

Evidence of early gene flow between Ashkenazi Jews and non-Jewish Europeans in mitochondrial DNA haplogroup H7

Doron Yacobi & Felice L Bedford, Ph.D.

Address for correspondence:

University of Arizona, P.O. Box 210068, Tucson, AZ 85721, USA

Abstract

To investigate European introgression into Ashkenazi Jewry, the European-dominant haplogroup H mitochondrial DNA was examined. The results provided genetic evidence that gene flow between Jewish and non-Jewish populations occurred early in Jewish settlement in Europe with isolation of the groups thereafter. We targeted branch H7 and found three Ashkenazi Jewish clades, two that were not previously recognized as Jewish (H7e, H7c2) and one newly identified group (tentatively H7j) characterized by 1700C and 152C transitions. A total of 100 new complete mitochondrial DNA sequences (mitogenomes) are reported, including the largest collection of H7e to date. H7e is a deeply nested clade with several subclades; more than 85% of the carriers had Ashkenazi maternal ancestry from such diverse areas as Germany and Austria in Western Europe, Poland, and the Baltic states in Central Europe, and Moldova, Ukraine and Belarus in Eastern Europe. Between 10% and 15% of the carriers had European non-Jewish ancestry which, strikingly, showed the greatest number of mutational differences from ancestral H7e. Moreover, there was no overlap with the Jewish-affiliated sequences other than at the ancestral node. Earlier research proposing early mixing followed by isolation has relied on less direct inferences. The smaller groups of H7c2 and H7j were exclusively Ashkenazi Jewish, with interesting sequence patterns. H7c2 consisted of a number of non-nested sister branches, reflecting recent expansion in a large population, while H7j showed a possible in-progress vanishing of the ancestral group, well on its way to mothering an orphan node. The severe bottleneck and subsequent population explosion in the Ashkenazim provide a unique opportunity to view haplogroups in all states of evolution and provide a window into the Mediterranean–Hellenistic world of antiquity.

Introduction

Over the last decade, evidence has accumulated that the genetic make-up of Ashkenazi Jewry is a combination of Levantine and European sources. Analyses of autosomal genes, reflecting a combination of paternal and maternal inheritance, have indicated a significant degree of European admixture among Ashkenazi Jews as well as a close relationship between most contemporary Jews and non-Jewish populations from the Levant (Atzmon et al., 2010; Behar et al., 2010). The source of the European contribution may come from the maternal line. Costa and colleagues (2013) argued that the majority of the Ashkenazi mitochondrial hap-

logroups, which are inherited only from the mother, were present in Europe long before the arrival of Jews. However, Behar and colleagues (2006) suggested that these same maternal haplogroups most likely originated in the Levant alongside paternally inherited Y chromosomes of Levantine origin (Atzmon et al., 2010; Ostrer & Skorecki, 2013).

When haplogroups have a notable presence in both the Near East and Europe, determining their geographic origins can be challenging and lead to differing interpretations. An example involves T2e, a haplogroup that harbors a couple of unique Jewish clades. Bedford (2012) reported prevalence of

T2e in Italy, Egypt, and parts of Saudi Arabia and favored a Near Eastern rather than European origin of the mutations that define T2e but left open the possibility that either locale could be the origin or recipient of migration. On the other hand, Pala et al. (2012), using similar geographic incidences, concluded that T2e's origin was European.

In principle, estimates of when mutations emerged can help resolve where they emerged. In practice, however, standard deviations of time estimates can extend across greater than a thousand years, and time estimates themselves can differ by an order of magnitude depending on the estimated mutation rate. In research on Jewish groups, we (Bedford et al., 2013; Bedford & Yacobi, 2014) reported on a Bulgarian Sephardic founding lineage (T2e1b), originally identified by Behar, which we found among both Ashkenazi and Sephardic Jews from diverse regions. Full genomic sequencing found much coding-region variability, with several haplotypes. Coalescence time for the sequences using a common mutation-rate estimate suggested that the shared mutation (9181G) predated the split between the Jewish groups and therefore likely arose in the Levant. However, a different, also justifiable mutation rate suggested the origin was much more recent, implicating gene flow in Europe after the split as the source as of the mutation common to both Sephardic and Ashkenazi populations.

Difficulty in distinguishing between Levantine and European sources for Ashkenazi mitochondrial haplogroups is further muddled by an often overlooked historical fact: that the boundaries of Europe and the Levant are a relatively recent historical construct dating back to the Arab conquest in the 7th century CE.

To further investigate the role of European maternal admixture into the Ashkenazi gene pool, we took a different approach than previous investigations. Rather than surveying a large number of haplogroups with ambiguous geographic origins, we conducted a detailed investigation into a haplogroup that is overwhelmingly European (e.g., Brotherton et al., 2013) yet still found among modern Ashkenazi Jews. Haplogroup H is the dominant European mtDNA haplogroup.

Its numerical success nears half the population in some countries, making it the most common haplogroup in Europe. Among Ashkenazi Jews, 23% have haplogroup H (Costa et al., 2013), yet despite being a "major" Ashkenazi haplogroup, it is often overlooked. When examining Ashkenazi H mitogenomes, Costa and colleagues found that most of them nest within west/central European subclades, with closely matching sequences in Eastern Europe. As such, haplogroup H's general European dominance may illuminate issues of introgression of European DNA into the Ashkenazi gene pool. Does haplogroup H reflect recent unions of non-Jewish women and Ashkenazi men, or does it point to events of more distant interest?

We focused on H7. While other choices were possible, we selected H7 as an understudied clade within haplogroup H that our pilot study suggested had an unexpected notable presence among the Ashkenazim. Finally, we also delved into Mediterranean and Jewish history to place the genetic results within their correct historical framework. A consideration of relevant Mediterranean and Jewish history is given in Appendix A. The combination of genetic results and accepted history may lead to a greater understanding of Jewish maternal lineages.

Materials and Methods

To identify Ashkenazi clusters within haplogroup H7, we initially selected two individuals with self-described Ashkenazi Jewish maternal lineages belonging to two different subclades of H7 from the customer base at Family Tree DNA (FTDNA; Houston, Texas, USA). FTDNA offers genetic testing services direct to individuals and has one of the largest databases in the world of individuals who have had their full mitochondrial genomes sequenced, including many with European and Ashkenazi Jewish roots. The data from FTDNA customers is increasingly being used as a scientific resource (Bedford, 2012; Bedford et al., 2013; Behar et al., 2012; Pike, 2006; Pike et al., 2010).

These two sequences were used as "kernels", or seeds, to search the FTDNA database for other full

mitochondrial sequences that differed by 0–3 mutations, as in our previous study (Bedford et al., 2013). These people were contacted by email and invited to be part of the research study. They were asked about 1) the additional mutations they carried in their mtDNA, 2) who their matches were within 0–3 genetic differences, and 3) their deep maternal ancestry. In this manner, a large number of different haplotypes belonging to both H7 subclades was identified, and a robust picture of all members of these Ashkenazi Jewish clusters was assembled.

Thereafter, the database of the H7 mtDNA genome project (“H7 MtGenome”), co-administered by one of us (Yacobi), was mined for additional sequences not uncovered by the above procedure. Within the H7 MtGenome project, 229 participants had tested their full mitochondrial genome at the time of this study. The H7 MtGenome project is open to anyone who has tested their mtDNA full genomic sequence with FTDNA and belongs to H7 or one of its subclades (https://www.familytreedna.com/public/mtdna_h7/). All members who were not contacted initially and whose data showed they belonged to one of the groups of interest (the two identified Jewish clades and any cluster which suggested Jewish presence) were also issued invitations to participate in the study and questioned as above.

In addition, for each Ashkenazi cluster found, a sister cluster was sought for comparison among project members without regard to ethnicity. Sister clusters were defined as two distinct branches deriving from the same mother node in the tree. Sequences will also be deposited in GenBank (see Supplementary Table 1).

We decided to use relative time origins, where appropriate, rather than ambiguous absolute time estimates.

Results

Three branches with a notable Jewish constituency were identified within haplogroup H7, for a total of 89 sequences. Two of these branches,

H7c2 and H7e, have been previously identified but not previously connected to Ashkenazi Jewish roots. The third branch is newly reported here; it is defined by a nucleotide transition from T to C at position 1700 in the coding region and by two additional mutations (152C, 573.1C), and thus was not identifiable from inspection of the first control region alone. We tentatively label this clade H7j, following standard mtDNA nomenclature (PhyloTree Build 17; Van Oven, 2015). The three branches likely represent three different maternal founders. In addition, two sister clades were identified for H7c2 among the project’s participants, namely H7c1 and H7c3, both documented branches of H7c. We did not find any sister clades to H7e or H7j in our data set. An overview of the five branches in relation to the H7 ancestral node is shown in Figure 1.

H7J

A total of 14 individuals belonging to newly identified H7j were found. Of these, nine agreed to participate. All nine participants reported Ashkenazi Jewish ancestry on their direct maternal line, with one noting additional possible ancient Sephardic Jewish roots. A notable pattern was observed in this small clade in which the most frequent sequence was not ancestral H7j, but rather a descendant branch (see Figure 1, bottom branch). There is no known positive selection pressure because its single change in the coding region (T11137C) is a synonymous mutation. The success of this branch within H7j may instead be due to random drift during the population explosion following the severe Ashkenazi bottleneck. We may be witnessing the in-progress disappearance of the mother node of H7j, which is becoming less prevalent than its daughter node, presumably an intermediate step before being lost entirely to history and producing breaks in the phylogenetic tree.

H7c2 and sister clades H7c1 and H7c3

A total of 25 people were found in H7c2, 17 of which responded to the invitation. All 17 reported Ashkenazi Jewish ancestry on the direct maternal line. We do not think this reflects sampling bias because public information available on individuals who did not respond pointed to Ashkenazi

Jewish ancestry as well. H7c2 consisted of individuals from regions of Austria, Hungary, Poland, Romania, and the Pale of Settlement.

Of the 25 individuals confirmed as belonging to H7c2, a large majority (20) belonged to the ancestral cluster (A13959T). The remaining five each had a unique haplotype. This is consistent with recent expansion in a large population, large enough for several branches to emerge contemporaneously. The deepest nesting was separated by two mutations from the ancestral H7c2, belonging to an individual of Hungarian Jewish ancestry (see Figure 1).

In contrast, the sister clade H7c1 (previously estimated to be over 3,000 years old; Behar et al., 2012) had a wider geographic distribution than Ashkenazi dominated locales, with our participants reporting ancestry from Egypt, Asia Minor, Italy, Germany, the British Isles, and the Ukraine. H7c1 is also found among the Druze of Israel (Shlush et al., 2008). One of our participants reported Sephardic Jewish ancestry, and the remaining participants denied any Ashkenazi Jewish ancestry. The current distribution of H7c1 may reflect population movements around the Mediterranean during and subsequent to the Roman era.

The second sister clade H7c3 (estimated by previous researchers to be 2440 years old) was distributed mainly in Northern and Eastern Europe with ancestry reported from Finland, Sweden, Russia, and Poland. As with H7c1, no individuals with Ashkenazi Jewish ancestry were reported despite the haplogroup being found in some of the areas heavily populated by Ashkenazi Jews, such as Galicia in Poland.

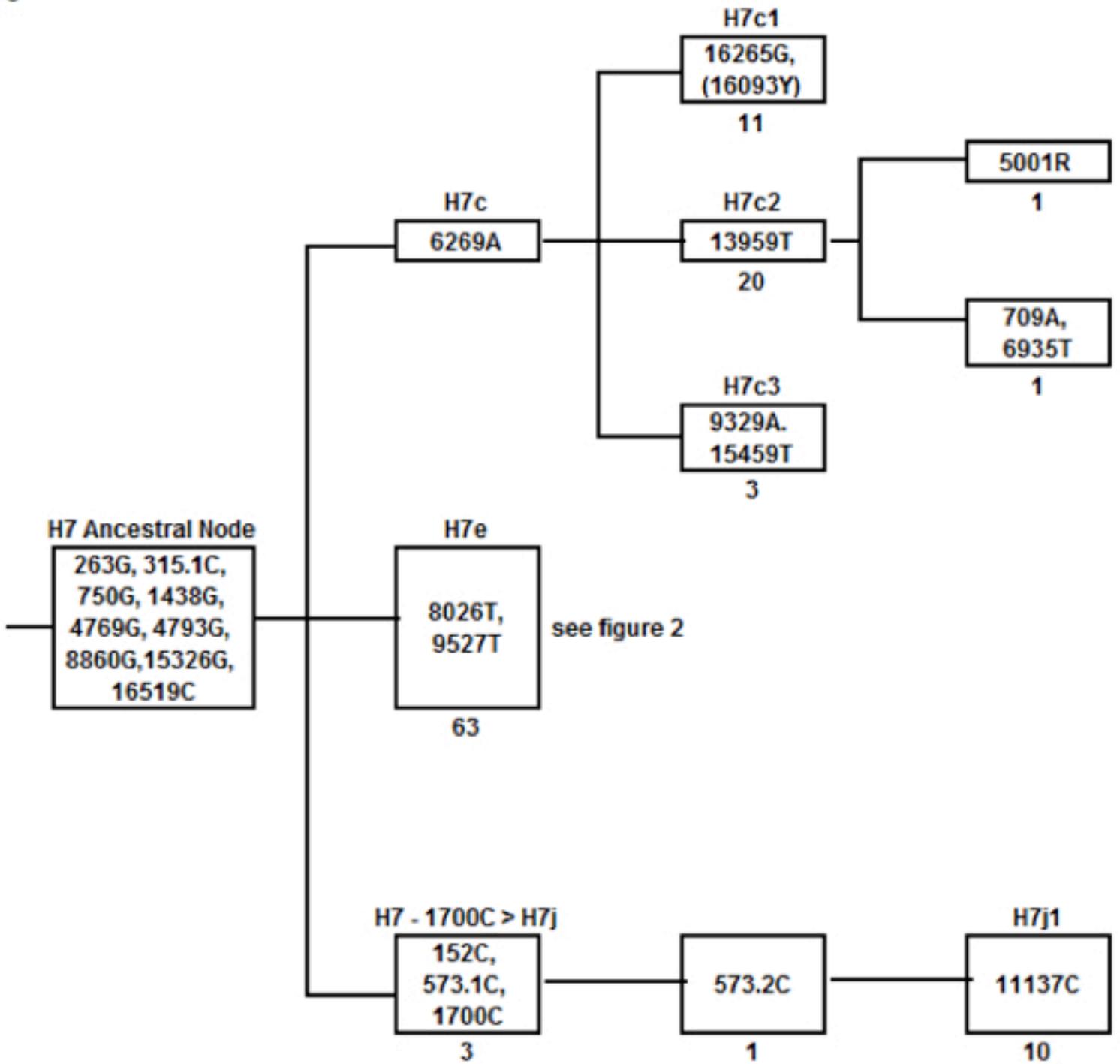
The Ashkenazi Jewish H7c2 appears to be a younger clade than sister H7c1 with one fewer mutation separating it from the mother haplogroup H7c and less rich nesting structure. H7c2 has been dated previously to 1,735 YBP (Behar et al. 2012), younger than the 3000+ YBP estimate for H7c1 and 2400+ YBP for H7c3. The relatively young cluster of H7c2, found here only in Ashkenazim (although among multiple diverse communities), favors a local European emergence in early Ashkenazi settlement predating their geographic dispersal. In view

of the wide geographic dispersal of the mother clade H7c in both Western Asia and Europe (estimated TMRCA of over 7,000 YBP; Behar et al., 2012), and the documented presence in the Levant of the daughter branch H7c1, which includes the Druze samples and at least one individual of Sephardic origin, a Levantine source for the precursor of H7c2 is a possibility. However, considering that the sister clade H7c3, as well as some of the H7c1 samples, trace their ancestry to Northern Europe, it is difficult to reach a conclusion based on this evidence. If the absolute time estimate for H7c2 is correct, this timing would also support a non-European origin for the maternal ancestress of the local Ashkenazi H7c2 mutation, because it dates to the early period of the Jewish diaspora (200–300 CE; i.e., it pre-dates 650 CE) when the vast majority of Jews were found outside of Europe (see Appendix). However, as noted, absolute time estimates from genetic mutations rates are problematic and cannot presently be relied upon to disambiguate origin. Brotherton and colleagues (2013), for example, using dated haplogroup H genomes to calculate mutation rates, found a mutation rate 45% higher than current estimates for human mitochondria.

H7e

In contrast with H7c2 and H7j, which were found to be exclusively Jewish, H7e included a few individuals of European ancestry with no known Jewish ancestry. H7e was also the largest of the predominantly Jewish clusters within H7, with 54 of the 63 individuals of self-described certain Ashkenazi Jewish. Behar and colleagues (2012) dated H7e to the 5th–6th Century CE, but, as with other examples noted, use of a different mutation rate or a high standard deviation means the cluster could either predate or postdate the critical 650 CE time boundary. We did not identify any individuals carrying only one of the defining mutations of H7e (8026T and 9527T), consistent with earlier work by Atzmon et al. (2010). H7 itself has been estimated to be 8890 years old (Behar et al., 2012), many thousands of years older than H7e. Overall, no conclusion can be drawn about the origin of H7e from looking at the haplotypes upstream.

Figure 1



Of the 63 individuals with H7e, 31 belonged to the ancestral cluster and carried only the defining mutations of the clade, 8026T and 9527T. In addition, 28 of these 31 individuals were either self-described certain Ashkenazi Jewish or were highly likely to have Ashkenazi roots based on the information provided about their direct maternal lines. For two individuals, there wasn't sufficient information to determine whether they had Ashkenazi roots, and one individual had no known Ashkenazi roots. None of those belonging to the Ashkenazi cluster were aware of Sephardic or other Jewish roots.

Ashkenazi Jewish H7e

In addition to the ancestral cluster in H7e, a number of distinct Ashkenazi clades within H7e were found. The cluster with the greatest internal diversity, which we tentatively labeled H7e1, was identified by the additional mutation 8994A in the coding region. All known members of H7e1 reported Ashkenazi ancestry on their maternal lines. The sequence most distant from the ancestral cluster had three additional mutations (Figure 2). The deep nesting provided evidence of the longevity of H7e among Ashkenazi Jews. An additional large Ashkenazi cluster, tentatively labeled H7e2, was identified by the mutation 12651A.

In total, 84% of the samples belonging to H7e had or highly likely had Ashkenazi Jewish roots on their direct maternal lines. The geographic distribution of these individuals in the ancestral cluster encompassed practically all of the countries in which Ashkenazi Jews lived at the beginning of the 20th Century, from Germany and Austria in Western Europe, through Poland and the Baltic states in Central Europe, to Moldova, Ukraine and Belarus in Eastern Europe. Furthermore, within the Ashkenazi subclades of H7e, distinct regional patterns of distribution were discernable, with disproportionate numbers reporting Lithuanian ancestry (60%) in H7e1 (8994A) and Polish ancestry (50%) in H7e2 (12651A).

The wide distribution of the ancestral cluster along with the more regional distribution of the

subclades indicate that H7e entered the Ashkenazi gene pool at a relatively early stage in the history of the haplogroup. The emergence most probably occurred no later than during the 9th and 10th centuries during the formative stages of Ashkenazi Jewry and prior to the movement eastwards to Central and finally to Eastern Europe.

Non-Jewish H7e

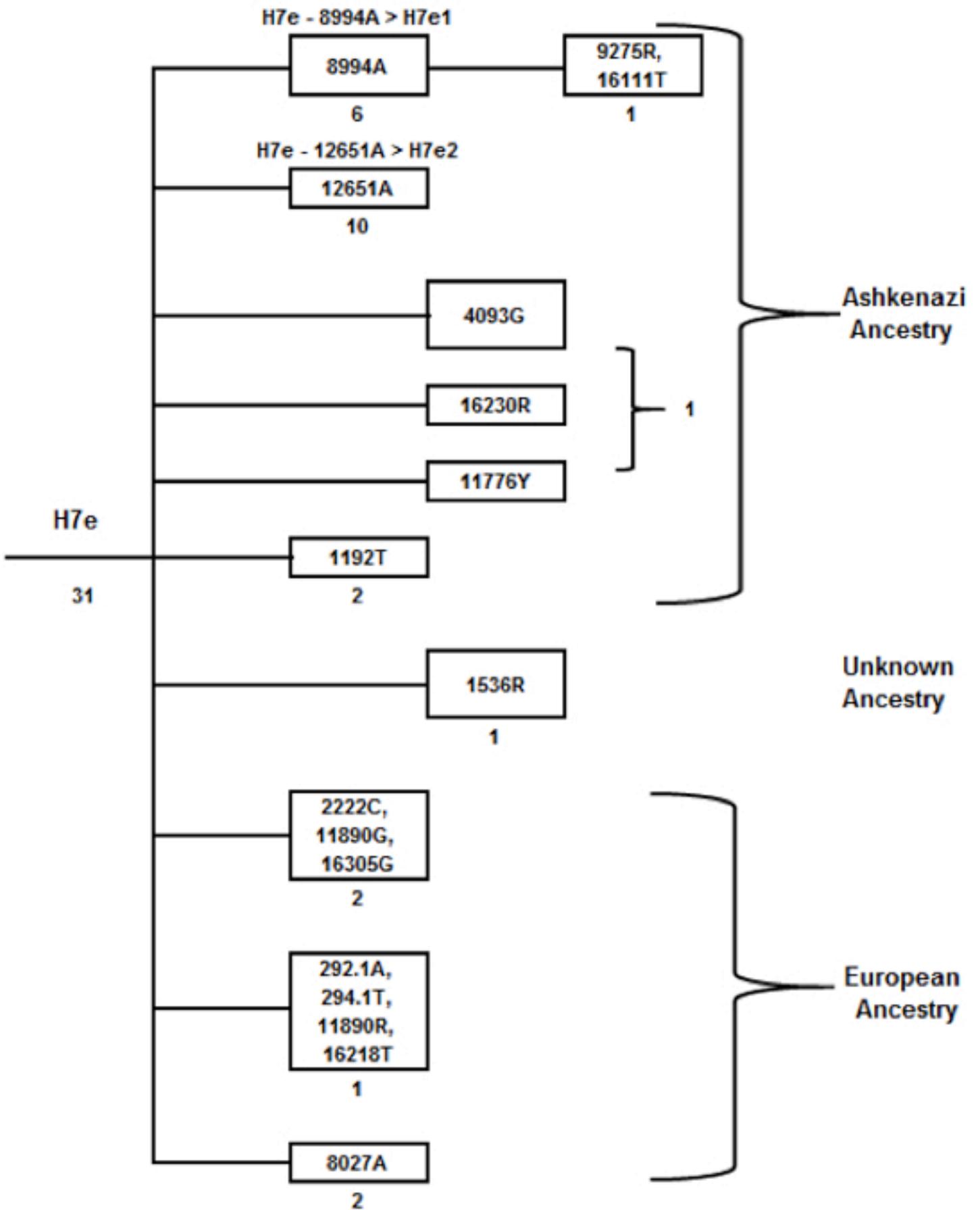
Of the 63 H7e individuals, six had no known Ashkenazi ancestry (~10%), including two who can trace their ancestry back to Germany and one to the island of Susak in Croatia. The remainder could not trace their ancestry beyond colonial America. Another three individuals are unlikely to have Ashkenazi ancestry (~5%).

A striking aspect about the non-Jewish H7e results is that they were found to be a considerable genetic distance from the ancestral cluster and separated by several mutations (see Figure 2). One sequence had four possible independent mutations (16218T, 292.1A, 294.1T, 11890R), and two sequences had three mutations (2222C, 11890G, 16305G). Furthermore, these clusters did not nest within the existing Jewish subclades of H7e, nor did those nearer to the ancestral cluster with no known Jewish roots. There seems to be a clear distinction between those belonging to the subclade with Ashkenazi Jewish roots and those without Ashkenazi Jewish roots, bar one member of the ancestral cluster with no known Jewish roots (< 4% of the ancestral cluster). The non-Jewish samples also show greater genetic diversity than the Jewish samples.

Discussion

The current work identified three clades and several subclades of H7 as predominantly Jewish. One of these (H7j) was previously undiscovered, and the others (H7e, H7c2) had not previously been identified as mainly Jewish. We focused on the European haplogroup H, rarely discussed within Ashkenazi genetics, to gain insight into early European Jewish maternal origins.

Figure 2



The largest group was H7e, with 63 individuals. This reflects the largest collection of complete H7e sequences reported to date; adding to the previous five sequences available on GenBank. At least two regionally distinct subgroups were newly found within H7e. The relatively large sample enabled several patterns to be revealed: 1) The bulk of H7e individuals have Ashkenazi maternal origins. 2) The geographic origins of Ashkenazi H7e encompassed all regions in which Ashkenazim were found including Germany and Austria in Western Europe, Poland and the Baltic states in Central Europe, and Moldova, Ukraine and Belarus in Eastern Europe, with regional subclades apparent. 3) Some H7e sequences were found in individuals who knew of no Jewish ancestry. 4) The Non-Jewish sequences showed rich nesting and several mutational differences from ancestral H7e. And, 5) the non-Jewish clusters showed no overlap with Jewish subclades. Taken together, these findings strongly implicate the introgression of a mitochondrial lineage either from or into the Jewish gene pool that occurred early in the settlement of European Jews. This was followed by no further genetic contact between the two groups. Genetic isolation led to separate expansions, especially among the Ashkenazi as they made their way deep into Eastern Europe.

One challenge facing research into Jewish maternal lineages has been their distinctiveness, which makes their origins difficult to determine. That is, many maternal lineages found among Jewish populations, despite having significant coding region variability, are restricted solely to the Jewish subgroup to which they are found in. In H7e, on the other hand, we found distinct evidence of both Ashkenazi Jewish and European non-Jewish maternal lineages with clear relationships based on coding region variability. Thus we can see genetic evidence of an oft-speculated but rarely seen early exchange, followed by independent development, in the gene pool between Jewish and non-Jewish groups.

But in which direction was the early genetic contribution? The dominance of haplogroup H as an early European rather than Near Eastern haplogroup may favor the hypothesis that one woman

belonging to Haplogroup H7e converted to Judaism and married into the Jewish community. The predominance of Jewish individuals within the ancestral cluster would, in this view, be explained by the Ashkenazi bottleneck and subsequent population boom (Carmi et al., 2014) which resulted in an inflated number of Ashkenazi Jewish women carrying the ancestral version of H7e than in the general European population.

One is also tempted to speculate that the non-Jewish European origin of H7e was German. This possibility is consistent with the fact that, of the few individuals without Jewish roots, two could trace their distant ancestry back to Germany. In addition, Ashkenazi Jewish history considers settlement in Germany to have occurred before expansion to Eastern European regions. If this is the case, then H7 is younger than previously thought, because there is practically no evidence of a Jewish presence during the 7th and 8th centuries in the Rhineland area (see Appendix).

A second possibility consistent with an older age for H7e is a European origin in Italy or Southern France. The Jewish presence in the Rhineland area, and later in central Europe, is considered the outcome of the migration of Jews from Southern Europe that began in the 9th and 10th centuries (Botticini & Eckstein, 2012). The gene flow, however, could have occurred in either direction: for example, non-Jewish French women marrying newly arriving Near Eastern Jewish men or Jewish women arriving to Italy from the Near East and leaving the Jewish community. Origin of H7e in Italy or Southern France would require an explanation for why all traces of the haplogroup have vanished from those areas. Such an explanation may not be hard to find. In general, many — perhaps most — haplogroups have likely vanished from existence; the unusual situation of the Ashkenazi extreme bottleneck and subsequent population explosion allowed otherwise extinguished haplogroups to survive in select demographics.

Finally, despite the predominance of haplogroup H in Europe and the other factors suggesting a European origin, we cannot definitively rule out the other extreme: that the ancestress of H7e was herself part of the Jewish community in antiqui-

ty. Regardless of where geographically the women were when the mutations of H7e arose, they still could have arisen in women whose ancestors were Jewish before leaving the Near East. H7 and other H clades could nonetheless have been in the Near East at the right times even if they predominately expanded in Europe. In this view, the small number of non-Jewish individuals belonging to H7e represents the descendants of women who left the Jewish community relatively early on in the history of the subclade. This would include the German, Croatian, and Colonial American participants in our study.

The present work also uncovered a small new clade tentatively labelled H7j and identified the previously known H7c2 group as Ashkenazi Jewish. Neither had any non-Jewish affiliation. The small sizes of the clusters may have precluded any minor non-Jewish presence from being detected, the small clusters may have vanished in all but the large Ashkenazi population, or the mutations characterizing these branches may simply have arisen among the isolated Ashkenazi communities while in Europe. We favor the latter hypothesis. Regardless, it is important to note that an ancient Near Eastern source for the precursors of H7c2 or H7j is possible under any of the hypotheses. We also found interesting patterns in the smaller H7j and H7c2 clusters. One cluster contained several, non-overlapping, shallow branches that emerged contemporaneously, reflecting a relatively new clade in a large population. The other pattern revealed a possible in-progress vanishing of the ancestral group, which may soon be lost to history and lead to missing links in the phylogenetic tree.

As analysis of H7 clades illustrates, determining the direction of gene flow with any degree of certainty is difficult, even when sequences belonging to non-Jewish populations are found (as for H7e). The problem is even greater when a mitochondrial lineage is restricted exclusively to Ashkenazi Jews, as often occurs. Consequently, it is notable that Costa and colleagues (2013) nonetheless concluded that 80% of Ashkenazi maternal ancestry is due to the assimilation of mtDNAs indigenous to Europe, most likely through conversion. We feel this conclusion is premature and goes beyond the

available evidence for several reasons: the intricacies of Jewish history are often overlooked, the methodology of looking at the immediate ancestral nodes is not always conclusive, time estimates that can be grossly inaccurate are often relied on too heavily, and confusion exists between where an individual lived when a *de novo* mutation arose and that person's origins. We provide an example and brief elaboration from the Costa et al., 2013 paper to illustrate. We belabor the point because of the importance of concluding such a definitive maternal origin for the vast number of Ashkenazi haplogroups.

The haplogroups surveyed by Costa and colleagues (2013) may have arisen in Europe between the last glacial period and the Neolithic as maintained. However, when, considering the complex history of migration within the Mediterranean basin over the last 3,000 years, as well as Jewish history (see Appendix), it is apparent that where a haplogroup first arose many thousands of years earlier need not have any bearing on where and when a specific distinctive mitochondrial haplogroup first emerged among Jewish populations. Furthermore, a sizeable portion of the Mediterranean–Hellenistic Jewry of antiquity was comprised of converts to Judaism rather than descendants of the Iron-Age Israelites. While the majority of these converted in the land of Israel prior to 65 CE, they undoubtedly included some descendants of merchants, colonists, and troops with roots tracing back to Mediterranean Europe, which could explain some of the European admixture found amongst the Jewish populations descending from the Mediterranean–Hellenistic Jewry of antiquity based in the Eastern Mediterranean.

For a specific example, consider the often discussed haplogroups K1a1b1a and K1a1b1a1 among Ashkenazi Jews. Costa and colleagues (2013) used maximum likelihood to estimate that K1a1b1a dates to approximately 4,400 YBP and K1a1b1a1 to 2,300 YBP. To place these results in their historical perspective, 2,300 YBP predates the dispersal of the Jewish population from the Levant to Europe, and 4,400 YBP predates by more than 1,000 years the earliest documented mention of the name "Israel" in historical record (the Mer-

neptah Stele, dated to 1209 BC). As they estimate the parent clade K1a1b1 to be over 10K years old, in the interim ~6,000 years between the appearance of K1a1b1 and the appearance of K1a1b1a, the maternal lineage could have migrated to and from the Levant on numerous occasions (in a manner similar to the movement pattern of H7c1). As noted earlier, prior to the Arab conquest in the 7th century CE the Western and Eastern sides of the Mediterranean basin were as well, if not better, connected to each other than the Western Mediterranean was to parts of Northwestern Europe. When considering the age of the haplogroup, its presence (however limited) among Sephardic Jews and its apparent absence in non-Jewish populations (Costa et al., 2013; Behar et al., 2006) all seem to indicate that a Levantine origin is far more likely for K1a1b1a than a European one, regardless of where K1a1b1 first originated.

Turning attention to mtDNA mutation rates, our finding of early exchange between the European and Jewish gene pools in haplogroup H mtDNA (H7e) suggests that the rates of mutations are much faster than commonly assumed. They are closer to those estimated using pedigrees. Madrigal and colleagues (2012) calculated a mutation rate of 1.24×10^{-6} per site per year in an analysis of individual family pedigrees from a well-documented population in Costa Rica, a rate three times faster than those commonly derived from phylogenies. The distinctiveness of Ashkenazi Jewish maternal lineages and their isolation from non-Jewish maternal lineages, coupled with a rapid population explosion and the relatively well-documented history of Ashkenazi Jewry, may provide a further basis for grounding the widely varying mutation rates offered by different sources.

Finally, we can reconsider the high degree of European admixture (30%–60%) observed among Ashkenazi, Sephardic, Italian, and Syrian Jews (Atzmon et al., 2010) in autosomal DNA studies, as well as the higher proportion of European admixture among North African Jews compared with non-Jewish North African populations (Campbell et al., 2012). Part of this clearly reflects limited more recent European admixture, hence

the elevated levels of European admixture when comparing Ashkenazi to Sephardic Jews or Moroccan to Djerban Jews. However, part undoubtedly reflects the legacy of the Mediterranean and the movement of peoples around the Mediterranean basin long before Christian Southern Europe become isolated from the Islamic Levant and North Africa, and results from conversions to Judaism prior to 65 CE in the Hellenistic and then Roman Levant and North Africa.

Little is known about the earliest days of settlement of the Ashkenazi Jews in Europe. Research into Jewish population genetics holds the promise of illuminating migrations and expansions that are poorly understood due to the scarcity of reliable historical sources. We believe we have provided one of the clearest views of this early period through a branch of maternally inherited mitochondrial DNA haplogroup H that strongly implicates gene flow between the Ashkenazi and non-Jewish European populations pre-dating the Ashkenazi expansion throughout Central and Eastern Europe. We focused on the most prevalent haplogroup in Europe, which also contains subclades found almost exclusively among Ashkenazi Jews, to provide further insight into the origins of the European Jewish communities. We found gene flow within haplogroup H7, evidence that will be beneficial in assessing the origin of other mitochondrial subclades found among Jewish groups.

Acknowledgments

The authors gratefully acknowledge the constructive critiques provided by Leah Larkin, Ian Logan, and two anonymous reviewers as well as the assistance of Jacques Beaugrand, administrator of the H7 MtGenome Project.

Appendix A. A Brief Consideration of Mediterranean and Jewish History

Historical considerations in mtDNA genetic studies tend to focus on prehistoric Europe because of the ages of many haplogroups and, in particular, the last glacial maximum and its impact on human migrations (Roostalu et al., 2007). Often overlooked, however, is that following these events many thousands of years ago, human migration continued unabated and, with it, the corresponding gene flow between different parts of Europe, Western Asia, and North Africa (e.g., Brotherton et al., 2013 re Haplogroup H in Europe).

One of the most important facilitators of migration between these geographical areas was the Mediterranean. As Abulafia (2003) pointed out, thanks to the ease of movement across the open sea, lands far removed from each other enjoyed vibrant trading, cultural, and political ties. Furthermore, from the Mediterranean, access could be gained to the European network of big rivers, such as the Danube and the Rhine, further facilitating the movement of goods and people from the Mediterranean basin inland into Central Europe. There is no doubt that this movement around the Mediterranean basin has very ancient roots. Archaeological sites in Israel reveal a Stone Age culture quite similar to that known in the Western Mediterranean from the limestone caves of Spain, France, and Northern Italy (Suano, 2003).

The Mycenaeans in the 14th century BCE were the first to start intensively traversing the Mediterranean carrying trade between the Aegean and the Levantine coastal cities, thus linking these regions to the central Mediterranean and, on occasion, Iberia. Permanent settlements of Mycenaeans have been identified on the coast of southern Italy, in Sicily, and in Sardinia (Torelli, 2003). The commercial traffic of the Mediterranean throughout the pre-Roman age was marked by colonial settlement as much as by mercantile contact. Following the collapse of the Mycenaean empire and the rise of classical Greece and Phoenicia, the trade rivalry between the Greeks and Phoenicians and the ensuing battle over the Mediterra-

nean trading routes between 1,000 BCE and 300 BCE led to the development of a wide ranging network of trading settlements and colonies. Colonies in Carthage and the ring of emporia in Libya, Motya, and Soluntum in Sicily; the harbors in Sardinia; and the bases and trading stations at Ibiza in the Baleric Islands, Cadiz beyond the straits of Gibraltar, and along the Moroccan Coast allowed the Phoenicians to dominate many of the trade routes straddling North Africa, Iberia, and the Levant. The Greeks as well as the Etruscans developed rival trading routes covering much of Southern Europe, the Adriatic, the Black Sea, and Asia Minor (Torelli, 2003).

The key period of Mediterranean unification occurred, however, under the rule of imperial Rome. For a period of roughly 800 years (300 BCE–500 CE) the whole Mediterranean was politically unified. As Rickman (2003) stated, “it is hardly surprising that a sea which the Romans, and the polyglot populations under their control had so thoroughly made their own should witness not just the circulation of goods, but also of people”. Military conquests during the Republic (300–100 BCE) and the expansion of the Roman Empire brought to the Italian peninsula significant economic migration of free immigrants as well as slaves from Gaul, Hispania, Germania, Magna Graecia, Asia Minor, Phoenicia, Egypt, and North Africa (Noy, 2000; Scheidel, 2004). Scheidel (2004) estimates that around 2 million people immigrated to Rome just during the last two centuries BCE while, according to Noy (2000), over 10% of foreigners buried at Rome came there from North Africa, and most were civilians rather than associated with the military (see Killgrove, 2010, 2013). The movements of people were not just to Rome. The names of the units stationed on Hadrian’s Wall reveal how widely Rome recruited its auxiliary regiments, from Spain, Gaul, Germany, the lands along the Danube, Asia Minor, Syria, and North Africa (Vindolanda, 2016).

Jewish history is intertwined with Mediterranean history. The formative stages of the Jewish diaspora occur during the period of the *Mare Nostrum* (or ‘our [Roman] sea’). There is a tendency to confuse the Iron-Age Israelites of the 8th and

9th centuries BCE with the Jewish population living in the Roman province of Judea nearly 1,000 years later just prior to the great revolt of 65–70 CE, however, while undoubtedly some of those living in Judea as Jews during the 1st century CE were the genetic descendants of the inhabitants of the ancient kingdoms of Israel and Judah, many others were not. The four centuries following the Babylonian conquest of Judah in 586 BCE had seen major political and demographic changes taking place in the land of Israel. Faust (2012) has persuasively shown that, based on the archaeological evidence, Judah experienced drastic demographic decline due to the war, subsequent famine, and epidemics that followed the conquest. Continuity in the following centuries with the Iron Age society of Judah was limited. There were survivors, and some of the population exiled to Babylon must have returned, but population recovery in the region must have also been triggered by new settlers from neighboring regions (Faust, 2012). Following its conquest by Alexander the Great in 332 BCE, Judea was no longer merely a buffer state between Egypt and Mesopotamia; it now formed the eastern edge of what was quickly becoming a pan-Mediterranean empire — the Roman ‘Mare Nostrum’. By 63 BCE, Judea was a client state of Rome and by 6 CE a Roman province.

In Goodman’s (1994) thorough research into proselytes and proselytizing to Judaism during the period of the Roman Empire, he concluded that there is evidence that prior to 65 CE, converts made up a significant proportion of the Jewish population and that Jews accepted as proselytes those gentiles who applied to join their number, although they did not feel compelled to encourage such conversions. As examples, Goodman (1994) referred to the spread of Jewish settlement in the diaspora, the increase in the population of Judea apparent from archaeological survey, and Josephus’ recording of the conversion en masse of neighboring populations such as the Idumeans and the Ituraeans by the Hasmonaean dynasty.

In the post-70 CE period, ambivalence by Rabbinical authorities towards the proselytization of gen-

tiles meant that conversion to Judaism was far less common, although there is some evidence of proselytes to Judaism all the way through into the medieval period (Goodman, 1994). This was especially true after the failed Bar Kokhba rebellion during Hadrian’s rule and the passage of legislation by Hadrian and his successors against the circumcision of non-Jews, the special Jewish tax (the *fiscus Judaicus*), and a series of Roman laws in the 4th and 5th centuries prohibiting conversion to Judaism, particularly by Christians. Furthermore, as Goodman (1994) pointed out, some conversions to Judaism probably took place to facilitate marriage. Considering the patriarchal nature of both Jewish and Roman societies, as well as the prohibition on circumcision that prevented men (but not women) from converting, many of the converts to Judaism to facilitate marriage were likely women.

How many of these conversions would have taken place in Europe? As can be seen in Table 1 based on the estimates of Botticini and Eckstein (2012), prior to 65 CE the majority of the Jewish population throughout the Middle East and the Mediterranean basin were located in the lands of Israel, Mesopotamia, Persia, and North Africa (mainly Egypt), while the number of Jews in Western Europe was relatively small and by 650 CE was negligible (~1,000). Thus the vast majority of conversion to Judaism during this period must have occurred outside of Europe in the Levant, Egypt, and Mesopotamia.

Furthermore, in a detailed study by Toch (2005) of Jews in Europe between 500–1050 CE, he concluded that between the mid-7th and mid-8th centuries, no source mentions Jews in Frankish lands (now France and Germany). Only in the 8th and 9th centuries was there evidence of growing numbers of Jews in the South of France, while in the 9th and early 10th centuries, brief hints attest to itinerant merchants in Germany. Toch (2005), therefore, concluded that no continuity could be assumed between the Jews of the Roman Empire and the Ashkenazi Jewish communities of the Middle Ages.

From a genetic perspective, based on this historical overview, maternal lineages restricted to Jew-

ish populations that pre-date 650 CE are highly unlikely to have originated in either Western or Eastern Europe, given the miniscule numbers of Jews in these regions during this period.

Table 1. Jewish population estimates in 65 CE and 650 CE (as per Botticini and Eckstein, 2012).

Region	c. 65 CE	c. 650 CE
Land of Israel	2,500,000	100,000
Mesopotamia and Persia (including the Arabian Peninsula)	1,000,000	700,000–900,000
North Africa (mainly Egypt)	1,000,000	4,000
Syria and Lebanon	200,000–400,000	5,000
Asia Minor and the Balkans	200,000–400,000	40,000
Western Europe (including Italy, France Germany, and Iberia)	100,000–200,000	1,000
Eastern Europe	–	–

References

Abulafia D (2003) What is the Mediterranean? In *The Mediterranean in History*, D Abulafia, ed. Getty Publications, Los Angeles, CA. pp. 11–32.

Atzmon G, Hao L, Pe'er I, Velez C, Pearlman A, Palamara PF, Morrow B, Friedman E, Oddoux C, Burns E, Ostrer H (2010) Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern ancestry. *American Journal of Human Genetics* 86: 850–859.

Bedford FL (2012) Sephardic signature in haplogroup T mitochondrial DNA. *European Journal of Human Genetics* 20: 441–448.

Bedford FL, Yacobi D, Felix G, Garza FM (2013) Clarifying mitochondrial DNA subclades of T2e from Mideast to Mexico. *Journal of Phylogenetics and Evolutionary Biology* 1: 121. doi:10.4172/2329-9002.1000121

Bedford FL, Yacobi D (2015) On two Jewish clades in mitochondrial DNA. *European Journal of Human Genetics* 23: 993-994. doi: 10.1038/ejhg.2014.231.

Behar DM, Metspalu E, Kivisild T, Achilli A, Hadid Y, Tzur S, Pereira L, Amorim A, Quintana-Murci L, Majamaa K, Herrnstadt C, Howell N, Balanovsky O, Kutuev

I, Pshenichnov A, Gurwitz D, Bonne-Tamir B, Torroni A, Villems R, Skorecki K (2006) The matrilineal ancestry

of Ashkenazi Jewry: portrait of a recent founder event. *American Journal of Human Genetics* 78: 487–497.

Behar DM, Yunusbayev B, Metspalu M (2010) The genome-wide structure of the Jewish people. *Nature* 466: 238–242.

Behar DM, Van Oven M, Rosset S (2012) A “Copernican” reassessment of the human mitochondrial DNA tree from its root. *American Journal of Human Genetics* 90: 675–684.

Botticini M, Eckstein Z. (2012) *The Chosen Few: How Education Shaped Jewish History*, 70–1492. Princeton University Press, Princeton, NJ.

Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, Richards SM, Der Sarkissian C, Ganslmeier R, Friederich S, et al. (2013) Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nature Communications* 4: 1764.

Campbell CL, Palamara PF, Dubrovsky M, Botigué LR, Fellous M, Atzmon G, Oddoux C, Pearlman A, Hao L, Henn BM, Burns E, Bustamante CD, Comas D, Friedman E, Pe'er I, Ostrer H (2012) North African Jewish and non-Jewish populations form distinctive, orthogonal clusters. *Proceedings of the National Academy of Sciences USA* 109: 13865–13870.

Carmi S, Hui KY, Kochav E, Liu X, Xue J, Grady F, Guha S, Upadhyay K, Ben-Avraham D, Mukherjee S, Bowen BM,

- Thomas T, Vijai J, Cruts M, Froyen G, Lambrechts D, Plaisance S, Van Broeckhoven C, Van Damme P, Van Marck H, Barzilai N, Darvasi A, Offit K, Bressman S, Ozelius LJ, Peter I, Cho JH, Ostrer H, Atzmon G, Clark LN, Lencz T, Pe'er I (2014) Sequencing an Ashkenazi reference panel supports population-targeted personal genomics and illuminates Jewish and European origins. *Nature Communications* 5: 4835. doi:0.1038/ncomms5835.
- Costa MD, Pereira JB, Pala M, Fernandes V, Olivieri A, Achilli A, Perego UA, Rychkov S, Naumova O, Hatina J, Woodward SR (2013) A substantial prehistoric European ancestry among Ashkenazi maternal lineages. *Nature Communications* 4: 2543 doi:10.1038/ncomms3543.
- Faust A (2012) *Judah in the Neo-Babylonian Period: The Archaeology of Desolation*. Society of Biblical Literature, Atlanta, GA, USA.
- Goodman M (1994) *Mission and Conversion: Proselytizing in the Religious History of the Roman Empire*. Oxford University Press, Oxford, England.
- Killgrove K (2013) Biohistory of the Roman Republic: the potential of isotope analysis of human skeletal remains. *Post-Classical Archaeologies* 3: 41–62.
- Killgrove K (2010) Identifying immigrants to Imperial Rome using strontium isotope analysis. In *Roman Diasporas: Archaeological Approaches to Mobility and Diversity in the Roman Empire*, ed. H Eckardt. *Journal of Roman Archaeology supplement* 78, ch. 9, pp. 157–174.
- Madrigal L, Castri L, Melendez-Obando M, Villegas-Palma R, Barrantes R, Raventos H, Pereira R, Luiselli D, Pettener D, Barbujani G (2012) High mitochondrial mutation rates estimated from deep-rooting Costa Rican pedigrees. *American Journal of Physical Anthropology* 148: 327–333.
- Noy D (2000) *Foreigners at Rome: Citizens and Strangers*. Classical Press of Wales, Swansea, Wales.
- Ostrer H, Skorecki K (2013) The population genetics of the Jewish people. *Human Genetics* 132: 119–127.
- Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, Reidla M, Tamm E, Karmin M, Reisberg T, Hooshiar Kashani B, Perego U A, Carossa V, Gandini F, Pereira JB, Soares P, Angerhofer N, Rychkov S, Al-Zahery N, Carelli V, Sanati MH, Houshmand M, Hatina J, Macaulay V, Pereira L, Woodward SR, Davies W, Gamble C, Baird D, Semino O, Villems R, Torroni A, Richards MB (2012) Mitochondrial DNA signals of late glacial recolonization of Europe from Near Eastern refugia. *American Journal of Human Genetics* 90: 915–924.
- Pike DA, Barton TJ, Bauer SL, Kipp EB (2010) mtDNA haplogroup T phylogeny based on full mitochondrial sequences. *Journal of Genetic Genealogy* 6: 1–24.
- Pike DA (2006) Phylogenetic networks for the human mtDNA haplogroup. *Journal of Genetic Genealogy* 2: 1–11.
- Rickman G (2003) *The Creation of Mare Nostrum: 300 BC–500 AD*. In *The Mediterranean in History*, D Abulafia, ed. Getty Publications, Los Angeles, CA. pp. 127–154.
- Roostalu U, Kutuev I, Loogvali EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R (2007) Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. *Molecular Biology and Evolution* 24: 436–448.
- Scheidel W (2004) Human mobility in Roman Italy, I: the free population. *Journal of Roman Studies* 94: 1–26.
- Shlush L., Behar D, Yudkovsky G, Templeton A, Hadid Y, Basis F, Hammer M, Itzkovitz S, Skorecki K (2008) The Druze: a population genetic refugium of the Near East. *PLoS ONE* 3: e2105. doi:10.1371/journal.pone.0002105.
- Suano M (2003) *The first trading empires: prehistory to c. 1000 BC*. In *The Mediterranean in History*, D Abulafia, ed. Getty Publications, Los Angeles, CA. pp. 67–98.
- Toch M (2005) *The Jews in Europe 500–1050*. In *The New Cambridge Medieval History*, vol. I, 500–1050, P Fouracre, ed. Cambridge University Press, Cambridge, England. pp. 547–570.
- Torelli M (2003) *The battle for the sea routes: 1000–300 BC*. In *The Mediterranean in History*, D Abulafia, ed. Getty Publications, Los Angeles, CA. pp. 99–126.
- Van Oven M (2015) *PhyloTree Build 17: Growing the human mitochondrial DNA tree*. *Forensic Science International Genetics* 5: e392–e394.
- Vindolanda Tablets Online database, <http://vindolanda.csad.ox.ac.uk/>, accessed 26 Jul 2016.