**INVESTIGATIVE REPORT**

**Novel and Recurrent PNPLA1 Mutations in Spanish Patients with Autosomal Recessive Congenital Ichthyosis; Evidence of a Founder Effect**

Uxia ESPERÓN-MOLDES1,2, Manuel GINAŘE VAL3, Laura RODRÍGUEZ-PAZOS4, Laura FACHAL1, José Manuel AZAÑA5, María Barberá FONS5, Monica VIEJO DIAZ1 and Ana VEGA1
1Galician Public Foundation of Genomic Medicine (FPGMX)-SERGAS, Genomics Medicine Group-USC, CIBERER, IDIS, 2Department of Forensic Sciences, Pathological Anatomy, Gynecology, Obstetrics and Pediatrics, USC, 3Dermatology Service of University Clinical Hospital of Santiago, Santiago de Compostela, 4Dermatology Service of University Clinical Hospital of Vigo, Vigo, 5Clinical Genetics Service of Central University Hospital of Asturias, Oviedo, Spain

Autosomal recessive congenital ichthyosis is a group of rare non-syndrome diseases that affect cornification. PNPLA1 is one of the 12 related genes identified so far. Mutation screening of this gene has resulted in the identification of 13 individuals, from 10 families, who carried 7 different PNPLA1 mutations. These mutations included 2 missense, 2 frameshift and 3 nonsense, 3 of them being novel. One of the identified variants, c.417_418delinsTC, was highly prevalent, as it was found in 6 out of 10 (60%) of our autosomal recessive congenital ichthyosis families with PNPLA1 mutations. Clinical manifestations varied significantly among patients, but altered sweating; erythema, palmar hyperlinearity and small whitish scales in flexor-extensor and facial areas were common symptoms. Haplotype analyses of c.417_418delinsTC carriers confirmed the existence of a common ancestor. This study expands the spectrum of the PNPLA1 disease, which causes variants and demonstrates that the c.417_418delinsTC mutation has founder effects in the Spanish population.

**Key words:** ARCI; Spanish population; PNPLA1; founder effects; c.417_418delinsTC.

Accepted May 22, 2019; E-published May 23, 2019

Acta Derm Venereol

Corr: PhD Ana Vega, Molecular Medicine Service, Fundación Pública Gallega de Medicina Xenómica, University Hospital Santiago de Compostela, floor -2, Santiago de Compostela, Spain. E-mail: ana.vega@usc.es

Novel and Recurrent PNPLA1 Mutations in Spanish Patients with Autosomal Recessive Congenital Ichthyosis; Evidence of a Founder Effect

**SIGNIFICANCE**

Due to the growing importance of PNPLA1 mutations in the development of autosomal recessive congenital ichthyosis (ARCI), we decided to sequence this gene in a large, well-characterized cohort of Spanish ARCI patients. The mutation analysis revealed 7 different PNPLA1 mutations, 3 of which were novel, in 13 individuals from 10 families. Interestingly, one of the identified mutations was present in 60% of the families (c.417_418delinsTC; p.Ser140Pro). Haplotype analysis of the c.417_418delinsTC carriers demonstrated that these families are descendants of a recent founder who lived in the XI century, implying that this recurrent mutation has founder effects in the Spanish population.

**ACTA Derm Venereol 2019; 99: XX–XX**

**ABCA12, ALOXE3, CYP4F22, NIPAL4, TGM1, LIPN** and more recently CERS3, PNPLA1, CASP14, SDR9C7 and SULT2B1 (1–3).

**ALOX12B, PNPLA1, PNPL1, SULF2**.

Numerous hereditary cornification disorders that are clinically and aetiologically heterogeneous and follow a Mendelian inheritance pattern are grouped under the term of inherited ichthyosis. Autosomal recessive congenital ichthyosis (ARCI) are a subgroup of non-syndrome ichthyosis, whose phenotypic spectrum ranges from harlequin ichthyosis (HI; OMIM 242500), to less severe phenotypes such as congenital ichthyosis-form erythroderma (CIE; OMIM 242100) or lamellar ichthyosis (LI; OMIM 242300) the last two being the most common phenotypes. To date, 12 genes have been implicated in the development of ARCI: ABCA12,
METHODS

Patients

A cohort of 91 Spanish patients (81 different families) with clinical suspicion of ARCI was studied. A dermatologist clinically characterized all affected individuals. Clinical and genealogical information was collected for each family. Pedigrees and available clinical photos of the families are presented in Fig. 1 and Fig. S1. This study, carried out at FPGMX, was approved by the Galician Ethical Committee for Clinical Research (Code 2013/056) and the procedures followed were in accordance with the Declaration of Helsinki. All participants provided written informed consent.

Mutation analysis

The PNPLA1 gene was examined by Sanger sequencing or targeted resequencing on SOLiD 5500xl or Ion Proton Platforms (Thermo Fisher Scientific; San Jose, CA, USA). Identified variants were confirmed in the proband and close relatives. The potential pathogenicity of each variant was assessed by using the Alamut® Visual 2.8.1 software (Interactive Biosoftware, Rouen, France).

Conserved Residue Analysis and Structural study of the c.417_418delinsTC mutation

The model published by Wilson et al. (13) was employed to represent the 3D human PNPLA patatin-like domain with the SWISS-MODEL tool (http://swissmodel.expasy.org). The effect of the mutation in the protein structure was evaluated with Swiss-PdbViewer (http://spdbv.vital-it.ch). The degree of conservation of the Ser140 was evaluated using the Clustal Omega tool (http://www.ebi.ac.uk/Tools/msa/clustalo).

Haplotype study

Haplotypes of c.417_418delinsTC mutation carriers, and 80 Spanish healthy controls were reconstructed by genotyping 8 microsatellite markers. These markers spanned a segment of 10Mb flanking the PNPLA1 gene locus. The age of the mutation was calculated with DMLE+ (http://www.dmle.org). The TMRCA was estimated by applying different linkage disequilibrium algorithms. Further details are available in Appendix S1.

RESULTS

Ten families, 8 new and 2 previously reported (5, 14, 15), harboured 7 different PNPLA1 mutations. These families encompassed 8 patients with LI, 4 patients with CIE and one individual without any available phenotypic information. Patient’s clinical and genotypic data is presented in Table I. The 7 mutations included 2 missense, 2 frameshift and 3 nonsense mutations. Three of the 7 mutations identified were novel: c.282dup; p.(Lys95*), c.729C>G; p.(Tyr243*) and c.892C>T; p.(Arg298*). The 4 known mutations comprised of 2 frameshift [c.1143del; p.(Pro382Alafs*74) and c.820del;p.(Arg274Glyfs*8)] previously described by Zimmer et al. (4) and Pichery et al. (6), respectively, and 2 missense [c.100G>A; p.(Ala34Thr) and c.417_418delinsTC; p.(Ser140Pro)] previously reported in a Spanish (13) and French family (4). Characteristics of the mutations found in the present study are described in detail in Table S1. The most prevalent mutation, c.417_418delinsTC, was found in approximately half of the families (60%). Analysis of the protein structure showed that the Ser125 residue (structurally equivalent to Ser140 in PNPLA1) is present in a loop region, formed by the highly conserved residues (Fig. 2), indicating that the local environment of Ser140 has probable structural consequences.

Clinical manifestations varied significantly among patients, but altered sweating, erythema, palmar hyperlinearity and small whitish scales that affected flexor and extensor surfaces and facial areas were common symptoms. However, prematurity and dark scales were infrequent. Ectropion, colloidion at birth, palmoplantar keratoderma, alopecia and big scales were occasionally seen. Seven out of 12 patients were taking topical reti-
Novel and recurrent PNPLA1 mutations in Spanish patients with ARCI

A clear correlation between the location/type of mutation and the resulting phenotype was not found. According to the data in Table I, the combination of nonsense and missense mutations led to LI phenotypes (Families 46, 98, 115), as well as frameshift and missense or homozygous frameshift variants (Families 62 and 137). The only carrier of a homozygous nonsense mutation showed CIE (Family 69). Patients with compound missense mutations had LI or CIE phenotypes indistinctly (Families 18, 23, 47, 119).

Reconstructed haplotypes of c.417_418delinsTC carriers are presented in Table SII. All carriers shared the same allele (7) for marker D6S439. A 4.1 Mb prevalent core haplotype (5–7–5) ranging from markers D6S497 to D6S1548 was identified. This core haplotype was found on 4 out of 9 carrier chromosomes (44%) versus 18 out of 160 (11%) found on control chromosomes. To date the TMRCA of these 6 families, we used different algorithms, which resulted in different estimates. Averaging all these estimates would place the TMRCA approximately 38 generations ago (95% CI 29-46). However, the age of the c.417_418delinsTC, ranged from 103 to 108 generations, dating the first appearance of the mutation to approximately 2,575–2,700 years ago (assuming 25 years per generation). TMRCA and mutation age estimates are shown in Table SII.

**DISCUSSION**

According to the localization of the reported PNPLA1 mutations, Zimmer et al. (4) proposed a highly conserved, extended patatin domain that ranges from amino acid positions 1 to 288, increasing the size of the former functional region (Ile16-Thr185).

Considering Zimmer’s proposal, 2 of the 3 novel mutations described in this study, c.282dup, p.(Lys95*) and c.729C>G, p.(Tyr243*) are located...
within this extended domain in exons 2 and 5, respectively. These variants create truncated mRNAs that would most likely be degraded by the process of nonsense-mediated decay. Our third novel nonsense mutation in exon 6, c.892C>T, p.(Arg298*), is the fifth variant reported, which is situated outside of the extended patatin domain (4, 5, 11). PNPLA1 protein contains a proline-rich domain (residues 326–451) in the C-terminal, which is conserved across species. Interestingly, all mutations located outside the extended patatin domain p.(Glu313Aspfs*49), p.(Pro370*), p.(Ser382Alafs*74), p.(Ala434Hisfs*10) and our novel variant p.(Arg298*) create premature stop codons before or within the proline-rich domain, with the exception of p.(Ser382Alafs*74). A schematic representation of PNPLA1, including novel and previously reported mutations, is shown in Fig. 3.

The 3D model of the PNPLA patatin-like domain showed that the c.417_418delinsTC mutation, which involves the substitution of 2 nucleotides (GT>CT), alters a conserved serine residue located in a beta sheet hairpin loop (Fig. 2), which has an impact on the close hydrogen-bond interactions with other residues affecting the protein structure.

The c.417_418delinsTC mutation has been described only in French and Spanish descents (4, 5). Due to the geographical proximity of these 2 countries, it is plausible that French and Spanish families share a common ancestor. The mutation could have arisen in Spain approximately 2,625 years ago (VII century BC) and brought to France by a Spanish carrier or vice versa.

To date, PNPLA1 mutations have been reported in different populations, French, Scandinavian, Iranian, Spanish, etc. (4, 10, 12, 14). The mutation frequency varies among the different populations, in Scandinavian countries mutations in PNPLA1 account for less than 1% of all ARCI-related mutations (10), whereas in highly consanguineous populations, such as Iran, the frequency reaches 15% (12). In our cohort, mutations in this gene represent 14% (13 out of 91) of the total ARCI cases. We propose that this unexpectedly high frequency of the mutated PNPLA1 could be due to the founder effect of the recurrent mutation, c.417_418delinsTC.
In conclusion, this study reports 3 novel mutations expanding the spectrum of the \textit{PNPLA1} disease causing variants and posits the possible mechanism underlying the pathogenicity of the c.417_418delinsTC mutation. Furthermore, we demonstrate that this mutation has founder effects in the Spanish population. TMRCA estimations suggest that carrier families could have descended from a single recent founder who lived in Spain during the XI century, while the mutation arose about 1,700 years prior.

**ACKNOWLEDGMENTS**

The authors are grateful to families for their cooperation and to Jamie Allen for proofreading the manuscript and to Andrés Etxeita for the photos of the family 23. We also would like to thank the Spanish Association of Ichthyosis (ASIC) for their cooperation. This work was partially supported by Ramón Areces Foundation project (Rare Diseases 2013-056); by Spanish Instituto de Salud Carlos III (ISCIII) (INT15/00070, INT16/00154, INT17/00133) and by Xunta de Galicia (IN607B). UE was supported by a pre-doctoral fellowship from Xunta de Galicia. The authors have no conflicts of interest to declare.

**REFERENCES**

**SUPPLEMENTARY MATERIAL AND METHODS**

**Linkage disequilibrium analysis and mutation age estimation**

The following microsatellite markers were chosen: D6S2446, D6S497, D6S439, D6S1548, D6S943, D6S2427, D6S1607 and D6S400 for TMRCA and mutation age estimations. Forward PCR primers were labelled with either FAM or HEX fluorescent dyes (Sigma-Genosys Ltd. Cambridgeshire, UK). DNA was specifically amplified by PCR with the labelled primers using 5 PRIME MasterMix (5Prime, Eppendorf, Hamburg, Germany). The amplification products were separated by capillary electrophoresis on an ABI3730xl sequencer and analyzed with GeneMapper v4 Software (Applied Biosystems, Foster City, CA, USA). All reactions were performed according to the manufacturer’s protocols. The phase of the haplotype was inferred using PHASE software (http://stephenslab.uchicago.edu/phase), and then, these phased haplotypes were used to calculate the age of the mutation and the time to the most recent common ancestor (TMRCA).

The age of TMRCA was calculated by employing two types of Linkage Disequilibrium (LD) methods: (i) Three single marker methods proposed by Bergman et al. (S1), Lander & Botstein (S2) and Labuda et al. (S3), respectively, applying the correction proposed by Labuda et al. (S4) for a growth rate of 0.0748. These methods are based on the fact that the mutation will be in LD with nearby marker alleles at polymorphic loci, and decay of this LD over time (through recombination) provides information about the TMRCA age. (ii) The gamma method published by Gandolfi et al. (S5) based on the idea that the size of the common haplotype region of the chromosome (genetic length of ancestral haplotypes) can be related to the number of generations that link the patients to their most recent common ancestor.

Single marker estimates were summarized by the mean of the results across 7 out of 8 available markers (D6S439 was not considered because it was not informative for the calculations since the same allele was present in all patients).

The most frequent allele outside the common region was considered as the founder allele, hence the conserved haplotype. The physical distances were converted into centimorgans (cM) assuming a sex averaged recombination rate of 1.3 cM equal to 1 Mb, which was calculated by using the estimates of recombination rates across the studied region according to deCODE genetic map available in the UCSC database (S6). Haldane’s mapping function (S7) was used for translation of map distances into recombination frequencies. For further details see **Table I**.

The time at which the c.417 418delinsTC mutation first occurred was also estimated by using DMLE+ (http://www.dmlr.org). This program takes into consideration map distances between the markers and the mutation site, haplotype data from affected patients and controls, the proportion of population sampled (f=2.95E-04, f=2.08E-04, f=2.49E-04) and the population growth rate (r=0.0748).

There are few data on the epidemiology of ARCI in Spain due to the lack of official reference centers for the disease and complete data from an official patient registry; consequently, it is not possible to know the exact number of patients diagnosed with ARCI. Therefore, we considered 3 different values of n (total number of patients) to calculate the proportion of population sampled (f): 144 which is the actual number of ARCI cases identified in Spain by Hernández-Martín et al. (S8) in 2011, 294 which is the estimated number of patients they have estimated in this same study, and the mean of both, 219.

**SUPPLEMENTARY REFERENCES**


Supplementary material to article by U. Espéron-Moldes et al. "Novel and Recurrent PNPLA1 Mutations in Spanish Patients with Autosomal Recessive Congenital Ichthyosis; Evidence of a Founder Effect"

Fig. S1. Pedigrees of families with PNPLA1 mutations.
Table SI. Characteristics of the \textit{PNPLA1} mutations identified in our cohort of ARCI patients

<table>
<thead>
<tr>
<th>Family</th>
<th>Nucleotide change</th>
<th>Aminoacidic change</th>
<th>Location</th>
<th>Domain</th>
<th>Resultant change</th>
<th>Mutation type</th>
<th>In Silico Prediction</th>
<th>TMRCA</th>
<th>MUTATION AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>18, 62</td>
<td>c.100G&gt;A</td>
<td>p.(Ala34Thr)</td>
<td>Exon 1</td>
<td>Patatin</td>
<td>Moderately conserved residue</td>
<td>Missense</td>
<td>Disease causing</td>
<td>111 (III.1)</td>
<td>111 (III.1)</td>
</tr>
<tr>
<td>46, 115</td>
<td>c.282dup</td>
<td>p.(lys95*)</td>
<td>Exon 2</td>
<td>Patatin</td>
<td>Truncated mRNA</td>
<td>Nonsense</td>
<td>Disease causing</td>
<td>119 (V.10)</td>
<td>119 (V.10)</td>
</tr>
<tr>
<td>Various</td>
<td>c.417_418delinsTC</td>
<td>p.(Ser140Pro)</td>
<td>Exon 2</td>
<td>Patatin</td>
<td>Highly conserved residue</td>
<td>Missense</td>
<td>Disease causing</td>
<td>119 (V.10)</td>
<td>119 (V.10)</td>
</tr>
<tr>
<td>137</td>
<td>c.820del</td>
<td>p.(Arg274Glyfs*8)</td>
<td>Exon 6</td>
<td>Extended patatin</td>
<td>Truncated mRNA</td>
<td>Frameshift</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>c.892C&gt;T</td>
<td>p.(Arg298*)</td>
<td>Exon 6</td>
<td>Patatin</td>
<td>Truncated mRNA</td>
<td>Nonsense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>c.11143del</td>
<td>p.(Pro382Alafs*74)</td>
<td>Exon 6</td>
<td>Patatin</td>
<td>Truncated mRNA</td>
<td>Frameshift</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mutation nomenclature: the Human Genome Sequence Variation guideline was followed. Reference sequences PNPLA1 (NM_001145717, NP_001139189) were used for naming the nucleotide and protein variations respectively. Available Minor Allele Frequencies (MAF) of European Non-Finnish population were revised in the gnomAD database (http://gnomad.broadinstitute.org/). References can be consulted in the manuscript.

Table SII. Haplotypes of the c.417_418delinsTC mutation carrier patients, TMRCA and mutation age estimations

<table>
<thead>
<tr>
<th>MARKER</th>
<th>47 (III.1)</th>
<th>47 (III.1)</th>
<th>98 (III.1)</th>
<th>115 (III.1)</th>
<th>46 (III.1)</th>
<th>119 (V.10)</th>
<th>119 (V.10)</th>
<th>23 (IV.5)</th>
<th>23 (IV.5)</th>
<th>Estimators</th>
<th>Generations</th>
<th>Population sampled</th>
<th>Generations</th>
</tr>
</thead>
<tbody>
<tr>
<td>D6S2446</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>13</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>Bergman</td>
<td>43 (29–57)</td>
<td>103 (77–143)</td>
<td></td>
</tr>
<tr>
<td>D6S497</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>Lander</td>
<td>37 (16–57)</td>
<td>108 (82–147)</td>
<td></td>
</tr>
<tr>
<td>D6S439</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>Labuda</td>
<td>26 (17–35)</td>
<td>104 (80–143)</td>
<td></td>
</tr>
<tr>
<td>D6S1548</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>Gandolf</td>
<td>45 (28–72)</td>
<td>105 (103–108)</td>
<td></td>
</tr>
<tr>
<td>D6S943</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6S2427</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6S1607</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6S4006</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td>38 (29–46)</td>
<td></td>
</tr>
</tbody>
</table>

Shaded boxes indicate the common haplotype shared among patients ranging from marker D6S497 to D6S1548. The 95% confidence interval of the estimations are in parentheses. \( f \) = proportion of population sampled. TMRCA: time to the most recent common ancestor.