ARTICLE

MtDNA evidence for a genetic bottleneck in the earl histor of the Ashkena i Jewish population

Doron M Behar[1], Michael F Hammer*,[2], Daniel Garrigan[2], Richard Villems[3], Batsheva Bonne-Tamir[4], Martin Richards[5]

$\sim$500 years, there was a period of rapid growth[3] culminating in an estimated population size of $\sim$8 million Ashkenazi Jews at the outbreak of the Second World War.[4]

tions. In our analysis, we take advantage of the ability to infer evidence for maternal population bottlenecks on the basis of comparative estimates of mtDNA sequence diversity.[13]

Subjects and methods

Samples

Blood or buccal swab samples were collected with informed consent from 565 unrelated individuals (404 male and 161 female subjects) 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 mother, grandmother, and in most cases, great grandmother. The samples were then classified into the follow-

in Ashkenazi Jews. To facilitate visual inspection of Hg frequency differences among populations, these 56 Hgs are combined into 28 Hgs as indicated in Figure 1. The most prevalent Ashkenazi Hgs were K (32%), H (21%), N1b (10%), and J1 (7%), followed by other Hgs at minor frequencies (supplementary material). Haplogroup diversity in the combined Ashkenazi sample, as well as for each community, is lower than that for either the European or Near Eastern non-Jewish populations (data not shown).

The high frequency of Hg K among the Ashkenazi Jews contrasts with its much lower frequency (~6%) reported in both Near Eastern and European non-Jews.[14] A notable exception among Near Eastern non-Jews is the Druze, with a Hg K frequency of 16%.[14,15] Based on HVS-1 sequences, there are four major subtypes of Hg K in the Ashkenazi Jewish sample. These subtypes and their frequencies within Hg K are: 223-224-234-311 (33%); 224-234-311 (24%); 093-224-311 (19%); and 224-311 (16%). Like the ancestral 224-311 haplotype of Hg K, its one-step derivative haplotype 093-224-311 is widespread (at low to moderate frequency) in both Near Eastern and European non-Jews. In contrast, the 224-234-311 haplotype is rare among European and Near Eastern non-Jewish populations. A single direct match was found in a sample of Palestinian Arabs.[14] However, HVS-1 sequence matches occur elsewhere at very low frequency in both Europe and the Near East, and therefore, its provenance remains unclear. Interestingly, the most frequent Ashkenazi haplotype, 223-224-234-311 (present at 10.6% in our Ashkenazi sample), is virtually absent in European and Near Eastern non-Jewish populations.[14] In Figure 1, we have subdivided these four Hg K haplotypes into two classes: K-Common includes those Hg K types that are geographically widespread (224-311 and 093-224-311), while K-Ashkenazi includes the two types (224-234-311 and 223-224-234-311) that are almost entirely restricted to Ashkenazi communities. The high frequency of haplotypes in Ashkenazi populations that are rare or absent in other populations may be an indication of the effects of strong genetic drift acting on the Ashkenazi population.

Evidence of the effects of accentuated genetic drift is also visible in the shape of gene trees and patterns of diversity within individual Ashkenazi mtDNA Hgs. When we perform median-joining network analysis[20] using HVS-1 sequence variation, the topologies of both the European and Near Eastern non-Jewish Hg K HVS-1 networks are star like (with a single most frequent haplotype), consistent with a history of population expansion from small initial size (Figure 2a, b, respectively). In contrast, the topology of Ashkenazi Hg K haplotypes is not star like, with the aforementioned four haplotypes occurring at intermediate frequency (Figure 2c). In addition, the second most frequent Hg found in Ashkenazi populations, Hg H (Figure 1), exhibits elevated frequencies of the derived nonsingleton tip nodes (Figure 2d), in contrast to the usual

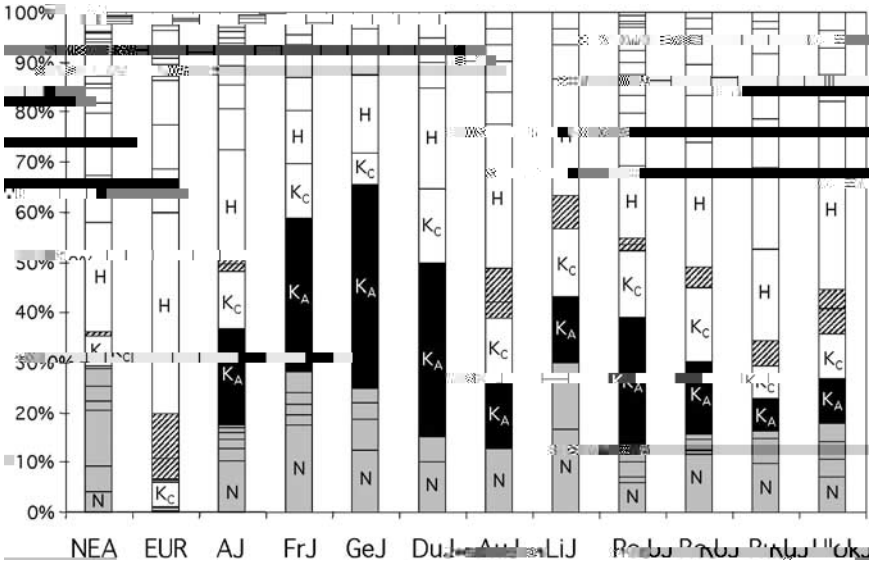

Figure 1 MtDNA Hg frequencies in Ashkenazi Jewish (AJ) and non-Jewish populations from Europe (EUR) and the Near East

‘star’ pattern of singletons at derived tip nodes seen in the Hg H networks of European and Near Eastern non-Jewish populations (data not shown).

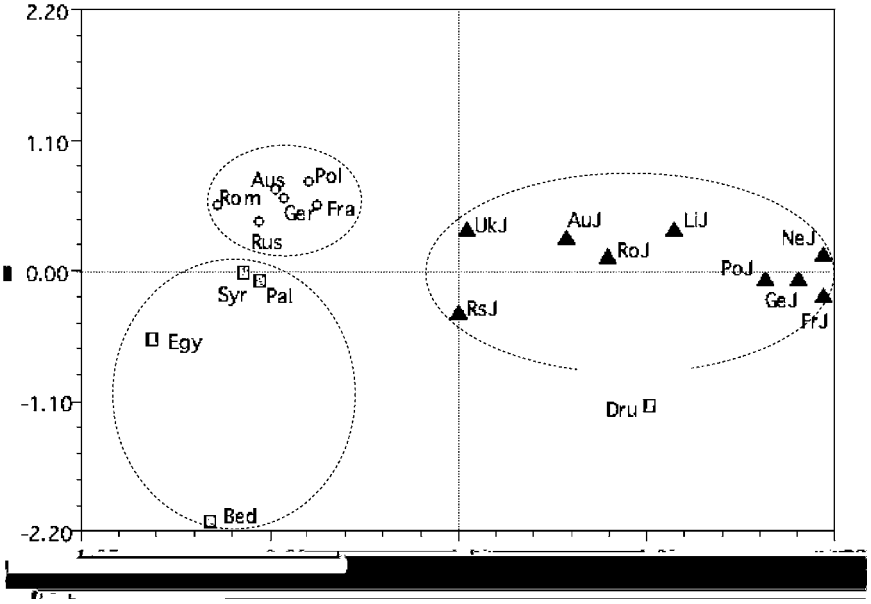

Figure 3 Multidimensional scaling plot for nine Ashkenazi (triangles), six European (open circles), and five Near Eastern (squares) non-Jewish populations based on mtDNA Hg frequencies. The 'stress' value (ie, the goodness of fit between the distances in the graphic configuration and the monotonic function of the original distances) is 0.08. For three letter population codes, see Table 1.

compared with both European and Near Eastern non-Jewish populations. The use of a two-tailed -test revealed that the diversity measure for all Ashkenazi Jewish mtDNAs is significantly lower than the corresponding diversity measures for European and Near Eastern non-Jewish mtDNAs (<0.0001). Nucleotide diversity (π) does not differ significantly among the groups (Table 1). This discrepancy results from a reduced number of HVS-1 haplotypes within Hgs that are fairly differentiated at the DNA sequence level.

The frequency distribution of polymorphisms also differs between Ashkenazi Jewish and non-Jewish populations. Table 1 demonstrates that nearly all of the Near Eastern and European non-Jewish populations yield significantly negative $_S$ values, consistent with an excess of rare mutations. Interestingly, all three western Ashkenazi $_S$ values are not statistically significant (ie, at the $\alpha=0.02$ level), and two of the six Eastern Ashkenazi values are not nearly as negative as those of other populations in Table 1.

Mismatch distributions

Visual inspection of the distribution of pairwise nucleotide differences reveals considerable insight into the recent demographic history of the Ashkenazi Jewish population. Figure 4a contrasts the mismatch distributions of large Near Eastern populations (Egyptians, Palestinians, and Syrians) with that of the Ashkenazi Jews. The means of both distributions (τ) are nearly identical ($\tau=5.5$ for Ashkenazi Jews and $\tau=6.2$ for Near Easterners), which reflects a shared period of ancient population growth (approximately 46 000–52 000 years ago). The Near Eastern estimate of τ was used in the coalescent simulations to generate $_0$ and $_1$ values under the null hypothesis of no population bottleneck and Pleistocene population growth. Despite the similarity of the τ estimates for the Ashkenazim and Near Eastern populations, the Ashkenazi genealogy harbors an increased rate of recent coalescent events, as evidenced by an increase in $_0$ and $_1$ of the mismatch distribution.[14] The 95% confidence interval for $_0$ in the Near Eastern dataset is 0.008–0.020, while the observed Ashkenazi Jewish value is 0.043. Similarly, the 95% confidence interval for the Near Eastern dataset $_1$ is 0.007–0.049, the Ashkenazi Jewish $_1=0.051$.

Figure 4b shows the mismatch distributions for both Eastern and Western Ashkenazi Jewish populations, which are identical, except for elevated $_0$ and $_1$ in the West. Figure 4c contrasts the Ashkenazi Jewish mismatch distribution with that of the Druze, a Near Eastern population with low effective size.[26] The Druze τ of 4.6 is similar to that of the Ashkenazi Jews, again suggesting a period of shared growth. However, the Druze show an increase in $_0$, but not $_1$, as do the Ashkenazi Jewish populations. Lastly, Figure 4d juxtaposes the Ashkenazi Jewish mismatch distribution with that of the European non-Jewish population. For European non-Jews, $\tau=3.4$, indicating a much more recent onset of population growth (approximately 20 000–25 000 years ago).

Coalescent simulations performed under the null hypothesis of no population bottleneck ($=1$) did not

produce a single replicate with values of $_0$ higher than that observed in the Ashkenazi data (<0.001). Similarly, simulations of a constant size model only produced 19 replicates with a value of $_1$ higher than that observed for the Ashkenazi data ($=0.019$). Based upon the assumption

Valley in the 8th century A.D., and (4) in the 12th century A.D., when migrations took place from western to eastern Europe. In addition, endogamy in combination with >100-fold population growth in the last 500 years[27] undoubtedly played a role in shaping patterns of variation in the Ashkenazi gene pool. While several authors posited that the high frequency of genetic conditions, such as Tay-Sachs disease, is the result of heterozygote advantage,[5,28–30] others have argued for an important role of genetic drift.[6,26] For example, Risch .[6] proposed that founder effects resulting from the dynamics of population growth in the 16–19th 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 was supported by the inference of a recent age of the single founder mutation (~350 years) that causes early-onset idiopathic torsion dystonia.[6] The much older estimated age of the factor XI type II mutation (~3000 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 particular to the Ashkenazi population.[7] This raises the possibility of either an ancient bottleneck in a population ancestral to the major Jewish groups, and/or positive selection on heterozygotes.[7] Comparisons of 'neutral' genealogies are particularly useful for drawing inferences regarding demographic history, and DNA variation at the uniparentally inherited regions of the genome are particularly sensitive to bottlenecks.[8] While mtDNA and the NRY cannot provide a complete picture of population history by themselves, the large database of mtDNA HVS-1 sequences from dozens of human populations and a large body of theoretical work on the effects of population expansion/contraction on mtDNA sequence variation[24,31] provide a comparative framework for inferring the demographic history of the Ashkenazi population. Population expansions and bottlenecks are known to shape patterns of variation in the genome in specific ways. After a period of time, rapid population growth from a small initial size is expected to produce a preponderance of recently derived low-frequency polymorphisms that often result in statistically significant negative values of the $_S$[18] statistic and unimodal peaks (waves) in mtDNA mismatch distributions with a near absence of zero and one mismatches among haplotypes.[19,32,33] Studies of mtDNA HVS-1 sequence diversity demonstrate that most human populations show evidence of remote (Pleistocene) expansion, and that such demographic signals in patterns of DNA variation may be lost due to recent bottlenecks.[13] For example, nonsignificant $_S$ values observed in some hunter-gatherer populations have been explained by recent bottlenecks that may have occurred in response to the Neolithic expansions of food-producing populations. Recently, Cordeaux .[25] demonstrated that several southern Indian tribal populations exhibit higher frequencies of the zero and one classes of mismatch distributions and more positive $_S$ values than do northern populations. These authors inferred that southern tribal groups experienced enhanced genetic drift because of small population sizes and/or bottlenecks that erased the signature of Pleistocene population expansion. Here, we demonstrate that the Ashkenazi population also exhibits the signature of reduced population size as seen in the higher $_0$ and $_1$ mismatch class frequencies and Fu's $_S$ values. However, the magnitude of the bottleneck in these populations was not so severe as to entirely erase the earlier signal of Pleistocene population expansion[34] and the shared history of population growth is clearly demonstrated by comparing the Ashkenazi Jewish mismatch distribution with those of Near Eastern populations

slower rate. The chosen rate places the time of the Pleistocene expansion at 50 000 years before the present, a time that is consistent with other estimates from both autosomal and mitochondrial data.[23,36–38]

We may also infer the effects of bottlenecks on some specific Near Eastern non-Jewish populations who are known to have small effective sizes such as the Druze (Figure 4c).[39] This suggests the possibility that contemporary Ashkenazi mtDNA diversity may derive, in part, from a small and subdivided ancestral mtDNA gene pool, and is consistent with the hypothesis that some high frequency disease alleles originated before the separation of Jewish communities in the Near East.[40,41] Indeed, estimates of the age of mutations causing Ashkenazi genetic diseases range from recent times (ie, during demographic upheavals within Europe in the past 500 years),[6,26,40,42] to times when ancestral Ashkenazi populations were first migrating to and within Europe,[43] to times before Jewish populations migrated out of the Near East.[40,44,45]

Conclusions

The combined mtDNA and disease mutation data suggest that Ashkenazi Jewish populations experienced a long period of accentuated genetic drift marked by an early bottleneck, perhaps beginning in the Near East. Prolonged periods of low effective population size can lead to the accumulation of slightly deleterious mutations throughout the genome.[46] Small founder populations derived from large ancestral populations are not always capable of purging these deleterious mutations. This may be the ultimate cause of the segregation of disease mutations in Ashkenazi Jews. However, this explanation does not preclude more proximal causes for the increase in frequency of disease mutations, such as those hypothesized by Risch .,[7] unequal contribution of a particular segment of the Ashkenazi Jewish community to the explosive population growth occurring in the Pale of Settlement approximately 25 generations ago. Low effective size may have enabled deleterious mutations to become established in the Jewish population, while the recent growth of affected segments of the community amplified these mutations to frequencies sufficiently high to form homozygotes.

Acknowledgements

I

I

(I

References

1 Ben-Sasson HH: . Cambridge: Harvard University Press; 1976.

2 Goodman RM: . Baltimore: The Johns Hopkins University Press; 1979.

3 Weinryb BD: . Philadelphia: The Jewish Publication Society of America; 1972.

4 Dellapergola S: Major demographic trends of world Jewry: the last hundred years. in Adam A (ed): . New York: Oxford University Press; 1992; **Chapter 1**: pp 3–30.

5 Jorde LB: Population diseases in the Ashkenazi population: evolutionary considerations. in Bonne-Tamir B, Adam A (eds): . New York, Oxford University Press; 1992, pp 305–312.

6 Risch N, de Leon D, Ozelius L : Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. 1995; **9**: 152–159.

7 Goldstein DB, Reich DE, Bradman N, Usher S, Seligsohn U, Peretz H: Age estimates of two common mutations causing factor XI deficiency: recent genetic drift is not necessary for elevated disease incidence among Ashkenazi Jews. 1999; **64**: 1071–1075.

8 Fay JC, Wu CI: A human population bottleneck can account for

23 Rogers AR: Genetic evidence for a Pleistocene population explosion. 1995; **49**: 608–615.
24 Excoffier L, Schneider S: Why hunter-gatherer populations do not show signs of pleistocene demographic expansions. 1999; **96**: 10597–10602.
25 Cordaux R, Saha N, Bentley GR, Aunger R, Sirajuddin SM, Stoneking M: Mitochondrial DNA analysis reveals diverse histories of tribal populations from India.