

# Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema

Aileen Sandilands<sup>1</sup>, Ana Terron-Kwiatkowski<sup>1</sup>, Peter R Hull<sup>1</sup>, Gráinne M O'Regan<sup>2</sup>, Timothy H Clayton<sup>2</sup>, Rosemarie M Watson<sup>2</sup>, Thomas Carrick<sup>3</sup>, Alan T Evans<sup>4</sup>, Haihui Liao<sup>1</sup>, Yiwei Zhao<sup>1</sup>, Linda E Campbell<sup>1</sup>, Matthias Schmuth<sup>5</sup>, Robert Gruber<sup>5</sup>, Andreas R Janecke<sup>6</sup>, Peter M Elias<sup>7</sup>, Maurice A M van Steensel<sup>8</sup>, Ivo Nagtzaam<sup>8</sup>, Michel van Geel<sup>8</sup>, Peter M Steijlen<sup>8</sup>, Colin S Munro<sup>9</sup>, Daniel G Bradley<sup>10</sup>, Colin N A Palmer<sup>11</sup>, Frances J D Smith<sup>1</sup>, W H Irwin McLean<sup>1,13</sup> & Alan D Irvine<sup>2,12,13</sup>

We recently reported two common filaggrin (*FLG*) null mutations that cause ichthyosis vulgaris<sup>1</sup> and predispose to eczema and secondary allergic diseases<sup>2</sup>. We show here that these common European mutations are ancestral variants carried on conserved haplotypes. To facilitate comprehensive analysis of other populations, we report a strategy for full sequencing of this large, highly repetitive gene, and we describe 15 variants, including seven that are prevalent. All the variants are either nonsense or frameshift mutations that, in representative cases, resulted in loss of filaggrin production in the epidermis. In an Irish case-control study, the five most common European mutations showed a strong association with moderate-to-severe childhood eczema ( $\chi^2$  test:  $P = 2.12 \times 10^{-51}$ ; Fisher's exact test: heterozygote odds ratio (OR) = 7.44 (95% confidence interval (c.i.) = 4.9–11.3), and homozygote OR = 151 (95% c.i. = 20–1,136)). We found three additional rare null mutations in this case series, suggesting that the genetic architecture of filaggrin-related atopic dermatitis consists of both prevalent and rare risk alleles.

Ichthyosis vulgaris (OMIM# 146700) is the most common disorder of keratinization, characterized by dry, scaly skin and often by features of atopic dermatitis (eczema) and related allergic phenotypes<sup>3</sup>. Recently, we have shown that ichthyosis vulgaris is a semidominant trait caused by null alleles of the *FLG* gene, which encodes filaggrin, a protein essential for epidermal barrier formation and hydration<sup>1,4,5</sup>. Notably, we established a strong association between two prevalent *FLG* null

alleles (R501X and 2282del4) and atopic dermatitis and between these null alleles and asthma occurring in the context of atopic dermatitis<sup>2</sup>. These results have been extensively replicated in several European populations<sup>6–12</sup>, as recently reviewed<sup>13</sup>. To date, no negative association study has been reported in European populations, supporting the view that *FLG* null alleles represent a very strong predisposing factor for atopic dermatitis and secondary allergic conditions<sup>13,14</sup>.

*FLG* is located within the epidermal differentiation complex (EDC) on 1q21.3, a gene cluster expressed late in epidermal differentiation<sup>15</sup>. Analysis of *FLG* is confounded by the large and highly repetitive exon 3, which also shows population size variation (12.7–14.7 kb)<sup>16</sup>. This exon encodes 10–12 full tandem repeats of the filaggrin protein that are almost 100% identical at the DNA sequence level, flanked by two partial repeats<sup>16</sup>. The hg18 release of the human genome sequence contains only the shortest, ten-repeat allele, which we here designate *FLG*<sup>N</sup> (RefSeq ID: NM\_002016). By long-range PCR, we isolated DNA fragments encompassing longer-size alleles and fully sequenced these by a unidirectional deletion strategy. This showed that two different 11-repeat alleles result from duplication of filaggrin repeat 8 (allele *FLG*<sup>8+</sup>, containing repeats 8.1 and 8.2) or repeat 10 (allele *FLG*<sup>10+</sup>, containing repeats 10.1 and 10.2). The 12-repeat allele (*FLG*<sup>8+10+</sup>) has both repeats 8 and 10 duplicated (Supplementary Fig. 1 online). From alignments of all known and newly identified repeats, we identified a relatively small number of bases that are absolutely specific to a given repeat and a second series of bases that are shared by only one other repeat (Supplementary Fig. 1). A further confounding factor is that some of these bases are polymorphic.

<sup>1</sup>Epithelial Genetics Group, Human Genetics Unit, Division of Pathology and Neuroscience, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. <sup>2</sup>Department of Paediatric Dermatology, Our Lady's Children's Hospital, Crumlin, Dublin 12, Ireland. <sup>3</sup>Department of Genetics and <sup>4</sup>Department of Pathology, Tayside University Hospitals NHS Trust, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. <sup>5</sup>Department of Dermatology, and <sup>6</sup>Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria. <sup>7</sup>Dermatology and Metabolism (Medicine) Services, Veterans Affairs Medical Center, San Francisco, California, USA. <sup>8</sup>Department of Dermatology, University Hospital Maastricht, Maastricht, The Netherlands. <sup>9</sup>Department of Dermatology, Southern General Hospital, Glasgow, UK. <sup>10</sup>Department of Genetics, Trinity College Dublin, Dublin 2, Ireland. <sup>11</sup>Population Pharmacogenetics Group, Biomedical Research Centre, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. <sup>12</sup>Department of Clinical Medicine, Trinity College Dublin, Dublin 2, Ireland. <sup>13</sup>These authors contributed equally to this work. Correspondence should be addressed to W.H.I.M. (w.h.i.mclean@dundee.ac.uk) or A.D.I. (irvine@tcd.ie).

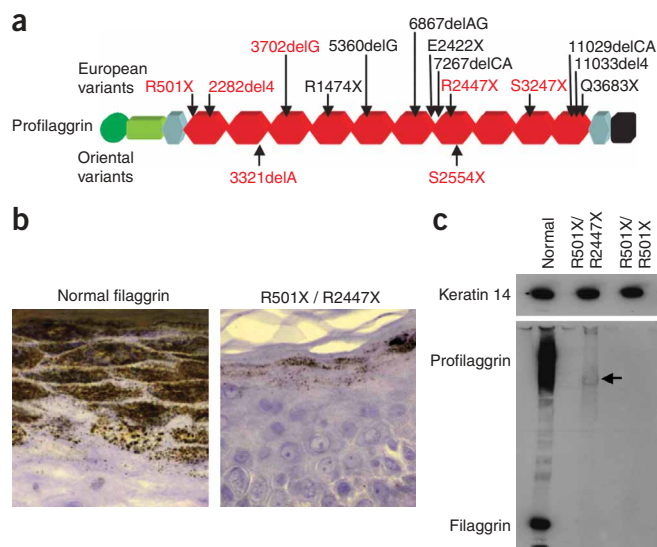


**Figure 1** Filaggrin variants and their biochemical consequences. (a) A schematic of the profilaggrin molecule (predicted from the ten-repeat allele *FLG*<sup>N</sup>) showing positions of known loss-of-function *FLG* mutations, all of which are either nonsense or out-of-frame deletion mutations. Variants shown in red are those known to be prevalent in at least some populations; those shown in black appear to very rare or family specific. The domain structure of the profilaggrin protein is as follows (from N terminus, left): green = S100 domain; light green = B domain; blue hexagons = partial filaggrin repeats; red hexagons = filaggrin repeats; black = unique C terminus. (b) Immunohistochemical staining of skin biopsy material for filaggrin (monoclonal antibody 15C10) shows abundant keratohyalin granules in normal epidermis, but in the skin of an R501X/R2447X compound heterozygote, there is appreciable but greatly reduced staining. Original magnification = 400×. (c) Immunoblot analysis shows that a greatly reduced amount of a truncated profilaggrin molecule (arrow) is expressed in the R501X/R2447X compound heterozygote compared with a normal control. We did not detect processed filaggrin, implying that the mutant proprotein cannot be processed normally. An R501X homozygote, lacking the 15C10 epitope, shows neither profilaggrin nor filaggrin staining. Keratin 14 staining of a duplicate blot was performed as a loading control.

Therefore, by testing multiple potential primer pairs, we developed a series of primer sets that reliably allows complete analysis of exon 3 in overlapping fragments (**Supplementary Fig. 2** online). Heterozygous SNPs within regions of overlap demonstrated that these primer sets showed minimal allele loss in routine mutation detection. The corresponding fragments can largely be sequenced using primers ending on specific bases, although cloning may be required for individuals heterozygous for the size variant alleles. The existence of further infrequent size variants remains a possibility.

Using this strategy, we analyzed 23 individuals originating from Ireland, Scotland, The Netherlands and Austria with severe ichthyosis vulgaris (that is, those individuals most likely to have a *FLG* mutation; **Supplementary Fig. 3** online). In addition to R501X and 2282del4, which were often found in the homozygous, compound heterozygous or heterozygous state in these individuals, and one uncommon mutation, 3702delG<sup>4</sup>, we identified seven additional nonsense or frameshift mutations, mostly in *trans* with R501X or 2282del4 (**Fig. 1**, **Supplementary Fig. 3** and **Supplementary Table 1** online). As part of a separate study, we identified two additional variants, 3321delA and S2554X, in Japanese individuals with atopic dermatitis and ichthyosis vulgaris. Both variants are prevalent in Japan<sup>17</sup>, and we found here that 3321delA is also prevalent in China (allele frequency = 0.018, *n* = 110). Of these 12 variants, we found five that were prevalent to varying degrees in 736 Irish population controls (combined minor allele frequency of 0.039). Limited SNP and microsatellite

analysis showed that the prevalent variants R501X, 2282del4, R2447X and S3247X each occur on a distinct haplotype in unrelated families from different European countries and probably represent ancient mutations (**Supplementary Table 2** online). We did not find the mutations identified in the Dutch and Austrian affected individuals in matched Dutch or Austrian control population sets or in Irish control populations, and the two Asian variants were absent from the Irish population. We found that all compound heterozygous mutations were present in *trans*, either by observing transmission or, in singletons, by cloning individual alleles<sup>1</sup>.

R501X and 2282del4, the originally identified *FLG* variants, occur in filaggrin repeat 1 and allow expression of the profilaggrin N terminus but not the filaggrin repeats<sup>1</sup>. The more 3' mutations identified here predict expression of up to ten filaggrin repeats, yet we observed ichthyosis phenotypes suggesting that these have similar or equal pathogenicity with respect to ichthyosis vulgaris as the repeat 1 mutations. Immunohistochemistry in compound heterozygotes for R501X with either R2447X in repeat 7 (**Fig. 1**) or 11033del4 in repeat 10 (**Supplementary Fig. 3**) demonstrated a greatly reduced epidermal granular layer. Ultrastructural analysis of an Austrian individual with R501X/R2447X showed greatly reduced keratohyalin granules; furthermore, there were abnormalities of the stratum corneum extracellular lipid matrix consistent with skin barrier impairment (**Supplementary Fig. 3**). Quantification by immunoblotting showed that although truncated profilaggrin species are made from 3' mutant alleles,

**Table 1** Case-control Pearson  $\chi^2$  analysis for individual and combined *FLG* variants

| Genotypes | R501X                         |                | 2282del4                      |                | R2447X                       |                | S3247X          |                | 3702delG        |                | Combined <i>FLG</i> null   |                |
|-----------|-------------------------------|----------------|-------------------------------|----------------|------------------------------|----------------|-----------------|----------------|-----------------|----------------|----------------------------|----------------|
|           | Population                    | AD case series | Population                    | AD case series | Population                   | AD case series | Population      | AD case series | Population      | AD case series | Population                 | AD case series |
| AA        | 717                           | 137            | 717                           | 152            | 734                          | 181            | 720             | 177            | 735             | 186            | 680                        | 103            |
| Aa        | 19                            | 51             | 19                            | 35             | 2                            | 7              | 16              | 11             | 1               | 2              | 55                         | 62             |
| aa        | 0                             | 0              | 0                             | 1              | 0                            | 0              | 0               | 0              | 0               | 0              | 1                          | 23             |
| Totals    | 736                           | 188            | 736                           | 188            | 736                          | 188            | 736             | 188            | 736             | 188            | 736                        | 188            |
|           | $(P = 7.8 \times 10^{-30})^a$ |                | $(P = 7.8 \times 10^{-17})^a$ |                | $(P = 1.7 \times 10^{-5})^a$ |                | $(P = 0.008)^a$ |                | $(P = 0.046)^a$ |                | $P = 2.12 \times 10^{-51}$ |                |

AD, atopic dermatitis.

<sup>a</sup>Although  $\chi^2$  analysis is not truly valid for tables with cells having counts of zero, as is the case here for the individual variants where the homozygote counts are mainly zero, we present this in parentheses for further information along with the raw genotype counts because it is common practice to present these regardless of cell size. However, the  $\chi^2$  analysis is fully valid for the combined genotypes. Note also that the minor allele label 'a' refers to the dominant mutant alleles throughout this analysis.

**Table 2 Fisher exact OR for the Irish atopic dermatitis case-control series**

| Dominant              | OR     | 95% c.i.       | P (Fisher exact test)  |
|-----------------------|--------|----------------|------------------------|
| R501X                 | 14.05  | 8.04–25.53     | $3.17 \times 10^{-23}$ |
| 2282del4              | 8.94   | 4.99–16.01     | $6.02 \times 10^{-14}$ |
| R2447X                | 14.19  | 2.92–68.90     | $3.25 \times 10^{-4}$  |
| S3247X                | 2.8    | 1.28–6.13      | 0.013                  |
| 3702delG              | 7.9    | 0.71–87.62     | 0.107                  |
| Combined dominant     | 10.02  | 6.75–14.89     | $5.36 \times 10^{-16}$ |
| Combined heterozygote | 7.44   | 4.9–11.3       | $2.1 \times 10^{-20}$  |
| Combined homozygote   | 151.85 | 20.29–1,136.46 | $9.56 \times 10^{-19}$ |

synthesis of mutant protein is greatly reduced, and processed filaggrin is undetectable (Fig. 1 and Supplementary Fig. 3). This is directly analogous to the *flaky tail* mouse mutant in which a truncating mutation in the region of repeat 6 is suggested and processed filaggrin is undetectable<sup>18</sup>. The two Asian mutations in repeats 2 and 7 also result in greatly reduced filaggrin expression<sup>17</sup>. Collectively, these data suggest that sequences close to the profilaggrin C terminus are necessary for processing; therefore, the more 3' mutations also prevent production of free filaggrin in the epidermis, although a trace of proprotein is still made, which may have some residual function.

We next investigated the expanded *FLG* mutation spectrum in atopic disease. We typed a case series of 188 Irish children with dermatologist-diagnosed moderate to severe atopic dermatitis (Supplementary Table 3 online) for all 12 variants. We found only five mutations that occurred in multiple affected individuals. Comparing the combined allele frequencies for the five recurrent mutations in the atopic dermatitis case series ( $n = 188$ ) versus the Irish control population ( $n = 736$ ) by  $\chi^2$  test gave  $P = 2.12 \times 10^{-51}$  (Table 1), with highly significant OR values (Table 2). The minor allele frequencies for the five *FLG* variants in affected individuals and controls are listed in Table 3. Complete sequencing of *FLG* in five selected individuals with atopic dermatitis with prominent features of ichthyosis vulgaris from this case series disclosed three additional variants (R1474X, 11029delCA and Q3683X), bringing the total number of known *FLG* variants to 15 (Fig. 1 and Supplementary Table 1). We identified these three variants only in single affected individuals, suggesting they are very rare. Including these rare variants, ~47% of individuals in the atopic dermatitis case series carry at least one filaggrin null allele.

Notably, the penetrance of the variants showed some variation; in particular, the repeat 9 variant S3247X seemed to be less penetrant (OR = 2.8) than R501X and 2282del4 (OR = 14.05 and 8.94, respectively; Table 2), despite similar minor allele population frequencies (Table 3). It should be noted that the unprocessed profilaggrin epitopes seen in granular cells of an R501X/R2447X compound heterozygote subsequently disappear in the stratum corneum (Fig. 1 and Supplementary Fig. 3), suggesting that a further degradation pathway exists in this tissue. Although the penetrance of more 3' mutations with respect to eczema seems to be lower than for repeat 1 mutations; it does not seem to be reduced in terms of ichthyosis vulgaris penetrance, as we have identified an R2447X homozygote, as well as many compound heterozygotes, who have the full, marked ichthyosis vulgaris phenotype, clinically indistinguishable from R501X homozygotes<sup>1</sup>. These preliminary observations suggest that the end

products of profilaggrin degradation within the stratum corneum might modulate atopy but not ichthyosis. Published evidence suggests that filaggrin is degraded to hygroscopic amino acids in the stratum corneum, contributing to epidermal moisture retention<sup>19</sup>, and, consistent with this, filaggrin expression also seems to be regulated by humidity in animal models<sup>20</sup>. Alternatively, filaggrin linker peptides have been found to be cross-linked components of the cornified cell envelope<sup>21</sup>, and these fragments could be derived from later profilaggrin processing events. An alternative hypothesis is that certain atopic dermatitis risk alleles of *FLG* may be in linkage disequilibrium with nearby protective factors, such as copy number variations known to affect the late cornified envelope gene cluster distal to *FLG*<sup>22</sup>. We note that filaggrin repeats show poor cross-species conservation (for example, ~35% mouse-human similarity). This could be indicative of an antimicrobial function whereby host-pathogen competition might drive rapid evolution. The ancillary functions of filaggrin, downstream of cytoskeletal aggregation, clearly require further functional analysis with respect to possible mechanisms of disease pathology.

Here we report three additional recurrent null alleles in *FLG* (R2447X, S3247X, 3702delG), each of which is significantly associated with an increased risk of eczema. Together with the previously reported prevalent *FLG* null alleles in the European population (R501X, 2282del4), these less common variants contribute to the overall population frequency of null alleles and reinforce their association with atopic phenotypes. Although these three alleles are less common, they have a cumulative population carrier frequency of 2.6% in Ireland, representing a significant contribution to the overall disease burden. In this now very well-studied population, a complex disease model is emerging for *FLG* null alleles and eczema susceptibility. Multiple recurrent alleles have emerged, each of which individually confers apparently differential strengths of disease risk but which in aggregate are found in a large percentage of affected individuals, where they confer highly significant susceptibility. In addition, the early direct sequencing data suggest that rare alleles also contribute to the overall disease burden. Calculating the exact contribution of these rare alleles can be achieved only by sequencing every individual not carrying the prevalent alleles in a well-characterized case series. In this regard, the genetic architecture of atopic dermatitis may be composed of substantial numbers of recurrent risk alleles accompanied by many individual rare or private risk alleles, as shown recently for low-high density lipoprotein (HDL) phenotypes<sup>23</sup>. Notably, the research group that initially mapped atopic dermatitis susceptibility to the 1q21 locus<sup>24</sup> has recently analyzed their childhood family study for the two common *FLG* mutations, which showed significant association<sup>12</sup>. When they removed the carriers of R501X and 2282del4 from their series, a residual lod score of 2.0 was still obtained for this locus<sup>12</sup>, so part of the residual linkage may be due to the other *FLG* alleles reported here. Comprehensive analysis of *FLG* in

**Table 3 Minor allele frequencies (MAF) of *FLG* variants in the Irish atopic dermatitis case-control series**

| Variant                  | MAF (controls) | 95% c.i.    | MAF (cases) | 95% c.i.    |
|--------------------------|----------------|-------------|-------------|-------------|
| R501X                    | 0.013          | 0.007–0.019 | 0.136       | 0.101–0.170 |
| 2282del4                 | 0.013          | 0.007–0.019 | 0.098       | 0.068–0.129 |
| R2447X                   | 0.001          | 0.000–0.003 | 0.019       | 0.005–0.032 |
| S3247X                   | 0.011          | 0.006–0.016 | 0.029       | 0.012–0.046 |
| 3702delG                 | 0.001          | 0.000–0.002 | 0.005       | 0.000–0.013 |
| Combined <i>FLG</i> null | 0.039          | 0.029–0.049 | 0.287       | 0.241–0.333 |

such a data set is necessary to evaluate the contribution (if any) of the many nearby *FLG*-related genes on 1q21 to atopic dermatitis risk.

This study raises questions of interest to the complex trait field. Would SNP tagging of these multiple, relatively rare alleles, with frequencies no greater than 0.013, have readily identified this particularly strong susceptibility gene? Although transmission studies were invaluable in identifying this susceptibility locus, they seem to lack the power of mutation screening by direct sequencing coupled with association studies to uncover the true contribution of this locus to the disease. Another notable question is why these null mutations are so prevalent. Genetic drift is one possible explanation; however, balancing selection is also possible, given the independent emergence of five recurrent risk alleles in the Irish population. Prevalent mutations have also been identified in non-European populations: for example, 3321delA is carried by 3.6% of the Chinese population. The presence of population-specific spectra of prevalent mutations might indicate a heterozygote advantage. It is tempting to speculate that increased *trans*-epidermal antigen transfer and exposure of host antigen presenting cells in individuals with skin barrier impairment may lead to some degree of 'natural vaccination' that could have conferred resistance to ancient pandemic plagues. This question should be tractable in the future by antigen or pathogen challenge of animal disease models.

The comprehensive mutation detection strategy reported here will enable efficient and detailed screening of other cohorts, including cohorts of different ancestries and regional backgrounds, for both family- and population-specific mutations in *FLG*. The true global cumulative contribution of both recurrent and family-specific *FLG* null alleles to risk of eczema and atopic phenotypes requires careful analysis.

## METHODS

**Ethical considerations.** In all cases, we followed the Helsinki guidelines and obtained informed written consent from parents or guardians of participants. The atopic dermatitis case-control study was approved by the Research Ethics Committee of Our Lady's Children's Hospital. Collection of additional individuals with ichthyosis vulgaris and atopic dermatitis was approved by local research ethics committees (Southern General Hospital, Ninewells Hospital, Innsbruck Medical University and University Hospital Maastricht).

***FLG* PCR amplification and sequencing.** Genomic DNA extracted from participants' peripheral blood was routinely amplified using the Expand High-Fidelity PCR System (Roche). Typically, 100 ng of DNA was amplified in a 25- $\mu$ l reaction containing 1.5 mM MgCl<sub>2</sub>, 10 pmol forward primer, 10 pmol reverse primer, 0.25 mM of each dNTP, DMSO (4%) and 1.3 units (U) *Taq*/*Tgo* polymerase mix. To sequence exon 3 of *FLG*, a series of overlapping PCR fragments were amplified and sequenced using repeat-specific primers. Specific priming sites for other PCR applications or sequencing can be found in **Supplementary Figure 1** or by contacting the authors for specific recommendations. The primer pairs used for amplification of the entire gene in overlapping fragments are listed in **Supplementary Figure 2**. DNA sequencing was performed using Applied Biosystems 3100/3730 DNA sequencers. For long-range PCR of *FLG* repeats 7–10, 100 ng of genomic DNA was typically amplified in a 25- $\mu$ l reaction volume using Platinum *Taq* PCR SuperMix High Fidelity (Invitrogen). PCR conditions were as follows: 94 °C for 5 min followed by 35 cycles of 94 °C (30 s), 61 °C (30 s) and 72 °C (6 min) and a final extension at 72 °C (7 min). Long-range PCR primers and conditions for amplification of exon 3 in one fragment have been reported previously<sup>1</sup>.

**Determination of the sequence of *FLG* size variants.** PCR products containing additional *FLG* repeats were cloned into the pCR2.1 vector (Invitrogen). Plasmid DNA was linearized by restriction enzyme digestion and then further digested with exonuclease III to generate a series of overlapping

unidirectional deletions (Promega Erase-A-Base System). Plasmid DNA containing overlapping deletions of the *FLG* gene was sequenced using M13 forward primer.

***FLG* mutation analysis.** Genomic DNA from affected individuals and controls was subjected to *FLG* mutation analysis. Proband DNA from 13 Irish, five Austrian and five Dutch ichthyosis vulgaris/atopic dermatitis families were fully sequenced, and *FLG* null alleles were identified in all cases. Genotyping for the R501X variant was performed using a TaqMan allelic discrimination assay as described previously<sup>2</sup>. *FLG* mutations 2282del4 and 3702delG were screened by sizing of fluorescently labeled PCR products as described previously<sup>1,4</sup>. Asian variants 3321delA and S2554X were screened as described previously<sup>17</sup>. Variant Q3683X does not alter any known restriction site and was not readily amenable to TaqMan analysis and was therefore screened in the Irish pediatric atopic dermatitis cohort by DNA sequencing of a short PCR fragment. *FLG* mutations 5360delG, 6867delAG, 7267delCA, 11029delCA and 11033del4 were screened by sizing of fluorescently labeled PCR products on an Applied Biosystems 3100 DNA sequencer. PCR primer sequences and expected product sizes are listed in **Supplementary Table 4** online. Typically, 50 ng of genomic DNA was amplified in a 10- $\mu$ l reaction containing 1.5 mM MgCl<sub>2</sub>, 0.25 mM of each dNTP, 5 pmol forward primer, 5 pmol reverse primer, 4% DMSO and 0.5 U AmpliTaq Gold polymerase (Applied Biosystems). *FLG* variants R1474X, E2422X and R2447X were screened by restriction enzyme digestion of PCR products. The primer and probe sequences and the sizes of the expected restriction fragments are listed in **Supplementary Table 4**. Briefly, 100 ng of genomic DNA was amplified in a 25- $\mu$ l reaction containing 1.5 mM MgCl<sub>2</sub>, 0.25 mM of each dNTP, 10 pmol forward primer, 10 pmol reverse primer, 4% DMSO and 0.5 U *Taq* polymerase (Promega). We performed restriction digest of 5  $\mu$ l of PCR products overnight at 37 °C with 2.5 U of *Nla* III in a 20- $\mu$ l reaction volume. Digests were run on 10% polyacrylamide/TBE gels (Invitrogen). To permit high-throughput screening for *FLG* variants R2447X and S3247X, TaqMan allelic discrimination assays were developed. Similarly, a TaqMan allelic discrimination assay was also developed for 2282del4 as a cheaper, higher-throughput alternative to fluorescent PCR sizing. Primer and probe sequences are listed in **Supplementary Table 4**. Variants 2282del4 and R2447X were screened using ABI 7700 (96-well) and 7900HT (384-well) sequence detection systems. We used standard TaqMan conditions using Applied Biosystems Universal PCR master mix and 500 nM primers with 300 nM probes. For R2447X, the inner forward primer 2 was included at 300 nM. Reaction volumes were 5  $\mu$ l for 384-well assays and 10  $\mu$ l for 96-well assays. *FLG* variant S3247X was screened using a Corbett Research Rotor-Gene 6000 PCR machine. For *FLG* variant S3247X, PCRs were performed using 20 ng of genomic DNA in a 20- $\mu$ l reaction volume containing 0.5 U Platinum *Taq* polymerase (Invitrogen), 6 pmol of each forward and reverse primer plus 3 pmol of 6-carboxyfluorescein (FAM)-labeled wild-type probe and 5 pmol VIC-labeled mutant-specific probe (Applied Biosystems MGB TaqMan). Amplification conditions were as follows: 95 °C for 2 min followed by 40 cycles of 95 °C (10 s) and 62 °C (30 s). DNA samples identified as positive for variants R2447X and S3247X by allelic discrimination assay were verified independently by restriction enzyme digestion (R2447X) and/or DNA sequencing.

**Statistical analysis.** All statistical analyses in the association studies were performed using SPSS for Windows, v.11.5 or Instat 3 for Macintosh (Graphpad Software). Allele frequencies were compared using Pearson  $\chi^2$  tests of the three genotype frequencies (AA, Aa and aa). OR values were calculated using Fisher's exact test. The variants did not significantly differ from the expected Hardy-Weinberg equilibrium (HWE) in the population samples, but the combined *FLG* null genotype was significantly different from HWE in the atopic dermatitis cohort ( $P = 0.0037$ ), adding further evidence of allele frequency distortion in the case cohort.

**Accession codes.** Entrez Gene RefSeq ID: ten-repeat allele of *FLG*, NM\_002016.

*Note: Supplementary information is available on the Nature Genetics website.*

## ACKNOWLEDGMENTS

We thank affected individuals and their families for their participation, which made this research possible, and J. McFarlane (Epithelial Genetics Group) for

clerical assistance. The Irvine group is supported by the Children's Medical and Research Foundation of Our Lady's Children's Hospital. The McLean/Smith group is supported by grants from the Dystrophic Epidermolysis Bullosa Research Association, the Pachonychia Congenita Project and the British Skin Foundation/National Eczema Society (W.H.I.M. and F.J.D.S.). The Palmer group is supported by the Biotechnology and Biological Sciences Research Council (award D13460), Scottish Enterprise Tayside and the Gannochy Trust and the Scottish Executive Genetic Health Initiative. The Innsbruck group is supported by the Medical Research Fund of the Innsbruck University Medical Center (Medical Finance Fund numbers 71 and 153) and the Austrian National Bank (Austrian National Bank number 10822).

#### AUTHOR CONTRIBUTIONS

The study was designed by W.H.I.M., A.D.I., A.S. and F.J.D.S. The atopic dermatitis case series was collected by A.D.I., G.M.O'R., R.M.W. and T.H.C.; Irish population controls were obtained by D.G.B. and additional clinical contributions were made by P.R.H., C.S.M., M.A.M.v.S., P.M.S., I.N. and M.S. DNA was extracted and stored by L.E.C. Clinical samples and case series were managed by P.R.H. Molecular biology was performed by A.S., A.T.-K., P.R.H., T.C., H.L., Y.Z., L.E.C., M.v.G., A.R.J. and R.G. Histological and ultrastructural work was done by A.T.E., W.H.I.M., M.S. and P.M.E. Statistical analysis was done by C.N.A.P. The manuscript was written by W.H.I.M., A.D.I., A.S., M.S. and F.J.D.S.

#### COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/naturegenetics](http://www.nature.com/naturegenetics).

Published online at <http://www.nature.com/naturegenetics>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

- Smith, F.J.D. *et al.* Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat. Genet.* **38**, 337–342 (2006).
- Palmer, C.N.A. *et al.* Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat. Genet.* **38**, 441–446 (2006).
- Judge, M.R., McLean, W.H.I. & Munro, C.S. Disorders of keratinization. in *Rook's Textbook of Dermatology* Vol. 2 (eds. Burns, T., Breathnach, S., Cox, C. & Griffiths, C.) 34.54–34.56 (Blackwell Scientific, Oxford, 2004).
- Sandilands, A. *et al.* Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. *J. Invest. Dermatol.* **126**, 1770–1775 (2006).
- Gruber, R. *et al.* Filaggrin mutations p.R501X and c.2282del4 in ichthyosis vulgaris. *Eur. J. Hum. Genet.* **15**, 179–184 (2007).
- Barker, J.N. *et al.* Null mutations in the filaggrin gene (*FLG*) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *J. Invest. Dermatol.* **127**, 564–567 (2007).
- Weidinger, S. *et al.* Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J. Allergy Clin. Immunol.* **118**, 214–219 (2006).
- Weidinger, S. *et al.* Filaggrin mutations strongly predispose to early-onset and extrinsic atopic dermatitis. *J. Invest. Dermatol.* **127**, 724–726 (2007).
- Ruether, A., Stoll, M., Schwarz, T., Schreiber, S. & Folster-Holst, R. Filaggrin loss-of-function variant contributes to atopic dermatitis risk in the population of Northern Germany. *Br. J. Dermatol.* **155**, 1093–1094 (2006).
- Marenholz, I. *et al.* Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J. Allergy Clin. Immunol.* **118**, 866–871 (2006).
- Stemmler, S., Parwez, Q., Petrasch-Parwez, E., Epplen, J.T. & Hoffjan, S. Two common loss-of-function mutations within the filaggrin gene predispose for early onset of atopic dermatitis. *J. Invest. Dermatol.* **127**, 722–724 (2006).
- Morar, N., Cookson, W.O., Harper, J.I. & Moffatt, M.F. Filaggrin mutations in children with severe atopic dermatitis. *J. Invest. Dermatol.* advance online publication, 15 February 2007 (doi:10.1038/sj.jid.5700739).
- Irvine, A.D. Fleshing out filaggrin phenotypes. *J. Invest. Dermatol.* **127**, 504–507 (2007).
- Irvine, A.D. & McLean, W.H.I. Breaking the (un)sound barrier: filaggrin is a major gene for atopic dermatitis. *J. Invest. Dermatol.* **126**, 1200–1202 (2006).
- Mischke, D., Korge, B.P., Marenholz, I., Volz, A. & Ziegler, A. Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J. Invest. Dermatol.* **106**, 989–992 (1996).
- Gan, S.Q., McBride, O.W., Idler, W.W., Markova, N. & Steinert, P.M. Organization, structure, and polymorphisms of the human profilaggrin gene. *Biochemistry* **29**, 9432–9440 (1990).
- Nomura, T. *et al.* Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. *J. Allergy Clin. Immunol.* **119**, 434–440 (2007).
- Presland, R.B. *et al.* Loss of normal profilaggrin and filaggrin in flaky tail (ft/ft) mice: an animal model for the filaggrin-deficient skin disease ichthyosis vulgaris. *J. Invest. Dermatol.* **115**, 1072–1081 (2000).
- Rawlings, A.V. & Harding, C.R. Moisturization and skin barrier function. *Dermatol. Ther.* **17**Suppl 1, 43–48 (2004).
- Katagiri, C., Sato, J., Nomura, J. & Denda, M. Changes in environmental humidity affect the water-holding property of the stratum corneum and its free amino acid content, and the expression of filaggrin in the epidermis of hairless mice. *J. Dermatol. Sci.* **31**, 29–35 (2003).
- Takahashi, M., Tezuka, T. & Katunuma, N. Filaggrin linker segment peptide and cystatin alpha are parts of a complex of the cornified envelope of epidermis. *Arch. Biochem. Biophys.* **329**, 123–126 (1996).
- McCarroll, S.A. *et al.* Common deletion polymorphisms in the human genome. *Nat. Genet.* **38**, 86–92 (2006).
- Cohen, J.C. *et al.* Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869–872 (2004).
- Cookson, W.O. *et al.* Genetic linkage of childhood atopic dermatitis to psoriasis susceptibility loci. *Nat. Genet.* **27**, 372–373 (2001).