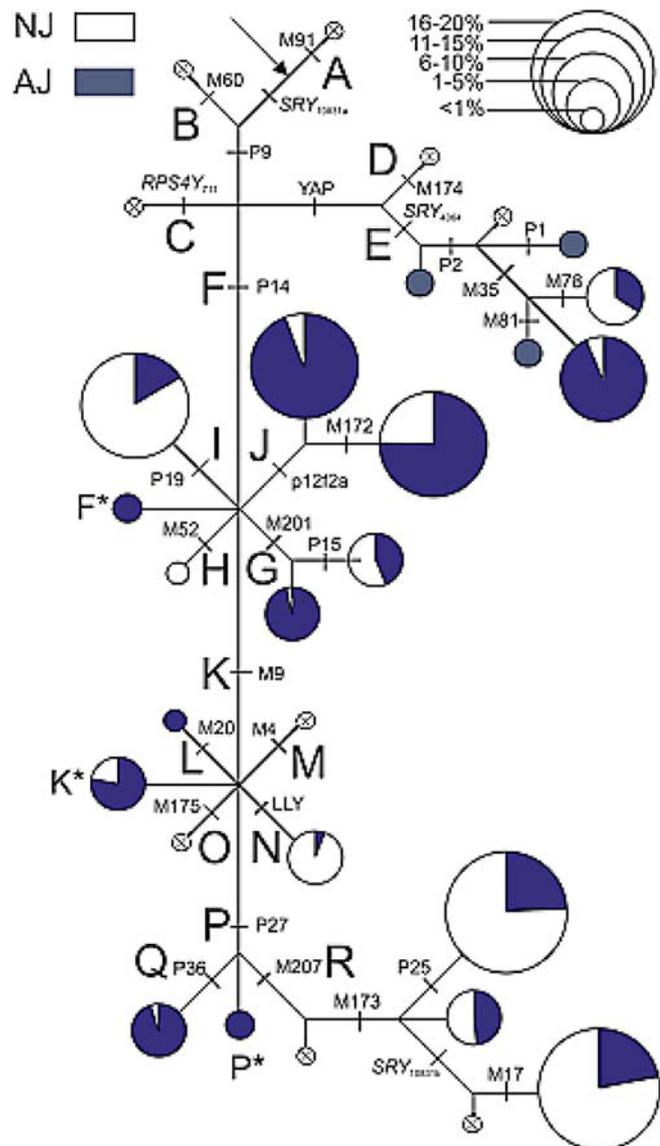


Fig. 1 Evolutionary tree for the 18 major NRY haplogroups denoted by capital letters (A–R) according to YCC (2002) recommendations. The root of the tree is indicated by an arrow; cross-hatches represent mutational events defining 29 haplogroups. The names of the 32 binary markers typed in this study are shown next to cross-hatches. Haplogroups are coded in blue and white for Ashkenazi Jews and European non-Jews, respectively. Open circles with an X denote absence in our sample. Pie charts represent the frequency of occurrence of a haplogroup within Ashkenazi Jewish and non-Jewish European populations (weighted by sample size). The overall size of each pie chart corresponds to one of five frequency classes (see insert) and represents the frequency of that haplogroup in the total sample of 790 Ashkenazi Jewish and European non-Jewish Y chromosomes

cial lineage-based naming system. We opted to omit the “x” notation and parenthetical convention for the short-hand mutation-based names used throughout the text. When no further downstream markers in the YCC (2002) tree were typed, we considered the most derived marker to define a haplogroup (also see Fig. 1). Table 1 gives a complete list of the lineage-based and mutation-based names of the 20 haplogroups/paragroups found in this study.

NRY marker analysis

We have followed the recommendations of Hammer and Zegura (2002) in typing NRY binary markers that define all 18 major haplogroups on the YCC (2002) tree. We chose a set of 32 relevant binary markers to type this set of 790 AJ and NJ chromosomes: M91, M60, *SRY*₁₀₈₃₁, *RPS4Y*₇₁₁, YAP, M174, *SRY*₄₀₆₄, P1, P2, M35, M78, M81, P14, M201, P15, M52, P19, 12f2b, M172, M9, M20, M4, LLY22g, Tat, M175, P27, P36, M207, M173, M17, and P25. The genotypes for these sites were determined by allele-specific polymerase chain reaction (PCR). After determining the location of a sample on the NRY tree defined by these markers, no further typing was performed. PCR protocols for detection of these polymorphisms have been previously reported by Karafet et al. (2002) and Underhill et al. (2001). For the microsatellite analysis, ten STRs (Y-STRs: *DYS19*, *DYS388*, *DYS389I*, *DYS389II*, *DYS390*, *DYS391*, *DYS392*, *DYS393*, *DYS426*, and *DYS439*) were genotyped in two multiplex reactions following the protocol of Redd et al. (2002). PCR products were electrophoresed on a 3100 Genetic Analyzer (Applied Biosystems) by using a 36-cm array and filter set D, and fragment lengths were converted to repeat number by means of allelic ladders. The data were analyzed with Genescan (v.3.7, Applied Biosystems) and Genotyper (v.1.1, Applied Biosystems). We define *DYS389CD* as equivalent to *DYS389I*, and we define *DYS389AB* as equivalent to *DYS389II* minus *DYS389I* (Rolf et al. 1998). Y-STR information was available for 774 out of 790 chromosomes described herein.

Statistical analysis

Measures of haplogroup diversity, including the number of haplogroups (k) and Nei's h (Nei 1987), were calculated by using the software package ARLEQUIN (Schneider et al. 2002), whereas the variance in STR allele size was calculated manually in an Excel spreadsheet. We also used ARLEQUIN to perform analysis of molecular variance (AMOVA). AMOVA produces estimates of

variance components and Φ -statistics (F-statistic analogs) reflecting the correlation of haplogroup diversity at different levels of hierarchical subdivision (Excoffier et al. 1992). We performed multidimensional scaling (MDS; Kruskal 1964) on a matrix of Nei's (1987) standard genetic distances based on haplogroup frequencies using the software package NTSYS (Rohlf 1998). The computer program ADMIX1_0 (Bertorelle and Excoffier 1998) was used to estimate admixture proportions (m_y) and their standard deviations based on 1,000 bootstrap runs. To infer NRY haplogroup frequencies of the Jewish parental population (P1), we used the average haplogroup frequencies of north African, Near Eastern, Yemenite, and Kurdish Jewish samples that were previously reported by Hammer et al. (2000). To obtain NRY haplogroup frequencies of the parental European population (P2), we used the haplogroup frequencies from the current data. The approach of Shriver (1997) was followed in selecting the haplogroups exhibiting the highest frequency differential (δ) between the two parental populations for use in the admixture analyses. The set of haplogroups used in the current survey defines a more fully resolved NRY tree compared with that reported previously (Hammer et al. 2000). Therefore, for purposes of comparison, we choose our J clade haplogroups (J-12f2b* and J-M172), R-*SRY*_{10831b}, and R-P25 and compared them with the equivalent Med, 1D, and 1L haplogroups in the same Hammer et al. (2000) database.

Results

Geographic distribution of NRY haplogroups in AJ and NJ populations

A total of 20 binary NRY haplogroups was observed in the sample of 790 chromosomes, 19 of which were present in AJ populations (Fig. 1, Table 2). Only seven of these haplogroups (E-M35*, G-M201*, J-12f2b*, J-M172, Q-P36,

Table 2 Haplogroup frequencies and within-haplogroup STR diversity for Ashkenazi and European non-Jewish populations

Lineage	Ashkenazi Jews					Non-Jewish Europeans				
	Freq ^a	n^b	k^c	DC ^d	Var ^e	Freq	n	k	DC	Var
E- <i>SRY</i> ₄₀₆₄ *	0.005	2	2	–	–	0.000	0	–	–	–
E-P1	0.002	1	1	–	–	0.000	0	–	–	–
E-M35	0.161	71	36	0.507	0.305	0.011	4	4	1.000	0.556
E-M78	0.027	12	10	0.833	0.342	0.052	18	16	0.889	0.252
E-M81	0.009	4	4	1.000	0.264	0.000	0	–	–	–
F-P14*	0.009	4	4	1.000	0.558	0.000	0	–	–	–
G-M201*	0.077	33	13	0.394	0.192	0.003	1	1	–	–
G-P15	0.020	9	9	1.000	0.344	0.026	9	9	1.000	0.267
I-P19	0.041	18	12	0.667	0.542	0.204	67	53	0.791	0.581
J-12f2b*	0.190	84	36	0.429	0.442	0.011	4	4	1.000	0.517
J-M172	0.190	84	36	0.429	0.279	0.060	18	18	1.000	0.370
K-M9*	0.020	9	7	0.778	0.375	0.006	2	2	–	–
L-M20	0.002	1	1	–	–	0.000	0	–	–	–
N-LLY*	0.002	1	1	–	–	0.034	12	12	1.000	0.569
P-P27*	0.005	2	2	–	–	0.000	0	–	–	–
Q-P36	0.052	23	3	0.130	0.016	0.003	1	1	–	–
R-M173*	0.014	6	6	1.000	0.197	0.014	5	5	1.000	0.500
R-M17	0.075	33	14	0.424	0.092	0.264	91	60	0.659	0.241
R-P25	0.100	44	30	0.682	0.309	0.307	96	75	0.781	0.314

^aFrequency of haplogroup in full sample

^bNumber of individuals with haplogroup typed for all 10 Y-STRs

^cNumber of 10 locus haplotypes

^dDiscrimination capacity (k/n)

^eAverage variance in allele

R-M17, and R-P25) were present at frequencies of $\geq 5\%$ (Table 2), accounting for 84.5% of AJ chromosomes. Haplogroups J and E were by far the most prevalent haplogroups in AJ populations. Haplogroup J was present at similar frequencies in western AJ (41.1%) and eastern AJ (37.0%) populations, whereas haplogroup E-M35 was present at lower frequencies in western AJ than in eastern AJ populations (7.1% versus 19.1%, respectively). Of note, the DuJ sample was distinguished from the other AJ populations by the presence of a relatively high frequency of the R-P25 haplogroup (26.1%) and the lowest frequency of J clade chromosomes (21.7%; supplementary data).

A total of 14 haplogroups was found in the sample of 349 NJ chromosomes (Fig. 1, Table 2). Only 5 haplogroups were present at frequencies $\geq 5\%$ (E-M78, I-P19, J-M172, R-M17, and R-P25), and accounted for 88.7% of NJ chromosomes (Table 2). The three most frequent NJ haplogroups, R-P25, R-M17, and I-P19, were present at 30.7%, 26.4% and 20.4%, respectively. Interestingly, the distribution of the two most frequent NJ haplogroups differed between western and eastern populations: R-P25 was the predominant western NJ haplogroup (54.3%) and R-M17 was the major eastern NJ haplogroup (37.0%). Whereas the R-M17 haplogroup was previously shown to predominate in eastern Europe, the most frequent NRY lineages in western Europe were represented by internal nodes on the NRY tree: paragroup R-M173* (Semino et al. 2000) and “haplogroup 1” (Rosser et al. 2000), which is now known to be a complex set of paraphyletic lineages in haplogroups Q and R (YCC 2002). Here, we demonstrate that the predominant western European Y chromosome lineage is a monophyletic group marked by the P25 mutation (haplogroup R-P25) consistent with the findings of Wilson et al. (2001).

Haplogroups that exhibited the greatest frequency differences between AJ and NJ populations included E-M35*,

G-M201*, I-P19, J-12f2*, J-M172, Q-P36, R-M17, and R-P25. Four of the seven major Jewish haplogroups (E-M35*, J-12f2*, G-M201*, and Q-P36) were at much lower frequencies or virtually absent from the NJ samples (Table 2). Within haplogroup E, AJ populations had higher frequencies of the E-M35* paragroup, whereas NJ populations had a higher frequency of the derived E-M78 haplogroup.

Population-level SNP and STR diversity

AJ and NJ population diversity statistics are presented for SNP haplogroups in Table 3 and for STR haplotypes in Table 4. In general, there was higher haplogroup diversity in AJ populations than in NJ populations (Table 3). Haplogroup diversity for the entire AJ population system was 0.877 ± 0.006 compared with 0.788 ± 0.011 for European non-Jews. For six of the seven comparisons between an Ashkenazi community and its matching non-Jewish population, haplogroup diversity was statistically significantly higher in the Ashkenazi sample (Table 3). Similar patterns of diversity were observed in the STR data (Table 4). The average variance in allele size for the Ashkenazi population system was 0.960, compared with a much lower value of 0.762 for non-Jewish Europeans. Again, for each of the seven comparisons between AJ populations and their European counterparts, the average variance in allele size was higher for Ashkenazi populations. This diversity difference at the population level was statistically significant in the Wilcoxon signed rank test ($P=0.008$). Figure 2 shows that there is a positive relationship between SNP h values and the average variance in allele size for AJ and NJ populations, and that AJ populations are generally more diverse than NJ populations. Interestingly, a different pattern was observed when discrimination capacity (i.e., the number of 10-STR haplotypes divided by the sample size)

Table 3 SNP haplogroup diversity in 10 Ashkenazi Jewish (AJ) and 7 non-Jewish (NJ) European populations

Population ^a	Ashkenazi Jews			Population	Non-Jewish Europeans			Difference ^d
	n	k^b	$h^c \pm SD$		n	k	$h \pm SD$	
Global	790	20	0.882±0.004	–	–	–	–	–
AJ	442	19	0.877±0.006	NJ	348	14	0.788±0.011	*
FrJ	50	11	0.840±0.031	Fre	64	9	0.593±0.068	*
GeJ	39	13	0.884±0.032	Ger	34	8	0.718±0.071	*
NeJ	23	8	0.885±0.036	–	–	–	–	–
AuJ ^e	39	9	0.873±0.022	Aus	31	6	0.748±0.051	*
AuJ ^e	39	9	0.873±0.023	Hun	56	9	0.778±0.031	*
BeJ	40	12	0.886±0.027	–	–	–	–	–
LiJ	37	9	0.865±0.023	–	–	–	–	–
PoJ	58	12	0.852±0.030	Pol	50	7	0.644±0.063	*
RoJ	47	11	0.872±0.023	Rom	54	10	0.829±0.028	*
RuJ	54	13	0.886±0.020	Rus	59	7	0.743±0.037	*
UkJ	55	12	0.867±0.037	–	–	–	–	–

^aAbbreviations are as defined in methods and Table 1

^bNumber of haplogroups

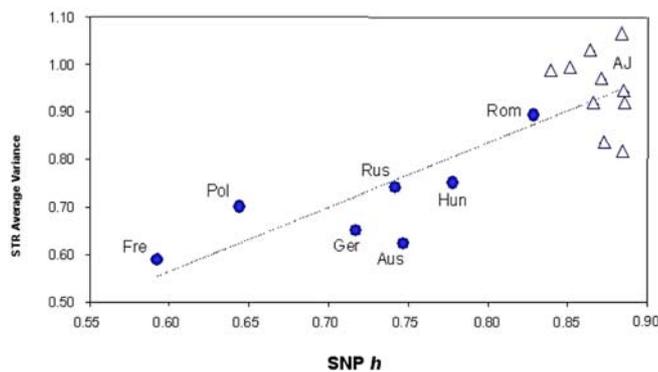
^cHaplogroup diversity (Nei 1987)

^dDifference in h between Ashkenazi Jews and non-Jewish Europeans, T test (* $P<0.001$)

^eNote that AuJ were compared once with Hun and once with Aus

Table 4 STR diversity in 10 Ashkenazi Jewish (AJ) and 7 non-Jewish (NJ) European Populations

Ashkenazi Jews					Non-Jewish Europeans				
Population	<i>n</i>	<i>k</i> ^a	DC ^b	Av var ^c	Population	<i>n</i>	<i>k</i>	DC	Av var
AJ	442	203	0.459	0.960	NJ	332	262	0.789	0.762
FrJ	50	34	0.680	0.983	Fre	48	44	0.917	0.594
GeJ	39	30	0.769	1.060	Ger	34	32	0.941	0.652
DuJ	23	21	0.913	0.812	–	–	–	–	–
AuJ ^d	39	32	0.821	0.833	Aus	31	30	0.968	0.625
AuJ ^d	39	32	0.821	0.833	Hun	56	53	0.946	0.750
BeJ	40	31	0.775	0.914	–	–	–	–	–
LiJ	37	30	0.811	1.028	–	–	–	–	–
PoJ	58	45	0.776	0.989	Pol	50	43	0.860	0.698
RoJ	47	44	0.936	0.967	Rom	54	51	0.944	0.892
RuJ	54	39	0.722	0.940	Rus	59	54	0.915	0.738
UkJ	55	40	0.727	0.915	–	–	–	–	–

^aNumber of 10 locus haplotypes^bDiscrimination capacity (*k/n*)^cAverage variance in allele size^dNote that AuJ were compared once with Hun and once with Aus**Fig. 2** Relationship between average variance in STR allele size and SNP haplogroup diversity (*h*) for seven European non-Jewish populations (solid circles) and ten Ashkenazi Jewish populations (open triangles). *Fre* French, *Pol* Polish, *Rus* Russian, *Rom* Romanian, *Ger* German, *Hun* Hungarian, *Aus* Austrian

was compared between AJ and NJ populations: all seven comparisons yielded lower haplotype diversity values for Ashkenazi Jews (Table 4).

Analysis of molecular variance

Table 5 presents variance components and Φ -statistics (Φ_{ST}) for AJ and NJ population groups for both Y-SNP and Y-STR data. When all 17 populations were considered as a single group, 86% and 94.3% of the total variance was found within populations for Y-SNPs and Y-STRs, respectively. The lower Φ_{ST} for Y-STRs is expected as a consequence of the elevated mutation rates and the stepwise mode of STR evolution that lower Φ_{ST} by increasing within-group variance (Jin and Chakraborty 1995). The SNP-based Φ_{ST} of 0.14 compares with a value of ~0.40 for global NRY datasets (Hammer and Zegura 2002). As expected, the among-group component of variance increased when AJ

Table 5 Analysis of molecular variance (AMOVA)

Grouping ^a	Number of populations	Number of groups	Within populations		Among populations Within groups		Among groups	
			Variance (%)	Φ_{ST} ^a	Variance (%)	Φ_{SC}	Variance (%)	Φ_{CT}
SNP haplogroups								
All populations	17	1	86.0	0.140				
AJ/NJ	17	2	78.8	0.212	3.5	0.043	17.6	0.176
WAJ/EAJ	10	2	96.4	0.036	0.8	0.008 ns	2.8	0.028
WNJ/ENJ	7	2	90.9	0.091	4.8	0.051	4.3	0.043
STR haplotypes								
All populations	17	1	94.3	0.056				
AJ/NJ	17	2	91.5	0.085	2.0	0.021	6.5	0.065
WAJ/EAJ	10	2	98.4	0.016	0.2	0.002 ns	1.4	0.014
WNJ/ENJ	7	2	94.4	0.056	2.2	0.022	3.5	0.035

^aAll Φ -statistics *P* values are ≤ 0.02 , except those indicated by *ns* (not significant)

and NJ populations were placed in different groups. For both Y-SNPs and Y-STRs, a good deal of this variation was partitioned among major groupings (i.e., with a minor fraction of the between-group genetic variation partitioned among-populations within groups). When AJ populations were subdivided into western and eastern groups, a small, but statistically significant, fraction of the between-group variance was partitioned between WAJ and EAJ populations, with no variation being found among-population within groups. When NJ populations were subdivided into western and eastern groups, roughly equal amounts of between-group variation were partitioned among-groups and among-populations within groups. The same patterns were observed for both Y-SNP and Y-STR data.

Multidimensional scaling

An MDS plot based on binary haplogroup frequencies and Nei's standard genetic distances for all 17 AJ and NJ populations is shown in Fig. 3. NJ populations formed two distinct clusters with western NJ populations in the upper-right and eastern NJ populations in the lower-right parts

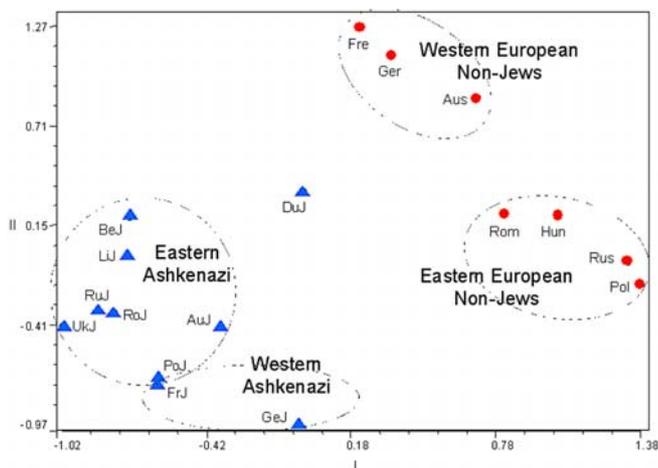


Fig. 3 MDS plot of ten Ashkenazi Jewish and seven non-Jewish European populations based on Nei's standard genetic distances for SNP haplogroups. *FrJ* French Jews, *GeJ* German Jews, *DuJ* Dutch Jews, *AuJ* Austro-Hungarian Jews, *BeJ* Byelorussian Jews, *LiJ* Lithuanian Jews, *PoJ* Polish Jews, *RoJ* Romanian Jews, *RuJ* Russian Jews, *UkJ* Ukraine Jews. Other labeling is as in Fig. 2

of the plot. Seven of the eight AJ populations formed a cluster in the lower left side of the plot. The Dutch Jews were intermediate between the AJ and the western NJ clusters. This is consistent with moderate levels of gene flow of non-Jewish Dutch Y chromosomes into the Dutch Jewish population (see below). In particular, the Dutch Jewish population had a relatively high frequency of the R-P25 haplogroup, which predominates in western European non-Jews.

Admixture estimates

Table 6 shows the haplogroups with the highest frequency differentials between European non-Jewish and non-Ashkenazi Jewish (Hammer et al. 2000) parental populations (see above) and a summary of the admixture estimates for AJ populations. Among the western AJ populations, haplogroups J-12f2b* and R-P25 were the most diagnostic for distinguishing the parental Jewish (P1) and the parental western NJ European population (P2W) components. Among the eastern AJ populations, haplogroups J-12f2b* and R-M17 were the most diagnostic for distinguishing the parental Jewish (P1) and the parental eastern NJ European population (P2E) components. All other haplogroups had δ values below 20% (data not shown). When these diagnostic haplogroups were used for analysis, the m_y value was $8.1\% \pm 11.4\%$, suggesting an even smaller contribution of European Y chromosomes to the Ashkenazi paternal gene pool than in the previous study by Hammer et al. (2000). Because of the apparently high level of admixture in Dutch Jews (m_y value of $46.0\% \pm 18.3\%$), we repeated the admixture calculation excluding the Dutch sample and found a lower estimate of admixture ($\sim 5\%$). Although not statistically significant, there was a higher level of admixture in eastern AJ versus western AJ populations. This is similar to differences in the levels of mtDNA introgression observed in western and eastern AJ populations (Behar et al. 2004).

Within-haplogroup Y-STR diversity

Table 3 lists the within-haplogroup STR diversity values (e.g., discrimination capacity and variance in allele size) for each haplogroup that is present in AJ populations, and the corresponding value for each haplogroup in European

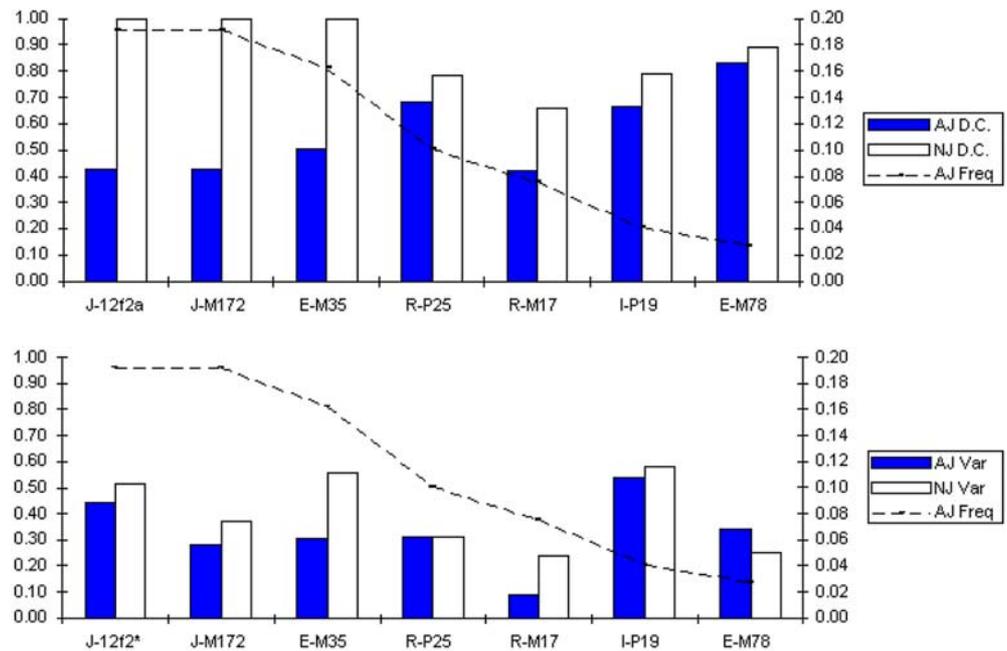
Table 6 Estimated admixture proportions (m_y) and parental haplogroup frequency differences (δ)

Population	P1	P2	Diagnostic	$\delta, \%$	m_y (diagnostic)	Bootstrap SD
All Ashkenazi	b	All NJ	J*, R1b*	30, 20	8.07	11.42
All Ashkenazi ^a	b	All NJ	J*, R1b*	31, 22	5.22	11.56
WAJ	b	All WNJ	J*, R1b*	36, 41	7.79	9.27
WAJ ^b	b	All WNJ	J*, R1b*	41, 44	-1.74	9.94
French AJ	b	French NJ	J*, R1b*	48, 52	-7.93	11.39
Dutch AJ	b	All WNJ	J*, R1b*	17, 28	46.03	18.27
EAJ	b	All ENJ	J*, R1a1*	28, 30	10.87	6.71
Polish AJ	b	Polish NJ	J*, R1a1*	44, 46	2.45	10.03

^aDutch AJ were excluded from the calculations

^bData of Hammer et al. (2000)

Fig. 4 Relative levels of Ashkenazi Jewish (*solid bar*) and European non-Jewish (*open bar*) within-haplogroup STR diversity for six SNP haplogroups occurring at >2% in the Ashkenazi population. *Top* Discrimination capacity (*D.C.*). *Bottom* Variance (*Var*) in allele size



NJ populations. The average within-haplogroup STR diversity was lower in AJ populations than in NJ populations for both discrimination capacity (0.662 versus 0.912, respectively) and variance in allele size (0.304 versus 0.417, respectively). Figure 4 provides a visual representation of the relative levels of AJ and NJ within-haplogroup diversity for haplogroups that are present at >2% in the AJ population (see Table 4 for their frequencies in NJ populations). The average variance in allele size was lower for six of the seven haplogroups shown in Fig. 4, bottom. Paragroup E-M35* and haplogroup R-M17 exhibited the greatest discrepancy in variance value, whereas haplogroup R-P25 STR variance in the AJ population was only slightly lower than that in the NJ population. The only exception to this pattern was haplogroup E-M78, which exhibited greater variance in AJ populations. All AJ haplogroups had lower discrimination capacities compared with those in NJ populations (Fig. 4, top).

It should be noted that there was no relationship between haplogroup frequency and STR variance for either the AJ population ($r^2=0.003$, $P=0.879$) or the NJ population ($r^2=0.126$, $P=0.314$). This was also true for the relationship between haplogroup frequency and discrimination capacity in the AJ population ($r^2=0.017$, $P=0.703$), but not for the NJ population ($r^2=0.881$, $P<0.001$).

Discussion

This survey of variation at 32 binary (SNP) and 10 STR markers in a sample of 442 Ashkenazi males from 10 different western and eastern Europe communities represents the largest study of Ashkenazi paternal genetic variation to date. In a previous study by Hammer et al. (2000), a set of 18 SNPs was typed in a diverse Jewish sample that included 113 Ashkenazim from the US (the European prove-

nance of these samples was unknown). This AJ sample was characterized by nine haplogroups that were also found in several other Jewish populations. Similarly, studies by Nebel et al. (2001) and Thomas et al. (2002) have included a modest Ashkenazi sample (i.e., <80 Y chromosomes) typed with a small set of SNPs (i.e., 13 and 10, respectively), and a set of six Y-STRs. The results of all three of these earlier surveys concur that Ashkenazi Jews (1) have been relatively isolated from host European non-Jewish populations, and (2) are closely related to non-Ashkenazi Jewish communities and some non-Jewish populations from the Near East. The phylogenetic resolution of these earlier studies was partly limited by the relatively small number of markers typed. For example, the identification of additional highly informative sublineages within the two most frequent Ashkenazi clades (E and J) was not possible because many recently discovered “downstream markers” were not available. The recent publication of highly congruent human Y-chromosome trees (Hammer et al. 2001; Underhill et al. 2001) and a standardized nomenclatural system for the resulting binary polymorphism-based consensus tree (YCC 2002) has provided an opportunity to understand paternal population origins, relationships, and dispersals with more phylogenetic and geographic resolution than was heretofore possible. Analyses of the higher resolution dataset presented here provide a better opportunity to infer the composition of the founding Ashkenazi paternal gene pool and to distinguish lineages that may have entered the Ashkenazi population after their arrival in Europe.

Origins of Ashkenazi NRY lineages

Based on the frequency and distribution of the 20 haplogroups observed in AJ and NJ populations, we subdi-

vided Ashkenazi Jewish lineages into the following three categories: major founder haplogroups, minor founder haplogroups, and shared haplogroups. The first two categories include those haplogroups likely to be present in the founding Ashkenazi population (and that now occur at high and low frequency, respectively). The latter category is comprised of haplogroups that either entered the Jewish gene pool recently as the result of introgression from European host populations, and/or that were present in both European and Jewish populations before the dispersal of ancestral Ashkenazim into Europe. We acknowledge that such categorization is complicated because current haplogroup distributions are the culmination of many past events. For example, haplogroups such as R-M17 and R-P25 that predominate in European populations today (see below) may have also been present in the Near East as part of the ancestral AJ gene pool. Similarly, haplogroups that predominate in AJ may have entered the European gene pool before AJ populations dispersed into Europe.

Paragroup EM35* and haplogroup J-12f2a* fit the criteria for major AJ founding lineages because they are widespread both in AJ populations and in Near Eastern populations, and occur at much lower frequencies in European non-Jewish populations. Because they have similar distributions as these major founder lineages, albeit at lower frequencies, we suggest that haplogroups G-M201 and Q-P36 are minor AJ founding lineages. Although J-M172 is also found at high frequency in AJ populations (and probably migrated to Europe with the original founding Ashkenazi population), its presence in European non-Jews at a frequency of 6% may reflect a more complicated history of migration to Europe (i.e., both before and during the Jewish Diaspora). This migration may have been mediated either by the diffusion of Neolithic farmers from the Near East between 4,000 and 7,500 years ago (Semino et al. 2000) or by sea-faring peoples in the Mediterranean region (Mitchell and Hammer 1996). Interestingly, M35* chromosomes (E3b*; or their evolutionary precursors E* and E3*) were previously hypothesized to have migrated to Europe with farmers in the Neolithic (Hammer et al. 1997; Rosser et al. 2000; Semino et al. 2000). However, because M35* chromosomes are rare in Europe, we instead hypothesize that the derived lineage, E-M78 (E3b1), is the more likely haplogroup reflecting Neolithic demic diffusion. Similarly, we suggest that G-P15 with its better representation in Europe, rather than its evolutionary precursor G-M201 (which is found mainly in AJ populations), is a better candidate marker for Neolithic migrations of farmers into Europe.

The best candidates for haplogroups that entered the AJ population recently via admixture include I-P19, R-P25, and R-M17. These haplogroups are thought to represent the major Paleolithic component of the European paternal gene pool, expanding from refugia populations after the Last Glacial Maximum more than 10,000 years ago (Rosser et al. 2000; Semino et al. 2000). Because haplogroups R-M17 and R-P25 are present in non-Ashkenazi Jewish populations (e.g., at 4% and 10%, respectively) and in non-Jewish Near Eastern populations (e.g., at 7% and 11%, re-

spectively; Hammer et al. 2000; Nebel et al. 2001), it is likely that they were also present at low frequency in the AJ founding population. The admixture analysis shown in Table 6 suggests that 5%–8% of the Ashkenazi gene pool is, indeed, comprised of Y chromosomes that may have introgressed from non-Jewish European populations. In particular, the Dutch AJ population appears to have experienced relatively high levels of European non-Jewish admixture. This is apparent in the MDS plot and by virtue of their elevated frequencies of haplogroups R-P25 (>25%) and I-P19 (>10%). These results are not surprising in view of the longstanding religious tolerance in this region. However, Dutch Jews do not appear to have increased levels of European mtDNA introgression (Behar et al. 2004), suggesting that admixture in this population is mainly the result of higher rates of intermarriage between Jewish women and non-Jewish men.

Diversity and bottleneck effects in the Ashkenazi paternal gene pool

The results presented here demonstrate that AJ populations have high levels of SNP diversity compared with their NJ counterparts. This is similar to the results of Thomas et al. (2002) who have interpreted elevated NRY diversity as evidence that AJ populations did not suffer a bottleneck. However, there are contrasting patterns when different measures of STR diversity are taken into account. By measures of average heterozygosity (Nei's h) and variance in allele size at the population level, AJ populations have higher diversity than NJ populations. On the other hand, by a measure of allelic diversity (e.g., discrimination capacity), we see a reduced number of alleles (haplotypes) per sample compared with NJ populations. We have also observed a reduced number of haplotypes within AJ haplogroups, and reduced within-haplogroup variance in allele size and heterozygosity. What do these contrasting patterns tell us about the possible role of a bottleneck in the AJ population?

When a population descends from only a small number of individuals, either because the population is initiated from a small number of founders (founder effect) or because a small number of individuals have survived in a particular generation (bottleneck), the process of genetic drift can be accentuated, leading to a reduced effective size of the population. For simplicity, we have referred to both of these processes as bottlenecks. Bottlenecks can lead to highly altered allelic frequencies relative to those in the ancestral population and to lower heterozygosity and fewer alleles. Because of their history of dispersal and endogamy for many generations, Ashkenazi Jews have probably experienced a strong reduction in effective population size. Evidence of accentuated genetic drift comes from the distribution of several recessive disease alleles that are found at unusually high frequencies in AJ populations, and from mtDNA data showing highly reduced diversity in Ashkenazi populations (Behar et al. 2004). One of the goals of the research presented here was to seek ev-

idence for the effects of a bottleneck in the paternally-inherited haploid compartment of the genome, the NRY.

When a highly variable locus such as an STR goes through a bottleneck, the loss of alleles occurs more rapidly than the loss of heterozygosity (Hedrick 2000). For instance, after a 100-fold bottleneck, the number of alleles at a highly polymorphic locus is expected to decline by ~65%, whereas heterozygosity at this locus is expected to change by less than 10% (data not shown). In addition, the probability of ancestral polymorphism remaining in a founder population depends on the founder size and the allele frequency at each locus. For low frequency alleles, there is a higher probability that they will not survive a bottleneck. This has important implications for understanding the dynamics of the survival of binary SNP alleles compared with multiallelic STR alleles, especially when considered together on non-recombining systems such as haplogroups or haplotypes. For example, the mean frequency of a SNP haplogroup (i.e., of the 19 found in our AJ sample) is 5%, whereas the mean frequency of a Y-STR haplotype (i.e., of the 203 present in our AJ sample) is 0.5%. Thus, during bottlenecks, STR haplotypes are expected to be lost at a much higher rate than SNP haplogroups. This is not to say that bottlenecks have no effect on moderately frequent SNP haplogroups; the main effect of genetic drift on these lineages will be to alter their frequencies relative to the ancestral population. Again, the loss of rare haplotypes will tend to reduce allelic diversity before there is a noticeable loss in heterozygosity.

Another important point to consider is the finding that Y-STR variation is typically structured more by NRY SNP haplogroup than by population (Bosch et al. 1999). In other words, SNP haplogroups usually represent genealogical units older than human subpopulations. This is also evident in the two population systems examined here. For example, AMOVA analysis indicated that only 1%–4% of the total STR variance is partitioned among AJ or NJ subpopulations, respectively, whereas 47% and 39% of the total STR variance was partitioned among the 19 and 11 SNP haplogroups present in AJ and NJ populations, respectively (data not shown). This means that a major portion of the STR diversity within a single AJ or NJ population represents between-haplogroup variation (that is older than the population itself) rather than between-population variation. Thus, estimates of population diversity based on Y-STRs may be fairly insensitive to reductions in effective population size, unless a bottleneck is so severe as to eliminate most SNP haplogroup variation. Instead, the diversity within SNP haplogroups may be a better indicator of recent reductions in population size resulting from bottlenecks. It should be noted that the Y-STR-based estimate of population diversity is itself dependent upon the level of molecular resolution of the haplogroups. Thus, for example, further binary site resolution within the J-M172 clade might yield separate haplogroups within which Y-STR diversity will be even lower. This is suggested by the presence of two seemingly different Y-STR founding haplotypes separated by four mutational steps within the J-M172 clade (data not shown).

Despite Ashkenazi Jews representing a recently founded population in Europe, they are probably derived from a large and diverse ancestral source population in the Near East, a population that may have been larger than the source population from which European non-Jews derived. This is consistent with our current finding that AJ populations contain higher levels of SNP and STR diversity than European NJ populations. The reduced allelic diversity within AJ haplogroups observed in AJ versus NJ populations and the reduced variance in allele size may be the signature of a bottleneck in AJ population history. A similar approach to identifying pronounced bottlenecks in Italian and Greece populations was recently taken by Di Giacomo et al. (2003). A prediction of this model is that Ashkenazi founder haplogroups should show more reduced within-haplogroup STR variation than shared haplogroups that introgressed more recently into the AJ population from European non-Jewish host populations. This prediction assumes that introgression of host Y chromosomes is random. Indeed, the diversity patterns of the two major founder lineages (E-M35* and J-12fa) and that of J-M172 support this prediction: they exhibit the lowest ratio of AJ to NJ within-haplogroup diversity (Table 2, Fig. 4). This is also evident in Fig. 5 where there is a statistically significant negative correlation between the ratio of AJ/NJ discrimination capacity and the ratio of AJ/NJ haplogroup frequency. Although the minor founder haplogroups G-M201 and Q-P36 are not shown in Fig. 5, because their frequencies were too low in European non-Jewish populations, they exhibit the two lowest discrimination capacity values and two of the three lowest variances in allele size. Haplotype diversity (h , Nei 1987) for Q-P36 is 0.30 (data not shown), a value that is much lower than that calculated (0.47) for the predominant Roma haplogroup H-M82. Gresham et al. (2001) concluded that this low h value is indicative of a profound bottleneck effect in the Roma population. It is also interesting that the admixed haplogroups I-P19 and R-P25 show almost no reduction in discrimination capacity in Fig. 5, as predicted by the bottleneck hypothesis (also see Fig. 4).

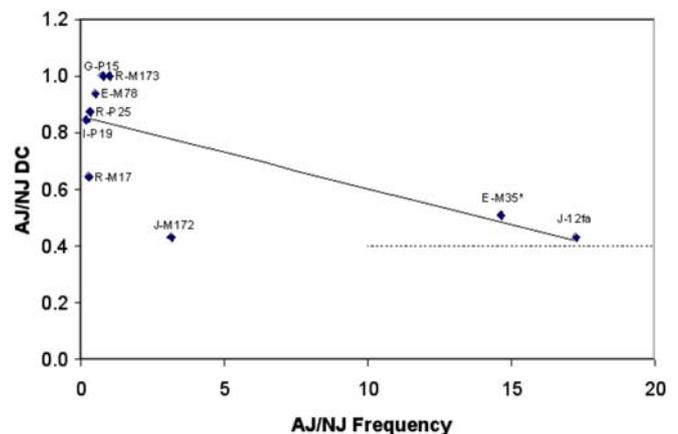


Fig. 5 Scatter plot of relative Ashkenazi Jewish/European non-Jewish haplogroup frequency (X axis) versus within-haplogroup discrimination capacity (Y axis) for nine SNP haplogroups

The current study provides an illustration of the importance of detailed combined analyses of both SNP-based haplogroups and associated allelic diversity at Y-STRs to obtain evidence for a demographic bottleneck effect, an effect that escaped attention in earlier studies based on simpler analyses of Ashkenazi NRY diversity. Whereas the results presented here are consistent with a bottleneck effect on the paternal lineage leading to modern AJ populations, further studies are still needed to estimate the timing and magnitude of such an event. Estimation of these demographic parameters is particularly important to elucidate the role of fluctuating historical demography in producing the excess of high frequency, deleterious, recessive mutations found among the AJ. Robust estimation of the timing and magnitude of a putative bottleneck requires knowledge of ancestral allele frequencies, which may be estimated from extant large Near Eastern populations. Such future lines of investigation promise to yield insights into the impact that non-equilibrium demographic histories can have on the genetic health of populations.

Acknowledgements We thank Dr. Marc-Alain Levy for help in collecting samples from the French Rhine Valley and Dr. Istvan Mucsi for donating the Hungarian samples. This work was supported by a grant from the National Institute of General Medical Sciences (GM53566-06) to M.H., a F.I.R.S.T. award grant from the Israeli Science Foundation to K.S., and kind donations from the Milin Charitable Foundation and the Jerome Tankin Foundation.

References

- Behar DM, Thomas MG, Skorecki K, Hammer MF, Bulygina E, Rosengarten D, Jones AL, Held K, Moses V, Goldstein D, Bradman N, Weale ME (2003) Multiple origins of Ashkenazi Levites: Y chromosome evidence for both Near Eastern and European ancestries. *Am J Hum Genet* 73:768–779
- Behar DM, Hammer MF, Garrigan D, Villemis R, Bonne-Tamir B, Richards M, Gurwitz D, Rosengarten D, Kaplan ME, DellaPerdola S, Quintana-Murci L, Skorecki K (2004) mtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *Eur J Hum Genet* (in press)
- Bertorelle G, Excoffier L (1998) Inferring admixture proportions from molecular data. *Mol Biol Evol* 15:1298–1311
- Bonne-Tamir B, Adams A (1992) Genetic diversity among Jews. Oxford University Press, Oxford
- Bosch E, Calafell F, Santos FR, Perez-Lezaun A, Comas D, Benchemsi N, Tyler-Smith C, Bertranpetit J (1999) Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am J Hum Genet* 65:1623–1638
- Chakravarti A, Chakraborty R (1978) Elevated frequency of Tay-Sachs disease among Ashkenazic Jews unlikely by genetic drift alone. *Am J Hum Genet* 30:256–261
- Di Giacomo F, Luca F, Anagnou N, Ciavarella G, Corbo RM, Cresta M, Cucci F, Di Stasi L, Agostiano V, Giparaki M, Loutradis A, Mammi C, Michalodimitrakis EN, Papola F, Pedicini G, Plata E, Terrenato L, Tofanelli S, Malaspina P, Novelletto A (2003) Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. *Mol Phylogenet Evol* 28:387–395
- Diamond JM (1994) Human genetics. Jewish lysosomes. *Nature* 368:291–292
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Goldstein DB, Reich DE, Bradman N, Usher S, Seligsohn U, Peretz H (1999) Age estimates of two common mutations causing factor XI deficiency: recent genetic drift is not necessary for elevated disease incidence among Ashkenazi Jews. *Am J Hum Genet* 64:1071–1075
- Goodman RM (1979a) Genetic disorders among the Jewish people. Johns Hopkins University Press, Baltimore
- Goodman RM (1979b) A perspective on genetic diseases among the Jewish people. In: Goodman RM, Motulsky AG (eds) Genetic diseases among Ashkenazi Jews. Raven, New York, pp 1–17
- Gresham D, Morar B, Underhill PA, Passarino G, Lin AA, Wise C, Angelicheva D, Calafell F, Oefner PJ, Shen P, Tournev I, Pablo R de, Kucinkas V, Perez-Lezaun A, Marushiakova E, Popov V, Kalaydjieva L (2001) Origins and divergence of the Roma (gypsies). *Am J Hum Genet* 69:1314–1331
- Hammer MF, Zegura SL (2002) The human Y chromosome haplogroup tree: nomenclature and phylogeography of its major divisions. *Annu Rev Anthropol* 31:303–321
- Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, Zegura SL (1997) The geographic distribution of human Y chromosome variation. *Genetics* 145:787–805
- 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 U S A* 97:6769–6774
- 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
- Hedrick PW (2000) Genetics of populations. Jones and Bartlett, Sudbury
- Jin L, Chakraborty R (1995) Population structure, stepwise mutations, heterozygote deficiency and their implications in DNA forensics. *Heredity* 74:274–285
- Jorde LB (1992) Population diseases in the Ashkenazi population: evolutionary considerations. In: Bonne-Tamir B, Adams A (eds) Genetic diversity among Jews. Oxford University Press, New York, pp 305–312
- Karafet TM, Osipova LP, Gubina MA, Posukh OL, Zegura SL, Hammer MF (2002) High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum Biol* 74:761–789
- Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29:1–27
- Mitchell RJ, Hammer MF (1996) Human evolution and the Y chromosome. *Curr Opin Genet Dev* 6:737–742
- Nebel A, Filon D, Brinkmann B, Majumder PP, Faerman M, Oppenheim A (2001) The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. *Am J Hum Genet* 69:1095–1112
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
- Ostrer H (2001) A genetic profile of contemporary Jewish populations. *Nat Rev Genet* 2:891–898
- Redd AJ, Agellon AB, Kearney VA, Contreras VA, Karafet T, Park H, Knijff P de, Butler JM, Hammer MF (2002) Forensic value of 14 novel STRs on the human Y chromosome. *Forensic Sci Int* 130:97–111
- Risch N, Leon D de, Ozelius L, Kramer P, Almasy L, Singer B, Fahn S, Breakefield X, Bressman S (1995) Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. *Nat Genet* 9:152–159
- Risch N, Tang H, Katzenstein H, Ekstein J (2003) Geographic distribution of disease mutations in the Ashkenazi Jewish population supports genetic drift over selection. *Am J Hum Genet* 72:812–822

- Rohlf FJ (1998) NTSYS-pc: numerical taxonomy and multivariate analysis system. Setauket, release 2.02H, NY
- Rolf B, Meyer E, Brinkmann B, Knijff P de (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, Hurler ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, Knijff P de, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526–1543
- Schneider S, Roessli D, Excoffier L (2002) Arlequin: a software for population genetic analysis. Release 1.1, Geneva
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290:1155–1159
- Shriver MD (1997) Ethnic variation as a key to the biology of human disease. *Ann Intern Med* 127:401–403
- 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
- Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Goldstein DB (1998) Origins of Old Testament priests. *Nature* 394:138–140
- Thomas MG, Weale ME, Jones AL, Richards M, Smith A, Redhead N, Torroni A, Scozzari R, Gratrix F, Tarekegn A, Wilson JF, Capelli C, Bradman N, Goldstein DB (2002) Founding mothers of Jewish communities: geographically separated Jewish groups were independently founded by very few female ancestors. *Am J Hum Genet* 70:1411–1420
- 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
- Weinryb BD (1972) *The Jews of Poland: a social and economic history of the Jewish community in Poland from 1100 to 1800*. Jewish Publication Society of America, Philadelphia
- Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB (2001) Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci USA* 98:5078–5083
- YCC (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12:339–348