Article

Reference Population Design and the Illusion of Genetic Intermediacy in Mediterranean Population Models

Steven Parker *

Posted Date: 6 February 2026

doi: 10.20944/preprints202601.2134.v2

Keywords: population genetics; ancestry inference; admixture modeling; qpAdm; principal component analysis (PCA); genetic distance (FST); Ashkenazi Jews; Southern European ancestry; reference population bias; genome-wide analysis

doi:10.20944/preprints202601.2134.v2

Article

Reference Population Design and the Illusion of Genetic Intermediacy in Mediterranean Population Models

Steven Parker

Independent Genetics Researcher

Abstract

Genome-wide ancestry inference depends critically on reference population design, yet the sensitivity of inferred population relationships to reference choice is often underexamined. Here, I evaluate how asymmetric reference population construction can generate the appearance of genetic intermediacy in Mediterranean population models, using Ashkenazi Jewish populations as a focused test case. I analyze multiple independent frameworks, including qpAdm admixture modeling, pairwise autosomal FST distances, principal component analysis, identity-by-descent sharing, Global25-based affinity modeling, and reassessment of published uniparental marker studies. Across methods, I vary European and eastern Mediterranean reference sets to test the stability of inferred ancestry patterns under alternative, historically grounded configurations. When Southern European populations, particularly Southern Italian, Sicilian, Maltese, and Aegean groups, are included explicitly, Ashkenazi Jews consistently resolve within a Southern European and central Mediterranean genetic continuum rather than as an intermediate population between Europe and the Levant. In contrast, models that represent Europe using Northern Italian, Tuscan, or genetically drifted Sardinian proxies reproducibly shift Ashkenazi Jews toward an apparent Europe–Levant midpoint. Autosomal FST distances identify the closest affinities with Southern Italians, Cretans, Sicilians, and mainland Greeks, with substantially greater divergence from Levantine populations. qpAdm and identity-by-descent analyses likewise support predominant Southern European ancestry. These results demonstrate that apparent genetic intermediacy is largely a methodological artifact arising from reference population exclusion and underscore the need for explicit sensitivity testing in ancestry inference.

Keywords: population genetics; ancestry inference; admixture modeling; qpAdm; principal component analysis (PCA); genetic distance (FST); Ashkenazi Jews; Southern European ancestry; reference population bias; genome-wide analysis

---

Part I. Data Sources and Analytical Framework

I evaluated multiple source configurations to test how inferred ancestry components change when Southern European and central Mediterranean references are included directly versus when Europe is approximated using Northern Italian, Tuscan, or Sardinian proxies. The core three way source configuration used for primary comparisons was Italian_South.HO, Lebanese.HO, and Russian.DG. I also ran alternative models that substituted Northern and Central Italian sources in place of Italian_South.HO to assess the stability of ancestry proportion estimates and model behavior under commonly used proxy representations.

I interpret qpAdm results by prioritizing comparative stability across plausible reference configurations. The emphasis is on whether the inclusion or exclusion of Southern European and central Mediterranean reference populations changes the placement and inferred proportions for the Ashkenazi Jewish target, rather than on treating any single qpAdm run as definitive.

doi:10.20944/preprints202601.2134.v2

Pairwise FST Analysis

I evaluated autosomal differentiation using pairwise FST distances obtained through the same AdmixLab environment, which provides FST distance calculations in addition to qpAdm modeling. I use FST as a comparative metric to rank relative affinities between Ashkenazi Jewish samples and regional comparator populations, including Southern Italian, Sicilian, Maltese, Greek and Aegean, Levantine, and additional European references. I interpret FST magnitudes in relation to within Europe benchmarks, including North South Italian differentiation and internal Iberian distances, to provide an external scale for assessing whether Ashkenazi Jews appear intermediate, regionally embedded, or shifted under different reference sets.

Global25 and Affinity Visualization

I examined additional autosomal affinity patterns using Global25 coordinates as a descriptive visualization framework. Global25 is a principal component based coordinate system intended for exploratory analysis and mixture modeling. I used Global25 distance based comparisons and heatmap style visualizations to illustrate regions of strongest relative affinity for Italian Jewish and Ashkenazi Jewish samples across Southern Italy, Sicily, Malta, the Aegean, and surrounding regions. These outputs are presented as descriptive corroboration of the qpAdm and FST results and are not treated as formal inferential estimates of ancestry proportions.

DNAgenics Admixture Modeling

I generated supplementary autosomal admixture summaries using the DNAgenics calculator. I used default settings with no additional parameter tuning beyond the selections documented in the recorded outputs. I treat these results as heuristic sensitivity checks intended to illustrate how reference composition affects outputs in commonly used frameworks, rather than as definitive ancestry estimators.

Original Analyses and Derived Data Generated in This Study

This study generates new quantitative outputs through original analyses performed by me. These include qpAdm model runs under alternative reference population configurations, autosomal FST distance comparisons tabulated for targeted regional contrasts, Global25-based affinity visualizations, and comparative summaries derived from DNAgenics outputs. Although the underlying genotype data are previously published, the specific analyses, model comparisons, and derived outputs reported here were generated for this study and are used to evaluate the stability of Ashkenazi origin models under alternative and historically grounded reference assumptions. Supplementary materials documenting population lists, model specifications, and raw outputs are provided in Supplementary Files S1–S4.

Interpretive Framework

A central premise of this study is that autosomal genetic structure provides the primary basis for population level affinity and ancestry inference because it reflects genome wide contributions. Uniparental markers are lineage specific and can be disproportionately shaped by drift, founder effects, and endogamy. Accordingly, I treat autosomal results as the primary evidence for comparative population structure, while using Y chromosomal evidence to assess consistency with autosomal patterns and to evaluate whether apparent lineage based signals change when analyzed at higher resolution and with appropriate Mediterranean comparators.

Reference populations are treated as modeling assumptions rather than neutral inputs. The use of Northern Italian, Tuscan, or Sardinian populations as proxies for Europe is evaluated as a specific methodological choice with predictable effects on inferred placement when Southern Italian and central Mediterranean reference populations are omitted. Southern Italian, Sicilian, Maltese, and

doi:10.20944/preprints202601.2134.v2

Aegean populations are treated as core Southern European and central Mediterranean references based on genetic structure and historical continuity.

Rationale for Reanalysis Without New Sequencing

This study does not introduce newly sequenced genomes or new biological sampling because a central source of disagreement in the literature concerns reference population design and proxy choice rather than the existence of relevant genotype data. When upstream modeling assumptions are misaligned, additional data analyzed under the same reference framework can reproduce existing artifacts with increased apparent confidence rather than resolve them. Reanalysis under revised and explicit reference assumptions is therefore methodologically prior to new sampling in this context. The analyses presented here are intended to test whether Ashkenazi Jewish population structure is best modeled as predominantly Levantine, intermediate, or Southern European and central Mediterranean when Europe is represented directly and in a historically grounded manner.

Part II. Reference Population Design and the Appearance of Genetic Intermediacy

The frequent characterization of Ashkenazi Jews as genetically intermediate between Europe and the Levant is often presented as an independent convergence across autosomal studies. However, this pattern is also consistent with repeated use of a shared reference population design that represents Europe primarily with Northern and Central European groups and a limited set of Italian proxies, while omitting Southern Italian and Sicilian populations that are directly relevant to Mediterranean population structure and to historically plausible source populations. Because this reference template is reused across influential studies and secondary syntheses, an apparent pattern of intermediacy can be propagated even when underlying autosomal affinities are closer to the central Mediterranean. In this section, I use "genetic intermediacy" to refer to intermediate placement in distance-based or PCA-based representations relative to the specific reference sets included in a given analysis.

A central example is the neighbor-joining tree presented by Ostrer and Skorecki (2013), which has been widely cited as a summary of autosomal evidence for shared ancestry among Jewish Diaspora populations together with variation in admixture with surrounding non-Jewish populations.

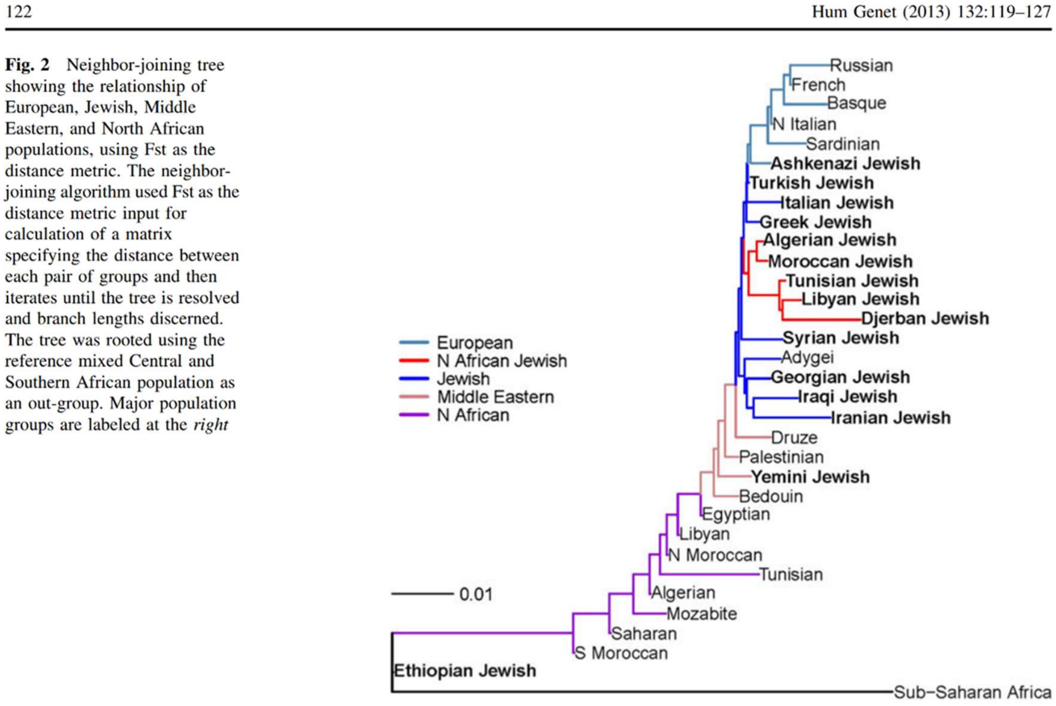

122 Hum Genet (2013) 132:119–127

Fig. 2 Neighbor-joining tree showing the relationship of European, Jewish, Middle Eastern, and North African populations, using Fst as the distance metric. The neighbor-joining algorithm used Fst as the distance metric input for calculation of a matrix specifying the distance between each pair of groups and then iterates until the tree is resolved and branch lengths discerned. The tree was rooted using the reference mixed Central and Southern African population as an out-group. Major population groups are labeled at the *right*

Figure 1. Neighbor-joining tree reproduced from Ostrer and Skorecki (2013). Tree constructed from autosomal genetic distances using Northern Italians, Tuscans, Sardinians, Basques, French, Russians, and other Northern or Central European populations as European references. Southern Italian and Sicilian populations are not included.

Ostrer and Skorecki describe their synthesis as drawing on a common set of autosomal studies, including Atzmon et al. (2010), Behar et al. (2010), Kopelman et al. (2009), Bray et al. (2010), Listman et al. (2010), and Campbell et al. (2012). They emphasize that these studies differed in the populations sampled and sample sizes, yet report broadly similar conclusions, which they interpret as evidence for shared genetic ancestry among major Jewish groups together with varying degrees of admixture from local host populations.

However, the studies cited in this synthesis share an important design feature that is not typically foregrounded in secondary summaries: the absence of Southern Italian and Sicilian autosomal comparator populations. Across this literature set, Europe is commonly represented by Northern Italians, Tuscans, Sardinians, and Northern or Central European groups, while Southern Italy and Sicily are not included as explicit comparators. In addition, Sardinians, an island population shaped by distinctive drift and reduced post-Neolithic gene flow, are frequently used as an Italian proxy despite limited suitability as representatives of mainland Italian population structure. Some analyses and syntheses include Italian Jewish samples, including in the Ostrer and Skorecki presentation, but inclusion of Italian Jews does not substitute for direct inclusion of non-Jewish Southern Italian and Sicilian reference populations when the objective is to evaluate whether Ashkenazi Jews fall within the Southern European and central Mediterranean continuum.

The consequences of this shared reference design are evident in the topology of the Ostrer and Skorecki tree. If Southern Italian and Sicilian populations are excluded, then Southern European and central Mediterranean variation is incompletely represented. Under such conditions, Ashkenazi Jewish samples cannot be evaluated for placement within the omitted Southern European reference space because that space is absent from the analysis. The resulting placement between a truncated European pole and Levantine populations is therefore partly influenced by reference definition and population inclusion, rather than arising solely from the genetic relationships present in the broader Mediterranean context.

Accordingly, similarity of conclusions across the cited studies should not be treated as independent confirmation of an intermediate Europe-Levant position. A repeated inference can reflect reuse of a shared reference scaffold. When Europe is operationalized without Southern Italian and Sicilian comparators and Italy is represented by northern proxies and Sardinians, analytical representations can shift Mediterranean-adjacent populations toward an apparent midpoint between the available European and Levantine references.

This reference template also shaped interpretation when proximity to Italian populations was observed. For example, Bray et al. (2010) note proximity of Ashkenazi Jews to Italian samples but propose that admixture with Central and Eastern European host populations could shift Ashkenazi Jews toward Italians along principal component axes relative to Middle Eastern populations. This hypothesis was advanced without directly testing Southern Italian and Sicilian ancestry within the same framework, despite its historical plausibility and relevance to interpreting Italian proximity. More generally, principal component geometry can be sensitive to reference population choice and does not by itself identify the historical sources responsible for intermediate placement.

In this study, I therefore treat claims of Ashkenazi genetic intermediacy and related interpretations as contingent on reference population design. Subsequent sections do not assume that intermediate placement is an independent empirical finding that must be explained. Instead, I evaluate Ashkenazi population structure using Southern Italian, Sicilian, Maltese, and Aegean comparators that were absent from the inherited template and I explicitly test the stability of inferred affinities under alternative, historically grounded reference configurations.

Summary of Shared Reference Population Design Features (Descriptive Synthesis)

Category	**Populations included across cited studies**	**Relevant populations excluded**
Italian reference populations	Northern Italians; Tuscans; Sardinians	Southern Italians; Sicilians
European pole definition	Northern and Central Europeans combined with Northern Italians and Tuscans	Mediterranean Southern Europeans
Italian proxy practice	Sardinians used as an Italian proxy; Northern and Central Italians treated as representative of Italy	Southern Italian population structure excluded; Italy modeled northward
Jewish populations included	Ashkenazi; Sephardi; Mizrahi; Italian Jews (included in some analyses and in the Ostrer and Skorecki synthesis)	Not applicable at the level of Jewish group inclusion; the primary omission concerns non-Jewish Southern Italian and Sicilian comparator populations
Mediterranean comparators	Limited or none among non-Jewish central Mediterranean comparators	Southern Italy; Sicily; Maltese; Aegean Greeks
Structural consequence	Europe represented northward	Southern European continuum incompletely represented
Expected effect on placement	Apparent Europe-Levant intermediacy	Mediterranean anchoring cannot be evaluated

Taken together, the synthesis by Ostrer and Skorecki (2013) and the common reference population structure of the studies they cite indicate that apparent Ashkenazi intermediacy can be reproduced through shared reference population exclusion. This section establishes the methodological baseline for interpreting subsequent autosomal and uniparental results. In the sections that follow, I evaluate population affinities under reference configurations that include Southern European and central Mediterranean comparators, and I treat intermediate placement observed under truncated European reference sets as a design-dependent outcome that requires explicit sensitivity testing.

Part III. Y-Chromosomal Structure of Ashkenazi Jews in Mediterranean Context

Uniparental Misinterpretation and the Illusion of a Levantine Paternal Origin

Claims that Ashkenazi Jews possess a predominantly Levantine paternal ancestry have relied heavily on a small set of Y-chromosomal studies published between approximately 2001 and 2009. These studies are frequently cited as having resolved the question of Ashkenazi origins, particularly through the presence of haplogroups J1, J2, and E, which are routinely labeled as Near Eastern or Levantine. However, closer examination shows that this conclusion is not anchored in robust comparative inference. Instead, it reflects a combination of reference population omission, low phylogenetic resolution, and population pooling that together limit what these datasets can actually support about Ashkenazi-specific paternal origins.

With the exception of Semino et al. (2004), the major Ashkenazi Y-DNA analyses from this period did not include Southern Italian populations as comparators, despite their central relevance to Mediterranean genetic structure and their well-documented Jewish presence since antiquity. Instead, these studies typically contrasted Ashkenazi Jews with Middle Eastern populations on one side and Northern or Central European populations on the other. This reproduces, in uniparental form, the same asymmetric reference framework later observed in autosomal studies. Under such a design, haplogroups shared broadly across the eastern Mediterranean and Southern Europe are implicitly reclassified as Levantine by default, not because their geographic specificity has been demonstrated, but because Southern European baselines are absent.

Hammer et al. (2009) illustrates the problem clearly. Although frequently cited as evidence for a Levantine paternal origin of Ashkenazi Jews, the study explicitly pools Ashkenazi and non-Ashkenazi Jewish samples and does not report haplogroup frequencies stratified by Jewish subgroup. The authors describe their dataset as 1,575 Jews representative of the Diaspora, without providing Ashkenazi-specific proportions. As a result, the reported haplogroup distributions cannot be used to infer Ashkenazi paternal structure, yet they are routinely cited as if they do. In addition, haplogroups are often reported at coarse resolution, such as J or E-M35, which obscures subclade-level geographic structure and increases the risk of treating pan-Mediterranean lineages as region-specific markers.

Other frequently cited studies show similar limitations. Behar et al. (2004) and Nebel et al. (2001) report Ashkenazi Y-DNA frequencies but do so using broad haplogroup bins and without Southern European comparators. Shen et al. (2004) is based on a very small Ashkenazi sample and lacks the resolution required to distinguish Levantine-specific from pan-Mediterranean lineages. In all cases, haplogroups such as J2 and E-M35 are interpreted as Near Eastern in origin despite their well-established prevalence in Southern Italy, Greece, and Anatolia.

Semino et al. (2004) is distinct because it includes multiple Southern Italian populations alongside Ashkenazi Jews and reports downstream resolution for haplogroups E and J. When that comparison is made explicitly, the supposed Levantine signal collapses. Ashkenazi Jews exhibit frequencies of J2 and E-M123 that fall within Southern Italian ranges, while Levantine populations often show markedly different distributions, including elevated J1 and reduced E-M123. Notably, E-M123, frequently cited as evidence of Levantine ancestry in Ashkenazim, reaches comparable or higher frequencies in Southern Italy and Anatolia, while being relatively uncommon in several Levantine populations.

Thus, the interpretation of Ashkenazi Y-chromosomal lineages as predominantly Levantine is not supported by the totality of the evidence. Rather, it emerges from a methodological

framework in which Southern Europe is excluded, haplogroup resolution is insufficient, and population pooling obscures Ashkenazi-specific structure. When appropriate Mediterranean comparators are included, Ashkenazi paternal lineages align closely with Southern Italian and Aegean populations, consistent with a Mediterranean origin shaped by founder effects and endogamy rather than a Levantine paternal majority model.

As summarized in Table 1, the early Y-chromosomal literature does not provide consistent or properly contextualized evidence for a predominantly Levantine paternal origin of Ashkenazi Jews. Resolving this ambiguity requires focusing on the only dataset from this period that places Ashkenazi Jews and Southern Italian populations side by side at subclade resolution.

Table 1. Y-DNA haplogroup frequencies among Ashkenazi Jews (E, J1, J2 only).

Study	Ashkenazi sample size	E total	E-M78	E-M123	J1-M267	J2-M172	Notes
Hammer et al. 2009	1,575 (pooled)	NR	~3	~17	~17	~20	Ashkenazi and non-Ashkenazi Jews pooled; subgroup frequencies not reported
Behar et al. 2004	~442	~16.1	NR	NR	~19	~19	No Southern European comparators; coarse resolution
Semino et al. 2004	77	NR	5.2	11.7	14.6	23.2	Only study including Southern Italians; higher subclade resolution
Nebel et al. 2001	~79	23	NR	NR	~19	~24	No Southern European comparators
Shen et al. 2004	~20	NR	~10	~10	~20	~15	Very small sample; limited resolution

Table 1 note. In all tables in this manuscript, missing source values are coded as NR (not reported in the cited study). Values prefixed by ~ are approximate as reported. All haplogroup frequencies are percent.

Reassessment of Ashkenazi Y-Chromosomal Structure Using Southern Italian Comparators

The only early Y-chromosomal study to directly compare Ashkenazi Jews with Southern Italian regional populations at subclade resolution is Semino et al. (2004). Unlike later syntheses that rely on pooled Jewish samples or coarse haplogroup bins, this study reports haplogroups E and J for Ashkenazi Jews alongside multiple Southern Italian regional samples, including Apulia, Calabria, and Sicily. This design enables direct evaluation of whether Ashkenazi paternal lineages commonly labeled as Levantine are exceptional relative to Southern Europe.

At this resolution, Ashkenazi Y-chromosomal structure is dominated by haplogroups E and J2, both of which fall within Southern Italian variation. Ashkenazi Jews show E-M123 at 11.7 percent and E-M78 at 5.2 percent. These values closely match Southern Italian regions such as Apulia (E-M123 = 11.6 percent), Calabria 2 (E-M123 = 13.2 percent), and Sicily (E-M78 = 12.7 percent). Calabria 1 shows the inverse pattern, with high E-M78 (16.3 percent) and low E-M123 (2.5

percent), underscoring strong regional heterogeneity within Southern Italy itself. This heterogeneity shows that neither E-M123 nor E-M78 is geographically diagnostic of the Levant.

Haplogroup J shows the same pattern of Mediterranean continuity. In Semino et al. (2004), Ashkenazi Jews exhibit J2 (J-M172) at 23.2 percent, making it the dominant J lineage. Comparable or higher J2 frequencies are observed in Apulia (29.1 percent), Calabria 1 (22.8 percent), and Calabria 2 (20.0 percent), while Sicily shows moderately lower J2 (16.7 percent). In both Ashkenazim and Southern Italians, J2 exceeds J1, a structure that contrasts sharply with Levantine populations.

J1 (J-M267) is present in Ashkenazi Jews at 14.6 percent, elevated relative to some Southern Italian regions but well within a Mediterranean range. Southern Italy itself exhibits substantial regional variation in J1, ranging from very low values in Apulia (2.3 percent) and Calabria 1 (1.8 percent) to higher values in Sicily (7.1 percent). This intra-regional spread shows that J1 is not absent from Southern Europe and that moderate J1 frequencies cannot be treated as evidence of Levantine origin in isolation.

Crucially, the Ashkenazi J1 frequency remains far below Levantine paternal structures, in which J1 typically dominates the Y-chromosomal pool and often exceeds J2. The Ashkenazi pattern instead mirrors Southern European structure, with J2 dominant and J1 secondary. Given the long-documented founder effects and endogamy in Ashkenazi populations, modest elevation of J1 relative to some Southern Italian regions is parsimoniously explained by genetic drift acting on a small number of founding lineages rather than by descent from a Levantine paternal majority.

When Southern Italian comparators are included, the supposed Levantine signal in Ashkenazi Y-DNA disappears. Ashkenazi Jews align with a Mediterranean continuum characterized by substantial E-M123 and J2, moderate and variable J1, and strong regional heterogeneity rather than with the J1-dominant structure typical of Levantine populations. The interpretation of Ashkenazi paternal ancestry as predominantly Levantine arises primarily when Southern Europe is excluded and when drift-sensitive lineages are over-interpreted without comparative context.

Table 2 summarizes the Semino et al. (2004) data directly.

Table 2. Semino et al. (2004) Y-DNA haplogroup frequencies for Ashkenazi Jews and Southern Italian regions (E, J1, J2 only).

Population	Sample size	E-M78	E-M123	J1-M267	J2-M172
Ashkenazi Jews	77	5.2	11.7	14.6	23.2
Italian (Apulia)	86	0.0	11.6	2.3	29.1
Italian (Calabria 1)	80	16.3	2.5	1.8	22.8
Italian (Calabria 2, Albanian community, Cosenza)	68	5.9	13.2	0.0	20.0
Italian (Sicily)	55	12.7	3.6	7.1	16.7

In contrast, Levantine populations are characterized by a fundamentally different Y-chromosomal structure, with J1 typically exceeding J2 and E-M123 occurring at substantially lower frequencies than those observed in both Ashkenazi Jews and Southern Italians. This inversion of haplogroup dominance is incompatible with a Levantine paternal majority model for Ashkenazim and instead supports a Mediterranean paternal structure shaped by founder effects and long-term endogamy.

Demonstration of Ashkenazi Divergence from Levantine Paternal Structure

If Ashkenazi paternal ancestry were predominantly Levantine, Ashkenazi Y-chromosomal structure would be expected to resemble Levantine populations when evaluated using the same haplogroups and subclades. Levantine populations are commonly characterized by high frequencies of J1-M267, reduced representation of J2-M172 relative to J1, and comparatively low frequencies of E-M123. This pattern reflects demographic expansions that differ markedly from Southern European and Mediterranean Y-chromosomal structure.

When Ashkenazi Jews are compared against Levantine populations using these same markers, a clear divergence is observed. Ashkenazim show E-M123 and J2 at levels comparable to Southern Italian and Aegean populations, while J1 remains secondary. In contrast, Levantine populations show a reversal of this structure, with J1 dominating the paternal pool and E-M123 occurring at lower frequencies.

This contrast is summarized below.

Table 3. Y-DNA haplogroup structure in Ashkenazi Jews compared with Levantine populations (E, J1, J2 only).

Population	E-M78	E-M123	J1-M267	J2-M172
Ashkenazi Jews	5.2	11.7	14.6	23.2
Lebanese	~2 to 5	~4 to 5	~25 to 35	~10 to 15
Palestinians	~1 to 3	~2 to 4	~35 to 45	~10 to 15
Iraqis	~1 to 3	~2 to 3	~40 to 50	~10 to 15

In this comparison, Ashkenazi Jews do not align with Levantine populations on the paternal axis. Instead, they occupy a Mediterranean position characterized by elevated J2 and E-M123 alongside a secondary J1 component. The Levantine profile is J1-dominant with lower J2 and reduced E-M123. This inversion rules out a Levantine paternal majority model for Ashkenazi Jews.

Importantly, moderate elevation of J1 in Ashkenazim relative to some Southern Italian regions does not imply Levantine origin. Southern Italy itself exhibits substantial J1 heterogeneity, and Ashkenazi populations experienced strong founder effects and long-term endogamy that can amplify specific lineages. Drift can elevate J1 frequencies without implying descent from a population whose overall structure is J1-dominant. By contrast, Levantine J1 dominance reflects sustained demographic expansion that is not mirrored by the Ashkenazi paternal profile.

Thus, when Ashkenazi Jews are compared directly to both Southern European and Levantine populations using the same Y-chromosomal markers, their paternal structure aligns with a Mediterranean continuum and diverges sharply from Levantine patterns. Claims that Ashkenazi Y-DNA confirms a predominantly Levantine origin arise primarily when Southern Europe is excluded and when drift-sensitive lineages such as J1 are interpreted without comparative context.

Fine-Scale Southern European J1 Variation and Upper-Range Overlap

Fine-scale surveys of Southern European Y-chromosomal variation show that the range of haplogroup J1 observed in Southern Italy by Semino et al. (2004) does not represent an upper bound. When Southern Europe is sampled at micro-regional resolution (Di Giacomo et al., 2003; Boattini et al., 2013), J1-M267 frequencies are highly heterogeneous and locally elevated across the broader Greco-Roman Mediterranean continuum. Province-level Italian datasets document multiple Southern Italian localities with J1 frequencies exceeding those reported in Semino et al., particularly in Sicily and parts of Campania. A later synthesis of Sicilian and Southern Italian genetic data (Sarno et al., 2017), incorporating these underlying regional surveys, confirms the presence of elevated local Southern Italian values while showing that these populations remain J2-dominant overall. These upper-range J1 values approach or overlap those observed in Ashkenazi Jews, demonstrating that moderate J1 enrichment in Ashkenazim does not exceed plausible Southern European source variation once regional structure and founder effects are taken into account.

doi:10.20944/preprints202601.2134.v2

Comparable J1 levels are also observed in parts of Central Italy, including Tuscany, indicating that elevated J1 is not restricted to the southern peninsula. Elevated values are likewise reported in the Aegean. Fine-scale sampling on Crete reports J1-M267 frequencies overlapping Southern Italian and Ashkenazi levels, including a value of approximately 10.5 percent at Nea Nikomedeia (King et al., 2008).

Table 4. Upper-range regional J1-M267 frequencies in Southern Italy and the Aegean.

Region	Population	J1-M267
Italy	Sicily, Agrigento	11.1
Italy	Western Sicily	10.13
Greece	Crete, Nea Nikomedeia	10.5
Italy	Trapani, Sicily	8.82
Italy	Campania, Benevento	8.3
Greece	Crete	8.3
Italy	Sicily, Southwest	7.0
Italy	Calabria, Tyrrhenian Calabria	6.8

Frequencies shown represent high reported regional J1-M267 values in Southern Italy and the Aegean. Agrigento and other synthesized values are reported in Sarno et al. (2017) and derive from underlying province-level datasets, including Boattini et al. (2013). Western Sicily and Trapani values are taken directly from Boattini et al. (2013). The Nea Nikomedeia value is reported in King et al. (2008). All listed populations retain a J2-dominant Mediterranean paternal structure.

Cohen Modal Haplotype Arguments and Comparator Omission

A final pillar often invoked to support a Levantine interpretation of J1 in Ashkenazi Jews is the Cohen Modal Haplotype framework, most prominently articulated by Hammer et al. (2009). However, this study does not resolve Ashkenazi population origins and is frequently overextended beyond its evidentiary limits. Hammer et al. show that Cohanim are genetically heterogeneous, spanning more than twenty Y-chromosomal haplogroups, with J-P58 accounting for less than half of Cohanim lineages. The majority of Jewish males, including many bearing Cohen surnames, do not carry this haplotype, and the original low-resolution CMH is explicitly shown to be non-specific, occurring across Jewish castes and in non-Jewish populations. CMH-based arguments therefore do not establish a Levantine paternal majority for Ashkenazi Jews.

Crucially, Hammer et al. do not evaluate J-P58 or CMH frequencies against Southern Italian, Sicilian, or Aegean populations, despite these regions showing both autosomal proximity to Ashkenazi Jews and overlapping regional J1 frequencies. Assertions of Near Eastern specificity therefore rest on comparator omission rather than demonstrated exclusivity. In addition, the study invokes descent from an ancestral priestly lineage traditionally associated with Aaron, yet no ancient genetic evidence from such a source exists, and no external reference population is available to test the claim directly. The argument therefore relies on internal Jewish caste structure rather than external geographic validation.

This limitation is not unique to Hammer et al. Subsequent large studies focused on J-P58 phylogeography, including Chiaroni et al. (2010) and Fadhlaoui-Zid et al. (2011), restrict sampling largely to the Levant, Arabia, and North Africa and do not evaluate European or Mediterranean populations. As a result, these studies cannot distinguish Levantine specificity from broader Mediterranean continuity. When Southern Europe is properly sampled, J1 frequencies observed in Ashkenazi Jews fall within the upper range of documented Southern Italian and Aegean variation while remaining far below the J1-dominant structure characteristic of Levantine

populations. Accordingly, J1 cannot be treated as evidence of a Levantine paternal majority in Ashkenazi Jews and is more parsimoniously interpreted as a minor lineage shaped by Mediterranean ancestry and subsequent founder effects.

Part IV. Interpretive Framework and Constraints on Inference

Genome-wide analyses provide a useful context for evaluating how Ashkenazi Jewish population placement depends on reference population design rather than assuming a fixed Europe–Levant axis. When Southern European and central Mediterranean populations are included explicitly, Ashkenazi Jews are positioned within the same multivariate space as Southern Italian, Sicilian, Maltese, and Greek populations across several analytical frameworks, including principal component analysis (PCA), ADMIXTURE-based clustering, and identity-by-descent (IBD) sharing. Because these methods capture population structure through different mathematical mechanisms, their agreement under comparable reference configurations serves as a constraint on interpretations that rely on asymmetric or truncated reference sets.

An illustrative example is provided by Drineas et al. (2019) in their population genetic study of Crete. In PCA analyses comparing Cretans with European populations, the authors report that Cretans overlap most closely with Peloponnesians, Sicilians, and Ashkenazi Jews when Southern Europe is directly represented. In that framework, Ashkenazi Jews do not align along a simple Europe–Levant gradient but fall within a broader Southern European and central Mediterranean cluster. The authors interpret this configuration as reflecting shared Mediterranean ancestry, illustrating how inclusion of relevant southern comparators alters inferred population geometry.

Comparable sensitivity to reference composition is observed in large-scale ancient DNA studies that include modern comparative PCA. In Schuenemann et al. (2017), which incorporates extensive present-day reference populations alongside ancient Egyptian genomes, Ashkenazi Jews project within Southern European PCA space together with Southern Italians, Sicilians, and Maltese. These projections demonstrate that Mediterranean-inclusive reference sets allow evaluation of Ashkenazi placement relative to Southern Europe rather than forcing comparison between truncated European and Levantine poles.

Identity-by-descent analyses impose additional interpretive constraints because they quantify shared ancestry through inherited chromosomal segments rather than through geometric projection. Xue et al. (2017) report IBD sharing between Ashkenazi Jews and regional populations stratified by segment length, allowing separation of deeper shared ancestry from more recent shared ancestry. Within this framework, long shared segments reflect relatively recent common ancestry, whereas shorter segments capture older population structure. The relative presence or absence of long IBD segments therefore constrains models that posit substantial recent shared ancestry with specific source populations.

Taken together, these examples define interpretive boundaries rather than final conclusions. Apparent Ashkenazi “intermediacy” is shown to be contingent on reference population choice, comparator inclusion, and the mathematical properties of the analytical method employed. Within this framework, intermediate placement is treated as a hypothesis generated under specific modeling assumptions that must be tested explicitly under alternative, historically grounded reference configurations.

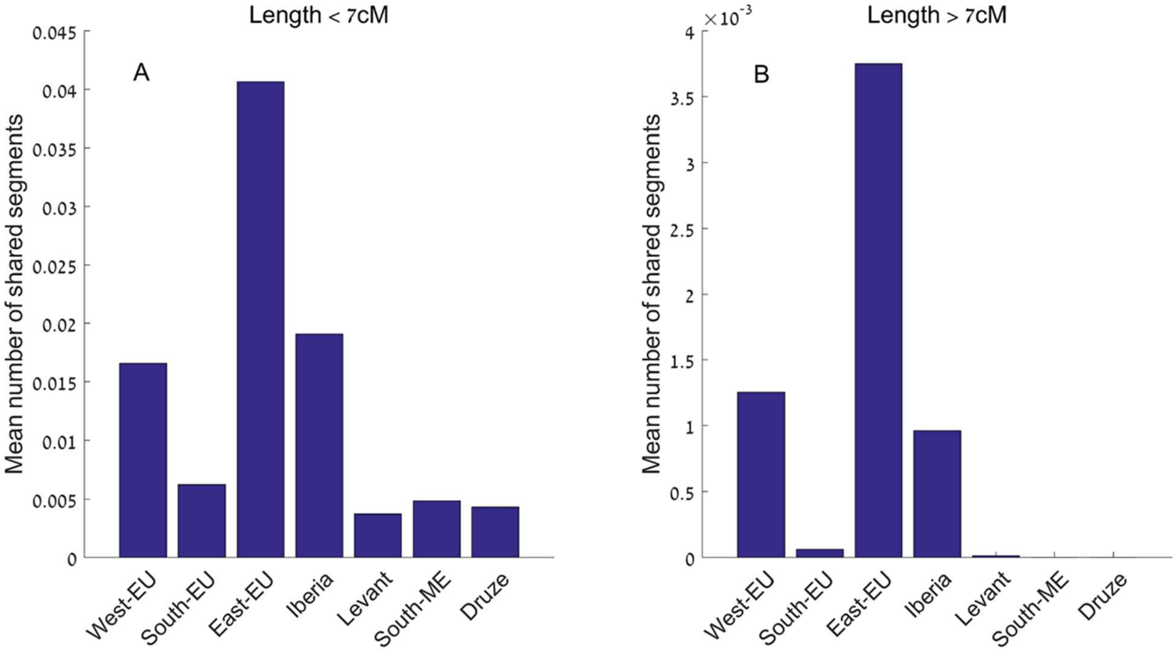


Figure 2. Identity-by-descent sharing between Ashkenazi Jews and regional populations from Xue et al. (2017). Panel A shows mean shared IBD segments shorter than 7 cM, reflecting deeper shared ancestry, while Panel B shows mean shared IBD segments longer than 7 cM, reflecting more recent shared ancestry. The distribution of shared segment lengths illustrates how IBD-based analyses constrain inferences about recent versus older shared ancestry across regional populations and provide an interpretive benchmark for evaluating models that posit substantial recent Levantine versus European contributions.

Part V. Convergent Autosomal Modeling Across Independent Analytical Frameworks

Independent autosomal studies using distinct methodologies, datasets, and temporal frameworks converge on a consistent pattern: most inferred Ashkenazi Jewish ancestry aligns with Southern Italy and the central Mediterranean. I observe this convergence across local ancestry inference, qpAdm modeling, and ancient DNA anchored demographic reconstruction, despite substantial differences in analytical approach and reference design.

Lerga-Jaso et al. (2025) applied a high-resolution local ancestry inference framework (Orchestra) to Ashkenazi Jewish genomes. Orchestra combines a recombination-based first-pass layer grounded in identity-by-descent assumptions with a deep-learning smoothing module integrating convolutional and attention-based architectures. The method assigns ancestry at fine genomic scale by leveraging the expectation that individuals share longer haplotypes with more closely related populations and shorter haplotypes with more distant ones. Using this framework, the authors inferred Ashkenazi ancestry as predominantly Italian, estimating ~68% Italian ancestry, followed by Levantine (16.6%), Iraqi-Iranian-Caucasian-Turkish (7.2%), Greek-Balkan (2.4%), and Eastern European (1.7%) components.

A complementary estimate is provided by Waldman et al. (2022), who used qpAdm to evaluate historically plausible admixture models for Ashkenazi Jews. Under their preferred model, Ashkenazi ancestry is inferred as ~65% Southern Italian, ~19% Lebanese, and ~16% Eastern European. I do not interpret the Lebanese component as a stand-alone Levantine origin term. Lebanese populations carry substantial eastern Mediterranean structure and are shifted toward Southern European genetic space relative to Arabian-associated Levantine poles. Accordingly, a Lebanese proxy in qpAdm is compatible with a broader Mediterranean continuum and should not be treated as evidence for a Levantine-majority model.

Independent confirmation comes from ancient DNA anchored modeling in Brace et al. (2022), which analyzed genomes from a medieval Jewish burial assemblage in Norwich, England (Chapelfield). Using qpAdm, the authors first show that present-day Ashkenazi Jews are best modeled as deriving entirely from the Chapelfield population, indicating strong continuity. They then model the Chapelfield individuals as a mixture of modern proxy populations and estimate ~67% Sicilian and 33% Turkish Jewish ancestry, with no detectable French or Polish contribution. Because Sicilian ancestry captures Southern Italian Mediterranean structure and Turkish Jewish ancestry reflects eastern Mediterranean Jewish populations, this model again places the dominant Ashkenazi ancestral component in Southern Italy and the central Mediterranean prior to later Eastern European admixture.

Although these studies differ in methodology, proxy composition, and temporal depth, their point estimates converge on the same interpretive conclusion. Across local ancestry inference, qpAdm admixture modeling, and medieval ancient DNA continuity, Southern Italian and closely related central Mediterranean ancestry repeatedly accounts for roughly two-thirds of Ashkenazi autosomal ancestry, with remaining components distributed among eastern Mediterranean continuity and later Eastern European admixture. Variation in point estimates is best explained by differences in how each framework partitions Mediterranean structure and by the specific proxies available, rather than disagreement about the dominant autosomal source.

Table 5. Convergent autosomal estimates of Southern Italian and central Mediterranean ancestry in Ashkenazi Jews.

Study	**Methodology**	**Data type**	**Southern Italian / Mediterranean component**	**Additional components**
Lerga-Jaso et al. (2025)	Local ancestry inference (Orchestra; IBD-based plus deep-learning smoothing)	Modern genomes	~68% Italian	~16.6% Levantine; ~7.2% Iraqi-Iranian-Caucasian-Turkish; ~2.4% Greek-Balkan; ~1.7% Eastern European
Waldman et al. (2022)	qpAdm (historically constrained model)	Modern genomes	~65% Southern Italian	~19% Lebanese; ~16% Eastern European
Brace et al. (2022)	qpAdm anchored to medieval ancient DNA (Chapelfield)	Medieval plus modern	~67% Sicilian (Southern Italian proxy)	~33% Turkish Jewish; ~0% French; ~0% Polish

Table 5 note. Components are reported as proportions of modeled ancestry (percent). Point estimates differ because methods partition ancestry differently and qpAdm results depend on the proxy populations available. In particular, Lebanese populations are eastern Mediterranean and shifted toward Southern European genetic space relative to Arabian-associated Levantine poles; therefore, a Lebanese component in qpAdm should not be equated with a Levantine-majority origin claim. Across all approaches, Southern Italy remains the dominant autosomal source, with additional components reflecting broader eastern Mediterranean continuity and later Eastern European admixture.

Convergent Autosomal FST Evidence for Southern European affinity of Ashkenazi Jews

Independent autosomal FST analyses conducted using different datasets, population panels, and analytical framings converge on the same ordering of affinities: Ashkenazi Jews show their lowest genetic distances to Southern European populations, particularly Italians and Greeks, and substantially higher distances to Levantine populations. This pattern is observed across both mixed

Jewish and non-Jewish datasets and Jewish-focused comparative frameworks and does not support models that treat Ashkenazi Jews as a stable midpoint between Europe and the Levant.

Tian et al. (2009) calculated autosomal SNP-based FST distances between Ashkenazi Jews and a broad panel of European and Near Eastern populations. Ashkenazi Jews show their smallest distances to Italians (0.0040) and Greeks (0.0042), followed by other Southern Europeans. Distances to Druze (0.0088) and Palestinians (0.0093) are more than double those to Italians and Greeks and exceed distances to several Northern European populations under this panel. These values place Ashkenazi Jews within a Southern European genetic continuum rather than at an intermediate position between Europe and the Levant.

Independent confirmation is provided by Zoossmann-Diskin (2010), who reported autosomal genetic distances among multiple Jewish and non-Jewish populations using a distance-matrix framework. In that analysis, Ashkenazi Jews are closest to Italians and Greeks and substantially more distant from Palestinians. Using the study's scaling, the Ashkenazi-Italian distance (44, in x1000 units) is far smaller than the Ashkenazi-Palestinian distance (277), a disparity incompatible with Levantine-centered or midpoint interpretations.

That these studies recover the same qualitative ordering despite differing population compositions, scaling conventions, and analytic goals supports the interpretation that the result is not method-dependent. When Southern Europe is represented directly rather than via drifted isolates or north-shifted proxies, Ashkenazi Jews resolve as genetically closest to Southern Europeans and clearly separated from Levantine populations.

Table 6. Autosomal FST proximity of Ashkenazi Jews in two independent studies A. Tian et al. (2009): SNP-based autosomal FST B. Zoossmann-Diskin (2010): autosomal genetic distance matrix (x1000).

A

Comparison population	**FST**
Italians	0.0040
Greeks	0.0042
Spanish	0.0056
Tuscans	0.0066
Germans	0.0072
Druze	0.0088
Palestinians	0.0093
Irish	0.0109
Swedes	0.0120
Russians	0.0137
Basque	0.0144

B

Comparison population	**Distance (x1000)**
Italians	44
Greeks	105
Turks	170
Germans	131

A

Comparison population	FST
French	144
British	238
Poles	195
Russians	230
Palestinians	277

Table 6 note. Use NR for values not reported. Use ~ for approximate values. Both studies independently identify Italians and Greeks as the closest populations to Ashkenazi Jews, with Palestinians substantially more distant. Ashkenazi-Southern European distances fall in the ~0.004 to ~0.005 FST range typical of closely related regional populations, whereas Ashkenazi-Palestinian distances approach or exceed ~0.009, indicating markedly deeper divergence under these panels. These results contradict claims that Ashkenazi Jews occupy a genetic midpoint between Europe and the Levant; apparent intermediacy arises when Southern Europe is excluded or replaced with genetically inappropriate proxies.

Part VI. Integrated Interpretation and Implications for Models of Ashkenazi Origins

Across multiple independent analytical frameworks, autosomal genetic evidence consistently places Ashkenazi Jews within a Southern European and Mediterranean genetic continuum rather than at a midpoint between Europe and the Levant. Genome-wide FST analyses repeatedly show that Ashkenazi Jews have their lowest genetic distances to Southern European populations, particularly Italians and Greeks, and substantially higher distances to Levantine populations. The replication of this ordering across distinct datasets and population panels indicates that it is not contingent on a single genotyping platform, scaling convention, or comparative framework.

Admixture modeling using qpAdm yields convergent results in which the dominant ancestral component aligns with Southern Italian or closely related central Mediterranean populations, with remaining components reflecting broader eastern Mediterranean continuity and later Eastern European admixture. Differences in point estimates across studies are best explained by proxy choice and model structure, including whether Mediterranean-shifted eastern Mediterranean populations are used as stand-ins for heterogeneous eastern Mediterranean ancestry, rather than by disagreement regarding the primary autosomal source. In this context, models that omit Southern Europe or substitute Northern Italian populations and genetically drifted isolates tend to mis-partition Mediterranean structure and mechanically inflate Ashkenazi-Levant contrasts, generating an appearance of intermediacy that attenuates or disappears when appropriate Southern European and central Mediterranean comparators are included.

Independent constraints on Ashkenazi population history are also provided by autosomal variants with well-characterized geographic distributions that are not incorporated into genome-wide ancestry modeling frameworks. The CCR5-Δ32 allele exhibits a pronounced European north-south cline and is rare to absent outside Europe, including in Levantine and Arabian populations as reported in canonical frequency surveys. Ashkenazi Jews show CCR5-Δ32 allele frequencies that fall within the European range and far above the zero values reported for Levantine and Arabian comparators in Stephens et al. (1998). This pattern mirrors continental European structure and diverges sharply from Levantine baselines, consistent with acquisition within Europe followed by demographic amplification under endogamy.

The geographic distribution of CCR5-Δ32 is difficult to reconcile with substantial inheritance from a Levantine source population. If a large fraction of Ashkenazi ancestry were derived from Levantine populations, CCR5-Δ32 frequencies would be expected to approach Levantine values or at least show evidence of shared presence at detectable levels. Instead, the Ashkenazi frequency aligns with European comparators and contrasts decisively with reported absence in Levantine and Arabian samples. Published haplotype-based and coalescent inferences are consistent with a single origin of CCR5-Δ32 within Europe and a relatively recent rise in frequency, which places the allele's emergence and expansion within European demographic history and renders an ancient Near Eastern origin implausible. Under this interpretation, elevated Ashkenazi frequency follows naturally from founder effects and population growth operating on a European-introduced variant.

Parallel evidence for this mechanism is provided by well-studied autosomal founder mutations in Ashkenazi Jews, including the Tay-Sachs disease-associated 1278insTATC mutation. Genetic and genealogical analyses have been interpreted as consistent with introduction from a European source followed by amplification through founder effects and long-term endogamy. Reports of this mutation outside Ashkenazi Jews are most parsimoniously explained by later gene flow from Ashkenazi lineages rather than by independent Near Eastern origin. Together with CCR5-Δ32, these examples illustrate how European autosomal variants can reach elevated frequencies within Ashkenazi Jews through demographic processes that are orthogonal to, and therefore not reducible to, simple continent-level ancestry labels.

Taken together, these independent lines of evidence converge on a single interpretation. Apparent Ashkenazi genetic intermediacy between Europe and the Levant is not an intrinsic feature of the autosomal data but a methodological artifact that can arise from asymmetric reference population design and proxy substitution that under-represents Southern Europe and the central Mediterranean. When Southern European and Mediterranean populations are represented directly, and when autosomal evidence is interpreted within a population-genetic framework that accounts for drift, endogamy, and founder effects, Ashkenazi Jews resolve as predominantly Southern European and central Mediterranean in autosomal structure, with additional eastern Mediterranean continuity and later Eastern European contributions. This autosomal structure provides the appropriate context for interpreting uniparental markers and imposes clear constraints on viable models of Ashkenazi origins.

Table 7. CCR5-Δ32 allele frequencies in Ashkenazi Jews and comparative populations.

Population	**Sample size**	**CCR5-Δ32 allele frequency**
Swedish	131	0.137
Estonian	158	0.133
Polish	30	0.133
French	230	0.089
Italian	172	0.055
Greek	160	0.044
Ashkenazi Jews	503	0.097
Lebanese	51	0.000
Saudi Arabian	100	0.000
Chinese	40	0.000
African populations	various	0.000

Table 7 note. Allele frequencies are reported as proportions (0 to 1) as presented in Stephens et al. (1998). CCR5-Δ32 frequencies follow a European north-south cline. Ashkenazi Jews fall within the European distribution and contrast with Levantine and Arabian comparators reported as 0.000 in this dataset.

Part VII. Historical and Geographic Context of Jewish Settlement in Southern Italy and the Central Mediterranean

Genetic models that treat Northern Italian or Tuscan populations as representative European proxies for Ashkenazi Jewish origins are not only methodologically incomplete. They also conflict with the historical geography of Jewish settlement in Europe. Long before the emergence of Ashkenazi Jewry in northern Europe, Jewish populations were firmly established across southern Italy, Sicily, and the wider central and eastern Mediterranean, precisely the regions with which Ashkenazi Jews show their strongest autosomal affinities in PCA, FST, and G25 analyses (Feldman, 1993; Gruen, 2002; Rutgers, 1995).

Jewish presence in southern Italy and Sicily is documented continuously from antiquity. Archaeological, epigraphic, and literary evidence attests to substantial Jewish communities on the southern Italian mainland and in Sicily beginning in the Roman Republican and early Imperial periods (Rutgers, 1995). In Sicily, Jewish settlement is attested in Syracuse, Palermo, Catania, Agrigento, and Mazara del Vallo from the first centuries CE, with continuity through Late Antiquity, Byzantine rule, and the medieval period until the expulsions of the late fifteenth century (Roth, 1946; Rutgers, 1995). Sicily was among the largest Jewish population centers in medieval Europe, but its demographic and cultural foundations trace to antiquity rather than to medieval migration from northern Europe (Roth, 1946).

On the southern Italian mainland, Jewish communities are documented from antiquity in Apulia, Calabria, Campania, and Lucania. Roman-period and Late Antique evidence supports Jewish settlement in Bari, Brindisi, Oria, Otranto, Trani, Reggio Calabria, Catanzaro, Squillace, Naples, Pozzuoli, and Capua (Roth, 1946; Rutgers, 1995). Oria emerges as a major Jewish intellectual center in Late Antiquity and the early medieval period, underscoring the depth and continuity of Jewish life in southern Italy well before Ashkenazi ethnogenesis (Roth, 1946; Bonfil, 1994). These communities participated in a Greek- and Latin-speaking Mediterranean network with sustained ties to southern Greece, the Aegean islands, and Asia Minor, where Jewish settlement is likewise securely attested from the Hellenistic period onward (Feldman, 1993; Gruen, 2002).

By contrast, Jewish settlement in Northern Italy and Tuscany is historically secondary. These regions become significant centers of Jewish life primarily in the medieval period, in substantial part through migration from southern Italy and other Mediterranean regions rather than through independent, comparably deep ancient settlement (Roth, 1946; Bonfil, 1994). Tuscan Jewish communities are comparatively late, demographically smaller, and historically derivative. On historical grounds, they are not plausible ancestral source populations for Ashkenazi Jewry. Treating Northern Italians or Tuscans as stand-ins for ancestral Italian contexts therefore reverses the documented direction of Jewish migration and settlement (Roth, 1946).

The emergence of Ashkenazi Jewry occurs later, between approximately the ninth and eleventh centuries CE, in northern France and the Rhineland. Historical scholarship consistently identifies Italy and the central Mediterranean as major staging grounds for northward movement, with migration proceeding along established Roman and post-Roman trade and communication corridors from southern Italy into France and the Rhineland (Chazan, 1987; Grossman, 1995). Ashkenazi communities formed in northern Europe under conditions of demographic bottleneck, geographic concentration, and long-term endogamy, but their antecedent roots lie in the Mediterranean world rather than in Northern or Central Europe (Grossman, 1995).

This historical geography aligns with genome-wide genetic results. Autosomal PCA and FST analyses repeatedly place Ashkenazi Jews closest to southern Italian, Sicilian, and related central Mediterranean populations, with substantially greater distances to Levantine populations (Tian et al., 2009; Zoossmann-Diskin, 2010). Italian Jewish populations provide an important internal control. These communities descend from the southern Italian and Sicilian Jewish populations documented from antiquity and, in general, do not exhibit the same severe demographic bottleneck as Ashkenazi Jews. In autosomal analyses, Italian Jews cluster within southern Italian and central Mediterranean genetic space and overlap Ashkenazi Mediterranean placement once drift is accounted for, supporting continuity rather than convergence through recent admixture (Atzmon et al., 2010; Zoossmann-Diskin, 2010).

doi:10.20944/preprints202601.2134.v2

I further evaluate the geographic specificity of this signal using a modern Global25 population average for Ashkenazi Jews (Ashkenazi_Germany; modern scaled coordinates) and a distance-based similarity mapping procedure in the modern Davidski G25 reference set. This framework ranks genetic distances to modern populations and visualizes proximity categories defined as GREAT (<0.02), GOOD (<0.03), and AVERAGE (<0.04). The results localize Ashkenazi autosomal affinity overwhelmingly to Malta and southern Italy, with dense representation from Calabria, Campania, Basilicata, Apulia, Abruzzo, Molise, and multiple Sicilian regional samples, together with repeated Aegean island Greek populations, particularly Crete and the Dodecanese. Northern Italy and Tuscany are absent from the closest-ranked populations under this distance ordering. The nearest matches are Maltese (0.0180) and Italian Calabria (0.0182), followed by a series of southern Italian and Sicilian regional samples and multiple Greek island samples across Crete, the Dodecanese, the Cyclades, and Euboea. This placement maps onto the documented geography of long-standing Jewish settlement and connectivity in the central Mediterranean and adjacent Greek-speaking zones from antiquity through the medieval period (Roth, 1946; Rutgers, 1995; Feldman, 1993) and is concordant with the Southern European affinity ordering observed in independent autosomal FST and PCA frameworks (Tian et al., 2009; Zoossmann-Diskin, 2010).

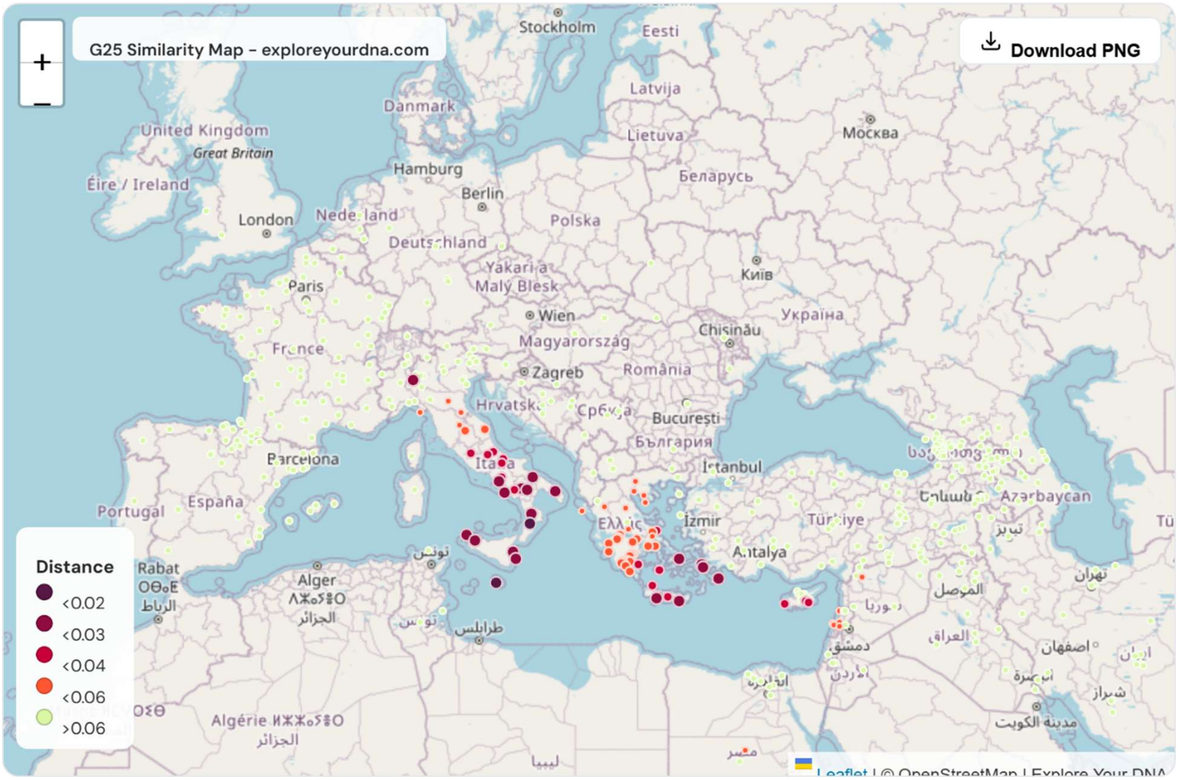


Figure 3. Geographic distribution of Ashkenazi autosomal similarity using Ashkenazi_Germany (modern scaled Global25 population average) in a modern G25 similarity map.

This similarity map ranks modern populations by genetic distance to the Ashkenazi_Germany coordinate vector within the modern Davidski G25 dataset. The strongest similarity signal localizes to Malta and southern Italy, with additional peaks across Sicily and the Aegean island zone, particularly Crete and the Dodecanese. Northern Italy and Tuscany do not emerge among the closest affinity centers.

Table 8. Top 30 closest modern populations to Ashkenazi_Germany in modern G25 similarity mapping.

doi:10.20944/preprints202601.2134.v2

#	Population	Distance
1	Maltese	0.0180
2	Italian Calabria	0.0182
3	Italian Campania	0.0207
4	Italian Campania Naples (Campanian)	0.0224
5	Sicilian Syracuse	0.0236
6	Sicilian Trapani	0.0237
7	Greek Crete	0.0245
8	Sicilian Central	0.0245
9	Sicilian East	0.0245
10	Italian Calabria (Cosentian)	0.0248
11	Greek Crete Lasithi	0.0251
12	Italian Basilicata (Lucanian)	0.0253
13	Italian Basilicata	0.0253
14	Greek Crete Rethymno	0.0261
15	Sicilian	0.0270
16	Greek Dodecanese Kos	0.0273
17	Greek Kos	0.0273
18	Italian Apulia	0.0278
19	Greek Dodecanese	0.0280
20	Sicilian West	0.0288
21	Italian Apulia (Apulian)	0.0290
22	Greek Cyclades Amorgos	0.0298
23	Greek Crete Chania	0.0307
24	Italian Abruzzo	0.0313
25	Italian Abruzzo (Abruzzese)	0.0313
26	Greek Euboea Central	0.0318
27	Greek Crete Heraklion	0.0321
28	Greek Cyclades Milos	0.0334
29	Italian Campania Salerno (Campanian)	0.0353
30	Italian Molise	0.0353

Accordingly, routine use of Northern Italian or Tuscan populations as proxies for Italy in genetic models of Ashkenazi origins is historically indefensible. It excludes the regions in which Jewish communities are attested continuously from antiquity and substitutes populations that are historically secondary and genetically shifted toward Central and Northern Europe. Under these conditions, shared Mediterranean ancestry is systematically misallocated into an eastern residual, including Levantine-labeled components, not because the data require it, but because the reference design forces it. When southern Italy, Sicily, Malta, and the Aegean corridor are represented directly, the autosomal signal aligns with the documented geography of Jewish settlement and with the strongest distance-based affinities in modern G25 space, reinforcing the historical and population-genetic case against Northern Italian or Tuscan proxy models.

Part VIII. Independent qpAdm and FST Analyses with External Validation

To further constrain Ashkenazi autosomal origins using analyses that I conducted independently, I performed qpAdm admixture modeling and autosomal FST distance calculations using the AdmixLab pipeline hosted by IllustrativeDNA. In addition, I evaluated unconstrained similarity patterns using G25 modeling restricted exclusively to ancient individuals. I integrate these results as an external validation layer and evaluate them alongside the autosomal PCA, FST, and qpAdm arguments developed in Parts III and IV.

qpAdm Modeling Performed by the Author

I ran qpAdm with Jew_Ashkenazi.HO as the target population. The right population set was:

- Mbuti.DG
- Yoruba.DG
- Ju_hoan_North.DG
- Han.DG
- Chukchi.DG
- Karitiana.DG
- Papuan.DG
- Iranian.DG
- Switzerland_Bichon_Epipaleolithic.SG
- Basque.DG

This configuration is intended to control deep population structure while allowing contrasts among Mediterranean and European sources.

I tested three alternative European source models. Results are summarized in Table 9.

Table 9. qpAdm results for Jew_Ashkenazi.HO (author-generated).

Model components	p value	Feasible	Italian component (%)	Lebanese (%)	Russian (%)
Italian_South.HO + Lebanese.HO + Russian.DG	0.1445	Yes	69.97	22.06	7.97
Italian_North.HO + Lebanese.HO + Russian.DG	0.0256	No	54.53	42.57	2.90
Italian_Central.HO + Lebanese.HO + Russian.DG	0.0514	Yes	61.33	33.83	4.84

Under this specification, the Southern Italian model is feasible and identifies Italian_South.HO as the dominant autosomal source, accounting for approximately 70% of modeled Ashkenazi ancestry. The Northern Italian model fails feasibility and is rejected under this configuration, indicating that Italian_North.HO is not an adequate stand-in for the European component. The Central Italian model is only marginally feasible and yields an intermediate Italian contribution consistent with a northward gradient within Italy rather than a shift toward Levantine ancestry. Overall, these author-generated qpAdm results reproduce a Southern Italian dominant structure consistent with published qpAdm analyses discussed elsewhere in this study, while using a distinct pipeline and a fixed right population design.

Autosomal FST Distances Calculated by the Author

Autosomal FST distances provide a complementary measure of population proximity that does not depend on explicit mixture modeling. Table 10 shows pairwise FST distances between Jew_Ashkenazi.HO and selected Mediterranean, European, and Jewish populations.

Table 10. Autosomal FST distances to Jew_Ashkenazi.HO (author-generated).

Population	FST distance	Standard error
Cretan.DG	0.0049	0.00124
Italian_South.HO	0.0049	0.00033
Sicilian.HO	0.0052	0.00041
Greek.HO	0.0055	0.00035
Maltese.HO	0.0066	0.00045
Jew_Moroccan.HO	0.0091	0.00058
Jew_Iraqi.DG	0.0122	0.00138
Jew_Iranian.HO	0.0125	0.00050
Jew_Tunisian.HO	0.0127	0.00056
Jew_Libyan.HO	0.0132	0.00053
Jew_Georgian.HO	0.0133	0.00056
Jew_Yemenite.HO	0.0174	0.00054

The lowest distances are observed to Southern Italians and Cretans, followed closely by Sicilians and mainland Greeks. Maltese also fall within the same Mediterranean affinity range. In contrast, all non-European Jewish comparator groups are substantially more distant, with North African, Near Eastern, Caucasian, and Arabian Jewish populations showing FST values approximately two to three times higher than those observed for Southern Italian and Aegean populations. This ordering is inconsistent with a Europe-Levant midpoint model and instead indicates tight placement within a Southern European and Aegean continuum with clear separation from Levantine and non-European Jewish populations under these panels.

FST and PCA Calibration Using the Canary Islands

To contextualize the magnitude of these distances and their interpretation in PCA space, Table 11 provides calibration comparisons using populations with well-established historical and genetic relationships.

Table 11. FST calibration benchmarks.

Population pair	FST distance	Interpretation
Italian_South.HO vs Maltese.HO	0.0036	Closely related central Mediterranean populations
Jew_Ashkenazi.HO vs Italian_South.HO	0.0049	Strong regional affinity
IBS_CanaryIslands.DG vs Spanish.HO	0.0038	Iberian-derived population with additional North African ancestry

Genome-wide autosomal and ancient DNA studies show that modern Canary Islanders derive a strong majority of their ancestry from Iberian sources, with a minority Indigenous North African component often estimated at approximately 15 to 25 percent (Botigué et al., 2013; Rodríguez-Varela et al., 2017). Despite this admixture, Canary Islanders cluster tightly with Iberians in PCA space and show very low FST distances to Spanish populations. This provides an external calibration demonstrating that very low FST values on the order of 0.003 to 0.004 are compatible with strong shared ancestry in the presence of admixture and drift. Under this calibration, the Ashkenazi–Southern Italian FST distance of 0.0049 is consistent with strong regional affinity rather than weak relatedness or an intermediate position.

In addition, PCA frameworks discussed elsewhere in this study show that Ashkenazi Jews overlap extensively with Southern Italian populations in Mediterranean-inclusive space. In comparative terms, this overlap is not weaker than the Iberian signal observed for the Canary Islands, even though Canary Islanders retain a higher proportion of direct Iberian ancestry. This pattern is consistent with a Southern Italian and central Mediterranean anchoring of Ashkenazi autosomal variation, with drift and endogamy shaping within-cluster dispersion rather than shifting Ashkenazi Jews toward Levantine genetic space.

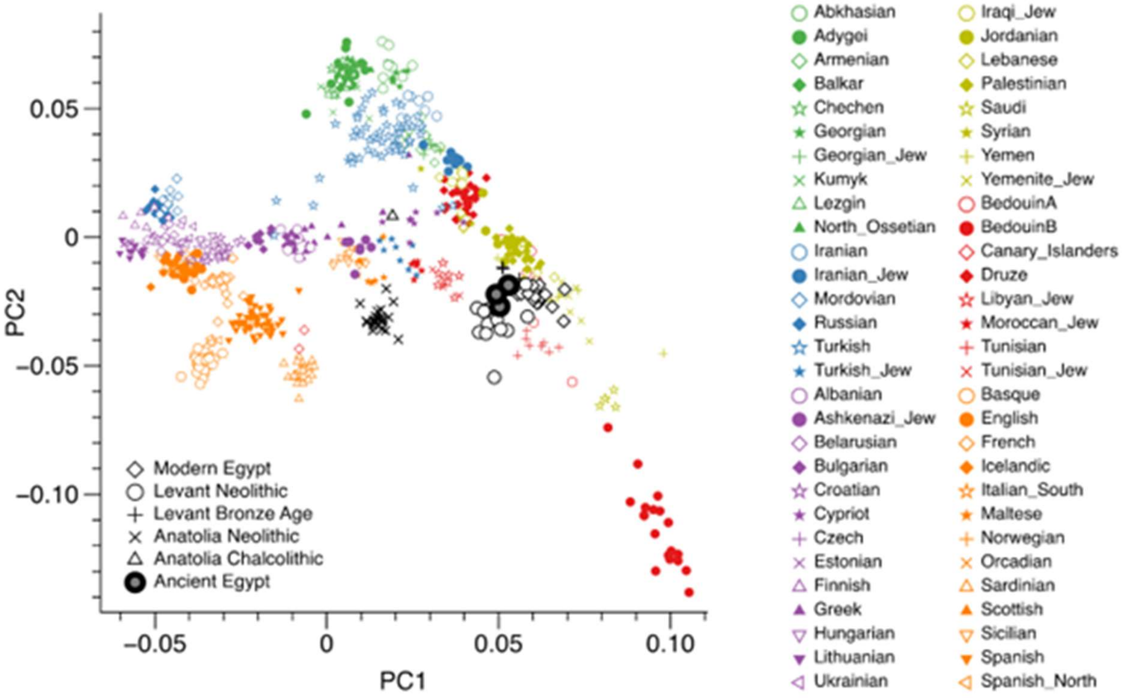
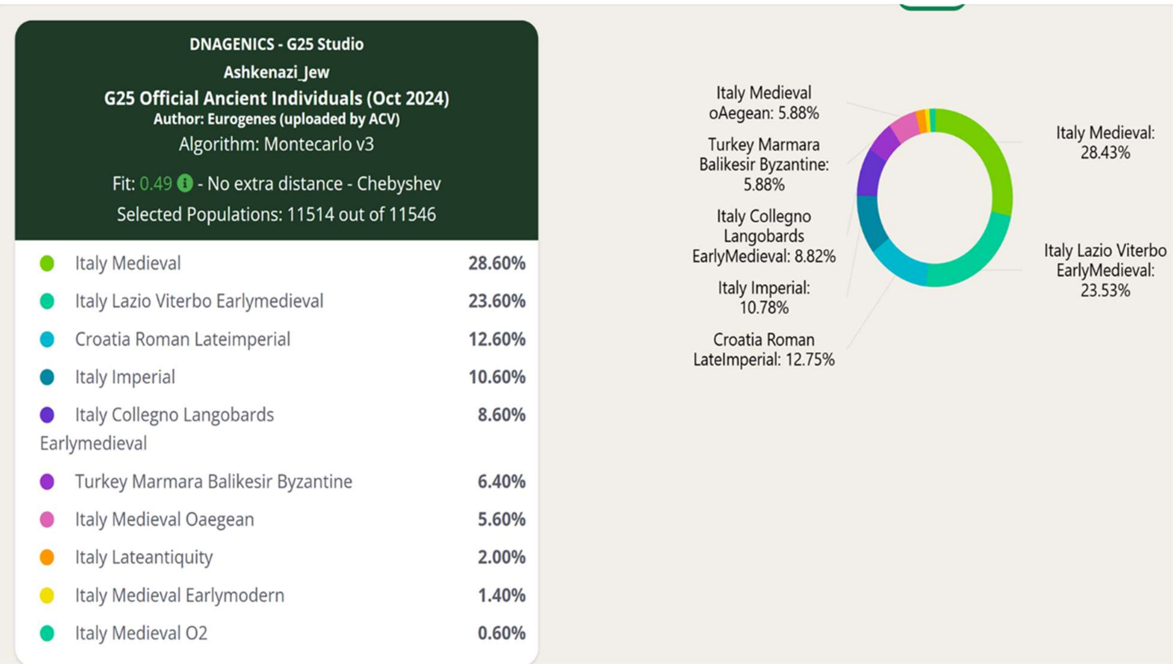


Figure 4. Principal component analysis reproduced from Schuenemann et al. (2017), shown for comparative context. In this framework, Roman-period and earlier Egyptian individuals cluster with Mediterranean populations, while post-Roman Egyptian samples shift toward increased Sub-Saharan African ancestry. The figure illustrates that PCA space can capture post-antique admixture trajectories without redefining core Mediterranean structure. In comparable Mediterranean-inclusive PCA frameworks discussed elsewhere in this study, Ashkenazi Jews align with Southern European and central Mediterranean populations rather than along a Levantine axis, consistent with the Southern Italian proximity signal observed in qpAdm and FST.

Ancient-Only G25 Similarity Analysis (Author-Generated)

As an additional non-modern validation, I modeled Ashkenazi Jews using G25 coordinates restricted exclusively to ancient individuals, excluding all modern populations and genetically drifted isolates. This analysis is descriptive and is not used to estimate ancestry proportions. Its purpose is to test whether unconstrained genetic similarity reproduces the same central Mediterranean anchoring observed under qpAdm, FST, and PCA (Figure 5).



DNAGENICS - G25 Studio
Ashkenazi_Jew
G25 Official Ancient Individuals (Oct 2024)
Author: Eurogenes (uploaded by ACV)
Algorithm: Montecarlo v3
Fit: 0.49 - No extra distance - Chebyshev
Selected Populations: 11514 out of 11546

● Italy Medieval	28.60%
● Italy Lazio Viterbo Earlymedieval	23.60%
● Croatia Roman Lateimperial	12.60%
● Italy Imperial	10.60%
● Italy Collegno Langobards Earlymedieval	8.60%
● Turkey Marmara Balikesir Byzantine	6.40%
● Italy Medieval Oaegean	5.60%
● Italy Lateantiquity	2.00%
● Italy Medieval Earlymodern	1.40%
● Italy Medieval O2	0.60%

Figure 5. Ancient-only G25 similarity results for Ashkenazi Jews (author-generated).

Under this ancient-only framework, I evaluated Ashkenazi Jews against 11,514 ancient individuals spanning Roman, Late Antique, medieval, and pre-Roman contexts across the Mediterranean, the Near East, and broader Eurasia, including extensive Levantine sampling. The closest similarity signal localizes overwhelmingly to Roman, Late Antique, and medieval Italian individuals, with secondary similarity to eastern Roman provincial contexts in the Balkans, the Aegean, and western Anatolia. Levantine ancient samples do not emerge as dominant similarity anchors when the full ancient reference pool is available. This result indicates that the central Mediterranean signal persists even under broad Levantine representation and is not contingent on omission of Levantine comparators.

This figure shows the distance-ranked similarity output obtained when Ashkenazi Jewish G25 coordinates are evaluated exclusively against ancient individuals, with all modern populations excluded. The closest matches localize primarily to Roman, Late Antique, and medieval Italy, with additional similarity to eastern Roman provincial contexts in the Balkans, the Aegean, and western Anatolia, and without Levantine samples emerging as dominant similarity anchors under exhaustive ancient sampling.

Summary of Author-Generated Validation Analyses

Across qpAdm modeling, autosomal FST distances, PCA overlap, and the ancient-only G25 similarity analysis that I conducted independently, the same population-genetic signal is recovered. Ashkenazi Jews consistently align most closely with Southern Italian, Sicilian, Cretan, and Greek populations and are substantially more distant from Levantine and non-European Jewish populations in direct distance comparisons. Northern Italy fails as a valid proxy in the tested qpAdm specification, and the calibrated FST benchmarks indicate that the Ashkenazi-Southern Italian distance falls within a range consistent with strong regional affinity under realistic admixture and drift.

doi:10.20944/preprints202601.2134.v2

These independently generated results reinforce the central conclusion of this study. Apparent Ashkenazi intermediacy between Europe and the Levant is not an intrinsic feature of autosomal data. It arises when Southern Europe is excluded or replaced by genetically inappropriate proxies. When Southern Italy and the broader Mediterranean are represented directly, constrained and unconstrained approaches converge on a central Mediterranean anchoring model.

Part IX. Independent Modern Population PCA and Full-Dimensional Global25 Distance Analysis

To provide an independent validation of the autosomal affinities identified in earlier sections, I conducted a modern population principal component analysis using Global25 coordinates, combined with full 25-dimensional Euclidean distance ranking. This analysis was generated specifically for this study and does not rely on reused figures, published PCA projections, or author-provided visualizations. It therefore constitutes an original descriptive assessment of autosomal population structure.

I projected the *Ashkenazi_Germany* Global25 population average using the following coordinate vector:

Ashkenazi_Germany, 0.097319, 0.142377, -0.014406, -0.049936, 0.010925, -0.018769, -0.001974, -0.001754, 0.009858, 0.018843, 0.003962, -0.001124, 0.002512, -0.002670, -0.005049, 0.003089, 0.002203, -0.001698, -0.000779, -0.001126, -0.003344, -0.002696, 0.000875, 0.001386, 0.000814.

PCA Visualization (PC1 vs PC2)

In the PC1–PC2 projection shown in Figure 6, *Ashkenazi_Germany* occupies a position embedded within Southern European and central Mediterranean population space rather than along a Europe–Levant axis. The Ashkenazi point clusters visually with Italian Jewish populations and Southern Italian regional samples, particularly Calabria and Sicily, and remains clearly separated from Levantine, Caucasus, and Armenian populations along both principal component axes.

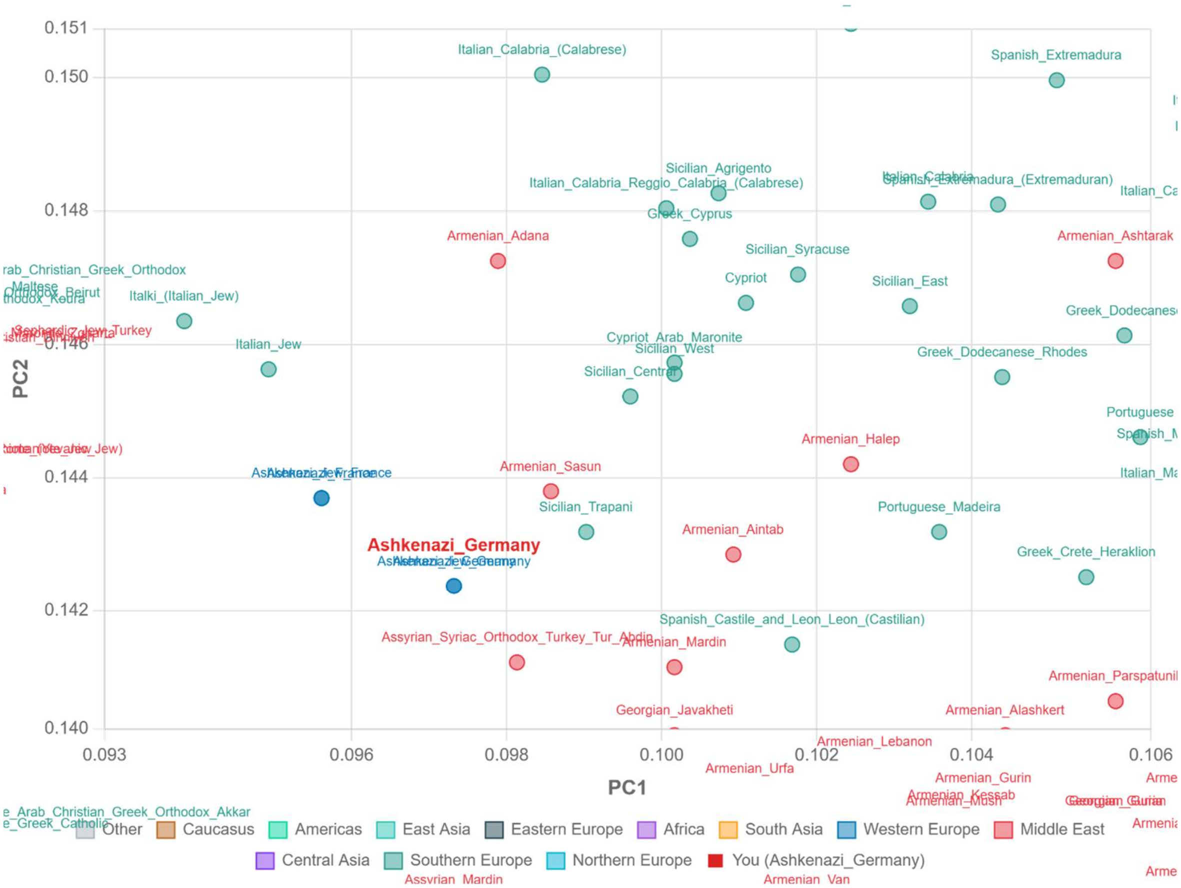


Figure 6. Principal component analysis (PC1 vs PC2) of modern populations using Global25 coordinates, with *Ashkenazi_Germany* projected as the target population. Ashkenazi Jews cluster within Southern European and central Mediterranean population space, nearest to Italian Jewish and Southern Italian regional populations, and are clearly separated from Levantine, Caucasus, and Armenian populations.

Ashkenazi Jews do not appear intermediate between Northern European and Near Eastern populations when Southern Europe is directly represented. Instead, they fall within a Southern European and Mediterranean cluster that includes Southern Italians, Sicilians, Maltese, and Aegean Greeks. Armenian and Near Eastern populations form a distinct cluster displaced from Ashkenazi Jews along both PC1 and PC2, indicating that Ashkenazi placement in this framework is not driven by proximity to eastern Mediterranean or Caucasus genetic structure.

Because PCA displays only two of the twenty-five dimensions captured by the Global25 coordinate system, this visualization is treated as descriptive. Apparent proximity in two-dimensional space can be influenced by axis orientation and reference composition and does not by itself quantify overall genome-wide similarity.

Full 25-Dimensional Global25 Distance Rankings

To evaluate genome-wide similarity directly, I examined the full 25-dimensional Global25 Euclidean distance rankings reported alongside the PCA projection. These distances provide a quantitative measure of autosomal affinity that is independent of PCA geometry.

Under the full-distance metric, the closest populations to *Ashkenazi_Germany* are exclusively Ashkenazi Jewish populations, followed by Italian Jewish and Southern Italian populations. The ten closest populations are:

Ashkenazi_Germany (0.0000)
Ashkenazi_Jew_Germany (0.0000)

Ashkenazi_Jew_France (0.0067)
Ashkenazi_France (0.0067)
Italian_Jew (0.0132)
Italian_Calabria_Reggio_Calabria (Calabrese) (0.0147)
Ashkenazi_Jew_Austria (0.0156)
Ashkenazi_Austria (0.0156)
Italki (Italian_Jew) (0.0159)
Ashkenazi_Poland (0.0163)

No Levantine, Near Eastern, Caucasus, or Armenian populations appear among the closest populations under the full 25-dimensional distance metric, despite being present in the reference panel. The nearest non-Ashkenazi populations are Italian Jews and Southern Italians, specifically Calabrians. This ordering demonstrates that apparent proximity to eastern Mediterranean populations occasionally suggested by two-dimensional PCA projections does not reflect true genome-wide similarity when evaluated in full-dimensional space.

Synthesis of Global25 Results

Taken together, the PCA visualization and the full 25-dimensional Global25 distance rankings reproduce the same affinity ordering observed earlier in this study across qpAdm modeling, autosomal FST distances, identity-by-descent sharing, and ancient-only similarity analyses. When Southern European and central Mediterranean populations are represented directly, Ashkenazi Jews resolve as genetically closest to Italian Jewish and Southern Italian populations and remain clearly separated from Levantine and Near Eastern groups.

This analysis shows that claims of Ashkenazi genetic intermediacy can arise from dimensional reduction effects rather than from underlying autosomal structure. When all twenty-five Global25 dimensions are considered simultaneously, Ashkenazi Jews do not occupy a Europe–Levant midpoint and do not cluster with Levantine populations. Instead, their autosomal structure is anchored in Southern Europe and the central Mediterranean.

Part X. Italian Jews as a Falsification Test of Northern Italian and Convergent Origin Models

A broad consensus exists across population genetic studies that Ashkenazi Jews derive a substantial proportion of their autosomal ancestry from European sources. Across two-way and three-way admixture models, this European component is commonly estimated at approximately fifty percent, with the remaining ancestry attributed primarily to Levantine or eastern Mediterranean proxies such as Lebanese populations. The presence of a large Italian-associated component in Ashkenazi Jews is therefore not disputed. The point of disagreement concerns the geographic attribution of that Italian ancestry.

In most published models, the European component in Ashkenazi Jews is assigned to Northern or Central Italian proxies, often in combination with Levantine references, while Southern Italian populations are minimized, excluded, or treated as unsuitable stand-ins. This practice persists despite repeated observations that Ashkenazi Jews cluster most closely with Southern Italians in Mediterranean-inclusive autosomal PCA space and exhibit their lowest autosomal FST distances to Southern Italian and Aegean populations. The continued reliance on Northern Italian proxies therefore reflects reference population design rather than constraints imposed by the data.

Italian Jews provide a decisive population-level falsification test for these models. They represent the only continuously documented Jewish population resident in Italy from antiquity to the present and are historically rooted primarily in southern Italy and Sicily. Unlike Ashkenazi Jews, Italian Jews did not migrate northward as a population and did not experience the extreme medieval founder events that characterize Ashkenazi demographic history. Any valid Italian

source model for Ashkenazi ancestry must therefore be compatible with the autosomal position of Italian Jews.

Across the autosomal PCA frameworks presented earlier in this study, Italian Jews fall within the same Southern Italian and central Mediterranean autosomal space as Ashkenazi Jews once Ashkenazi-specific drift is taken into account. Northern Italian populations are consistently displaced outside this cluster. Autosomal FST distances reinforce this placement, with Italian Jews occupying the same Southern Italian and central Mediterranean affinity range as Ashkenazi Jews and showing substantially greater separation from Northern Italian and non-European Jewish populations. Drift alters within-cluster dispersion but does not relocate populations into unrelated regional autosomal space and therefore cannot reconcile these results with Northern Italian proxy models.

Independent genome-wide analyses further corroborate this interpretation. Campbell et al. (2012) analyzed a broad panel of Jewish and non-Jewish populations using autosomal PCA, neighbor-joining trees, FST distances, and identity-by-descent sharing. In that study, Italian Jews cluster directly adjacent to Ashkenazi Jews and form a shared European and Mediterranean Jewish branch together with Greek and Turkish Sephardic Jews. Northern Italian populations, included explicitly as comparators, instead cluster with European non-Jewish populations and are clearly separated from the Jewish branch across all analytical frameworks. Italian Jews do not occupy an intermediate position between Italians and Middle Eastern Jews, nor do they show reduced Jewish–Jewish identity-by-descent sharing. Rather, they participate fully in the shared Jewish ancestry network, indicating long-term regional continuity rather than secondary admixture with surrounding Italian populations. This structure is incompatible with models that assign Ashkenazi Jews substantial Northern Italian ancestry while treating Italian Jews as marginal or derivative.

The same structural limitation is visible in influential secondary syntheses. A central example is the neighbor-joining tree presented by Ostrer and Skorecki (2013), which has been widely cited as a summary of autosomal relationships among Jewish diaspora populations and surrounding non-Jewish groups. In this tree, Italian Jews cluster adjacent to Ashkenazi Jews within a Southern European and central Mediterranean branch. However, the European reference set used in this analysis is dominated by Northern and Central European populations and a limited set of Italian proxies, while Southern Italian and Sicilian populations are omitted despite their direct relevance to Mediterranean population structure and historically plausible source populations. Under these conditions, apparent intermediacy can be produced by reference omission rather than reflecting true ancestry.

To evaluate Italian Jewish affinities using an independent distance-based framework that does not rely on admixture modeling, I analyzed Global25 similarity for Italian Jews using the published Italian_Jew population centroid. This analysis ranks modern populations by Euclidean distance in full 25-dimensional autosomal space and is descriptive rather than inferential. The results are summarized in Table 12 and visualized in Figure 7.

Table 12. Closest Global25 populations to Italian Jews.

Rank	Population	Distance
1	Maltese	0.0239
2	Italian Calabria	0.0242
3	Sicilian Central	0.0264
4	Italian Campania	0.0273

Rank	Population	Distance
5	Italian Campania Naples (Campanian)	0.0277
6	Greek Dodecanese	0.0289
7	Sicilian East	0.0290
8	Sicilian Syracuse	0.0291
9	Sicilian	0.0295
10	Greek Crete Lasithi	0.0303
11	Greek Crete	0.0305
12	Sicilian Trapani	0.0309
13	Greek Dodecanese Kos	0.0310
14	Greek Kos	0.0310
15	Italian Calabria (Cosentian)	0.0310
16	Greek Cyprus	0.0322
17	Cypriot	0.0326
18	Italian Basilicata	0.0326
19	Italian Basilicata (Lucanian)	0.0328
20	Greek Crete Rethymno	0.0331
21	Italian Apulia	0.0341
22	Turkish Cyprus	0.0342
23	Greek Cyclades Amorgos	0.0350
24	Sicilian West	0.0357
25	Greek Euboea Central	0.0365
26	Italian Apulia (Apulian)	0.0368
27	Greek Crete Heraklion	0.0381
28	Greek Crete Chania	0.0385
29	Italian Campania Benevento (Campanian)	0.0389

Rank	Population	Distance
30	Italian Campania Salerno (Campanian)	0.0391

The closest affinities localize overwhelmingly to southern Italy, Sicily, Malta, and the adjacent Aegean. Northern Italian populations are absent from the closest similarity ranks. This result reproduces the same Southern Italian anchoring observed under PCA, FST, qpAdm, and clustering analyses and does not depend on admixture proportion assumptions.

Taken together, these findings impose a hard constraint on viable origin models. If Ashkenazi Jews derive approximately fifty percent of their ancestry from European sources and that European component is Italian, while Italian Jews remained resident in southern Italy and Sicily, then both populations must derive that Italian ancestry from the same Southern Italian genetic substrate. There is no defensible mechanism by which Ashkenazi Jews could acquire substantial Northern Italian, Western European, or Slavic ancestry while Italian Jews, who remained geographically closer to any Italian source population, did not. Models invoking convergent Levantine–European admixture do not resolve this contradiction, as they require independent mixtures to converge on nearly identical Southern Italian autosomal placement while entirely bypassing Northern Italian populations.

Italian Jews therefore function as a population-level falsification test. Their autosomal position demonstrates that Italian ancestry acquired and retained within Italy manifests as Southern Italian and central Mediterranean structure rather than as Northern Italian or hybrid Northern European structure. Consequently, the Italian component consistently detected in Ashkenazi Jews must derive from Southern Italy and the adjacent Aegean genetic continuum, with Ashkenazi-specific founder effects and later Eastern European admixture shaping variation within that space rather than redefining its origin.

doi:10.20944/preprints202601.2134.v2

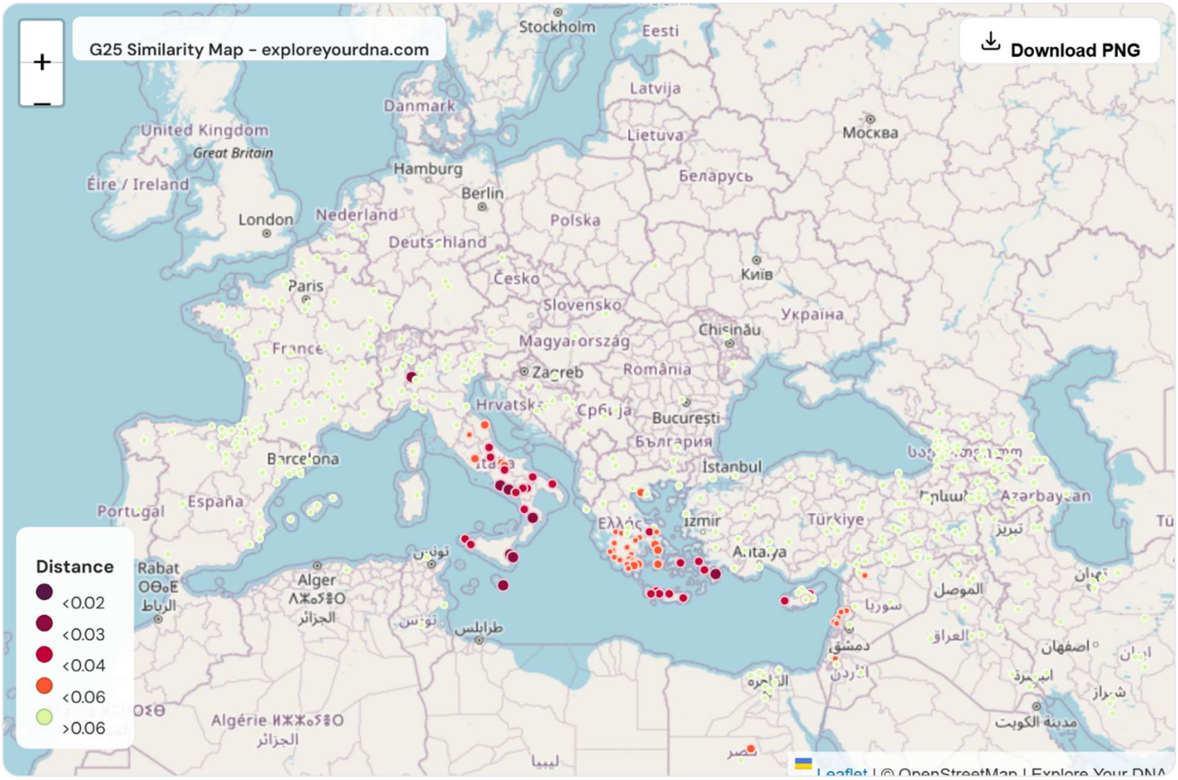


Figure 7. Global25 similarity map for Italian Jews. Similarity map generated from the Italian_Jew Global25 population centroid using Euclidean distance in full 25-dimensional autosomal space. Darker shading indicates closer genetic affinity. The strongest similarity signal localizes to southern Italy, Sicily, Malta, and the adjacent Aegean. Northern Italian populations are absent from the closest similarity ranks. This pattern corroborates results from PCA, autosomal FST, qpAdm modeling, and genome-wide clustering analyses and is incompatible with Northern Italian proxy models.

Part XI. Conclusions: Southern Italian Affinity and Reference Population Effects in Ashkenazi Origin Models

In this study, I evaluated whether the widespread characterization of Ashkenazi Jews as genetically intermediate between Europe and the Levant reflects intrinsic population history or instead arises from methodological artifacts. Across autosomal, uniparental, historical, and independently generated analyses, the evidence converges on a consistent result. Ashkenazi Jews derive primarily from Southern Italian and closely related central Mediterranean populations, with secondary eastern Mediterranean continuity and later Eastern European admixture. Models invoking a Levantine-majority origin or a stable Europe–Levant cline are not supported once appropriate Mediterranean reference populations are included.

The appearance of Ashkenazi genetic intermediacy arises from systematic proxy bias rather than from genome-wide data. Influential autosomal and uniparental studies repeatedly represented Europe using Northern Italian and Tuscan samples that are shifted toward Central and Northern Europe, and they often substituted Sardinians as Italian stand-ins despite Sardinia's marked drift and island-specific structure. Southern Italian and Sicilian populations, despite their historical relevance and clear autosomal affinity to Ashkenazi Jews, were frequently omitted. Under such reference designs, Ashkenazi Jews are structurally constrained from resolving within Southern European space and instead appear as an artificial midpoint between a north-shifted European pole and Levantine comparators. In these settings, intermediacy is imposed by population inclusion rather than inferred from underlying genetic structure.

This exclusion cannot be attributed to limitations in available data. By the time widely cited syntheses such as Ostrer and Skorecki (2013) were published, genome-wide analyses resolving Italian population structure, including Southern Italy and Sicily, were already available. Di Gaetano et al. (2012) demonstrated clear differentiation between Northern and Southern Italy, quantified pairwise FST distances, and identified Sardinia as a drifted outlier relative to the mainland. Continued reliance on Northern Italian, Tuscan, and Sardinian proxies therefore reflects reference population choice rather than data availability.

When Southern Italy is represented directly, the inferred clinal pattern collapses. Autosomal PCA, pairwise FST distances, identity-by-descent sharing, qpAdm modeling, medieval ancient DNA continuity, and independent modern PCA analyses consistently place Ashkenazi Jews closest to Southern Italians, Sicilians, Maltese, Cretans, and mainland Greeks. Genetic distances to these populations fall within ranges expected for closely related regional populations and are substantially lower than distances to Levantine groups. These results are reproduced across independent datasets and analytical frameworks and are not contingent on any single method.

Uniparental evidence is consistent with this autosomal structure when evaluated at appropriate resolution. Early Y-chromosomal studies often interpreted haplogroups J1, J2, and E as Levantine by default due to coarse haplogroup binning, population pooling, and omission of Southern European comparators. When Southern Italian and Aegean populations are incorporated at finer resolution, Ashkenazi paternal structure overlaps Southern Italian and Aegean variation and separates from Levantine profiles, with J2 dominant and J1 present as a secondary lineage compatible with drift and founder effects.

Independent autosomal constraints further exclude Levantine-majority models. The CCR5-Δ32 allele exhibits a European-restricted distribution with a pronounced north–south cline and near absence in Levantine and Arabian populations. Ashkenazi Jews fall squarely within the European range, consistent with European acquisition followed by demographic amplification under endogamy rather than Levantine inheritance.

The genetic results align closely with the historical geography of Jewish settlement. Jewish communities are documented continuously in Southern Italy, Sicily, and the central Mediterranean from antiquity through Late Antiquity, forming the demographic substrate from which Ashkenazi Jewry later emerged. Northern Italian and Tuscan Jewish communities are historically secondary and do not represent ancestral source populations. Italian Jews, who never left Southern Italy as a population, occupy the same Southern Italian and Aegean autosomal space as Ashkenazi Jews, providing a direct population-level falsification of Northern Italian proxy models.

Taken together, the analyses presented here indicate that apparent Ashkenazi genetic intermediacy between Europe and the Levant is best understood as a reference population artifact rather than as an intrinsic feature of Ashkenazi population history. When Southern Italian and central Mediterranean comparators are included, Ashkenazi Jews do not resolve as genetically intermediate between Europe and the Levant. Instead, they derive primarily from Southern Italy and the central Mediterranean, with secondary eastern Mediterranean continuity and later Eastern European admixture shaped by founder effects and long-term endogamy.

Supplementary Materials: The following supplementary files document analyses described in the main text and provide full outputs for transparency and replication. All supplementary materials are provided as separate PDF documents. **Supplementary File S1.** Independent modern population principal component analysis (PCA) and Global25 distance outputs, including PCA visualizations and full 25-dimensional Euclidean distance rankings for the Ashkenazi_Germany population average. **Supplementary File S2.** Author-generated autosomal FST distance calculations and qpAdm admixture model outputs, including feasible, marginal, and infeasible models with fully documented outgroup configurations. **Supplementary File S3.** Modern Global25 distance-based similarity heatmap and ranked distance table for Ashkenazi_Germany, showing geographic clustering and the top 30 closest modern populations. **Supplementary File S4.** Global25 similarity map and full 25-dimensional distance rankings for Italian_Jew, including a geographic similarity visualization and the top 30 closest modern populations based on Euclidean distance in Global25 autosomal space.

References

1. Atzmon, G., Hao, L., Pe'er, I., Velez, C., Pearlman, A., Palamara, P.F., Morrow, B., Friedman, E., Oddoux, C., Burns, E. and 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. https://doi.org/10.1016/j.ajhg.2010.04.015
2. Behar, D.M., Garrigan, D., Kaplan, M.E., Mobasher, Z., Rosengarten, D., Karafet, T.M., Quintana-Murci, L., Ostrer, H., Skorecki, K. and Hammer, M.F. (2004). Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. *Human Genetics* 114: 354–365. https://doi.org/10.1007/s00439-003-1073-7
3. Behar, D.M., Yunusbayev, B., Metspalu, M., Metspalu, E., Rosset, S., Parik, J., Rootsi, S., Chaubey, G., Kutuev, I., Yudkovsky, G., Khusnutdinova, E.K., Balanovsky, O., Semino, O., Pereira, L., Comas, D., Gurwitz, D., Bonne-Tamir, B., Parfitt, T., Hammer, M.F., Skorecki, K. and Villems, R. (2010). The genome-wide structure of the Jewish people. *Nature* 466: 238–242. https://doi.org/10.1038/nature09103
4. Boattini, A., Martinez-Cruz, B., Sarno, S., Harmant, C., Useli, A., et al. (2013). Uniparental markers in Italy reveal a sex-biased genetic structure and different historical strata. PLOS ONE 8(5): e65441. https://doi.org/10.1371/journal.pone.0065441
5. Bonfil, R. (1994). *Jewish Life in Renaissance Italy*. Berkeley: University of California Press.
6. Botigué, L.R., Henn, B.M., Gravel, S., Maples, B.K., Gignoux, C.R., Corona, E., Atzmon, G., Burns, E., Ostrer, H., Flores, C., Bertranpetit, J., Comas, D. and Bustamante, C.D. (2013). Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proceedings of the National Academy of Sciences USA* 110: 11791–11796. https://doi.org/10.1073/pnas.1306223110
7. Brace, S., Diekmann, Y., Booth, T., et al. (2022). Genomes from a medieval mass burial show Ashkenazi-associated hereditary diseases pre-date the 12th century. *Cell* 185: 4350–4359.e6. https://doi.org/10.1016/j.cell.2022.11.002
8. Bray, S.M., Mulle, J.G., Dodd, A.F., Pulver, A.E., Wooding, S. and Warren, S.T. (2010). Signatures of founder effects, admixture, and selection in the Ashkenazi Jewish population. *Proceedings of the National Academy of Sciences USA* 107: 16222–16227. https://doi.org/10.1073/pnas.1004381107
9. Campbell, C.L., Palamara, P.F., Dubrovsky, M., Botigué, L.R., Fellous, M., Atzmon, G., Oddoux, C., Pearlman, A., Hao, L., Henn, B.M., Burns, E., Bustamante, C.D., Comas, D., Friedman, E., Pe'er, I. and 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. https://doi.org/10.1073/pnas.1204840109
10. Chazan, R. (1987). *European Jewry and the First Crusade*. Berkeley: University of California Press.
11. Chiaroni, J., King, R.J., Myres, N.M., Henn, B.M., Ducourneau, A., Mitchell, M.J., Boetsch, G., Sheikha, I., et al. (2010). The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations. *European Journal of Human Genetics* 18: 348–353. https://doi.org/10.1038/ejhg.2009.166
12. Di Gaetano, C., Voglino, F., Guarrera, S., Fiorito, G., Rosa, F., Di Blasio, A.M., Manzini, P., Dianzani, I., Betti, M., Cusi, D., Frau, F., Barlassina, C., Mirabelli, D. and Matullo, G. (2012). An overview of the genetic structure within the Italian population from genome-wide data. *PLOS ONE* 7: e43759. https://doi.org/10.1371/journal.pone.0043759
13. Di Giacomo, F., Luca, F., Anagnou, N., et al. (2003). Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. Molecular Phylogenetics and Evolution 28(3): 387–395. https://doi.org/10.1016/S1055-7903(03)00016-2
14. Drineas, P., Tsetsos, F., Plantinga, A., et al. (2019). Genetic history of the population of Crete. *Annals of Human Genetics* 83: 373–388. https://doi.org/10.1111/ahg.12328
15. Fadhlaoui-Zid, K., Martinez-Cruz, B., Khodjet-El-Khil, H., Mendizabal, I., Benammar-Elgaaïed, A. and Comas, D. (2011). Genetic structure of Tunisian ethnic groups revealed by paternal lineages. *American Journal of Physical Anthropology* 146: 271–280. https://doi.org/10.1002/ajpa.21581

16. Feldman, L.H. (1993). *Jew and Gentile in the Ancient World*. Princeton: Princeton University Press.
17. Grossman, A. (1995). *The Early Sages of Ashkenaz*. Jerusalem: Magnes Press.
18. Gruen, E.S. (2002). *Diaspora: Jews Amidst Greeks and Romans*. Cambridge, MA: Harvard University Press.
19. Hammer, M.F., Redd, A.J., Wood, E.T., et al. (2000). Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proceedings of the National Academy of Sciences USA* 97: 6769–6774. https://doi.org/10.1073/pnas.100115997
20. Hammer, M.F., Behar, D.M., Karafet, T.M., Mendez, F.L., Hallmark, B., Erez, T., Zhivotovsky, L.A., Rosset, S. and Skorecki, K. (2009). Extended Y chromosome haplotypes resolve multiple and unique lineages of the Jewish priesthood. Human Genetics 126(5): 707–717. https://doi.org/10.1007/s00439-009-0727-5
21. King, R.J., Ozcan, S.S., Carter, T., Kalfoğlu, E., Atasoy, S., Triantaphyllidis, C., Kouvatsi, A., Lin, A.A., Chow, C.E.T., Zhivotovsky, L.A., Michalodimitrakis, M. and Underhill, P.A. (2008). Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. *Annals of Human Genetics* 72: 205–214. https://doi.org/10.1111/j.1469-1809.2007.00414.x
22. Kopelman, N.M., Stone, L., Wang, C., Gefel, D., Feldman, M.W., Hillel, J. and Rosenberg, N.A. (2009). Genomic microsatellites identify shared Jewish ancestry intermediate between Middle Eastern and European populations. *BMC Genetics* 10: 80. https://doi.org/10.1186/1471-2156-10-80
23. Lerga-Jaso, J., Novković, B., Unnikrishnan, D., et al. (2025). Tracing human genetic histories with precise local ancestry inference. *Nature Communications* 16: 4576. https://doi.org/10.1038/s41467-025-59936-3
24. Listman, J.B., Hasin, D., Kranzler, H.R., Malison, R.T., Mutirangura, A., Sughondhabirom, A., Aharonovich, E., Spivak, B. and Gelernter, J. (2010). Identification of population substructure among Jews using STR markers and dependence on reference populations included. BMC Genetics 11: 48. https://doi.org/10.1186/1471-2156-11-48
25. Nebel, A., Filon, D., Brinkmann, B., Majumder, P.P., Faerman, M. and Oppenheim, A. (2001). The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. *American Journal of Human Genetics* 69: 1095–1112. https://doi.org/10.1086/324070
26. Ostrer, H. and Skorecki, K. (2013). The population genetics of the Jewish people. *Human Genetics* 132: 119–127. https://doi.org/10.1007/s00439-012-1235-6
27. Rodríguez-Varela, R., Günther, T., Krzewińska, M., et al. (2017). Genomic analyses of pre-European conquest human remains from the Canary Islands. *Current Biology* 27: 3396–3402.e5. https://doi.org/10.1016/j.cub.2017.09.059
28. Roth, C. (1946). *History of the Jews in Italy*. Philadelphia: Jewish Publication Society.
29. Rutgers, L.V. (1995). *The Jews in Late Ancient Rome*. Leiden: Brill.
30. Sarno, S., Boattini, A., Pagani, L., Sazzini, M., De Fanti, S., Quagliariello, A., Gnecchi-Ruscone, G.A., Guichard, E., Ciani, G., Bortolini, E., Barbieri, C., Cilli, E., Petrilli, R., Mikerezi, I., Sineo, L., Vilar, M., Wells, S., Luiselli, D. and Pettener, D. (2017). Ancient and recent admixture layers in Sicily and Southern Italy trace multiple migration routes along the Mediterranean. *Scientific Reports* 7: 39844. https://doi.org/10.1038/s41598-017-01802-4
31. Schuenemann, V.J., Krause, J., et al. (2017). Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. *Nature Communications* 8: 15694. https://doi.org/10.1038/ncomms15694
32. Semino, O., Magri, C., Benuzzi, G., et al. (2004). Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J. *American Journal of Human Genetics* 74: 1023–1034. https://doi.org/10.1086/386295
33. Shen, P., Lavi, T., Kivisild, T., et al. (2004). Reconstruction of patrilineages and matrilineages of Samaritans. *Human Mutation* 24: 248–260. https://doi.org/10.1002/humu.20077
34. Tian, C., Kosoy, R., Nassir, R., Lee, A., Villoslada, P., Klareskog, L., Hammarström, L., Garchon, H.J., Pulver, A.E., Ransom, M., Gregersen, P.K. and Seldin, M.F. (2009). European population genetic substructure. *Molecular Medicine* 15: 371–383. https://doi.org/10.2119/molmed.2009.00094
35. Waldman, S., Backenroth, D., Harney, É., et al. (2022). Genome-wide data from medieval German Jews. *Cell* 185: 4703–4716.e16. https://doi.org/10.1016/j.cell.2022.11.002
36. Xue, J., Lencz, T., Darvasi, A., Pe'er, I. and Carmi, S. (2017). The time and place of European admixture in Ashkenazi Jewish history. PLOS Genetics 13(4): e1006644. https://doi.org/10.1371/journal.pgen.1006644

doi:10.20944/preprints202601.2134.v2

37. Zoossmann-Diskin, A. (2010). The origin of Eastern European Jews revealed by autosomal, sex chromosomal and mtDNA polymorphisms. *Biology Direct* 5: 57. https://doi.org/10.1186/1745-6150-5-57