

Original research

Sephardic origins revealed for rare skin disorder, recessive dystrophic epidermolysis bullosa, in individuals carrying the unique c.6527insC mutation

Emily Mira Warshauer ^{1,2}, Paul A Maier,³ Goran Runfeldt,³ Ignacia Fuentes,^{4,5,6} Maria José Escamez,⁷ Laura Valinotto,^{8,9} Monica Natale,⁸ Graciela Manzur,¹⁰ Nuria Illera,⁷ Marta Garcia,⁷ Marcela Del Rio,⁷ Angeles Mencia,⁷ Almudena Holguin,⁷ Fernando Larcher,^{7,11} Garrett Hellenthal,¹² Adam R Brown,¹³ Liliana Consuegra,¹⁴ Carolina Rivera,^{14,15} Inês Nogueiro,¹⁶ Jean Tang,¹⁷ Anthony Oro,^{17,18} Peter Marinkovich,¹⁸ Francis Palisson,^{5,6} Matthias Titeux,^{19,20} Alain A Hovnanian,^{19,20,21} Eli Sprecher,^{22,23} Karl Skorecki,^{24,25} David Norris,^{1,26} Anna Bruckner,²⁶ Igor Kogut,^{1,26} Ganna Bilousova,^{1,26} Dennis Roop^{1,26}

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/jmg-2025-110967>).

For numbered affiliations see end of article.

Correspondence to
Dr Emily Mira Warshauer;
emwarshauer@gmail.com

Received 1 June 2025
Accepted 8 September 2025

ABSTRACT

Background Recessive dystrophic epidermolysis bullosa (RDEB) is a rare and severe blistering skin disorder caused by loss-of-function mutations in the type VII collagen gene (*COL7A1*). The *COL7A1* c.6527insC mutation is curiously prevalent among individuals with RDEB and is found worldwide in Europe and the Americas. Previous research has suggested the possibility of a Sephardic Jewish origin of the mutation; however, individuals with RDEB are not known to have predominant Jewish ancestry.

Methods In this study, a global cohort of individuals with RDEB with the c.6527insC founder mutation from Spain, France, Argentina, Chile, Colombia and the USA were investigated by autosomal genotyping, pairwise identical-by-descent matching and a local ancestry analysis. Age estimation analysis was performed to determine when Jewish founders introduced the c.6527insC mutation into Iberian and Native American populations (~900 CE and 1492 CE, respectively).

Results Sephardic ancestry was identified at the haplotype spanning the c.6527insC mutation in 85% of the individuals, despite mixed ancestry elsewhere in the genome and no known recent Sephardic ancestry. Identical-by-descent matching between this RDEB subpopulation and a known crypto-Jewish community in Belmonte, Portugal was also ascertained, providing support for crypto-Jewish ancestry in this RDEB subpopulation.

Conclusion The identification of this unique RDEB subpopulation unified by the single most prevalent c.6527insC mutation holds great potential to facilitate promising new RDEB therapies using CRISPR Cas 9 gene and base editing. The identification of a single guide RNA allowing efficient and safe editing of this variant would represent a unique drug to treat a large cohort of patients with the same founder mutation.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Shared ancestral origins have been identified in distinct Recessive Dystrophic Epidermolysis Bullosa (RDEB) populations.

WHAT THIS STUDY ADDS

⇒ Our study elucidates common Sephardic ancestry for individuals with RDEB carrying the c.6527insC mutation in Spain, France, Argentina, Chile, Colombia and the USA.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ We report the most comprehensive study to date of individuals with RDEB carrying a single mutation (c.6527insC) with a unique shared history, findings which hold great potential to accelerate promising new RDEB therapies.

BACKGROUND

Recessive dystrophic epidermolysis bullosa (RDEB) is a rare and severe blistering skin disorder caused by mutations in the type VII collagen gene (*COL7A1*).^{1–3} RDEB exhibits pronounced mutational heterogeneity, and most *COL7A1* mutations recur quite infrequently,⁴ yet the *COL7A1* c.6527insC mutation is intriguingly prevalent.⁵ While the exact mechanism underlying the c.6527insC mutation has not been fully elucidated, it is known that premature termination codons mediate mRNA decay, leading to null alleles and a deficiency of functional type VII collagen.^{1,6} The great majority of individuals with RDEB with the c.6527insC mutation exhibit a very severe phenotype, except for rare cases with a mild phenotype.⁷ The recent investigation of several RDEB Hispanic populations from Spain and other Spanish-speaking populations in Europe and the Americas suggested common ancestry, inherited at least in part through Sephardic migration, among individuals carrying the c.6527insC mutation.⁸



© Author(s) (or their employer(s)) 2025. No commercial re-use. See rights and permissions. Published by BMJ Group.

To cite: Warshauer EM, Maier PA, Runfeldt G, *et al.* *J Med Genet* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jmg-2025-110967

Jewish people have faced complex waves of migrations during their extensive and dynamic history, resulting in elaborate population genetic patterns, and providing insight into general patterns of health and disease in society.^{9,10} The majority of Sephardic Jews emigrated from Spain during the time of the Spanish Inquisition to other European countries, North Africa, the Middle East and the Americas, along with a number of pathogenic mutations. Many who did not leave Spain during the Inquisition converted to Catholicism, while continuing to observe Judaism in secrecy, creating crypto-Jewish communities. The presence of surviving crypto-Jewish descendants has been established in an earlier study of paternal lineages across the Iberian Peninsula.¹¹

Individuals with RDEB carrying the c.6527insC mutation with otherwise unknown Jewish ancestry from Spain, France, Argentina, Chile, Colombia and the USA were investigated in detail and considerable Sephardic ancestry at the region of the c.6527insC mutation was unambiguously identified. The findings in this study strongly support crypto-Jewish roots as part of the RDEB history.

METHODS

Participant selection and sample collection of individuals with the c.6527insC mutation

A total of 132 RDEB homozygous and compound heterozygous patient samples from Spain, France, Argentina, Chile, Colombia and the USA with the c.6527insC mutation in *COL7A1* were used in this study. Their genotypes were previously identified using various sequencing technologies and subsequently confirmed by Sanger sequencing (online supplemental table S1).^{5,12–14} Out of these, 126 were used in the final analysis (see section ‘Autosomal genotyping and kinship estimation’ below). In addition, five Sephardic individuals (two kept in the final analysis) from the endogamous community of Belmonte, Portugal were included to evaluate the genetic relationship of RDEB-carrying individuals with a uniquely preserved crypto-Jewish community with thriving modern Sephardic people.

Autosomal genotyping and kinship estimation

Autosomal genotyping was performed by Gene by Gene, Ltd, on a customised version of the Infinium Global Screening Array-24 V3.0 BeadChip, and analysed with the Family Finder autosomal DNA test.¹⁵ This array includes approximately 700 000 SNPs for the 22 pairs of autosomal chromosomes and chromosome X. We used KING V2.2.4¹⁶ to estimate kinship coefficient between each pair of 137 samples initially considered in the study. Individuals were removed from further analysis if they were third-degree relatives or closer, resulting in a final total of 126 c.6527insC-carriers and two Belmonte individuals.

Local ancestry analysis

Local population ancestry was estimated for the 126 unrelated individuals carrying one or more copies of the c.6527insC mutation. Samples were first imputed to the Illumina Human OmniExpress BeadChip array SNP set for compatibility with all reference data. We used IMPUTE2¹⁷ and the 1000 Genomes¹⁸ reference panel with a union of SNPs from both chip types, and an imputation probability threshold of 0.9. Next, we used Eagle V2.4.1¹⁹ for long-range phasing of haplotypes and to impute any remaining no-calls. For a reference panel, we used an extensive collection (104,521) of FamilyTreeDNA (hereafter ‘FTDNA’) customers spanning each major worldwide population. With a genotyping rate of 0.988 and all SNPs present in

Table 1 Reference populations used for local ancestry classification and their sources

Target population*	Proxy population†	N‡	Citation§
Sub-Saharan African	Nigerian	53	Siva 2008 ¹⁸
Native American	Native American groups in South America: Cachi, Colla, Quechua, unknown tribal affiliation in Ecuador, Peru and Bolivia	53	Eichstaedt <i>et al</i> 2014; Pagani <i>et al</i> 2016; Mallick <i>et al</i> 2016; Maier <i>et al</i> 2021 ^{20–23}
East Asian	South Han Chinese	53	Maier <i>et al</i> 2021 ²³
South Asian	Dravidian	53	Siva 2008 ¹⁸
North European	British	53	Maier <i>et al</i> 2021 ²³
South European	Iberian	53	Maier <i>et al</i> 2021 ²³
Ashkenazi Jewish	Ashkenazi Jewish	53	Maier <i>et al</i> 2021 ²³
Sephardic Jewish in Turkey	Sephardic Jewish in Turkey	53	Maier <i>et al</i> 2021 ²³
Sephardic Jewish in Morocco	Sephardic Jewish in Morocco	38	Maier <i>et al</i> 2021 ²³

*Continental or regional population of interest.
†Specific reference population used in local ancestry analysis.
‡Sample size; all non-Jewish reference populations were capped to the largest Jewish sample size.
§Study from which reference samples were drawn.

the phasing panel, we phased and imputed 637 645 SNPs across chromosomes 1 through 22. One positional Burrows-Wheeler transform phasing iteration using the default autoselection process.

A reference panel for local ancestry classification was constructed to represent population ancestries from each major continent, as well as various Jewish and non-Jewish populations in Europe. We included nine reference populations in total: Sub-Saharan African, Native American, East Asian, South Asian, North European, South European, Ashkenazi Jewish, Sephardic Jewish from Turkey and Sephardic Jewish from Morocco (table 1).^{18,20–23} Sample size disparities can adversely influence machine learning accuracy, so we randomly chose an approximately equal number of samples from each population. The two smallest sample sizes available were Turkish Sephardic (n=53) and Moroccan Sephardic (n=38), and inordinately small sample sizes can also be problematic. Therefore, we compromised on a sample size of 53 for all populations and used all available samples of Moroccan Sephardic.

We used a method similar to RFmix²⁴ for local classification but adapted by FTDNA for their proprietary tool myOrigins V3.0.²³ This method for classifying phased segments into populations was found to be more accurate than RFmix, and the subsequent step of smoothing out classification errors with a conditional random field is identical to that described in RFmix. We classified each 500 SNP segment, using a sliding window to move each overlapping segment across each chromosome in increments of 200 SNPs. A hidden Markov model was used to correct switch errors implicit in imperfect phasing, by predicting the true (hidden) diploid phase of each maternal and paternal segment, given the observed order of population labels.²³ The transition probabilities used were as follows: no strand flip (0.850; eg, 1/2 to 1/2), strand flip (0.000; eg, 1/2 to 2/1), partial overlap (0.128; eg, 1/2 to 1/3), partial overlap and strand flip (0.017; eg, 1/2 to 3/1), other (0.005; eg, 1/2 to 3/4). Results were processed and plotted using custom scripts with R V4.1.2.²⁵

Pairwise identical-by-descent matching

We used the Family Finder algorithm to perform a SNP-wise comparison for each pair of all 128 individuals to find shared segments, defined as runs of SNPs sharing at least one allele. Seed segments were formed from at least 900 adjacent matching SNPs and extended in both directions to merge with adjacent segments. Unqualified segments were then discarded if they did not contain at least 480 SNPs, a genetic distance of 2 centimorgans (cM), or a density of 105 SNPs/cM. We used the linkage map (build 37) from the 1000 Genomes Project.¹⁸ Given the potentially very distant identical-by-descent (IBD) matches expected among descendants of a centuries-old Jewish diaspora, we used custom filters to determine whether two individuals ‘match’. A match is defined as two people with at least 5 cM in common and allowing for shared segments ≥ 2 cM. We summarised pairwise matches overall, and those with segments overlapping the c.6527insC mutation, in R V.4.1.2.²⁵

Triangulation

We used triangulation as an additional validation to more conservatively assess IBD matching between all 128 samples. Triangulation²⁶ is the transitive property of shared DNA: segments that are shared IBD between three or more individuals from a recent common ancestor should match pairwise between all three, at an overlapping genetic location. Any segments with a total overlap between group members of < 2 cM were removed. We constructed a dendrogram based on a hierarchical structure analysis of total shared triangulated DNA to find clusters of more closely related individuals. This assumes all individuals have recent common ancestry through just one genealogical line, through which they inherited the c.6527insC mutation. We used the `hclust` and `dendrogram` functions in base R, with the complete linkage method to find similar clusters.²⁵ Distances were computed as $1 - (\text{cM}/\text{cM}_{\text{max}})$.

Age estimates for c.6527insC-containing haplotypes

We used two independent methods to estimate the date that c.6527insC was introduced, first from Sephardic people into Iberia, and later into the Americas. First, we applied the ‘Gamma’ method.^{27–28} The basic theory is that, under the Haldane recombination model,²⁹ the length between recombination breakpoints is a random Poisson process, and exponentially distributed with rate of 1.0. This simple model does not account for interference (inhibition of one crossover by another) and is not perfectly accurate. However, at shorter genetic distances, the likelihood (and effect) of multiple recombination events is negligible. The logic is as follows: a mutation is inherited within a haplotype, which is iteratively broken into smaller segments with each successive meiosis. The mean distance from mutation to one of its two flanking recombination breakpoints is $1/\tau$ Morgans, or $100/\tau$ cM, where τ is the time to most recent common ancestor (TMRCA) in units of generations. Note: the TMRCA for a group of mutation-carriers may be (substantially) younger than the TMRCA of the mutation itself, since it only considers a small subset of all mutation-carriers, past and present. Therefore, the length of the entire segment is distributed as *Gamma* ($2, \tau$), with expected length $2/\tau$ Morgans, or $200/\tau$ cM.

The method of Gandolfo²⁸ corrects for the bias of small sample size of segment lengths l_i , with a bias correction factor, $b(n) = (2n - 1)/2n$. Using this correction, they estimate the TMRCA as:

$$\hat{\tau} = \frac{b(n)2n}{\sum_{i=1}^n l_i}$$

Exact confidence intervals are then calculated from τ/τ . This calculation assumes that all mutation-carriers are unrelated (‘independent’) since their TMRCA, that is, no close cousin relationships, which would break the distributional assumption of independent and identically distributed samples. A second ‘correlated’ version of the calculation accounts for the possibility of reticulating history, using the bias correction factor $b(n^*) = 2n/(1 + (n - 1)\rho)$, where ρ is the mean pairwise correlation between segment lengths, and \bar{l} and s^2 are respectively the mean and variance of segment lengths:

$$\hat{\rho} = \frac{\bar{n}l - s^2(2n + 1)}{\bar{n}l + s^2(n - 1)}$$

For the Sephardic-Spanish time estimate (T_1), we used the population segment edges derived from myOrigins V.3.0. Segment boundaries were defined as start and end points of a Sephardic segment spanning the mutation at position 48 611 297 on chr3.³⁰ For compound heterozygotes, we only used the haplotype carrying the c.6527insC mutation. This haplotype was determined by IBD matching to c.6527insC-homozygous individuals and selecting the best match (ie, having the fewest nucleotide mismatches). For the American time estimate (T_2), we used the maximum of pairwise IBD matching segments derived from Family Finder. Only segments from New World individuals were used for T_2 .

The second age estimate method we applied was Alder V.1.03,³¹ which uses patterns of LD decay to compute time since an admixture event. Alder builds on previous methods,^{32–33} which show that admixture LD scales with time since admixture (t), genetic distance (d) and the initial difference (D_0) in allele frequencies between mixing populations:

$$D_t \approx e^{-td} D_0$$

The method fits a least-squares curve to patterns of exponential LD decay, to solve for the number of generations since admixture, and then calculates confidence intervals by jackknifing each chromosome.

We used Alder to estimate T_1 as the admixture date between Sephardic and Iberian references, with Old World RDEB individuals as the admixed population. We then estimated T_2 as the admixture date between Native American references and Old World RDEB individuals, with New World RDEB individuals as the admixed population. For both estimations, we considered a range of genetic distances between 0.05 and 50 cM. We converted g generations with generation time gt to years before present as $gt \times (g + 1)$, to account for the unknown age of participants.

Importantly, each of the events we tested (T_1 and T_2) actually consists of two potentially different time periods: (1) the TMRCA of an IBD segment, and (2) the potentially later time of population admixture. For example (T_2), the ancestral population of Sephardic-Iberian mutation donors could have lived in Europe for some time prior to colonisation of the Americas, and subsequent intermixing with Native Americans. Gamma gets at (1) whereas Alder gets at (2). However, (1) and (2) should converge on the same time period if the population size of donors was small. Although we hereafter refer to T_1 and T_2 as each being a single time period, we note the possibility that Gamma could overestimate the admixture time compared with Alder.

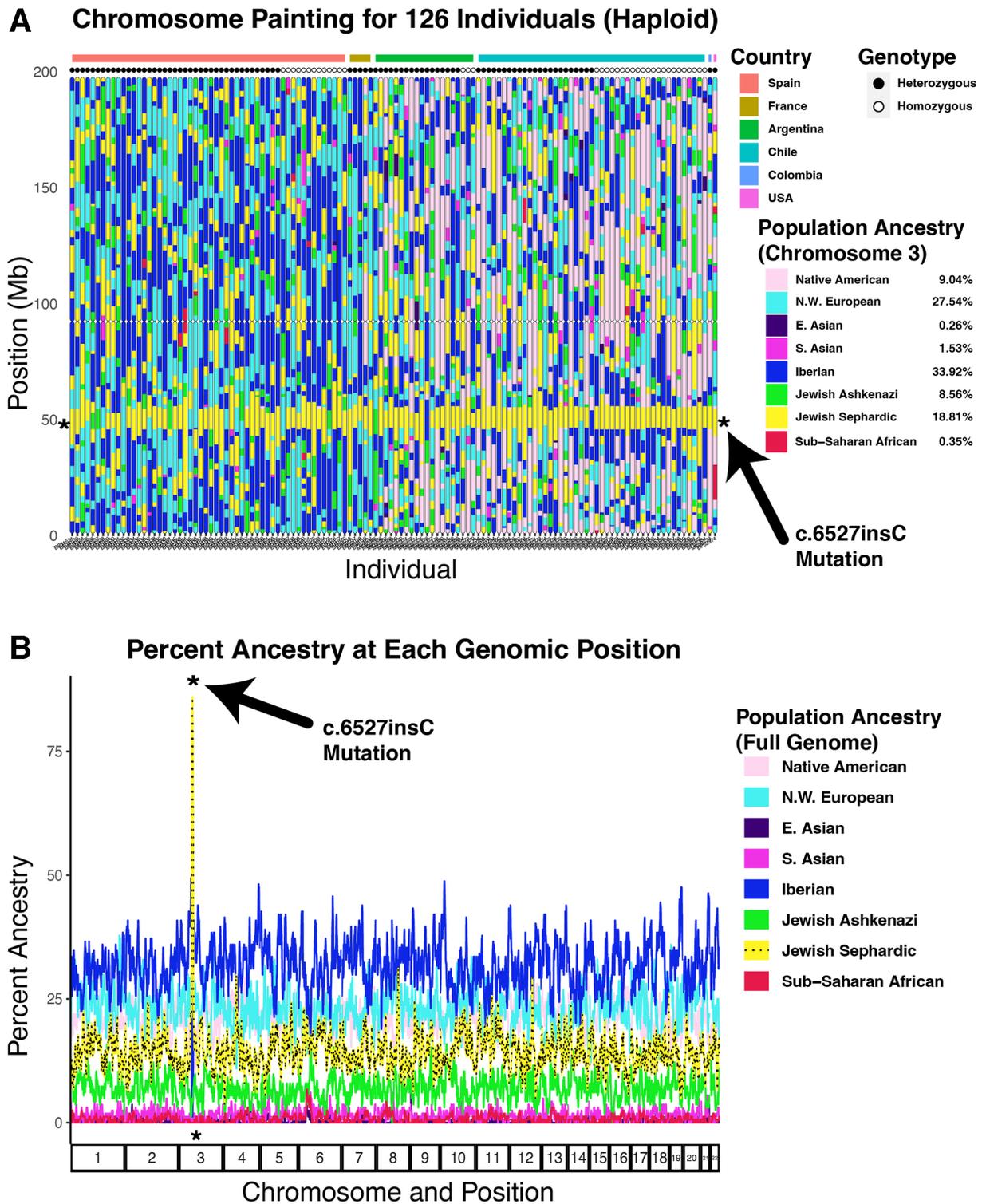


Figure 1 Sephardic Local Ancestry Spans c.6527insC mutation. (A) Results of myOrigins V.3.0 analysis showing local and overall population ancestry for chromosome 3. All 126 unrelated carriers of the c.6527insC mutation are shown for the haplotype that is most Sephardic at position 48,611,297. Turkish and Moroccan Sephardic ancestry are merged into one group for display due to lower sample sizes and genetic similarity. Country of origin is denoted by rainbow colours, and genotype is denoted by circles. (B) Genome-wide population ancestry aggregated across all 126 samples. Both haplotypes are included for homozygotes, and one haplotype (best match to homozygotes) is included for heterozygotes. A peak of 83% Sephardic ancestry is found at chr3, position 48,611,297.

RESULTS

Local ancestry analysis

myOrigins V.3.0 analysis found a strong signal of Sephardic population ancestry spanning the c.6527insC mutation (figure 1). Out of 164 haplotypes carrying the mutation (38 homozygotes; 88

heterozygotes), 136 (83%) were found to be Sephardic, and 142 (87%) were found to be Jewish more generally. In contrast, only 15% (95% CI: 8% to 22%) of the entire genome was Sephardic, and only 22% (95% CI: 13% to 30%) of the entire genome was Jewish (yellow line in figure 1B). Many heterozygotes did

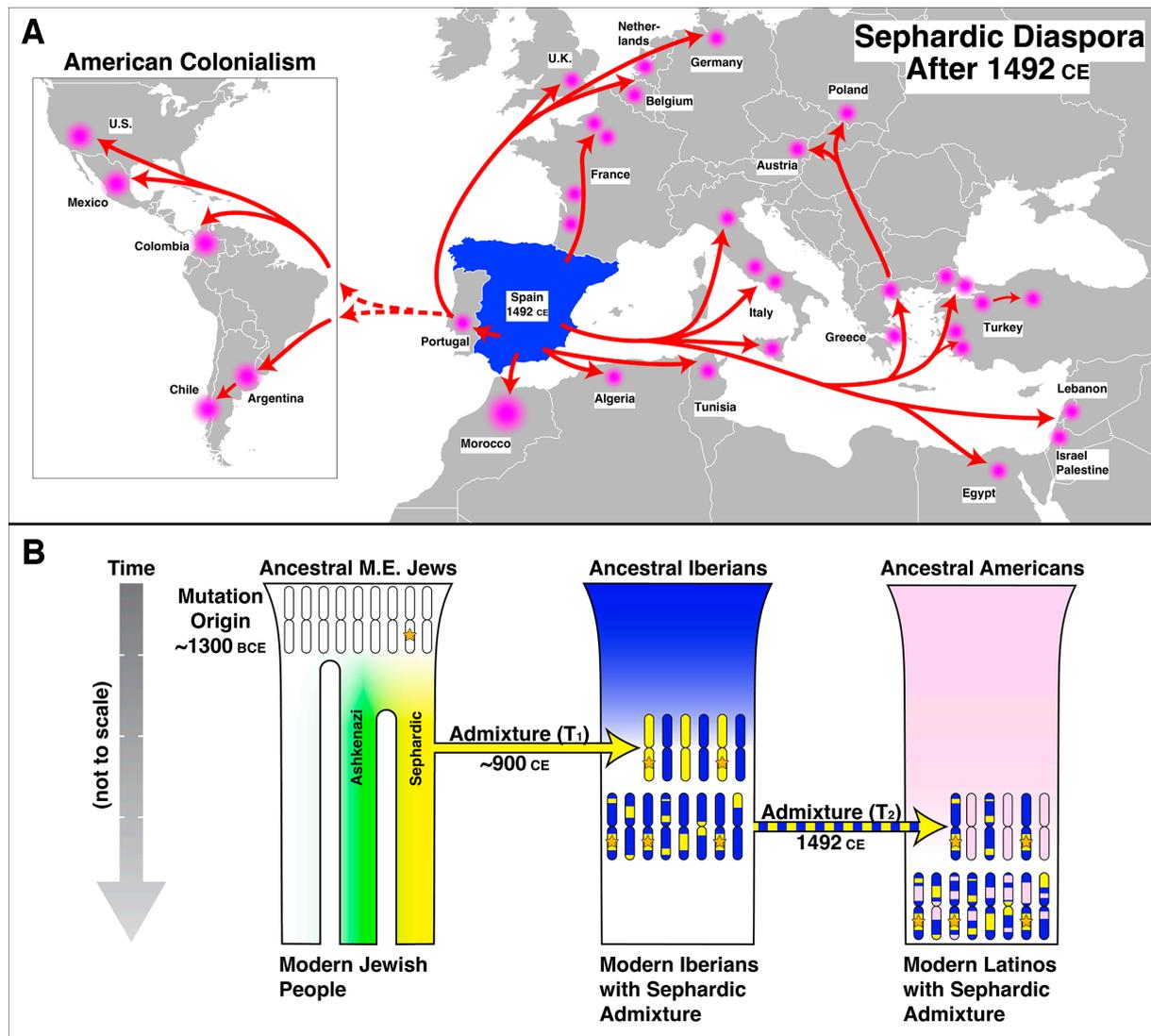


Figure 2 Sephardic diaspora and the c.6527insC mutation. (A) Geographic representation of Sephardic Diaspora during post-Columbian history (after 1492 CE). Routes and locations are only approximate and likely incomplete. (B) Conceptual historical model of the c.6527insC mutation, with hypothesised times to be tested. First (T_0), the mutation arose in a single ancestor during the Iron Age.¹⁴ Next (T_1), a group of Sephardic individuals with high prevalence of the mutation migrated to Iberia where they admixed with local Spaniards. This may have occurred during the Golden Age of Jews in Spain.^{8,14} Finally (T_2), many post-Columbian diaspora ensued following the Spanish Inquisition, including converso Sephardic individuals who migrated to the Americas. Importantly, this model predicts the mutation to be overlapped by primarily Sephardic DNA segments in both Old and New World individuals.

not have overlapping Sephardic ancestry on their second haplotype (online supplemental fig. S1), which included a range of other RDEB-causing mutations (online supplemental table S1), suggesting that the c.6527insC mutation in particular shows this pattern. As expected, the main difference between New and Old World individuals was a high prevalence of Native American segments in the New World, but also higher Iberian ancestry. No other Sephardic spikes of ancestry were observed elsewhere in the genome (figure 1B). This is consistent with our two-step model of Sephardic heritage for this mutation (figure 2B).

Pairwise IBD matching

Out of 8128 possible pairs of matching participants ($128 \times 127/2$), 1203 pairs (15%) were found to have matching segments after applying a 5 cM minimum IBD threshold (figure 3). Given the potentially long time elapsed since the cohort shared a common ancestor, and the wide variance of IBD overlap after just a few generations, many null matches were expected. Out of 15%

of matching pairs, 1035 (86%) of 1203 matched at a segment spanning the c.6527insC mutation. We performed a one-sample Wilcoxon ranked-sign test to determine that 1035 matches spanning c.6527insC are significantly higher than the genome-wide mean of 30% or 3% ($W=0$, $p < 0.005$, $n=100$). The Chilean samples showed the highest incidence of pairwise matching, probably owing to higher post-Columbian shared ancestry and/or endogamy. Notably, the two included Belmonte Sephardic samples (with no mutation; last two columns of figure 3) matched many mutation-carrying samples (with no previously known Sephardic ancestry).

Triangulation

We found a general pattern of autosomal clustering within countries of origin for most study participants (figure 4). Proximate countries in the New World (Argentina, Chile and Colombia) and Old World (Spain, Portugal, France) were found on many adjacent branches of the hierarchical tree, suggesting shared

Pairwise IBD Matching for 128 Individuals

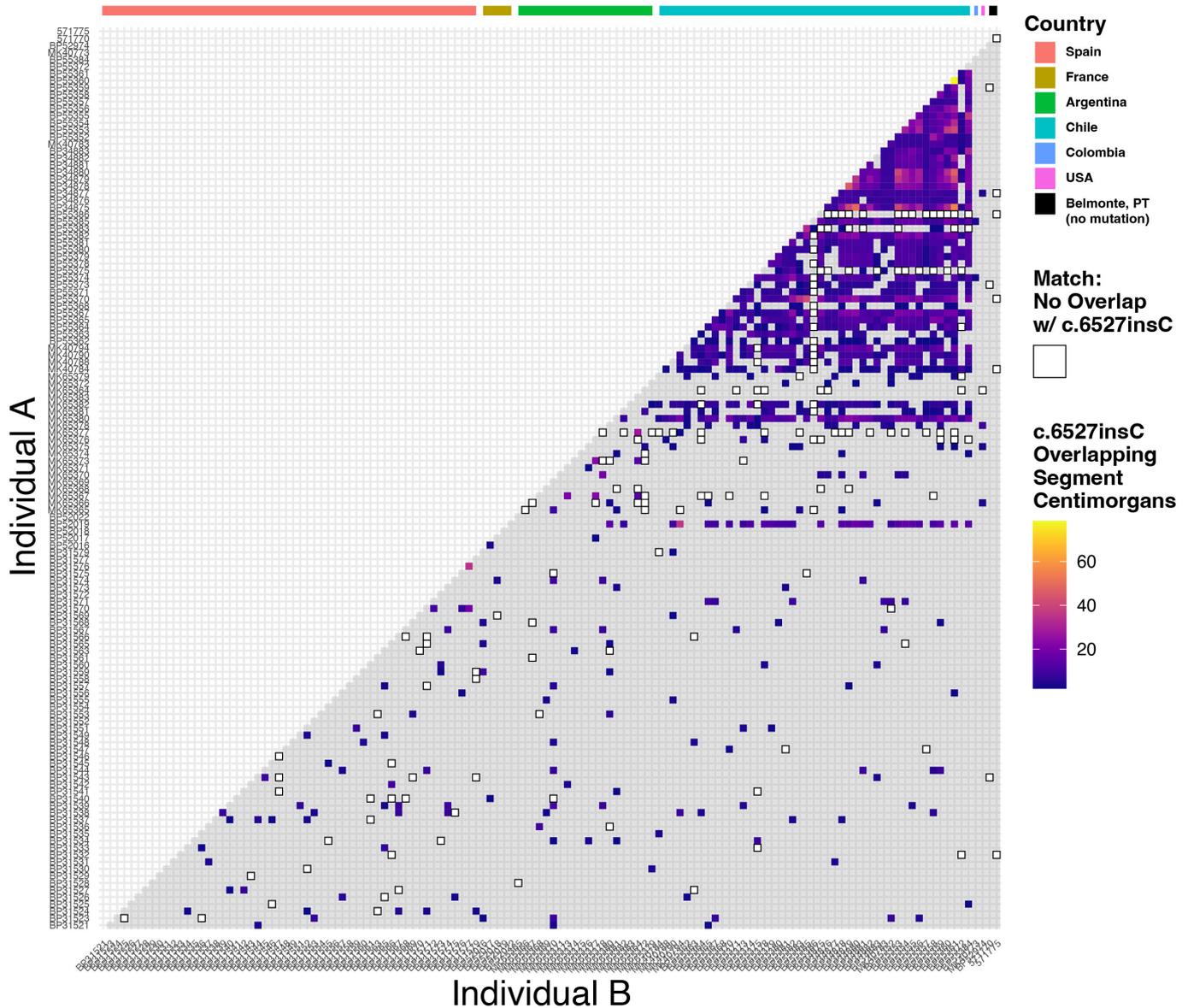


Figure 3 Pairwise IBD matching. Family finder pairwise matching with relevant thresholds applied between all pairs of 128 samples (including two Belmonte samples). Any IBD match spanning the *c.6527insC* mutation is shown by plasma colours denoting the segment length in cM. If a pair matches but without spanning the mutation, an empty box is shown. Rainbow colours denoting country of origin are identical to those in figure 1, except that Belmonte samples are shown in black. IBD, identical-by-descent.

post-Columbian ancestry between geographically close individuals. Interestingly, one French individual was found nested in a clade of Chilean individuals (BP52019), whereas one Chilean (MK40783) and two Argentinean samples were found nested in Spanish clades (MK65370 and MK65374). Most New World (particularly Chilean) clades were much more recently related than clades in other countries, consistent with our IBD results.

Age estimates for *c.6527insC*-containing haplotypes

Both Gamma and Alder methods of age estimation supported a pre-Columbian introduction of *c.6527insC*-containing haplotypes into Iberia (T_1), and subsequent post-Columbian introduction by this admixed population into the Americas (T_2 ; table 2). Given the evidence from IBD and triangulation analysis of recent shared ancestry in the Americas, we favour

the ‘correlated’ genealogy approach of the Gamma method; however, we report ‘independent’ values for completeness. Generation times have varied considerably between eras and cultures, and we report several calendar dates based on generation times of 20, 25 and 30 years per generation. However, here we assume 30 years, because recent estimates suggest a value between 26 and 30.³⁴

The Gamma method estimated T_1 to occur in 990 (95% CI: 371 to 1610) CE, whereas Alder estimated an earlier time of 767 (95% CI: 308 to 1226) CE. Although there is a 200 year difference in these mean estimates, their 95% CIs are highly overlapping (online supplemental fig. S2), and both estimates coincide with the so-called Golden Age of Jews in Spain. As noted in the methods, the Gamma method estimates the TMRCA of an IBD segment, whereas Alder estimates the potentially later time

Hierarchical Clustering of Total Triangulated >2 cM

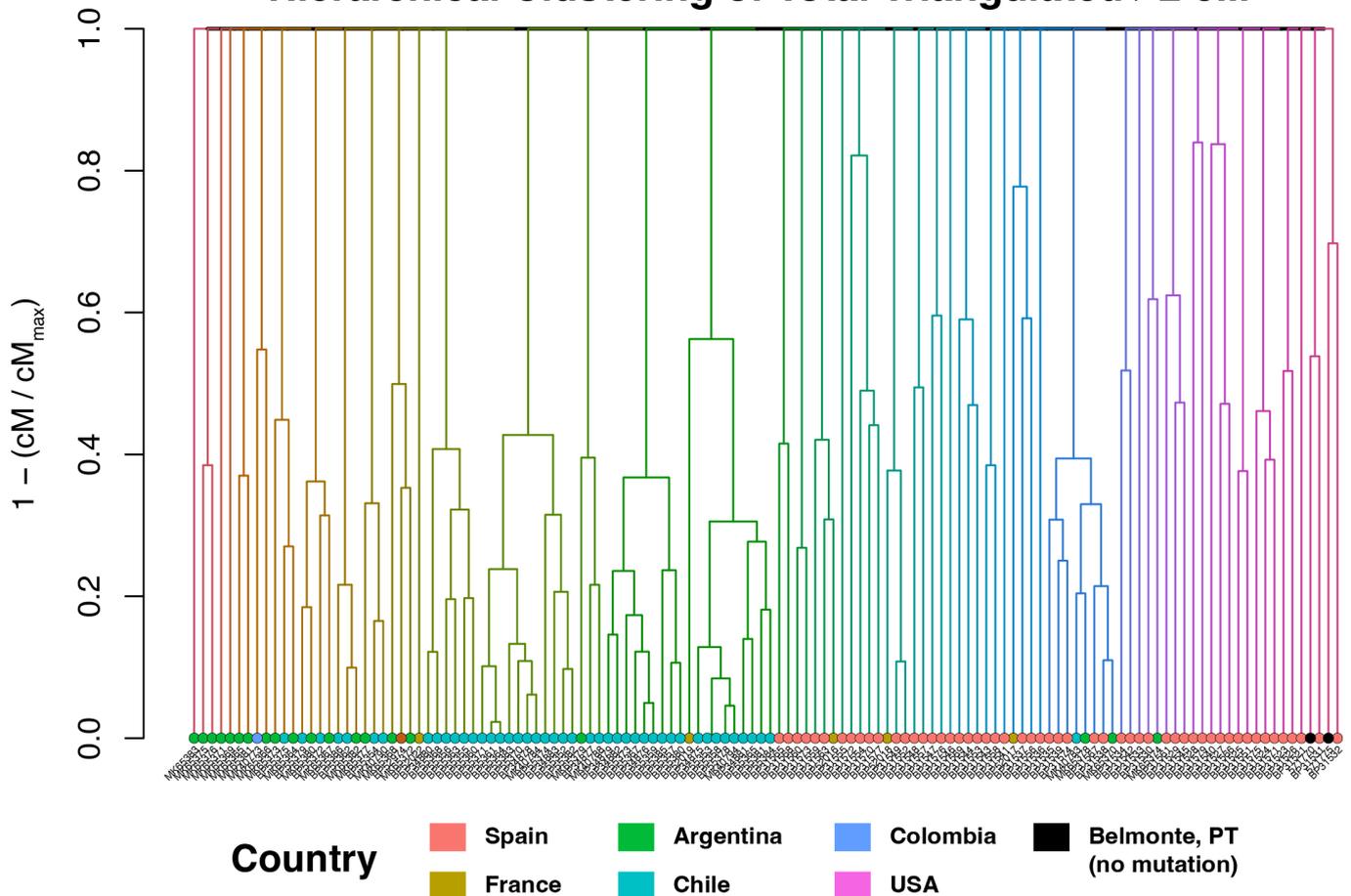


Figure 4 Autosomal relationships of study participants. Hierarchical clustering tree showing proximity of the relationship for all 128 individuals in the study. Clustering was performed using total cM for all triangulated DNA segments, that is, those matching between three or more individuals. The vertical axis is normalised between 0 and 1. Rainbow colours for circles denote country of origin and are identical to those in figure 3. Rainbow colours for tree branches denote clusters.

of population admixture. Convergent estimates by these two approaches likely imply a small donor population size.

The Gamma method estimated T_2 to occur in 1702 (95% CI: 1539 to 1859) CE, whereas Alder estimated an earlier time of 1665 (95% CI: 1650 to 1679) CE. These estimates are more consistent and have tighter CIs, but are noticeably later than the post-Columbian era beginning in 1492 CE. This leaves the possibility that Jews carrying the c.6527insC mutation in Europe first assimilated there before migrating to the Americas generations later. This is consistent with our hypothesis (figure 2b) and

supported by the fact that more European than Native American segments flank the c.6527insC segment (figure 1A). A one-tailed binomial test confirmed that flanking segments to the mutation were more European than expected by chance ($p = 2.62 \times 10^{-12}$).

CONCLUSION

Genetics, culture, history and religion have unified the Jewish people since their origins in the Middle East more than 5000 years ago.³⁵ The c.6527insC mutation in modern Jews originally

Table 2 Age estimates for admixture of c.6527insC mutation into Iberia and Americas

Admixture event*	Method†	Assumption‡	Generations BP§	Calendar date (20 Y/G)¶	Calendar date (25 Y/G)¶	Calendar date (30 Y/G)¶
Sephardic+Spanish	Gamma	'Independent' genealogy	29.7 (26.4–33.5)	1408 (1333–1475) CE	1254 (1160–1338) CE	1101 (988–1201) CE
Sephardic+Spanish	Gamma	'Correlated' genealogy	33.4 (12.8–54.1)	1334 (922–1748) CE	1162 (647–1679) CE	990 (371–1610) CE
Sephardic+Spanish	Alder	–	40.9 (25.6–56.2)	1186 (880–1492) CE	976 (594–1359) CE	767 (308–1226) CE
American+Span-Seph	Gamma	'Independent' genealogy	8.9 (7.5–10.6)	1824 (1790–1853) CE	1774 (1732–1810) CE	1725 (1674–1767) CE
American+Span-Seph	Gamma	'Correlated' genealogy	9.7 (4.5–15.1)	1809 (1701–1914) CE	1755 (1620–1886) CE	1702 (1539–1859) CE
American+Span-Seph	Alder	–	10.9 (10.4–11.4)	1784 (1774–1794) CE	1725 (1712–1737) CE	1665 (1650–1679) CE

*Which modelled admixture event, either T_1 or T_2 .

†Age estimation method.

‡Assumption of Gamma method.

§Mean estimated age in generations before present, with 95% CI.

¶Mean estimated age in calendar year (see assumed generation time), with 95% CI.

dates to ~1300 BCE with successive admixture among ancient Iberians occurring in ~900 CE and with Native Americans ~1492 CE (figure 2A and B). The c.6527insC mutation is still present in Spain today. Endogamy historically predominated in Jewish populations, giving rise to a high prevalence of genetic diseases. Previously unknown as prevalent in Jewish populations, RDEB was only recently recognised to have an association with Sephardic communities.⁸ The RDEB population in this study from Spain, France, Argentina, Chile, Colombia and the USA, previously unknown to have Jewish origins, unambiguously demonstrates considerable Sephardic ancestry at the region surrounding the c.6527insC mutation (figure 1A and B). An unequivocal Sephardic lineage in this population can be explained by shared ancestors who underwent forced conversions to Catholicism on the Iberian Peninsula during the Spanish Inquisition. Crypto-Jewish communities emerged at this time of massive persecution, so families could secretly maintain their Jewish faith while outwardly professing adherence to Catholicism. Descendants of crypto-Jews frequently lost their Jewish identity over time; however, genetic diseases present in these communities consistently survived. The c.6527insC mutation undoubtedly has Jewish origins, arriving from the Middle East ~900 CE (estimates ranging from 767 to 1101 CE; table 2) during the 'golden age', when Jewish life was flourishing on the Iberian Peninsula (figure 2B). Thus, our results are broadly consistent with our hypothesised time course (figure 2B). The mixing of individuals with RDEB with Native American individuals has estimated dates ranging from 1665 to 1725 CE (table 2), during a post-Columbian era that saw the emergence of many crypto-Jewish communities.

The Spanish Inquisition is among the fiercest examples of a system enforcing religious intolerance and propagating genetic homogeneity, with consequences clearly still relevant to our world today.³⁶ It is estimated that persecuted Jewish individuals during the Inquisition represented approximately one-third of early Spanish immigrants and the propagation of a rare genetic disease, such as RDEB, in Iberian exile communities can be attributed, at least in part, to a founder effect.^{37,38} To date, there is scarce knowledge about Sephardic crypto-Jewish descendants, due to limited data from the original source of Sephardic Jews on the Iberian Peninsula combined with the challenge of identifying the existent communities today.³⁹ Communities in Portugal (Belmonte and Bragança) and Mallorca (Chuetas) have known crypto-Jewish ancestry. Interestingly, individuals from the Belmonte community were shown to distantly match individuals with RDEB carrying the c.6527insC mutation, substantiating crypto-Jewish ancestry (figures 3 and 4).

The migration patterns of individuals with RDEB with the c.6527insC mutation notably resemble the Sephardic Diaspora map after 1492 CE (figure 2A). While Sephardic ancestry predominates overwhelmingly at the region surrounding the c.6527insC mutation in this RDEB population, traces of Ashkenazi ancestry are also evident (figure 1A). The combination of both Sephardic and Ashkenazi ancestry in this region suggests that the mutation predates Sephardic and Ashkenazi admixture during the Middle Ages.¹⁰ Jewish people have lived on the Iberian Peninsula as far back as the Roman Empire, and the c.6527insC mutation can be traced back even further to pre-Roman communities on the Iberian Peninsula.¹⁴

Waves of Jewish people, including RDEB ancestors, migrated to the Americas from Europe after 1492 CE and greatly shaped the modern day Latin American population structure.³⁷ The European, Native American and Sephardic admixture represented by the c.6527insC mutation reflects the migration of

Jewish-Iberian people escaping the Iberian Peninsula during the Inquisition.^{38,40} The c.6527insC mutation is widespread in the Americas and recently was identified as the most frequent RDEB variant in Brazil.⁴¹ While many individuals did leave for the Americas, the majority in this RDEB subpopulation are Spanish, indicating a strong example of Jewish families remaining in Spain throughout the duration of the Inquisition. Many of those individuals who did leave Spain, settled in remote areas of Europe and the Americas where they were less likely to be discovered and exposed. A number of mutations were brought to the Americas via these routes, including the pathogenic growth hormone receptor mutation in Laron syndrome, observed in the isolated Lojano community in Ecuador, a population known to have the influence of Sephardic crypto-Jewish ancestry.³⁸

One particular region in North America well-known to attract immigrants who wished to maintain their Jewish practices under less scrutiny in the New World while demonstrating a veneer of Catholicism was the isolated San Luis Valley in northern New Mexico and southern Colorado.⁴² This region of the USA supports a population structure with greater frequencies of genetic diseases well-known in Jewish populations, including Pemphigus Vulgaris, Bloom Syndrome and BRCA1/BRCA2 associated breast cancer, suggesting those fleeing the Inquisition may have brought Jewish founder mutations to contemporary Hispanic populations.^{40,43} The 185delAG mutation in BRCA1 was identified at a surprisingly high frequency in non-Jewish Colorado families originating in the San Luis Valley and similar to the RDEB population, who were found to have Jewish ancestry.^{43,44}

Interestingly, one of the individuals with RDEB is heterozygous for c.6527insC and c.7485+5G>A, a mutation predominantly found in the San Luis Valley, harbouring multiple RDEB mutations potentially integrated in crypto-Jewish communities. Further studies will elucidate the extent of RDEB mutations associated with Jewish ancestry.

This unique group of individuals with RDEB carrying the c.6527insC mutation unambiguously exhibits collective Sephardic ancestry and may also represent the largest set of individuals with RDEB ever reported with a single mutation in an otherwise remarkably heterogeneous disease. The recognition of this unique RDEB subpopulation highlights the patterns of this rare genetic disease and illuminates the genetic architecture of the Sephardic Jewish population. Furthermore, recognition of the RDEB subpopulation unified by the single most prevalent c.6527insC mutation will enhance the efficient implementation of CRISPR Cas 9 gene and base editing therapies. The identification of a single guide RNA allowing efficient and safe editing of this variant would represent a unique drug to treat a large cohort of patients with the same founder mutation. At present, the therapies accomplished in preclinical settings, including the ex vivo and in vivo correction of c.6527insC by multiple gene editing strategies, inspire great optimism for the future of promising treatments for RDEB and other rare diseases.⁴⁵⁻⁵⁰

Author affiliations

¹Charles C. Gates Center for Regenerative Medicine, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, Colorado, USA

²Pediatric, Adolescent, and Adult Dermatology, Point Pleasant, New Jersey, USA

³Gene by Gene, FamilyTreeDNA, Houston, Texas, USA

⁴Centro de Genética y Genómica, Universidad del Desarrollo Facultad de Medicina Clínica Alemana, Santiago, Santiago Metropolitan Region, Chile

⁵Fundacion DEBRA Chile, Santiago, Chile

⁶Departamento de Bioingeniería, Celular y Molecular, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Santiago, Chile

⁷Universidad Carlos III de Madrid, Departamento de Bioingeniería (UC3M), Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Centro de

Investigación Biomédica en Red de Enfermedades Raras (CIBERER-ISCIII), Instituto de Investigación Sanitaria Fundación Jiménez Díaz (IIS-FJD), Madrid, Spain

⁸Center for Research in Genodermatosis and Epidermolysis Bullosa (CEDIGE), Buenos Aires, Argentina

⁹National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

¹⁰Dermatology, Hospital de Clínicas José de San Martín, Buenos Aires, Buenos Aires, Argentina

¹¹Epithelial Biomedicine Division, Centre for Energy, Environment and Technology Research (CIEMAT), Madrid, Spain

¹²UCL Genetics Institute (UGI), Department of Genetics, Evolution and Environment, UCL, London, UK

¹³Avotaynu Foundation, Englewood, New Jersey, USA

¹⁴Fundacion DEBRA Colombia, Bogotá, Colombia

¹⁵Medical Genetics, Universidad del Rosario, Bogotá, Bogota, Colombia

¹⁶University of Porto Institute of Molecular Pathology and Immunology, Porto, Porto District, Portugal

¹⁷Dermatology, Stanford University, Stanford, California, USA

¹⁸Program in Epithelial Biology and Department of Dermatology, Stanford University School of Medicine, Stanford, California, USA

¹⁹Imagine Institute, Paris Cité University, Paris, France

²⁰Laboratory of Genetic Skin Diseases, INSERM U1163, Paris, France

²¹Department of Genetics, Necker Hospital for Sick Children, AP-HP, Paris, France

²²Dermatology, Tel Aviv Sourasky Medical Center, Tel Aviv-Yafo, Tel Aviv District, Israel

²³Human Molecular Genetics, Tel Aviv University Sackler Faculty of Medicine, Tel Aviv-Yafo, Tel Aviv District, Israel

²⁴Genetics & Developmental Biology, Technion Israel Institute of Technology The Ruth and Bruce Rappaport Faculty of Medicine, Haifa, Haifa District, Israel

²⁵Nephrology, Technion Israel Institute of Technology, The Ruth and Bruce Rappaport, Faculty of Medicine, Haifa, Haifa District, Israel

²⁶Dermatology, University of Colorado, Anschutz Medical Campus, School of Medicine, Aurora, Colorado, USA

Acknowledgements We thank all the patients and their families for helping us to carry out this study. We also honour the memory of Stephen Berman, MD, the Founding Director of the Epidermolysis Bullosa Center of Excellence at Children's Hospital Colorado, who cared for Hispanic patients with RDEB in Colorado for over three decades and originally suggested to us that these patients may be of crypto-Jewish ancestry. With appreciation for all of our great mentors, a very special thanks to Bruce Lee Warshauer MD, a true inspiration and brilliant mentor in Dermatology.

Contributors EMW (guarantor): conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualisation, writing—original draft, writing—review and editing. PAM: data curation, formal analysis, investigation, methodology, resources, software, validation, visualisation, writing—original draft, writing—review and editing. GR: data curation, formal analysis, software. IF: data curation, formal analysis, methodology, writing—review and editing. MJE: data curation, formal analysis, methodology, resources, visualisation, writing—review and editing. LV: investigation, resources, writing—review and editing. MN: investigation, resources, writing—review and editing. GM: data curation, resources. NI: investigation, resources. MG: investigation, resources. MdR: investigation, resources. ÁM: investigation, resources. AH: investigation, resources. FL: data curation, resources. GH: methodology, writing—review and editing. AB: project administration, supervision. LC: data curation, resources. CR: data curation, resources. IN: data curation, visualisation. JT: resources. AO: resources. PM: resources. FP: data curation, resources. MT: data curation, resources. AH: data curation, formal analysis, investigation, resources. ES: visualisation. KS: conceptualisation, visualisation. DN: funding acquisition, resources. AB: resources. IK: data curation, investigation, resources. GB: data curation, investigation, resources. DR: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, visualisation, writing—original draft, writing—review and editing.

Funding This study was supported by the Avotaynu Foundation, Epidermolysis Bullosa Research Partnership, Cure EB, Epidermolysis Bullosa Medical Research Foundation, National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institute of Health (NIH) (R01AR059947 and U01AR075932), the Department of Defense (DOD) (W81XWH-18-1-0706), Dystrophic Epidermolysis Bullosa Research Association (DEBRA) International, and the Gates Frontiers Fund. The Spanish team is supported by grants from the Spanish Ministry of Science and Innovation and European Regional Development Fund (PID2020-119792RB-I00), the Institute of Health Carlos III (RD21/0001/0022, Spanish Network of Advanced Therapies; TERAV-ISCIII) and DEBRA Austria (APR 2021-12 Carlos Leon_014). The funders had no involvement in the study design, collection, analysis and interpretation of the data, writing of the report or decision to submit the paper for publication with the exception of the Avotaynu Foundation, providing expertise in the local ancestry analysis.

Competing interests None declared.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants and was approved by Spain: UC3M Ethics Committee (Approval number: CEI21_15); France: Necker Comité de Protection des Personnes (Clinical Trials Reference number: NCT01874769); Argentina: Ethics Committee of the Ricardo Gutierrez Children's Hospital of Buenos Aires (Registry number: 16.38); Chile: Comité Ético Científico, Facultad de Medicina, Clínica Alemana—Universidad del Desarrollo (Project number: 2013-145); USA: Colorado Multiple Institutional Review Board (COMIRB no: 09-0192) and Stanford University (Project number: 30586) and Colombia: Universidad del Rosario (CIE-UR DVO005 1149-CV1192). Participants gave informed consent to participate in the study before taking part. Informed written consent was obtained from all patients in concordance with Institutional Review Board approval from Spain: UC3M Ethics Committee (Approval number: CEI21_15); France: Necker Comité de Protection des Personnes (Clinical Trials Reference number: NCT01874769); Argentina: Ethics Committee of the Ricardo Gutierrez Children's Hospital of Buenos Aires (Registry number: 16.38); Chile: Comité Ético Científico, Facultad de Medicina, Clínica Alemana—Universidad del Desarrollo (Project number: 2013-145) USA: Colorado Multiple Institutional Review Board (COMIRB no: 09-0192) and Stanford University (Project number: 30586) and Colombia: Universidad del Rosario (CIE-UR DVO005 1149-CV1192).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iD

Emily Mira Warshauer <http://orcid.org/0000-0002-4055-5196>

REFERENCES

- Hovnanian A, Rochat A, Bodemer C, *et al*. Characterization of 18 new mutations in COL7A1 in recessive dystrophic epidermolysis bullosa provides evidence for distinct molecular mechanisms underlying defective anchoring fibril formation. *Am J Hum Genet* 1997;61:599–610.
- Has C, Bauer JW, Bodemer C, *et al*. Consensus reclassification of inherited epidermolysis bullosa and other disorders with skin fragility. *Br J Dermatol* 2020;183:614–27.
- Bardhan A, Bruckner-Tuderman L, Chapple ILC, *et al*. Epidermolysis bullosa. *Nat Rev Dis Primers* 2020;6:78.
- van den Akker PC, Jonkman MF, Rengaw T, *et al*. The international dystrophic epidermolysis bullosa patient registry: an online database of dystrophic epidermolysis bullosa patients and their COL7A1 mutations. *Hum Mutat* 2011;32:1100–7.
- Escámez MJ, García M, Cuadrado-Corrales N, *et al*. The first COL7A1 mutation survey in a large Spanish dystrophic epidermolysis bullosa cohort: c.6527insC disclosed as an unusually recurrent mutation. *Br J Dermatol* 2010;163:155–61.
- Uitto J, Hovnanian A, Christiano AM. Premature termination codon mutations in the type VII collagen gene (COL7A1) underlie severe recessive dystrophic epidermolysis bullosa. *Proc Assoc Am Physicians* 1995;107:245–52.
- Chacón-Solano E, León C, Carretero M, *et al*. Mechanistic interrogation of mutation-independent disease modulators of RDEB identifies the small leucine-rich proteoglycan PRELP as a TGF- β antagonist and inhibitor of fibrosis. *Matrix Biol* 2022;111:189–206.
- Warshauer EM, Brown A, Fuentes I, *et al*. Ancestral patterns of recessive dystrophic epidermolysis bullosa mutations in Hispanic populations suggest sephardic ancestry. *Am J Med Genet A* 2021;185:3390–400.
- Behar DM, Metspalu E, Kivisild T, *et al*. Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS ONE* 2008;3.
- Oddoux C, Guillen-Navarro E, Ditiuoli C, *et al*. Mendelian diseases among Roman Jews: implications for the origins of disease alleles. *J Clin Endocrinol Metab* 1999;84:4405–9.
- Adams SM, Bosch E, Balaesque PL, *et al*. The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet* 2008;83:725–36.
- Natale MI, Manzur GB, Lusso SB, *et al*. Analysis of COL7A1 pathogenic variants in a large cohort of dystrophic epidermolysis bullosa patients from Argentina reveals a new genotype-phenotype correlation. *Am J Med Genet A* 2022;188:3153–61.
- Cuadrado-Corrales N, Sánchez-Jimeno C, García M, *et al*. A prevalent mutation with founder effect in Spanish Recessive Dystrophic Epidermolysis Bullosa families. *BMC Med Genet* 2010;11.

- 14 Sanchez-Jimeno C, Cuadrado-Corralles N, Aller E, *et al.* Recessive dystrophic epidermolysis bullosa: the origin of the c.6527insC mutation in the Spanish population. *Br J Dermatol* 2013;168:226–9.
- 15 Hu R, Maier P, Runfeldt G, *et al.* FAMILY FINDER MATCHING 5.0 Matching Algorithm and Relationship Estimation. 2021.
- 16 Manichaikul A, Mychaleckyj JC, Rich SS, *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* 2010;26:2867–73.
- 17 Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009;5.
- 18 Siva N. 1000 genomes project. *Nat Biotechnol* 2008;26:256.
- 19 Loh PR, Palamara PF, Price AL. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet* 2016;48:811–6.
- 20 Eichstaedt CA, Antão T, Pagani L, *et al.* The Andean adaptive toolkit to counteract high altitude maladaptation: genome-wide and phenotypic analysis of the Collas. *PLoS One* 2014;9:e93314e93314.
- 21 Mallick S, Li H, Lipson M, *et al.* The simons genome diversity project: 300 genomes from 142 diverse populations. *Nature New Biol* 2016;538:201–6.
- 22 Pagani L, Lawson DJ, Jagoda E, *et al.* Genomic analyses inform on migration events during the peopling of Eurasia. *Nature New Biol* 2016;538:238–42.
- 23 Maier P, Hu R, Runfeldt G, *et al.* MYORIGINS 3.0: Combining Global and Local Methods for Determining Population Ancestry. 2021.
- 24 Maples BK, Gravel S, Kenny EE, *et al.* RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet* 2013;93:278–88.
- 25 Team RC. R: a language and environment for statistical computing. 2022. Available: <https://www.r-project.org>
- 26 Ferrando-Bernal M, Morcillo-Suarez C, de-Dios T, *et al.* Mapping co-ancestry connections between the genome of a Medieval individual and modern Europeans. *Sci Rep* 2020;10:6843.
- 27 McPeck MS, Strahs A. Assessment of linkage disequilibrium by the decay of haplotype sharing, with application to fine-scale genetic mapping. *Am J Hum Genet* 1999;65:858–75.
- 28 Gandolfo LC, Bahlo M, Speed TP. Dating rare mutations from small samples with dense marker data. *Genetics* 2014;197:1315–27.
- 29 Haldane JBS. The combination of linkage values, and the calculation of distances between the loci of linked factors. *J Genet* 1919;8:299–309.
- 30 The human gene mutation database. n.d.
- 31 Loh P-R, Lipson M, Patterson N, *et al.* Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* 2013;193:1233–54.
- 32 Moorjani P, Patterson N, Hirschhorn JN, *et al.* The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet* 2011;7.
- 33 Patterson N, Moorjani P, Luo Y, *et al.* Ancient admixture in human history. *Genetics* 2012;192:1065–93.
- 34 Moorjani P, Sankararaman S, Fu Q, *et al.* A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc Natl Acad Sci U S A* 2016;113:5652–7.
- 35 Atzmon G, Hao L, Pe'er I, *et al.* Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry. *Am J Hum Genet* 2010;86:850–9.
- 36 Drelichman M, Vidal-Robert J, Voth HJ. The long-run effects of religious persecution: Evidence from the Spanish Inquisition. *Proc Natl Acad Sci U S A* 2021;118:118.
- 37 Velez C, Palamara PF, Guevara-Aguirre J, *et al.* The impact of Converso Jews on the genomes of modern Latin Americans. *Hum Genet* 2012;131:251–63.
- 38 Gonçalves FT, Fridman C, Pinto EM, *et al.* The E180splice mutation in the GHR gene causing Laron syndrome: witness of a Sephardic Jewish exodus from the Iberian Peninsula to the New World? *Am J Med Genet A* 2014;164A:1204–8.
- 39 Nogueiro I, Teixeira JC, Amorim A, *et al.* Portuguese crypto-Jews: the genetic heritage of a complex history. *Front Genet* 2015;6.
- 40 Ostrer H. The origin of the p.E180 growth hormone receptor gene mutation. *Growth Horm IGF Res* 2016;28:51–2.
- 41 Ota VK, Mariath LM, Torrelío RMF, *et al.* Genetic profiling of epidermolysis bullosa in a large Brazilian cohort. *Clin Exp Dermatol* 2025.
- 42 Alberro S. Crypto-jews and the mexican holy office in the seventeenth century. In: *The Jews and the expansion of Europe to the West 1450–1800*. 2.2001.
- 43 Mullineaux LG, Castellano TM, Shaw J, *et al.* Identification of germline 185delAG BRCA1 mutations in non-Jewish Americans of Spanish ancestry from the San Luis Valley, Colorado. *Cancer* 2003;98:597–602.
- 44 Makriyanni I, Hamel N, Ward S, *et al.* BRCA1:185delAG found in the San Luis Valley probably originated in a Jewish founder. *J Med Genet* 2005;42.
- 45 García M, Bonafont J, Martínez-Palacios J, *et al.* Preclinical model for phenotypic correction of dystrophic epidermolysis bullosa by *in vivo* CRISPR-Cas9 delivery using adenoviral vectors. *Mol Ther Methods Clin Dev* 2022;27:96–108.
- 46 Bonafont J, Mencía A, Chacón-Solano E, *et al.* Correction of recessive dystrophic epidermolysis bullosa by homology-directed repair-mediated genome editing. *Mol Ther* 2021;29:2008–18.
- 47 Bonafont J, Mencía Á, García M, *et al.* Clinically relevant correction of recessive dystrophic epidermolysis bullosa by dual sgRNA CRISPR/Cas9-mediated gene editing. *Mol Ther* 2019;27:986–98.
- 48 Mencía Á, Chamorro C, Bonafont J, *et al.* Deletion of a pathogenic mutation-containing exon of COL7A1 allows clonal gene editing correction of RDEB patient epidermal stem cells. *Mol Ther Nucleic Acids* 2018;11:68–78.
- 49 Peking P, Koller U, Duarte B, *et al.* An RNA-targeted therapy for dystrophic epidermolysis bullosa. *Nucleic Acids Res* 2017;45:10259–69.
- 50 Chamorro C, Mencía A, Almarza D, *et al.* Gene editing for the efficient correction of a recurrent COL7A1 mutation in recessive dystrophic epidermolysis bullosa keratinocytes. *Mol Ther Nucleic Acids* 2016;5.