Genome-wide data from medieval German Jews show that the Ashkenazi founder event pre-dated the 14th century

Waldman et al, *Cell*, 2022

Relevant Photos

Frequently asked questions

Author: Shai Carmi

Last update: Dec 2, 2022

1. What is known about Ashkenazi Jewish history?

Ashkenazi Jews formed as a distinct cultural unit in the Rhineland (Western Germany) around the 10th century. Despite many speculations, not much is known about their earlier origins beyond connections to Jews in France and Italy. In the following centuries, Ashkenazi culture spread to other communities in Germany and Central Europe. By the end of the Middle Ages, following persecutions, Ashkenazi Jews began migrating to Eastern Europe. While suffering at times, they also flourished there culturally and demographically, reaching millions by the 19th century. Following major migration events over the past 150 years, which were intensified after the holocaust, most Ashkenazi Jews today reside in Israel and the USA.

2. What did genetics tell us so far about Ashkenazi history?

Starting from the early days of human genetics, it became clear that Ashkenazi Jews (AJ) carry a particularly high burden of recessive disease-causing variants. Such variants can lead to serious and even lethal childhood diseases when inherited from both parents. The high prevalence of these variants suggested that the Ashkenazi population underwent a “founder event”, or a “bottleneck”, whereby the population started out as extremely small. As the population grew, pathogenic variants that were carried by the founders became widespread.

Further evidence for a founder event came from studies of mitochondrial DNA sequences, which are maternally inherited. A landmark 2006 paper found that four mitochondrial sequences (representing four ancestral mothers) are carried by as many as 40% of present-day AJ. This result was interpreted to suggest a very small number of maternal founders in early Ashkenazi history. Similar observations were made for the Y chromosome, which is paternally inherited.

Genome-wide studies of AJ have appeared since 2008, repeatedly showing that AJ are genetically distinct from other defined human populations, including neighboring European populations and other Jewish groups (examples: 1, 2, 3, 4, 5, 6). Virtually all pairs of Ashkenazi individuals, even unrelated, were found to have long blocks of identical DNA sequences shared between them. This is again due to the descent from a small number of founders, which makes all AJ individuals appear related via these founders. The Ashkenazi population is genetically highly homogeneous, with no discernable differences between AJ from different (current or previous) countries. The only subgroup with slightly distinct ancestry was that of AJ with origins in France (Alsace) and Germany (see also here).

3. What was still unknown prior to our study?

Existing studies remained ambiguous regarding the founder event. When exactly did it happen? Was it a single catastrophic event or a continuous decline over centuries? Where did the founders live? Information on the origin of AJ is also lacking. Where did the AJ founders and their ancestors come from? Did early AJ descend from Judean Jews, or were they converts from other areas in the Mediterranean or elsewhere in Europe?

Finally, the Ashkenazi population is genetically homogeneous today. Was it equally uniform in the past? Were Jews from different communities in Northern Europe related only culturally or also genetically? Did the AJ gene pool change over the years due to intermarriage within Jews groups or with non-Jews?

4. What was the purpose of our study?

DNA from present-day individuals encodes information on past demographic events. However, DNA from people who have lived *during* the events, or *ancient DNA*, can be orders of magnitude more informative. Ancient DNA can document migrations and replacements as they occur, or demonstrate the continuity of populations. It is also informative about the size of ancient populations and about marriage patterns.

Given that no DNA sequences existed for historical AJ, we sought to generate ancient DNA data for this population. Our hope was to fill the gaps in our understanding of AJ early history. Naturally, we did not expect a single dataset to address each and every open question. Nevertheless, we hoped to illuminate some aspects of Ashkenazi Jewish demography during the Middle Ages.

5. Is it possible to study the DNA of deceased Jewish individuals?

In rabbinical Jewish law, exhumation of corpses and disturbance of the dead are prohibited under almost all circumstances. Therefore, excavating Jewish graves for research is not permitted. However, when skeletal remains have *already* been extracted, typically as part of a rescue excavation, DNA testing may be permissible.

Naturally, the question of DNA sequencing of Jewish skeletal remains was not discussed by rabbinical authorities prior to the recent development of the technology. This has changed most notably in the middle 2010’s in the context of the Yemenite Children Affair. The affair involved the fate of hundreds of children, most of whom born to Yemeni Jewish immigrants, who disappeared or died in the early years of the state of Israel (1948-1954). Given that parents were often not properly informed of their children's death, doubt lingered over the true identity of their children's remains. Following a heated debate, the state received green light by religious authorities (1, 2) and permitted the excavation and DNA testing of the remains. In more recent writing on the subject, it was suggested that ancient DNA sequencing should be permissible also for research, provided that only detached teeth are used.

6. What did we find at the medieval Jewish cemetery in Erfurt?

In Waldman et al, we report genomic data for 33 individuals from the medieval Jewish cemetery in Erfurt, Germany. The Erfurt Jewish community existed between the late 11th century to 1454, with a short gap following a 1349 massacre (pogrom) that nearly wiped out the entire community. At times, it was a wealthy community and one of the largest in Germany, serving as a local hub for the entire region. In fact, the oldest intact synagogue in Central Europe is found in Erfurt. Following the expulsion of all Jews in 1454, the city built a large granary on top of the Jewish cemetery, just outside the old city walls. In 2013, when the granary stood empty, the city permitted its conversion into a parking garage. The conversion required the construction of a ramp, and was therefore preceded by an archaeological rescue excavation.

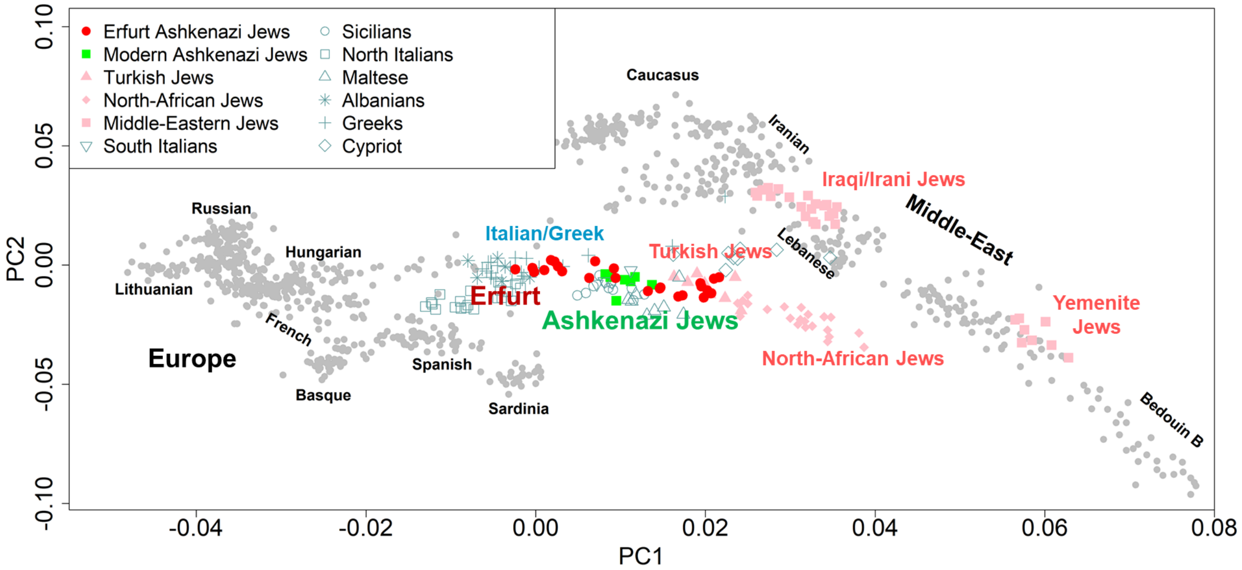
The excavation discovered 47 graves, densely buried in an area of just 16 x 12 square meters. All individuals except one were laid with their legs facing Jerusalem, following Jewish rabbinical tradition, and only one showed signs of a violent death. Eventually, all skeletons were reburied in the 19th-century Jewish cemetery. However, just before reburial, we sought to sample material for performing a DNA study. After discussions with the Jewish community of Thuringia, in light of the above mentioned rabbinical thought development, the community approved our study. We consequently collected detached teeth from 38 individuals.

7. What data did we generate?

Out of the 38 teeth, we generated genome-wide DNA sequences for 33 individuals, 19 females and 14 males, many of them children. While DNA quality varied, only six genomes had data for less 50,000 markers, and the median number of markers was about 400,000. Interestingly, we found three families: a mother and two children, a father and a daughter, and three second-degree relatives; family members were typically buried in proximity to one another. We further dated a subset of the teeth using carbon-14 to the 14th century. We next sought to use our sample of late medieval German Jews to study the origin of AJ.

8. What did the data tell us about the ancestry of Erfurt Jews?

Our first question was simple: do medieval Erfurt Jews and modern Ashkenazi Jews belong to the same genetic population? To answer that, we used a method called Principal Components Analysis (PCA). PCA is a way of transforming the entire DNA sequence of an individual (in our case, hundreds of thousands of markers) into a small number of coordinates that are most informative on differences between individuals. When using just the first two coordinates, individuals can be placed on a 2D "map", with individuals from the same or closely-related populations typically appearing nearby. The PCA for Ashkenazi Jews is presented below. In the figure, each dot represents a single individual, and some populations are highlighted. The plot shows that AJ (green) have an intermediate position in the PCA map between Middle Eastern and European populations, close to Turkish Jews and to other Mediterranean populations. Erfurt Jews (red) overlap the position of modern AJ, indicating that, broadly, medieval and modern AJ were part of the same genetic population. Our simulations showed that the AJ gene pool changed very little between the 14th century and the present, with a replacement of no more than 2-4% by outside influences from Eastern Europe during this 600-year period.



While broad ancestry patterns remained stable since the Middle Ages, a closer inspection of the PCA revealed a striking finding: Erfurt Jews were noticeably more diverse than modern AJ, spanning a wider range in the left/right (European/Middle Eastern) axis. An even closer inspection revealed that the Erfurt population was divided into two groups: one with more European ancestry compared to modern AJ, and one with more Middle Eastern ancestry. The presence of two distinct groups was supported by several statistical methods.

We next performed a series of analyses to characterize the two Erfurt groups. The group with the greater Middle Eastern ancestry was genetically most similar to modern AJ with origin in France and Germany. The group with the greater European ancestry had clear genetic influences from Eastern Europe. It can be speculated that some members of this group were migrants: records from the second half of the 14th century document Erfurt individuals with Slavic given names and with surnames indicating origin in Bohemia, Moravia, and Silesia. Further, the two groups had distinct levels of oxygen isotopes, indicating the usage of different water sources during childhood. Nevertheless, individuals from the two groups were buried side by side in the cemetery, suggesting no social segregation.

Can we speculate on the historical circumstances that created the two medieval AJ genetic groups? It turns out that in the Middle Ages, AJ were divided geographically, culturally, and linguistically into two major groups. Western Jews from the Rhineland (where AJ had first settled) spoke German dialects and used religious rites and given names different from those in the East. This community may correspond to the Erfurt group with the greater Middle Eastern ancestry. Eastern Jews, who inhabited eastern Germany, Austria, Bohemia, Moravia, and Silesia, spoke Old Czech. This group may correspond to the Erfurt group with the greater Eastern European ancestry.

Modern AJ do not show the medieval genetic heterogeneity; instead, they can be modeled as a nearly even mixture of the two Erfurt groups. Studies of names, dialects, and religious rites suggest that Western and Eastern AJ eventually merged and formed a single Ashkenazi culture, consistent with the genetic results. Erfurt was at the geographic boundary between the two medieval AJ communities, and in the 14th century, it was likely a home to Jews from both communities. These may be the historical circumstances leading to the detection of two genetically distinguishable groups in the Erfurt medieval cemetery.

What about more ancient sources of ancestry? Here, the picture is less clear. We attempted to model Erfurt Jews as a genetic mixture of several source populations. The best fitting model had three components: Southern European, particularly from South-Italy (explaining about 65% of the Erfurt ancestry), Middle Eastern (explaining about 20%), and Eastern European (another 15%). However, these numbers should not be overinterpreted. First, these estimates are based on present-day source populations. This generates technical problems (which we will not cover here), but, more importantly, a conceptual problem: the ancestry of individuals from any given region (say, South-Italy) today is likely different from what it was two millennia ago. Indeed, levels of Middle Eastern ancestry in Italy fluctuated widely throughout history. This implies that the estimated proportions of ancestry, particularly from the Middle East, may be substantially under- or over-estimated. Additionally, it is difficult to pinpoint the precise source populations, as we could fit the Erfurt data also with several other Southern European and Middle Eastern putative sources. In summary, this analysis suggests strong genetic links between medieval AJ and Southern European Mediterranean populations (supporting previous research), but the precise identity and influence of each source remain unclear. In contrast, the identification of a minor Eastern European ancestry component, already in 14th-century AJ, was robust.

9. What did the data tell us about the founder event?

Our next question was whether Erfurt Jews show signs of the founder event (“bottleneck”) that is so clearly evident in modern AJ. The answer was a definitive yes, based on three lines of evidence. The first was analyses of mitochondrial DNA. One sequence, carrying the cryptic name K1a1b1a, is found in 20% of modern AJ. It was also found in over a third (!) of the Erfurt individuals (11 out of 31 (after excluding children from mother/child pairs)). In other words, the early AJ population was so small that a third of the Erfurt individuals descended (through their maternal genealogical lines) from a single woman. The most likely time when she has lived is the first millennium CE, although we cannot rule out earlier periods. Other Erfurt Jews also carried mitochondrial sequences that are common in AJ today.

The second line of evidence was the presence of mutations (variants) that are common in AJ today but are extremely rare in other populations. This is a hallmark of a founder event, because variants carried by a founder can become common in a newly founded population, even if they had been extremely rare before. We found several Ashkenazi “founder variants” in Erfurt, in the expected amount if modern AJ and Erfurt AJ had shared the same bottleneck. Further, we found

16 *pathogenic* variants, several of them part of routine genetic testing in AJ. Recessive variants cause diseases only when inherited from both parents and are tested in AJ couples before they are about to have children. We identified variants for diseases such as retinitis pigmentosa, Gaucher, Usher syndrome, and factor XI deficiency. Dominant variants cause (or increase risk of) a disease even when inherited from just a single parent. Variants we identified include one in the *BRCA1* gene that increases risk for breast and ovarian cancer. Testing of this variant is freely available to all Israeli women of AJ descent.

We note that pathogenic founder variants were also found in another medieval AJ group, in a paper published a few months after our preprint. In that work, the remains of six individuals from 12th-century Norwich, UK, were rescued and their DNA sequenced. The individuals were found to have ancestry closely related to that of modern AJ, and they carried four Ashkenazi pathogenic variants. This convergence of evidence strengthens the case that late medieval AJ already experienced the founder event.

The final evidence for a founder event came from high levels of "runs of homozygosity". Whenever parents are related, they share identical DNA sequences in sections of their genome inherited from the common ancestor. In their children, the corresponding sections may be entirely "homozygous", namely having the same maternal and paternal sequences. This is precisely what we see in Erfurt. An analysis of 16 Erfurt individuals showed that nearly all carried multiple "homozygous" genetic blocks, implying relatedness between their parents. The lengths of the homozygous blocks suggested that the parents were not necessarily close relatives, but rather that the population has been extremely small for many generations.

To estimate the size of the Ashkenazi population throughout history, we used both modern and ancient DNA. In the modern data, we searched for DNA segments that are identical between individuals due to inheritance from common ancestors from the past ≈10-50 generations ago. Using these segments, we estimated that the effective population size of medieval AJ (which can be thought of as the number of "genetic" ancestors) has been maintained at approximately 1000-2000 for nearly 20 generations, in line with previous estimates (e.g., 1, 2). However, estimates based on runs of homozygosity in Erfurt suggested that their population size was approximately 3-fold smaller. To reconcile these estimates, we inferred that in addition to Erfurt, medieval AJ must have included at least one additional subgroup in which the bottleneck was less pronounced. This model fitted the data well. Nevertheless, we urge caution in making historical interpretations, as the model was based on several simplifying assumptions and other models may fit the data equally well. In summary, the Erfurt population has definitively undergone a founder event shared with present-day AJ, and it might have even descended from a slightly smaller set of founders.

A note. The effective population size is a concept used by geneticists, and it may not directly correspond to the actual census size. Typically, it is lower. To begin with, only reproducing individuals are counted, and these form only about a third of the population. The effective size is measured as the size of an idealized population that would generate the same patterns of genetic variation as observed in the real population. The idealized models usually assume that mating is random and is possible between all individuals, and that all individuals are equally likely to have children. When estimating bottleneck parameters, the estimated effective size can also change depending on how the bottleneck is modeled (e.g., a sudden vs a gradual event). While these are important limitations, the effective population size is still useful for two reasons. First, it provides geneticists with a convenient estimate of the expected degree of genetic diversity in the population (as larger populations are more genetically diverse). Second, while the effective size may not directly measure the census size, it can still capture major demographic trends in the history of populations, particularly bottlenecks and expansions. In the case of AJ, with an estimated effective size around 1000-2000 individuals, it is conceivable that the population size during the Middle Ages was in the thousands or at most tens of thousands, as proposed by past demographic studies.

10. What are the main insights we gained on Ashkenazi history?

Put together, our work generated several new insights that we summarize below.

1) Ashkenazi Jews have already acquired their main sources of genetic ancestry by the 14th century, including from Eastern Europe. There was little change in those ancestry components in the 600 years that followed, likely due to maintenance of very low levels of gene flow into the population.
2) In contrast, the internal genetic structure of AJ has changed over the years. Medieval AJ are best viewed not as a single homogeneous community (as it came to be at the present), but as an “archipelago” of communities, differentially affected by founder events and mixture with local populations.
3) Specifically, we identified a division within Erfurt between one group genetically similar to present-day AJ from Western Europe, who may represent descendants of Rhineland AJ, and another group with Eastern European ancestry, who may represent medieval AJ from Central and Eastern Europe who were at the time culturally and linguistically distinct from Western AJ.
4) A key source of pre-medieval AJ ancestry is related to people living today in Southern Europe, particularly Italy.
5) AJ were very small in number during the first centuries of the second millennium. Consequently, late medieval AJ already carried disease-causing and other variants that drifted to higher frequencies in AJ compared to neighboring populations.

11. What remains unknown?

We highlight below the main angles of AJ demographic history where our results remain uncertain or where the literature as a whole is still inconclusive.

1) Our work is based on a single cemetery, representing a single city in a single century. This leaves uncertainty regarding the generalizability of our inferences to the medieval Ashkenazi population as a whole. Are the two genetic groups we identified specific to 14th-century Erfurt? Will they be observed in other cities at the boundary between medieval Western and Eastern Jews? Will we detect additional AJ groups?
2) We are still lacking precise estimates of the AJ population size during the Middle Ages and beyond. Our current estimates are prone to multiple sources of technical errors and

are based on a long list of simplifying assumptions. More precise estimates will draw a clearer picture of the medieval AJ demographic depletion and the subsequent rapid growth.

3) The earlier origins of AJ, from before their settlement in Germany, are not well understood, with various competing theories and ambiguous genetic inferences. The scarcity of genetic data from the relevant time periods makes this question particularly challenging. Open questions include the route(s) of migration into Germany, the role of migrants from the Levant or the greater Middle East, relation of early AJ to other Jewish populations, and the places where AJ have intermarried.
4) We do not understand the process that formed the two Erfurt genetic groups. Did the two groups originate from a single source? When did they split? Did they mix before the 14th century and to what extent? Where exactly did the Erfurt-EU group mix with non-AJ populations? The role of minor gene flow into Erfurt-EU from the Caucasus or East Asia is also unclear.

12. What are the next steps?

Researchers have already generated genome-wide data for thousands of AJ individuals, and an order of magnitude more genomes are available at genetic testing companies. However, genomes of historical AJ exist only from two sites: Erfurt (Germany, 14th century, this study) and Norwich (UK, 12th century). Therefore, we expect that future progress in AJ population genetics will rely mostly on ancient DNA from additional sites. We expect the two ancient DNA studies just published to stimulate a discussion on associated cultural and religious issues, and thereby provide a firm ethical basis to future studies of ancient Jewish DNA.

It will be particularly important that the next ancient DNA studies will include burials from non-Ashkenazi Jews (from, e.g., Southern France, Italy, or Spain; both from antiquity and from the Middle Ages) and from Eastern Europe. Such sites could document the formation of Jewish communities in Europe, the relation between Jewish groups, and the recent AJ expansion. Sampling additional sites from medieval Central Europe would shed light on the history of the medieval AJ subgroups.

13. A recurring question from AJ readers: Were the 14th-century Erfurt Jews my ancestors? Will I be able to confirm this by comparing my DNA to theirs?

We showed that there was genetic continuity between the medieval Erfurt individuals and modern AJ. This raises the question of whether the Erfurt individuals were the ancestors, or close relatives of the ancestors, of any particular modern AJ individual. Also, we made public the genotypes of the Erfurt individuals. Therefore, interested readers who have genotyped their own DNA could in principle compare their DNA to that of the medieval Erfurt Jews. This can be performed, e.g., in GEDmatch. But what would the results imply? What if you find a section of your genome that’s identical with an Erfurt individual? (Known as identical-by-descent (IBD) segment.) What type of relationship to the ancient individual would that suggest?

To answer these questions, we use results from a few classic papers published ~20 years ago (e.g., 1, 2, 3). These papers derived very interesting mathematical properties of human genealogies. Consider a population consisting of N individuals each generation. Then consider people who have lived about log(N) generations ago, or about 500 years ago for typical humans. The mathematical modeling discovered the following. When going back 500 years ago, approximately 20% of the people back then have **zero descendants today**. In contrast, the remaining 80% are the **genealogical ancestors of the entire population**.

What do these results mean? Consider a famous medieval person, say, a famous rabbi or an individual that happened to be part of an ancient DNA study. If that person has **any** descendants today, then he or she is the **ancestor of the entire population today**. In other words, no single individual today is special in being a descendant of that person. In Erfurt, some of the individuals we sequenced were children, and they of course did not leave descendants. Other Erfurt individuals may have not left descendants today as well. But the ones who did are the genealogical ancestors of most AJ today.

How much genetic similarity do we expect to observe between a modern AJ individual and an Erfurt Jew? It is important to understand that each particular modern AJ *may or may not* inherit DNA from any particular Erfurt Jew (or from any other particular medieval person, for that matter). Let's explain why. Very roughly, each AJ today inherits DNA fragments from about 2000 14th-century ancestors. This happens because DNA is transmitted between generations in large chunks, which are pruned by rare recombination events. Therefore, despite the billions of letters in the DNA, each of us inherits DNA only from a much smaller number of recent ancestors.

In contrast to the small number of DNA ancestors, our number of **genealogical ancestors** grows exponentially fast: 2 after one generation, 4 after two, 8 after three, and so on, reaching millions already in the Middle Ages. Therefore, for each modern AJ, the 2000 medieval DNA ancestors are only a tiny fraction of the millions of *genealogical ancestors* that he or she has from 600 years ago. Of course, if the entire population consisted of only 10,000 people, the genealogy will reach each of the medieval ancestors through thousands of different genealogical paths.

So each present-day individual inherits DNA from only about 2000 14th-century ancestors. But who will be those ancestors? From the viewpoint of a living person today, these ancestors will be chosen at random by following a random genealogical path 600 years back. Importantly, the probabilities with which these ancestors will be selected are uneven. The more prolific an ancestor (and his/her children and grandchildren and so on) has been, the more genealogical paths will lead to that ancestor, and the more likely we are to inherit DNA from that particular ancestor. But now the key point. When looking back 600 years in the past, the mathematical models suggest that **we all have nearly the exact same genealogy**. Therefore, the number of genealogical paths that reach each ancestor will be the same for all present-day individuals. Consequently, all present-day individuals from a given population choose ancestors to inherit DNA from with the exact same probabilities.

Going back to Erfurt, these arguments imply that some AJ will inherit DNA from particular 14th-century Erfurt individuals. Some will inherit from other Erfurt individuals. Some may inherit from none. (And the same goes for inheriting DNA from any famous medieval person.) This is because any particular medieval person may or may not be one of those 2000 medieval ancestors each of us selects at random to inherit DNA from. But the probabilities that govern which ancestor is selected **are the same for everyone**. Therefore, seeing a particular shared DNA segment does not imply any unique genealogical relatedness between a modern AJ and an Erfurt Jew.

To summarize, all (or most) modern AJ are expected to be genealogical descendants of each and every Erfurt Jew who have descendants today. However, whether IBD segments are shared with any particular ancestor is determined at random, and it does not imply the presence or absence of genealogical relatedness to that ancestor. All modern AJ should be equally genealogically related to all medieval Ashkenazi Jews.

Of course, all of the above is assuming that there is a single population in which all individuals can marry all others with the same likelihood. In reality, populations are often geographically divided, such that individuals are less likely to have children with geographically remote spouses. Very likely there were also a few individuals with non-AJ ancestry historically joining the population. These will lead to deviations from the above predictions. For example, a group of modern AJ with origin in a given region may systematically share more (or less) DNA with the medieval Erfurt individuals. Similarly, individuals with recent non-AJ ancestry may share less than the average. Uncovering such patterns (which are expected to be subtle, given the overall genetic homogeneity of modern AJ), will be the subject of future research.