Dystrophic Epidermolysis Bullosa: *COL7A1* Mutation Landscape in a Multi-Ethnic Cohort of 152 Extended Families with High Degree of Customary Consanguineous Marriages

Hassan Vahidnezhad[1,2,11], Leila Youssefian[1,3,11], Sirous Zeinali[2,4], Amir Hossein Saeidian[1], Soheila Sotoudeh[5], Nikoo Mozafari[6], Maryam Abiri[2,3,4], Abdol-Mohammad Kajbafzadeh[7], Mohammadreza Barzegar[6], Adam Ertel[8], Paolo Fortina[8,9] and Jouni Uitto[1,10]

Dystrophic epidermolysis bullosa is a heritable skin disease manifesting with sub-lamina densa blistering, erosions, and chronic ulcers. *COL7A1*, encoding type VII collagen, has been identified as the candidate gene for dystrophic epidermolysis bullosa. In this study, we have identified *COL7A1* mutations in a large multi-ethnic cohort of 152 extended Iranian families with high degree of consanguinity. The patients were diagnosed by clinical manifestations, histopathology, and immunoepitope mapping. Mutation detection consisted of a combination of single nucleotide polymorphism-based whole-genome homozygosity mapping, Sanger sequencing, and gene-targeted next-generation sequencing. A total of 104 distinct mutations in *COL7A1* were identified in 149 of 152 families (98%), 56 (53%) of them being previously unreported. Ninety percent of these mutations were homozygous recessive, reflecting consanguinity in these families. Three recurrent mutations were identified in five or more families, and haplotype analysis suggested a founder effect in two of them. In conclusion, *COL7A1* harbored mutations in the overwhelming majority of patients with dystrophic epidermolysis bullosa, and most of them in this Iranian cohort were consistent with autosomal recessive inheritance. The mutation profile attests to the impact of consanguinity in these families.

Journal of Investigative Dermatology (2017) **137,** 660–669; doi:10.1016/j.jid.2016.10.023

INTRODUCTION

The heritable forms of epidermolysis bullosa (EB) comprise a heterogeneous group of disorders with the clinical hallmark of blistering of the skin and the mucous membranes (Fine et al., 2014; Uitto et al., 2016b). The diagnosis is characteristically made at birth or shortly thereafter, but the subsequent progression of the disease and the overall prognosis can be highly variable depending, in part, on the underlying molecular defects, which have now been identified in as many as 18 different genes expressed in the dermal-epidermal junction (Uitto et al., 2016a, 2016b). EB has been classified into four broad categories based on the topographic location of the blistering within the cutaneous basement membrane zone. The phenotypic, both intra- and interfamilial, variability in these different forms of EB reflects the physiological functions of the proteins encoded by the mutated genes, the types and combinations of the mutations, and the influence of the environmental factors, particularly trauma and infections. In the most severe cases, EB can be lethal during the early days or weeks of life, whereas in some forms the disease manifests with protracted, life-long blistering tendency without affecting the longevity of the individual.

One of the four subtypes of EB is the dystrophic variant (DEB) (OMIM: 131750 and 226600) characterized by sub-lamina densa blistering leading to erosions, poorly healing

[1]*Department of Dermatology and Cutaneous Biology, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, Pennsylvania, USA;* [2]*Molecular Medicine Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran;* [3]*Department of Medical Genetics, Tehran University of Medical Sciences, Tehran, Iran;* [4]*Kawsar Human Genetics Research Center, Tehran, Iran;* [5]*Department of Dermatology, Children's Medical Center, Pediatric Center of Excellence, Tehran University of Medical Sciences, Tehran, Iran;* [6]*Skin Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran;* [7]*Paediatric Urology Research Center, Department of Urology, Children's Hospital Medical Center, Tehran University of Medical Sciences, Tehran, Iran;* [8]*Department of Cancer Biology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, Pennsylvania, USA;* [9]*Department of Molecular Medicine, Sapienza University, Rome, Italy; and* [10]*Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania, USA*

[11]*These authors contributed equally to this work.*

Correspondence: Jouni Uitto, Department of Dermatology and Cutaneous Biology, Sidney Kimmel Medical College at Thomas Jefferson University, 233 South 10th Street, Suite 450 BLSB, Philadelphia, Pennsylvania 19107, USA. E-mail: Jouni.Uitto@Jefferson.edu

Abbreviations: DEB, dystrophic epidermolysis bullosa; EB, epidermolysis bullosa; Mb, mega-base pair; NGS, next-generation sequencing; RDEB, recessive dystrophic EB; SNP, single nucleotide polymorphism; WES, whole-exome sequencing

Received 20 September 2016; revised 17 October 2016; accepted 18 October 2016; accepted manuscript published online 27 October 2016; corrected proof published online 27 October 2016

chronic ulcers, and extensive scarring, particularly at the sites of trauma on the hands and feet. Blistering can also occur in cornea and the gastrointestinal epithelium, and patients often experience scarring of the esophagus, which has clinical consequences with respect to oral feeding (Fine and Mellerio, 2009). Patients frequently develop an aggressive squamous cell carcinoma, leading to early death from metastases (Montaudie et al., 2016). DEB is caused by mutations in *COL7A1*, the gene encoding type VII collagen, a major protein component of the anchoring fibrils that play a critical role in securing the attachment of the dermal-epidermal basement membrane to the underlying dermis (Chung and Uitto, 2010). Electron microscopic examination of the anchoring fibrils in patients with DEB can show morphological alterations, reduction in their number, or complete absence, which correlate to a certain degree with phenotypic severity of the disease. DEB can be inherited in either an autosomal dominant (DDEB) or autosomal recessive (RDEB) fashion, the clinical manifestations in RDEB being characteristically more severe, reflecting in many cases complete absence of the anchoring fibrils. RDEB has been divided into generalized severe (previously known as the Hallopeau-Siemens type) and generalized other (formerly called the non—Hallopeau-Siemens type) (Fine et al., 2014).

In this study, we examined a large multi-ethnic cohort of DEB patients in Iran, a country of approximately 80 million inhabitants, where consanguineous marriages are frequent. In fact, approximately 38% of all Iranian marriages are consanguineous, with a mean inbreeding coefficient (alpha) of 0.0185 (Saadat et al., 2004). First cousin marriages (28%) are the most common form of consanguineous unions (Saadat et al., 2004). The Iranian population includes several ethnic groups, including Persians, Turks, Kurds, Lurs, Baluchis, Arabs, Gilakis, Tabari, and Taleshis, all with different ancestries, distinct linguistics, defined cultures, and limited geographic areas of habitation. We studied the *COL7A1* mutation profile in this cohort, consisting of 238 patients representing 152 extended families with different forms of DEB. The purpose of the study was to determine the mutation landscape with special reference to the effect of consanguinity.

RESULTS

Patient population and DEB diagnosis

As part of our efforts to characterize a large cohort of EB patients in Iran, 152 extended families with 238 patients whose clinical presentation, immunoepitope and/or homozygosity mapping were suggestive of DEB, were included for *COL7A1* mutational analysis. Patients with the diagnosis of EB seen in a number of medical centers in Iran were examined for subclassification as having DEB by several criteria. First, characteristic clinical signs consisted of blistering and erosions, mutilating scarring of the hands and feet, corneal involvement, poor dentition, and development of squamous cell carcinoma (Figure 1a—h). In some of these patients, radiological examination showed esophagus atresia (Figure 1i). In most patients suspected of having DEB, immunofluorescence epitope mapping was performed, which showed sub-lamina densa blistering, and in many of these patients staining for type VII collagen was markedly reduced or absent (Figure 1j).

Examination of the pedigrees suggested an autosomal recessive inheritance pattern in most families but an autosomal dominant inheritance pattern in only one family (Figure 1k). Among 152 families, 139 families had evidence of consanguinity, 106 of them being first cousin marriages. In some families, when skin biopsy sample analysis was not informative or available from the affected individual, homozygosity mapping was performed with a single nucleotide polymorphism (SNP)-based microarray, followed by alignment of the homozygosity blocks with positions of 21 genes known to be involved in different forms of EB or other skin fragility syndromes, which are in their differential diagnosis (Figure 1l). In a subset of patients, homozygosity mapping was performed using six short tandem repeat markers linked to *COL7A1*. Although the proband(s) showed a homozygous haplotype pattern, their parents were heterozygotes (see Supplementary Figure S1 online).

Demographic data showed the presence of multiple Iranian ethnicities in our cohort of DEB patients, but the contribution of Lurish ethnicity, an ancestral population residing primarily in the central western part of Iran, with a strong tradition of consanguineous marriages, was significantly higher in the DEB cohort than expected on the basis of their share of the total population (Figure 1m). Severity of DEB among our inbred cohort was markedly different from outbred populations, such as in the United States. Although in the US patient cohort the distribution of three major forms of DEB (i.e., RDEB generalized severe, RDEB generalized other, and dominant dystrophic EB) were 9%, 33%, and 58%, respectively, in our Iranian cohort this distribution was dramatically different and consisted of 90%, 8%, and 2%, respectively. The rate of consanguinity among the parents of the Iranian cohort of DEB patients was 92%, whereas the average rate of consanguinity of the total population in the country is about 38% (Saadat et al., 2004). The first cousin and nonconsanguineous marriages were dramatically overrepresented and underrepresented among the DEB patients in our cohort (69% and 8%, respectively) compared with the general population in Iran (28% and 61%, respectively) (Figure 1o).

Mutation analysis

Type VII collagen mutations were first sought by PCR amplification of all 118 exons and flanking intronic sequences, followed by bidirectional Sanger sequencing of 73 families. By this approach, *COL7A1* mutations were found in 69 of these families. Subsequently, a *COL7A1*-targeted array of next-generation sequencing (NGS) was designed. This design had a 99% theoretical coverage of all coding regions of the gene, which spans approximately 30 kilo-base pairs of the genome in the short arm of chromosome 3 (Figure 2a). Bioinformatics analysis with appropriate filtering showed mutations in 95 families of the 98 tested by this approach (96.9%). Fifteen families initially genotyped by Sanger sequencing of the entire *COL7A1* gene were included in this panel and served as controls for the gene-targeted NGS sequencing. All mutations disclosed by *COL7A1*-targeted sequencing were verified by bidirectional Sanger sequencing, as shown in Figure 2c.

In the total cohort of 152 families with DEB, 104 distinct mutations in *COL7A1* were detected in 149 families (detection

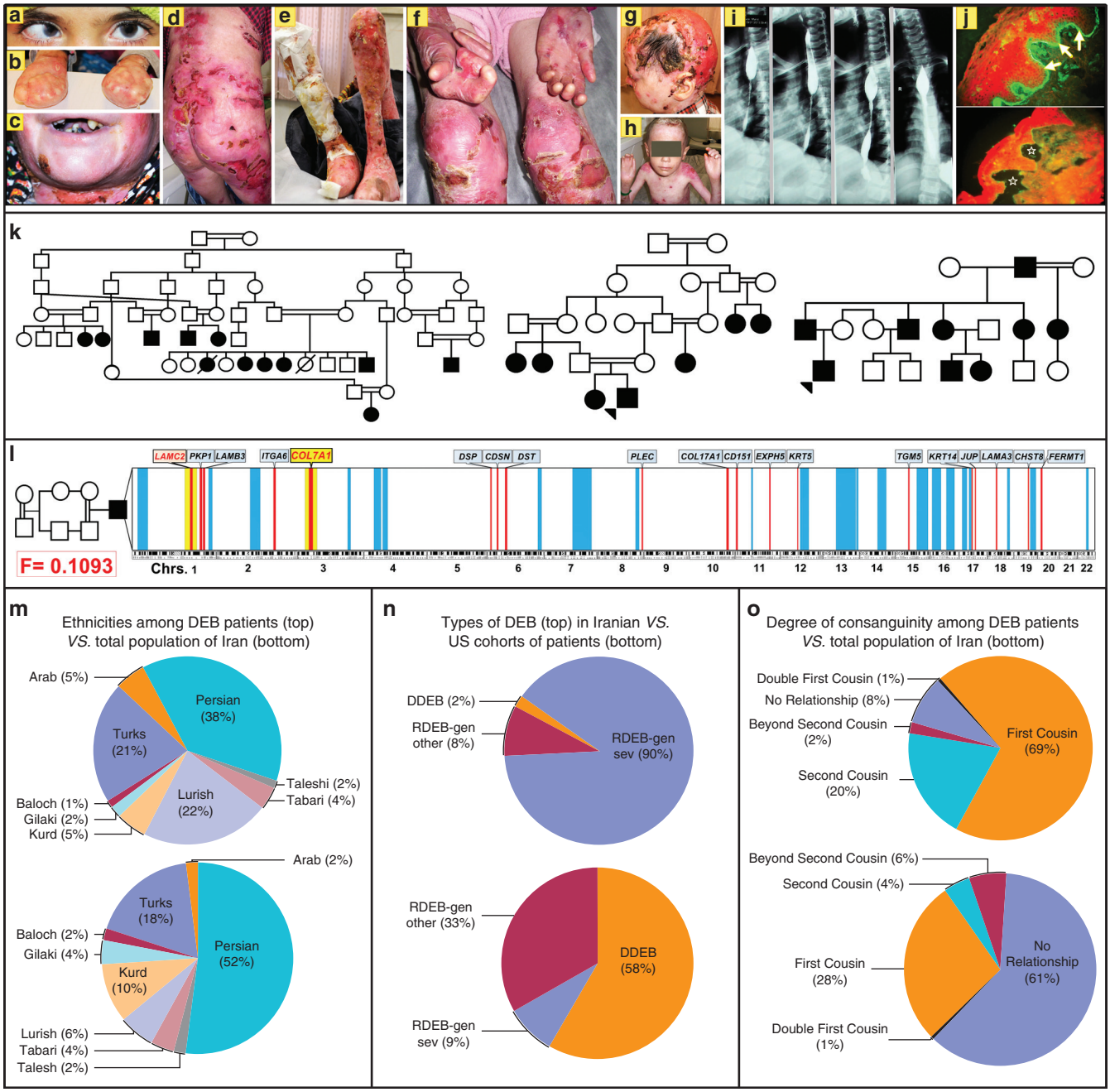


Figure 1. Candidate gene identification in families with DEB by homozygosity mapping. (**a**–**j**) Clinical features of patients with DEB, (**k**) examples of multiplex families, (**l**) homozygosity mapping in a family with RDEB, (**m**) distribution of ethnicity of DEB patients and total population of Iran, (**n**) types of DEB in Iranian versus US cohorts, and (**o**) degree of consanguinity among DEB patients versus the total population in Iran. (**a**–**h**) Note the characteristic phenotype with erosions and scarring with mitten deformities of the hands and feet as well as corneal erosions and hypodontia. (**i**) Radiography of barium swallowing shows esophageal strictures. (**j**) Immunofluorescence for type VII collagen shows markedly reduced or absent staining in RDEB patients' skin (lower panel, the stars indicate blistering) compared with normal skin with linear staining at the dermal epidermal junction (arrows). (**k**) Note the presence of multiple affected individuals in extended families, consistent with autosomal recessive (two left pedigrees) and autosomal dominant (right pedigree) patterns. Probands are indicated by arrowheads. (**l**) Genome-wide homozygosity mapping identifies multiple homozygosity blocks of >2 Mb (blue) in the autosome of a patient (in **f**) with consanguineous parents with coefficient of inbreeding (of 0.1093). Alignment of multiple genes associated with heritable blistering diseases (red lines) shows an overlap with *LAMC2* and *COL7A1* with a homozygosity block in chromosome 3, identifying them as candidate genes in this family. (**m**) Among the Iranian ethnicities, the Lurish have a higher percentage of DEB patients than expected from their presence in the total population in Iran. (**n**) The Iranian DEB population studied in this cohort consists predominantly of patients with RDEB generalized severe (90%) disease, whereas only 2% of patients have dominant dystrophic EB. This distribution dramatically differs from that in the United States. (**o**) The impact of consanguinity on DEB is illustrated by the percentages of first and second cousin parents of the Iranian DEB patients (89%), as opposed to the degree of consanguinity in the total population in Iran, where 61% of married couples have no family relationship. DEB, dystrophic epidermolysis bullosa; EB, epidermolysis bullosa; RDEB, recessive dystrophic epidermolysis bullosa; Mb, mega-base pair.

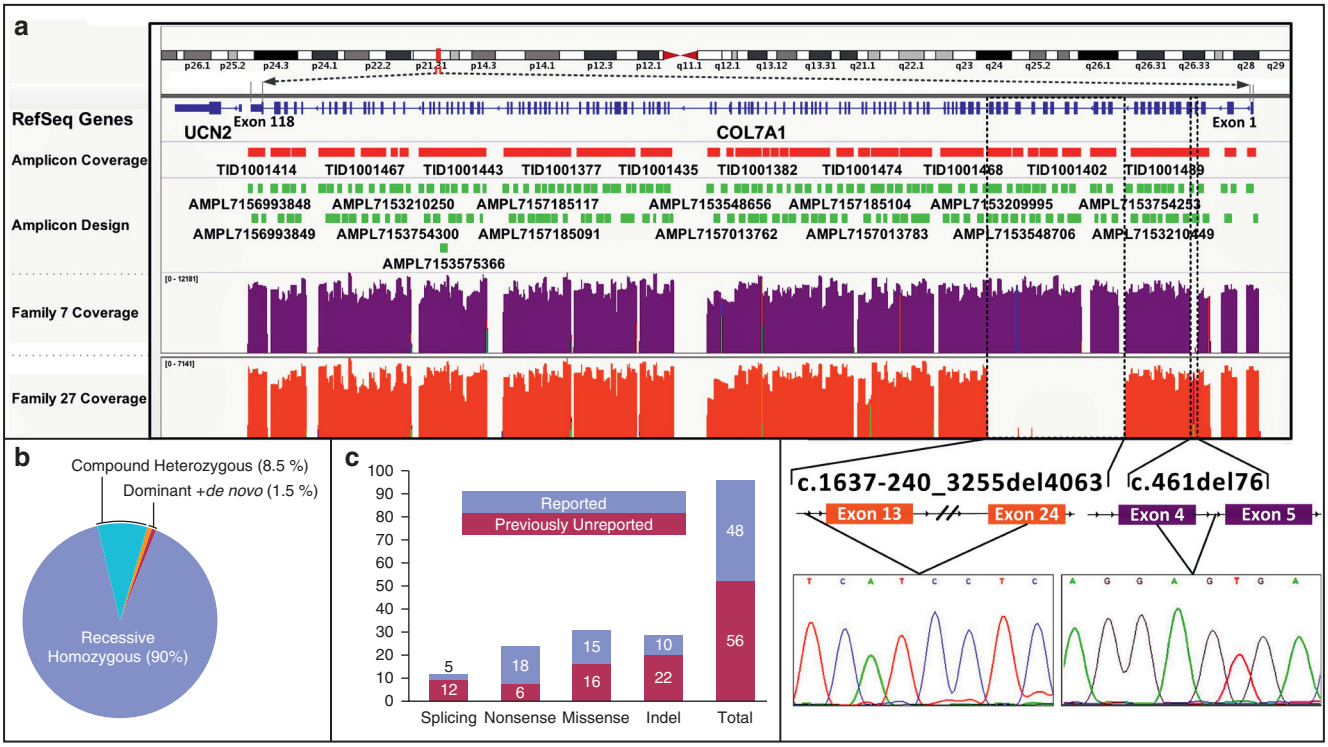


Figure 2. Schematic design of *COL7A1*-targeted next-generation sequencing array. (**a**) *COL7A1* consists of 118 exons in the chromosomal region 3p21.31. Amplicon design resulted in approximately 99% coverage of the coding regions and of 100 base pairs of flanking intronic sequences. The coverage of families 7 and 27 is shown as an illustration of mutation detection, which shows c.461del76 and c.1637-240_3255del4063 deletion mutations spanning from exon 4 to intron 4 and from intron 12 to exon 24, respectively. (**b**) This approach identified recessive homozygous mutations in 90% of the families. (**c**) These mutations were confirmed by Sanger sequencing. Of the 104 mutations identified by this array or by Sanger sequencing of PCR amplified exons and flanking intronic sequences, 56% were previously unreported; distribution of splicing, nonsense, missense, and indel mutations is indicated.

rate of 98%) by combined Sanger sequencing and gene-targeted NGS (Figure 3, and see Supplementary Table S1 online). The genetic variants included missense (31), nonsense (24), and splice junction (17) mutations, as well as large and small indel mutations (32) resulting in premature termination codon (Figures 2c and 3, and see Supplementary Tables S2 and S3 online). The missense mutations substituting glycine residues in the collagenous domain are likely to result in destabilization of the triple helical conformation (Christiano et al., 1996), and the mutations within the NC1 domain may interfere with the interactions of type VII collagen with type IV collagen and laminin 332, components of the cutaneous basement membrane zone (Brittingham et al., 2006; Chen et al., 1999). Among the 104 distinct mutations identified, 56 were previously unreported to our knowledge, and 48 have been reported before. Most nonsense mutations (75%) were previously reported while most indel mutations (69%) were previously unreported to our knowledge (Figure 2c, and see Supplementary Table S1). Among the 149 families in which mutations in *COL7A1* were detected, 134 families had biallelic mutations consistent with autosomal recessive inheritance (90%) (Figure 2b). One family had a constellation and mutation analysis, heterozygous c.5264G>A, p.Gly1755Asp (family no. 77, Figure 1k, right panel), and in one family (no. 95) the heterozygous *COL7A1* mutation c.6119G>T, p.Gly2040Ser was de novo, not present in the parents of the proband, consistent with dominant dystrophic EB (Christiano et al., 1996). Thus, the overwhelming majority (90%) of DEB families in this Iranian cohort showed homozygous recessive mutations, apparently reflecting consanguinity in this population (Figure 2b).

Among 152 families examined, consanguinity was evident in 139 families. Mutation analysis showed homozygous recessive mutations in 132 of them; in four consanguineous families, the mutations were compound heterozygous; and in three families no *COL7A1* mutations were detected. In 13 families, there was no obvious relationship between the parents of the proband, and nine of them had compound heterozygous mutations, only two of them being homozygous recessive. These observations re-emphasize the influence of consanguinity in determining the types of genetic variants disclosed in rare heritable diseases, as illustrated here by RDEB (see Supplementary Table S2).

Recurrent mutations and evidence for founder effect

A number of mutations detected in *COL7A1* in our patient cohort were recurrent, and 30 of them were found in two or more independent, unrelated families. The three most frequent recurrent mutations were c.6269_6270delC, p.Pro2090LeufsX116; c.4233delT, p.P1411PfsX298; and c.553C>T, p.Arg185X, which were detected in 18, five, and five distinct unrelated families, respectively (a total of 19% of all families) (Figure 3, and see Supplementary Table S1). Geographic distribution of the mutations within the Iranian national borders showed that the c.6269_6270delC mutation, found in 18 families, was concentrated on the western

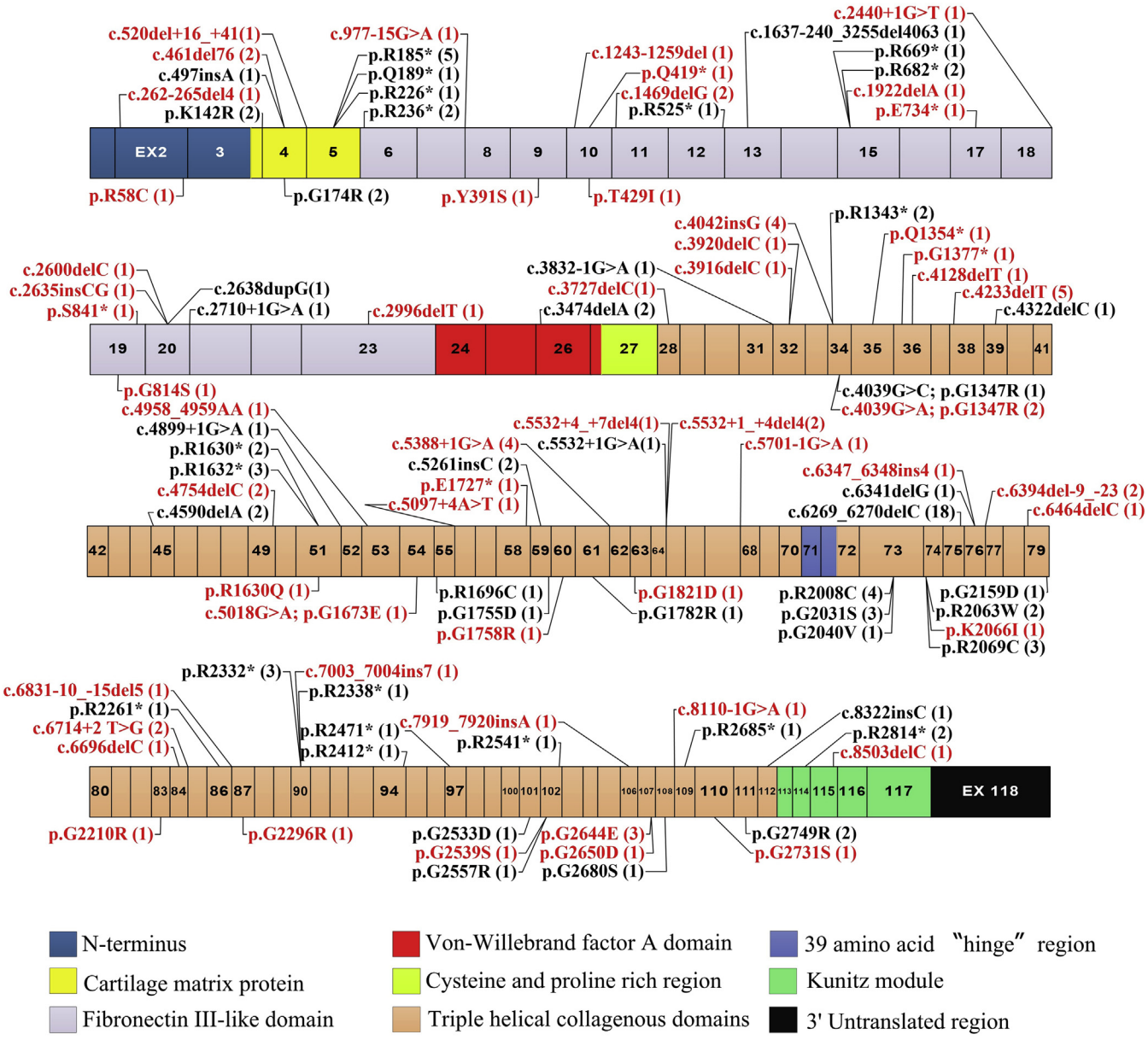


Figure 3. Organization of type VII collagen cDNA consisting of 118 exons, with positions of the mutations identified in this study. Missense mutations are below and premature termination causing mutations are above the cDNA. The mutations in red are previously unreported to our knowledge, whereas mutations in black have been reported previously. The number in parentheses after each mutation indicates the number of families harboring this mutation. The protein domains of type VII collagen encoded by *COL7A1* are identified by color code below the gene.

part of the country, extending from northwest to southwest within a distance of about 960 miles (1,500 km) (Figure 4). These 18 families belong to either Lurish (eight), Turkish (eight), or Persian (two) ethnicities. The c.4233delT mutation was also concentrated on the western part of Iran and was observed exclusively in Lurish ethnicity, whereas the p.Arg185X did not show any specific pattern of distribution and was seen in both Persian and Turkish ethnicities (Figure 4).

To address the question about the origin of these recurrent mutations and to specifically investigate whether the high frequency of these mutations was due to a founder effect and/or a mutational hotspot, we explored evidence of shared common haplotypes in these families by genotyping 33 SNPs within and in flanking regions of *COL7A1* (Figure 5). A homozygous conserved haplotype in a 4.62 mega-base pair (Mb) genomic interval, spanning from rs12489865 to rs614288, was observed in 13 families harboring the c.6269_6270delC mutation. Recombinational events in three additional families shortened the conserved homozygous haplotype to 1 Mb genomic interval, spanning from rs6784820 to rs11797. This common 1-Mb haplotype in these families was absent in our in-house, ethnic-specific control population, an observation that supports the notion that c.6269_6270delC is a common founder mutation (Figure 5). However, in one Persian family this haplotype was not conserved, and this same mutation has been previously reported in a Chinese family (Pfendner et al., 2003),

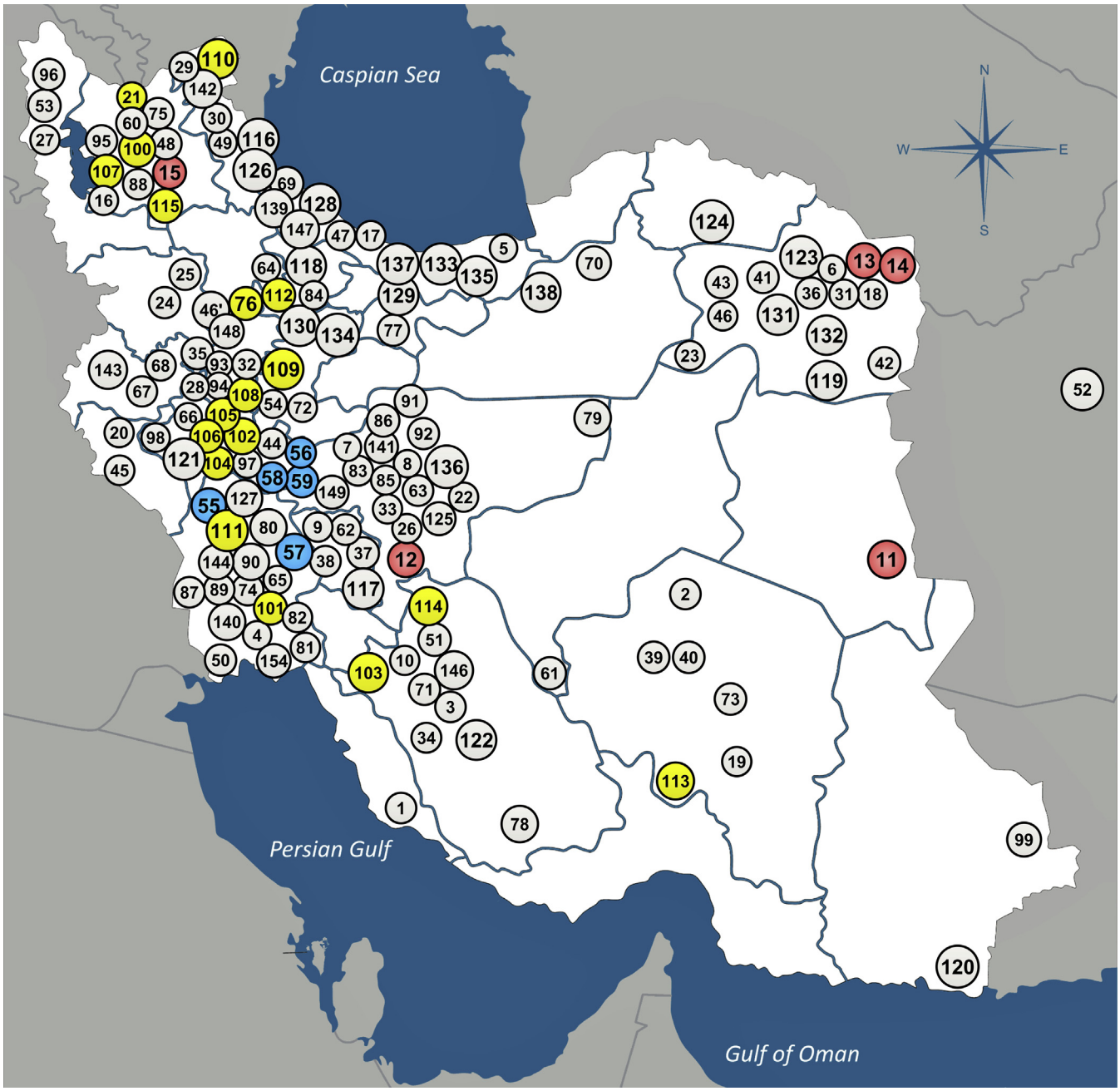
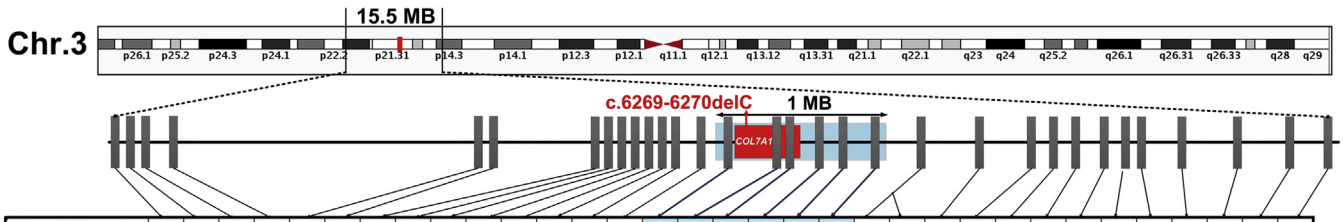


Figure 4. The geographic location of families in Iran with *COL7A1* mutations identified in this study. The families harboring the recurrent mutations c.6269_6270delC, c.4233delT, and c.553C>T are identified by yellow, blue, and red colors, respectively.

suggesting that the mechanism of this mutation may involve slippage of replication due to mispairing in the tandem five cytosines in the sequence, which can occur independently in different families. A common conserved haplotype was observed in all five Lurish families harboring the c.4233delT mutation, suggesting that this mutation also originated from a common ancestor. In contrast, two different haplotypes were observed in the five families tested with the p.Arg185X mutation. Four families had a common conserved haplotype, but one family had a different haplotype, suggesting that the p.Arg185X mutation is a hotspot event in these families (data not shown). Supporting this conclusion is the fact that this mutation reflects a cytosine-to-thymine transition, presumably as a result of methylation of cytosine to 5-methylcytosine, a common mechanism for human mutations (Youssoufian et al., 1988). In fact, this mutation mechanism has been previously reported in *COL7A1* in patients with DEB (Hovnanian et al., 1994).

DISCUSSION

This study has examined the mutation database in an exceptionally large cohort of patients with DEB, a rare Mendelian disorder inherited either in an autosomal dominant or autosomal recessive pattern. This multi-ethnic cohort consisted of 238 patients in 152 extended Iranian families with a high degree of customary consanguineous marriages.



SNP Location (Chr.3)		40.08M	40.3	40.45	41.2	45.79	45.81	47.49	47.52	47.56	47.61	47.82	47.84	48.03	48.34	48.46	48.5	48.5	48.68	48.97	49.41	49.422	49.422	50.6	51.21	51.53	52.18	52.52	52.8	52.97	53.43	54.23	55.03	55.54
Family No.	Ethnicity	rs9838815	rs2278928	rs6801859	rs1798802	rs2251109	rs2531748	rs13059519	rs2101247	rs12489865	rs1014228	rs3772399	rs17290490	rs5012971	rs12486944	rs11797	rs1264194	rs2228561	rs4858798	rs4974078	rs6784820	rs1464567	rs1464566	rs175634	rs4619816	rs6773188	rs614288	rs13303	rs4481150	rs9846976	rs516362	rs9878916	rs3773553	rs1167245
101	Turks	C	C	G	G	T	A	C	A	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	T	C
110	Lurish	C	C	G	A	T	A	C	A	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	T	T/C
115	Turks	C/T	C/T	G/T	A	T	A/G	C	A	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	T	C	T	T	T	T	A	C	T	T
103	Turks	C	C	G	A	C	A	C	A	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	G	C	C	T/C
21	Turks	T	C	G	G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	T	T/C
109	Turks	C	C	G	G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	C	C
107	Turks	C	T	T	G	C	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	T	C
111	Lurish	T	T	T	G	C	G	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	C	C
112	Turks	C/T	C/T	T	A	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	T	C	T	T	T	T	A	C	C	C
108	Lurish	T	C	G	A/G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	C	C	C	A	C	C	C
102	Lurish	C	C/T	G/T	G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	C	C	C	A	C	C	T/C
104	Lurish	C/T	C	G	G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	C	C	C	A	C	C	C
105	Lurish	C	C	G	G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	T	C	C	C	A	TC	TC	TC
114	Persian	C/T	T	T	A/G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	G	C	T	G	C	C	C	C	C	G	C	T	T
100	Turks	T	T	T	G	C	A	T	G	T	C	C	C	C	A	C	C	A	A	T	G	G	C	T	G	C	T	T	T	T	A	C	C	T
106	Lurish	T	C/T	G/T	G	T	A	T	G	A	T	G	T	T	G	C	C	A	A	T	G	C	C	T	G	C	T	C/T	C	T	A	C	C	C
113	Persian	C/T	C	G	G	C	A	C	A	A	T	G	T	T	G	C	T	G	G	C	A	C	T	T	G	C	T	C	T	T	G	T	T	C

Figure 5. Haplotype analysis of 17 families with SNP marker array on chromosomal locus 3p21.31 on and surrounding the *COL7A1* locus. Sequences within a 1 Mb haplotype block indicated conservation consistent with founder effect in patients primarily of Lurish and Turkish ethnicity. One patient with Persian background (Family 113) had a different haplotype. A, adenine; C, cytosine; Chr., chromosome; G, guanine; MB, mega-base pair; SNP, single-nucleotide polymorphism; T, thymine.

Considering the prevalence of EB, we have estimated that there are about 200 identified families with DEB in Iran. Thus, this study consisting of 152 families represents most of the total DEB patient population. Mutation detection identified 104 distinct mutations in *COL7A1*, most these mutations being previously unreported to our knowledge. Thus, this study significantly expands the mutation database in this disease. The mutation detection consisted of two complementary strategies allowing comparison of their efficiency and costs. First, bidirectional Sanger sequencing of PCR amplified exons and flanking intronic sequences was applied to a cohort of 73 DEB patients, and 69 of these patients (94.9%) were found to harbor mutations in *COL7A1*. Subsequently, we developed a *COL7A1*-targeted NGS array, which allowed parallel sequencing of 98 patients. Using this approach, *COL7A1* mutations were found in 95 of these patients (96.9%), and all these mutations were verified by bidirectional Sanger sequencing. Thus, both techniques were equally efficient in identifying mutations in these patients, with the overall mutation detection rate being 98%. Considering the costs of mutation detection by these two approaches, it is clear that the targeted gene array approach is more cost and time effective. Specifically, the cost of PCR amplification of 118 exons and flanking intronic sequences, followed by Sanger sequencing, in our laboratory is approximately $1,000 per sample (PCR material, labor, and sequencing costs). In contrast, application of the *COL7A1*-targeted NGS array, when applied to batches of up to 96 samples, yields results with comparable efficiency at a cost of approximately $250 per sample. Several recent studies have also advocated the use of whole-exome sequencing (WES) for identification of mutations in heritable skin diseases (South et al., 2015; Takeichi et al., 2013). Although the cost of WES has recently decreased, and continues to do so, the current cost for mutation analysis by this technique in patients with EB is several-fold higher compared with the gene-targeted NGS approach (Rehm, 2013). WES bioinformatics analysis of the results requires significantly more time and effort than interrogation of a single gene. The complexity of the WES procedures may explain, at least in part, the observation that in one study, only approximately 25% of putative mutations in heritable diseases could be disclosed by this approach, the number being somewhat higher, approximately 35%, in autosomal recessive disorders (Yang et al., 2013, 2014). Thus, a *COL7A1*-targeted array approach is an expedient and cost-effective way of finding mutations in patients in whom the diagnosis has been ascertained on a clinical and/or immunoepitope mapping basis. At the same time, in cases of EB patients with new phenotypic presentations with no obvious candidate genes, or those DEB patients in whom our approaches failed to find *COL7A1* mutations, extended EB gene-targeted array or WES could be successful in identifying mutations at least in some of these patients. In our study, the gene-targeted approach

did not identify *COL7A1* mutations in three families despite a clear-cut clinical diagnosis of RDEB and homozygosity mapping focusing on *COL7A1* genomic region. Failure to detect mutations in these three families may reflect the presence of mutations in the genetic areas not covered by the NGS panel, including far upstream regulatory elements or deep intronic sequences.

Iran is a multi-ethnic society with a number of ancestral ethnicities. Occurrence of DEB was found in all major ethnic groups, in most cases proportional to their presence in the overall population of Iran. An exception to this was the Lurish population, which represents approximately 6% of individuals within Iranian national borders. However, 22% of patients with DEB were from Lurish ancestral background, a finding that may be explained, at least in part, by the recurrence of certain mutations in this population. Specifically, a distinct mutation, c.6269_6270delC was found in 18 independent families concentrated in the western part of the country. Haplotype analysis of these families indicated conservation of a homozygosity block surrounding the mutation, suggesting a founder effect. This mutation, as well as another recurrent mutation, c.4233delT, was particularly prevalent in Lurish ethnicity.

Consanguineous marriage, defined as an intrafamilial marriage between people who are second cousins or closer, is customary in different geographic regions, particularly in the Middle East and North and Sub-Saharan Africa (Hamamy et al., 2011). Also, many migrant communities from these regions now reside in Western countries with similar customs. It is estimated that globally at least 20% of the human population (>1.1 billion individuals) live in communities with a preference for consanguineous marriage and that at least 8.5% of children have consanguineous parents (Modell and Darr, 2002). In addition to the degree of relationship between the parents, the increase in risk of having an affected child born to consanguineous parents depends on population frequency of the disease allele in inversely proportional order. The less common the disease allele causing a recessive disorder is in the gene pool, the greater the probability that the parents of an affected individual are consanguineous. EB is an example of a rare Mendelian disorder. As noted in our study, 92% of the patients with DEB had consanguineous parents (the incidence of all types of DEB in the US population = 6.5 per one million live births) (Pfendner et al., 2001). For more prevalent Mendelian disorders, the rate of parent consanguinity is less. For example, the consanguinity rate among the parents of autosomal recessive nonsyndromic deafness patients from Iran and Turkey is 73% and 70%, respectively (incidence = 280 per one million live births in the US population) (Bademci et al., 2016). Similarly, in an Iranian cohort of 515 patients affected by a heterogeneous group of primary immunodeficiency disorders, the average consanguineous marriage rate among parents of ataxia-telangiectasia patients was 81% (incidence = 10 per million live births in the US population) (Rezaei et al., 2006), whereas in Chediak-Higashi syndrome, 100% of the parents are of consanguineous union (about 200 cases of the conditions have been reported worldwide) (Rezaei et al., 2006). Another example of the effect of consanguinity in Iranian population is Kindler syndrome, a form of EB, for which all recently reported cases are a product of consanguineous unions (Youssefian et al., 2015), and EB simplex, which is usually an autosomal dominant disease but for which approximately 40% of Iranian patients are autosomal recessive (Vahidnezhad et al., 2016).

Identification of specific mutations in families with DEB, as well as other heritable disorders, has several implications for diagnostics and prognostication of the disease. In the case of DEB, most children are born with blistering and erosions or develop them within a few hours after birth. Considering the spectrum of EB and the phenotypic heterogeneity in the four principal subtypes of this group of disorders, initial subclassification is difficult, hampering the prognostication of the disease outcome. However, identification of specific mutations in the candidate genes allows prognostic considerations in general terms, which often dictate the types of approaches to be taken to manage the patients and counteract the complications of the disease (Hammersen et al., 2016; Mellerio et al., 2016; Tamai and Uitto, 2016). Furthermore, although EB is currently an intractable disorder, recent preclinical studies have opened a number of new avenues for potential treatment of these disorders, and some of these approaches have already entered early clinical trials (Uitto et al., 2016a). For many of these approaches, knowledge of the mutant genes and identification of the specific mutations are prerequisites for the application of allele-specific precision medicine (Uitto et al., in press). At the level of extended families at risk for recurrence of heritable blistering diseases, the precise knowledge of the mutations can be used for identification of heterozygous carriers and can form the basis for prenatal testing and preimplantation genetic diagnosis, coupled with appropriate genetic counseling (Pfendner et al., 2003). At the national and even global levels, an approach to reduce the risk of rare autosomal recessive disorders is to discourage consanguineous marriages. Although this approach has been promoted in the Middle East (Modell and Darr, 2002), there is recognition of the fact that consanguineous kinship patterns are, in many of these countries, integral to the structure of the society, and appropriate approaches for genetic counseling must work within such societal conventions.

MATERIALS AND METHODS

Patient data

This study was approved by the institutional review board of the Pasteur Institute of Iran, and all subjects and parents of underaged patients gave written informed consent to participate in research and gave their permission to publish their images. In this study, 152 extended families, including 238 DEB patients, were included for *COL7A1* mutational analysis from a large cohort of Iranian EB patients. Criteria for inclusion were clinical presentation and immunoepitope mapping. In consanguineous families, when clinical presentation or immunoepitope mapping were not informative, we used SNP-based whole-genome microarrays and short tandem repeat marker-based homozygosity mapping for detection of homozygosity of the 3p21.31 genomic region, which harbors *COL7A1*. The details of immunoepitope mapping, short tandem repeat homozygosity mapping, as well as PCR and Sanger sequencing, are provided in the Supplementary Materials online.

SNP microarrays

Whole-genome SNP-based homozygosity mapping was performed by Illumina Infinum Exome-24 Bead Chip (240,000 SNP markers) using 200 ng of genomic DNA as described by the manufacturer (Illumina, San Diego, CA). B-allele frequency and log2R ratio were analyzed with Illumina Genome Studio software (Illumina). The software identified runs of homozygosity based on the presence of homozygosity in the B-allele frequency but no change in the log2R ratio, thus excluding regions of hemizygosity due to a deletion. Genotyping errors and run of homozygosity interrupted by homozygous deletions were manually adjusted. The SNP data were analyzed by using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/ibdibs.shtml#homo), and runs of homozygosity of less than 2 Mb were eliminated. Superimposing 21 skin fragility related genes with homozygosity blocks identified potential candidate genes for mutational analysis.

COL7A1 custom panel design, library preparation, and NGS

Targeted sequencing of *COL7A1* was performed on the Ion Torrent PGM platform, using an AmpliSeq 3-pool primer design intended to capture all *COL7A1* exons plus 100 base pairs of flanking intronic sequences. (For details of NGS using custom panel-targeted sequencing on the Ion Torrent PGM platform, see Cao et al. [2014] and Millat et al. [2014a, 2014b]). The panel consisted of three separate PCR primer pools, which covered 99.9% of the 25,114 target bases and produced a total of 187 amplicons. An aliquot of each DNA sample (20 ng) was used for library construction by an Ion AmpliSeq Library Kit v2.0 (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. Samples were distinguished by unique barcodes (Ion Xpress Barcode Adapters Kit; Life Technologies) and then pooled in equimolar concentrations. Genomic DNA from 14 samples were loaded on an Ion 318 Chip (Thermo Fisher Scientific, Waltham, MA). Alignment to hg19 genome and germline variant calling were performed using the Torrent Suite software (Applied Biosystem, Foster City, CA) with variant caller plugin, version 4.4. Validation studies were performed by bidirectional Sanger sequencing for the pathogenic variants. Genome annotations (hg19 coordinates for *COL7A1*: Chromosome 3: 48601506-48632593) and predicted variant effects were appended to variant results using Annovar software (http://annovar.openbioinformatics.org). In addition, SIFT (sift.bii.a-star.edu.sg) and Polyphen2 (genetics.bwh.harvard.edu/pph2) were used to estimate evolutionary conservation and the effects of the amino acid substitutions on the structure and function of the protein (see Supplementary Table S3). Sequence variants were compared with The Database of Short Genetic Variation, Build 147 (www.ncbi.nlm.nih.gov/projects/SNP), 1000 Genomes (www.1000genomes.org), and mutations reported in the current literature (www.col7.info and www.ncbi.nlm.nih.gov/pubmed/). Mutations not reported previously were screened by our in-house ethnic-specific control panel to rule out the presence of these SNPs in the normal population. Mutations were annotated using nomenclature recommended by the Human Genome Variation Society (www.hgvs.org/mutnomen), and the variants were reported in reference to NM_000094 (cDNA) and NP_000085 (protein).

Haplotype analysis

We used HapMap data to choose tag SNPs, representative SNPs with high linkage disequilibrium that represent the haplotype, spanning a region of about 15.5 Mb encompassing locations upstream and downstream of *COL7A1* (http://snpinfo.niehs.nih.gov/snpinfo/snptag.htm) and (https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/). We also selected two common intragenic polymorphisms as tagSNPs, exon 14: c.1784C>T, p.Pro595Lys (rs2228561) and exon 21: c.2844A>G, p.Pro939Pro (rs1264194), which were previously detected by sequencing in the course of the mutational screening of subjects and also were observed in the Iranian control population (Minor allele frequencies: 0.10 (adenine) and 0.32 (thymine), respectively). Altogether, 33 SNPs spanning 15.5 Mb of the *COL7A1* region were genotyped by PCR and bidirectionally sequenced using an automated sequencer (3730; Applied Biosystems, Foster City, CA).

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

The authors thank the patients and their families for participation. Mina Tabrizi, Cecilia Kim, Maryam Daneshpazhooh, Seyyed Mohammad Ghahestani, Anna Isaian, Mohammad Hamid, Hamid Galehdari, and Javad Mohammadi-asl referred patients to this study. Mohsen Siavashi, Mona Tehrani, Mojgan Jalilipour, Sara Afsharaalam, Fatemeh Vand-Rajabpour, Romina Dastmalchi, Masoomeh Torabi, Fatemeh Golnabi, Ali Davoudian (Rare Disease Foundation of Iran), Sara Norouzzadeh, Niloufar Amiri, Zohreh Sharifie, Azam Ahmadi, Hamideh Bagherian, and Malihe Shaiganynahad assisted in the collection and processing of samples and the clinical data. Theodore Schurr assisted in interpretation of the haplotype data, Kevin McCormick contributed to NGS data analysis, and Hakon Hakonarson assisted in homozygosity mapping. Carol Kelly assisted in manuscript preparation. This work was supported by DEBRA International (JU) and by an Institutional grant from the Sidney Kimmel Cancer Center of Thomas Jefferson University (PF, AE). This study is in partial fulfillment of the PhD thesis of HV.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at http://dx.doi.org/10.1016/j.jid.2016.10.023.

REFERENCES

Bademci G, Foster 2nd J, Mahdieh N, Bonyadi M, Duman D, Cengiz FB, et al. Comprehensive analysis via exome sequencing uncovers genetic etiology in autosomal recessive nonsyndromic deafness in a large multiethnic cohort. Genet Med 2016;18:364–71.

Brittingham R, Uitto J, Fertala A. High-affinity binding of the NC1 domain of collagen VII to laminin 5 and collagen IV. Biochem Biophys Res Commun 2006;343:692–9.

Cao YY, Qu YJ, Song F, Zhang T, Bai JL, Jin YW, et al. Fast clinical molecular diagnosis of hyperphenylalaninemia using next-generation sequencing-based on a custom AmpliSeq panel and Ion Torrent PGM sequencing. Mol Genet Metab 2014;113:261–6.

Chen M, Marinkovich MP, Jones JC, O'Toole EA, Li YY, Woodley DT. NC1 domain of type VII collagen binds to the beta3 chain of laminin 5 via a unique subdomain within the fibronectin-like repeats. J Invest Dermatol 1999;112:177–83.

Christiano AM, McGrath JA, Tan KC, Uitto J. Glycine substitutions in the triple-helical region of type VII collagen result in a spectrum of dystrophic epidermolysis bullosa phenotypes and patterns of inheritance. Am J Hum Genet 1996;58:671–81.

Chung HJ, Uitto J. Type VII collagen: the anchoring fibril protein at fault in dystrophic epidermolysis bullosa. Dermatol Clin 2010;28:93–105.

Fine JD, Bruckner-Tuderman L, Eady RA, Bauer EA, Bauer JW, Has C, et al. Inherited epidermolysis bullosa: Updated recommendations on diagnosis and classification. J Am Acad Dermatol 2014;70:1103–26.

Fine JD, Mellerio JE. Extracutaneous manifestations and complications of inherited epidermolysis bullosa: part I. Epithelial associated tissues. J Am Acad Dermatol 2009;61:367–84.

Hamamy H, Antonarakis SE, Cavalli-Sforza LL, Temtamy S, Romeo G, Kate LP, et al. Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. Genet Med 2011;13:841–7.

Hammersen J, Has C, Naumann-Bartsch N, Stachel D, Kiritsi D, Soder S, et al. Genotype, clinical course, and therapeutic decison making in 76 infants with severe generalized junctional epidermolysis bullosa. J Invest Dermatol 2016;136(11):2150–7.

Hovnanian A, Hilal L, Blanchet-Bardon C, de Prost Y, Christiano AM, Uitto J, et al. Recurrent nonsense mutations within the type VII collagen gene in patients with severe recessive dystrophic epidermolysis bullosa. Am J Hum Genet 1994;55:289–96.

Mellerio JE, Robertson SJ, Bernardis C, Diem A, Fine JD, George R, et al. Management of cutaneous squamous cell carcinoma in patients with epidermolysis bullosa: best clinical practice guidelines. Br J Dermatol 2016;174:56–67.

Millat G, Chanavat V, Rousson R. Evaluation of a new high-throughput next-generation sequencing method based on a custom AmpliSeq library and Ion Torrent PGM sequencing for the rapid detection of genetic variations in long QT syndrome. Mol Diagn Ther 2014a;18:533–9.

Millat G, Chanavat V, Rousson R. Evaluation of a new NGS method based on a custom AmpliSeq library and Ion Torrent PGM sequencing for the fast detection of genetic variations in cardiomyopathies. Clin Chim Acta 2014b;433:266–71.

Modell B, Darr A. Science and society: genetic counselling and customary consanguineous marriage. Nat Rev Genet 2002;3:225–9.

Montaudie H, Chiaverini C, Sbidian E, Charlesworth A, Lacour JP. Inherited epidermolysis bullosa and squamous cell carcinoma: a systematic review of 117 cases. Orphanet J Rare Dis 2016;11:117.

Pfendner E, Uitto J, Fine JD. Epidermolysis bullosa carrier frequencies in the US population. J Invest Dermatol 2001;116:483–4.

Pfendner EG, Nakano A, Pulkkinen L, Christiano AM, Uitto J. Prenatal diagnosis for epidermolysis bullosa: a study of 144 consecutive pregnancies at risk. Prenat Diagn 2003;23:447–56.

Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. Nat Rev Genet 2013;14:295–300.

Rezaei N, Pourpak Z, Aghamohammadi A, Farhoudi A, Movahedi M, Gharagozlou M, et al. Consanguinity in primary immunodeficiency disorders; the report from Iranian Primary Immunodeficiency Registry. Am J Reprod Immunol 2006;56:145–51.

Saadat M, Ansari-Lari M, Farhud DD. Consanguineous marriage in Iran. Ann Hum Biol 2004;31:263–9.

South AP, Li Q, Uitto J. Next-generation sequencing for mutation detection in heritable skin diseases: the paradigm of pseudoxanthoma elasticum. J Invest Dermatol 2015;135:937–40.

Takeichi T, Nanda A, Liu L, Salam A, Campbell P, Fong K, et al. Impact of next generation sequencing on diagnostics in a genetic skin disease clinic. Exp Dermatol 2013;22:825–31.

Tamai K, Uitto J. Stem cell therapy for epidermolysis bullosa—does it work? J Invest Dermatol 2016;136(11):2119–21.

Uitto J, Atanasova VS, Jiang Q, South AP. Precision medicine for heritable skin diseases—the paradigm of epidermolysis bullosa. J Invest Dermatol, in press.

Uitto J, Bruckner-Tuderman L, Christiano AM, McGrath JA, Has C, South AP, et al. Progress toward treatment and cure of epidermolysis bullosa: Summary of the DEBRA International Research Symposium EB2015. J Invest Dermatol 2016a;136:352–8.

Uitto J, Christiano AM, McLean WHI, McGrath JA. Novel molecular therapies for heritable skin disorders. J Invest Dermatol 2012;132:820–8.

Uitto J, Has C, Vahidnezhad H, Youssefian L, Bruckner-Tuderman L. Molecular pathology of the basement membrane zone in heritable blistering diseases: The paradigm of epidermolysis bullosa [e-pub ahead of print]. Matrix Biol 2016b; http://dx.doi.org/10.1016/j.matbio.2016.07.009.

Vahidnezhad H, Youssefian L, Saeidian AH, Mozafari N, Barzegar MR, Sotoudeh S, et al. KRT5 and KRT14 mutations in epidermolysis bullosa simplex with phenotypic heterogeneity, and evidence of semi-dominant inheritance in a multiplex family. J Invest Dermatol 2016;136:1897–901.

Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 2013;369:1502–11.

Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, et al. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA 2014;312:1870–9.

Youssoufian H, Antonarakis SE, Bell W, Griffin AM, Kazazian Jr HH. Nonsense and missense mutations in hemophilia A: estimate of the relative mutation rate at CG dinucleotides. Am J Hum Genet 1988;42:718–25.

Youssefian L, Vahidnezhad H, Barzegar M, Li Q, Sotoudeh S, Yazdanfar A, et al. The Kindler syndrome: a spectrum of FERMT1 mutations in Iranian families. J Invest Dermatol 2015;135:1447–50.