



ELSEVIER

Sequence analysis of filaggrin gene by novel shotgun method in Japanese atopic dermatitis

Takashi Sasaki^{a,b}, Jun Kudoh^{a,b}, Tamotsu Ebihara^c,
Aiko Shiohama^{a,b}, Shuichi Asakawa^a, Atsushi Shimizu^a,
Atsushi Takayanagi^a, Itaru Dekio^c, Chieko Sadahira^c,
Masayuki Amagai^c, Nobuyoshi Shimizu^{b,*}

^a Department of Molecular Biology, Keio University School of Medicine,
35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

^b Advanced Research Center for Genome Super Power, Keio University, 2 Okubo,
Tsukuba, Ibaraki 300-2611, Japan

^c Department of Dermatology, Keio University School of Medicine, 35 Shinanomachi,
Shinjuku-ku, Tokyo 160-8582, Japan

Received 27 October 2007; received in revised form 4 February 2008; accepted 21 February 2008

KEYWORDS

Filaggrin;
Atopic dermatitis;
DNA sequencing;
Shotgun

Summary

Background: Recent reports indicated that nonsense mutations in filaggrin (FLG) found in ichthyosis vulgaris (IV) patients are predisposing factors for atopic dermatitis (AD) with asthma. The exon 3 of FLG contains tandemly repeated, highly homologous, 11–13 sequence units of 972 or 975 bp, each of which corresponds to the coding sequence of the processed filaggrin with slight sequence difference. This unique gene structure has hampered the precise DNA sequence determination.

Objective: We developed a novel DNA sequencing method “FLG-shotgun” to directly characterize the mutations in Japanese AD patients.

Methods: We examined 24 Japanese AD patients with “FLG-shotgun” method.

Results: Multiple units of FLG were amplified by PCR using several sets of common primers for the conserved regions, and DNA sequences of each cloned PCR product were determined. Multiple reads of DNA sequences in both alleles were aligned and re-constructed to cover the entire coding regions. We found three major genotypes (A, B, and C) which represent different numbers (11–13) of homologous sequence units. Furthermore, we found two novel nonsense mutations; one mutation 8666-8667CC>GA on the unit 9 of allele B that causes a nonsense mutation S2899X in two patients and the other mutation 9887C>A on the unit 10 of allele B that causes a nonsense mutation S3296X in two patients.

* Corresponding author. Tel.: +81 29 865 1641; fax: +81 29 865 1641.
E-mail address: smzgsp@dmb.med.keio.ac.jp (N. Shimizu).

Conclusion: We found two novel FLG mutations by directly analyzing Japanese patients with AD. FLG-shotgun will provide a valuable tool to further define the nature of the AD phenotype associated with FLG mutations.

© 2008 Japanese Society for Investigative Dermatology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Filaggrin (FLG) is a major epidermal protein and expressed as a profilaggrin of >500 kDa in the prominent keratohyalin granules of outer epidermis [1,2]. The profilaggrin is processed by protease to filaggrin of ~35 kDa and the resulting filaggrins are cross-linked by transglutaminase to build skin barrier [3]. The skin barrier is indispensable to prevent invasion of micro-organisms, chemical compounds and allergens, and to keep moisture level of the skin [2], therefore, the reduced filaggrin production would disturb skin barrier formation.

Ichthyosis vulgaris (IV; OMIM no. 146700) is one of the most common genetic skin disorder [4]. The clinical features of IV include palmar hyperlinearity, keratosis pilaris and scaling at lower abdomen, arms and legs [5]. Thirteen kinds of FLG gene mutations were found in the IV and atopic dermatitis (AD) patients from European population [6–8], and two different kinds of mutations (S2554X and 3321delA) were found in the Japanese IV patients ([9], Nomura et al.). These results indicated that FLG gene mutations produce reduced amounts of filaggrin protein which would generate IV phenotypes.

AD frequently coincides with keratosis pilaris, allergy, and asthma, and AD patients have increased recently in the developed world. Importantly, AD and keratosis pilaris are commonly found in IV. Actually, recent reports indicated that nonsense mutations in FLG gene found in IV patients are predisposing factors for AD with asthma [6–13]. These results strongly suggest that predisposing factors for AD include not only immune-deficiency but also skin barrier deficiency.

The FLG is located on human chromosome 1q21.3 as a member of the large gene cluster “epidermal differentiation complex (EDC)” involved in the terminal differentiation. The EDC gene cluster is composed of FLG, six FLG-related genes, and genes for S100 family, small proline-rich protein family, loricrin, involucrin and so on [14]. The FLG gene consists of three exons and produces a transcript of 12,747 nt [15,16]. Interestingly, relatively large exon 3 (12,048 bp) contains all the coding sequence to encode a large profilaggrin of >500 kDa and in fact it consists of

~11 homologous sequence units of 972 or 975 bp, each of which corresponds to the coding sequence of the processed filaggrin with slight sequence difference. Hence, precise DNA sequence determination has been hampered by these multiple units of homologous sequence.

In this study, we developed a shotgun DNA sequencing method “FLG-shotgun” specialized for FLG. As a result of FLG-shotgun of 24 Japanese AD patients, we isolated three major genotypes (A–C) containing 11, 13, and 12 FLG homologous sequence units, respectively. Furthermore, we found two novel nonsense mutations on type B allele. These results indicated efficiency of FLG-shotgun for DNA sequence of FLG in AD patients.

2. Materials and methods

2.1. Clinical materials

Blood samples were obtained from 24 Japanese AD patients, who have been suffering from severe, itchy, chronically relapsing, inflammatory skin condition for a long time, and observation of total IgE amount incensement. The diagnosis of atopic dermatitis was made using the AD diagnostic criteria by Japanese Dermatological Association. Among these 24 Japanese AD patients, there is no IV patient who shows palmar hyperlinearity and typical scaling at arms and legs. This study was approved by the committee for medical ethics at Keio University.

2.2. FLG-shotgun

Genomic DNA was isolated from peripheral bloods or saliva by PAX gene Blood DNA kit (Qiagen). The homologous sequence region in FLG was amplified by Expand High Fidelity PCR polymerase (Roche) with four distinct primer sets: FLG set1 FLGREP-24F and mixture of FLGREP-965R1/R2/R3, FLG set2 FLGREP-487F and FLGREP-463R, FLG set3 FLGREP-257F and FLGREP-215R, FLG set4 FLGREP-683F and FLGREP-669R (see Supplementary Table 1 for primer DNA sequence). Thus, segments of 900–950 bp can be amplified with small overlaps. PCR amplification condition was as follows, 1 cycle of 94 °C for 2 min,

35 cycles of 94 °C for 30 s, 60 °C for 1 min, and 72 °C for 4 min, and 1 cycle of 72 °C for 5 min. The PCR products were purified with Wizard plus SV mini-preps DNA purification system (Promega) and ligated into TA cloning vector pGEM-tEasy (Promega). *E. coli* was transformed with these DNAs and isolated colonies were inoculated into 384-well plates. The plasmid DNA in the *E. coli* of each well was amplified by TempliPhi DNA Amplification kit (GE healthcare bioscience), and was sequenced by each primer for PCR reaction using Big Dye ver. 3.1 and 3730 DNA analyzer (Applied Biosystems). Phred/Phrap/Consed was used for assembly of DNA sequences.

Exon 2 was amplified by PCR with FLGexon2F and FLGexon2R. Both terminal regions in exon 3 (Fig. 2) were amplified by PCR with FiLLR2F and FLGREP-215R, and FLGREP-487F and FiLLR1R. These PCR products were sequenced as described above.

2.3. Long PCR of whole homologous sequence region in FLG

Whole homologous sequence region in FLG was amplified by Long PCR as described previously with slight modification, namely, KOD FX (TOYOBO) was used as DNA polymerase [6].

3. Results

3.1. FLG-shotgun DNA sequencing

The FLG consists of three exons and relatively large exon 3 (12,048 bp) contains all the coding sequence to encode a large profilaggrin of >500 kDa and in detail it consists of ~11 homologous sequence units (Fig. 1A and B). These homologous sequence units correspond to the coding sequence of each of the processed filaggrins with slight sequence difference. In fact, the self dot matrix analysis of FLG cDNA sequence (Ref seq ID: NM_002016) revealed 11 homologous sequence units within a region (892–11,773 nt) and a small sequence of 184 bp after 11th unit (Fig. 1B and C). Those homologous sequence units are extremely similar to each other, and their length is commonly 972 bp except that units 2 and 6 are 3 bp larger than others (Fig. 1D). Therefore, it was hard to precisely analyze each homologous sequence unit with high accuracy.

To solve this problem, we applied the shotgun DNA sequencing method to genomic DNA for ~20-kbp FLG and developed a “FLG-shotgun” method (Fig. 2). The first step of FLG-shotgun involves PCR-amplification of all homologous sequence units using

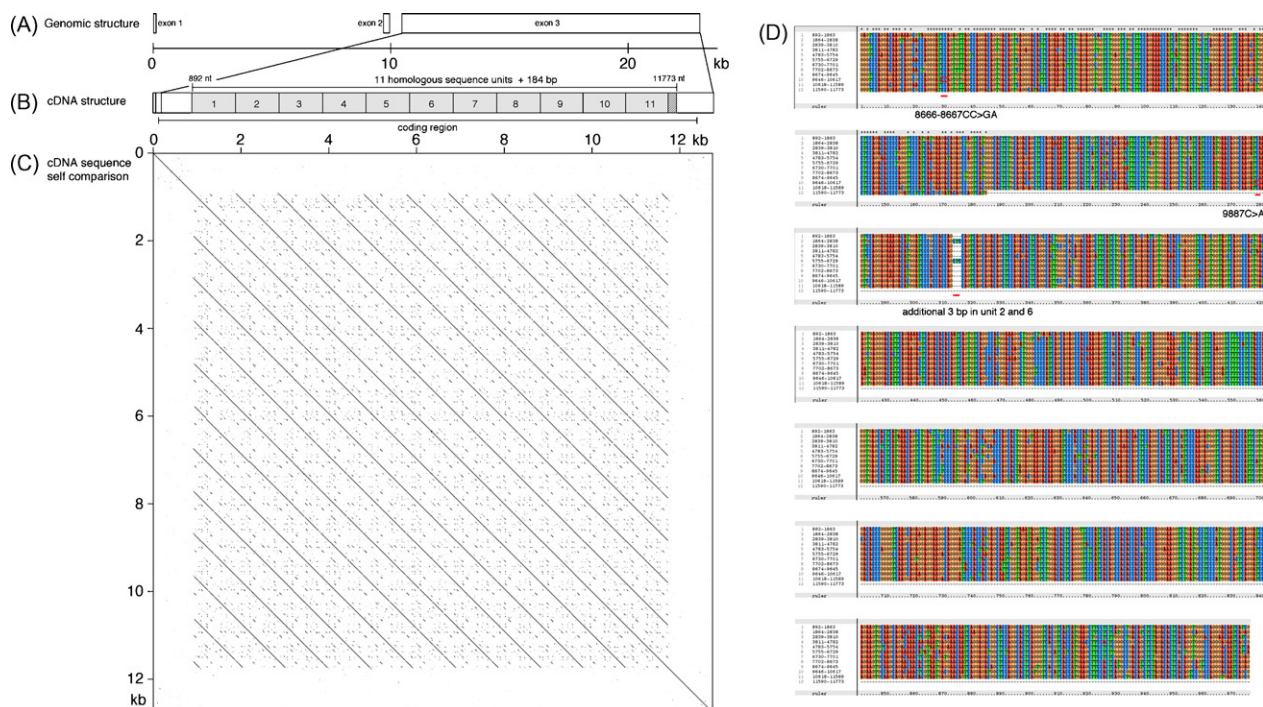


Fig. 1 Features of human filaggrin (FLG) gene. (A) Genomic structure. The FLG gene consists of three exons. (B) cDNA structure (NM_002016). Exon 3 is composed of eleven homologous sequence units of 972 or 975 bp each of which encodes filaggrin. These units are located in tandem at a region 892–11,773 nt. (C) The self dot matrix analysis of FLG cDNA sequence (NM_002016). This figure indicates eleven homologous sequence units of 972 or 975 bp and one small unit of 184 bp. (D) The DNA sequence alignment of FLG cDNA (NM_002016). Three-base insertion and two novel nonsense mutations can be seen.

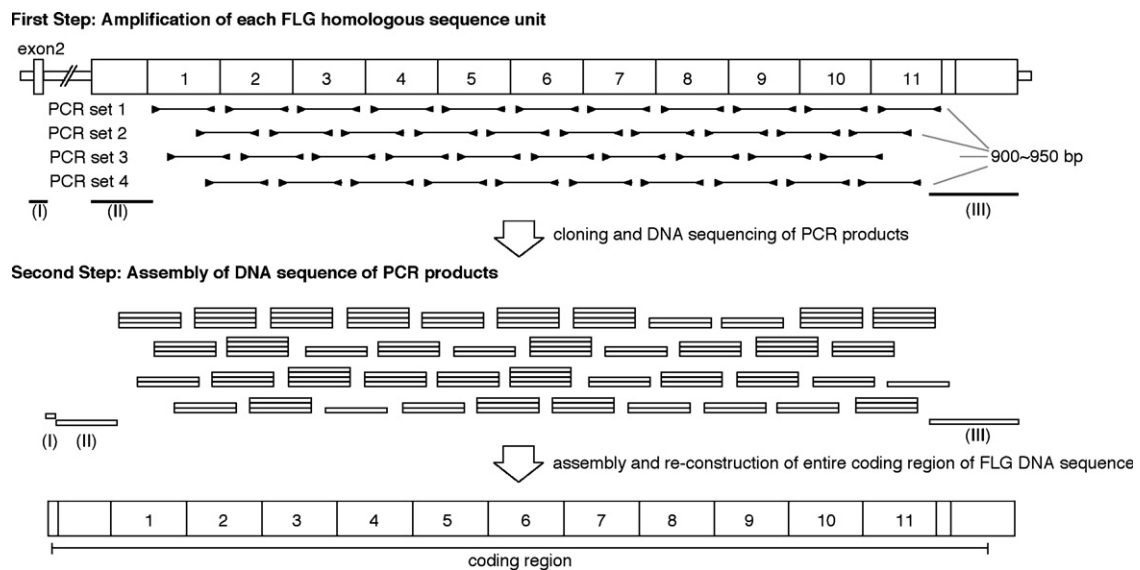


Fig. 2 Strategy of FLG-shotgun DNA sequencing and data assembly. First step involves the PCR-amplification of each homologous sequence unit using four sets of primers which amplify every 900–950 bp. Also, primer sets for exon 2 (I), the distal (II) region of unit 1, and proximal (III) region of unit 11 are indicated as thick lines. Second step involves the assembly of DNA sequences to assemble each unit and finally re-construct entire ORF region of FLG DNA sequence. The amount of PCR products varies depending on the genotypes.

four different sets of PCR primers that cover 900–950-bp segments with slight overlaps. The PCR products were cloned into plasmids and individually sequenced. The second step involves assembly of DNA sequences to determine each homologous sequence units and eventually re-construct the entire FLG sequence. At this stage, the amount of PCR products varied depending on the patient genotypes as illustrated in Fig. 2. Furthermore, three additional PCR-primer sets (I–III) were used to amplify exon 2 and both terminal region of exon 3.

3.2. Identification of three FLG genotypes in Japanese AD patients

We determined FLG sequence of 24 Japanese AD patients using FLG-shotgun method described above. At the first step of assembly of DNA sequence, we noticed that assembled FLG sequences were classified into two groups. The first group consists of nine cases and has almost identical sequence with no or a few heterozygous nucleotides in both alleles indicating homozygote for FLG sequence, while the second group consists of 15 cases has apparently heterozygous FLG sequences with many variations between two alleles. We first determined homozygous FLG sequences and they were classified into three major genotypes (A–C) as described below.

Homozygous genotype A was found in 3 AD patients (P006, P008, and P011) and was matched with reference sequence of FLG cDNA sequence

(NM_002016) containing 11 homologous sequence units except a few SNPs. Fig. 3A shows a representative DNA sequencing data for the patient P008. In this case, we could re-construct the continuous entire FLG sequence with only two heterozygous nucleotides.

Homozygous genotype B consists of 13 homologous sequence units and was found in 4 Japanese AD patients (P005, P007, P017, and P023). We precisely determined FLG sequence of genotype B with four primer sets, and a representative DNA sequencing data for the patient P023 was shown in Fig. 3B.

Homozygous genotype C consists of 12 homologous sequence units and was found in 2 Japanese AD patients (P001 and P022). We precisely determined FLG sequence of genotype C with four primer sets (Fig. 3C). In both these cases, only a single continuous FLG sequence was assembled.

In other 15 AD patients, we assembled over 20 homologous sequence units first. It apparently indicated that they have two different FLG genotypes on distinct alleles. In these cases, we re-constructed two FLG sequences on distinct alleles based on the SNPs between two alleles. As a result, 28 out of 30 re-constructed FLG sequences were classified into genotype A, B, or C except two cases derived from one allele of P019 and P021 (Table 1). Although FLG sequences of these two alleles were basically matched with genotype B FLG, two homologous sequence units were lost, resulting in 11 homologous sequence units (Fig. 4A). We designated this genotype as type Bs.

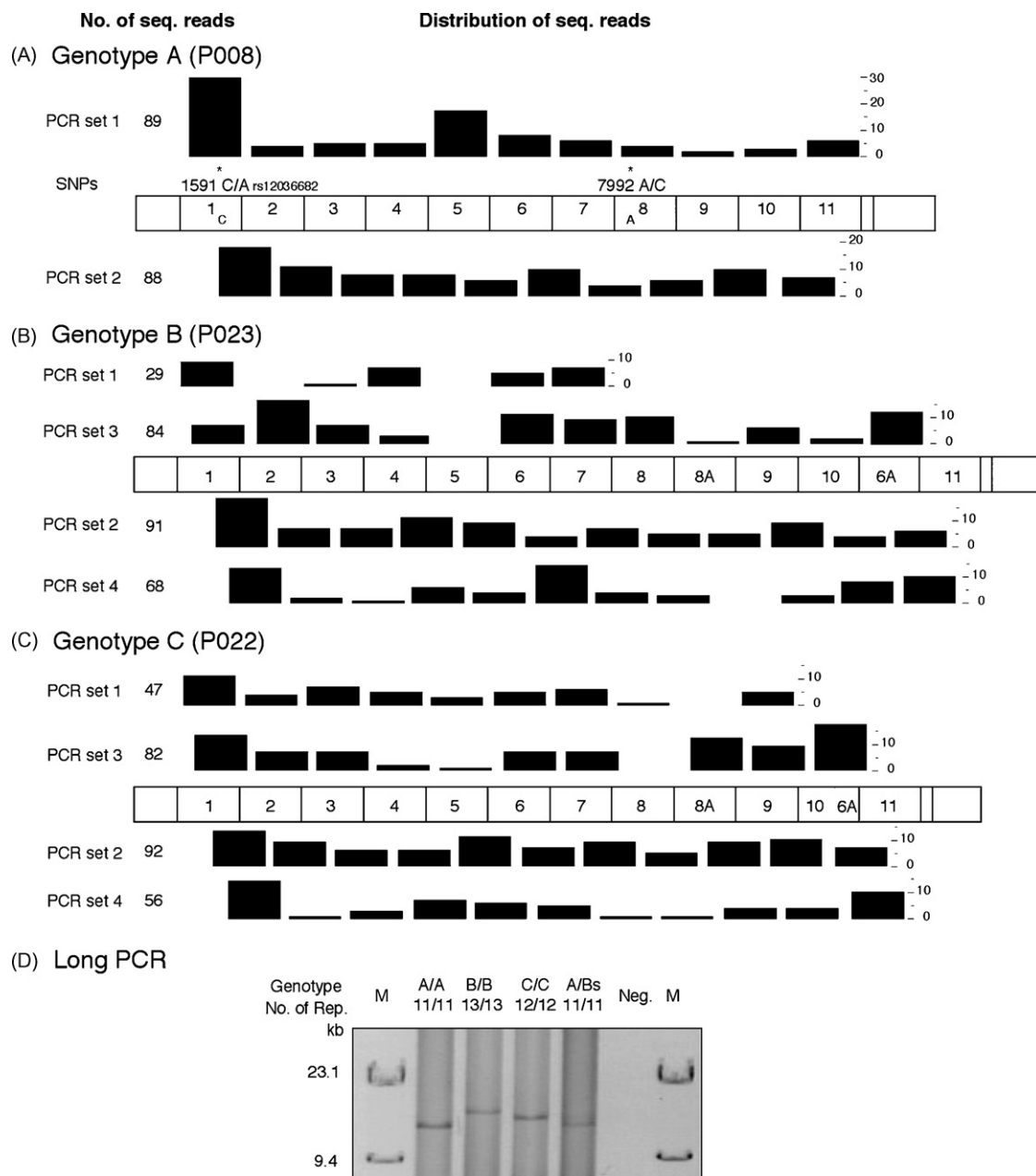


Fig. 3 Schematic representation of data analysis for genotypes A, B, and C of FLG. The frequency of sequence-reads is indicated by thickness of black rectangles. (A) The genotype A of FLG. The FLG DNA sequence of genotype A was re-constructed by two sets of PCR (set 1 and set 2). Two heterozygous nucleotides (1591 C/A and 7992 A/C) were found in genotype A of FLG derived from P008. (B) The genotype B of FLG. The FLG DNA sequence of genotype B was re-constructed by four sets of PCR (sets 1–set 4). No heterozygous nucleotides was found in genotype B of FLG derived from P023. (C) The genotype C of FLG. The FLG DNA sequence of genotype C was re-constructed by four sets of PCR (set 1–set 4). No heterozygous nucleotides was found in genotype C of FLG derived from P022. (D) Long PCR analysis of genotypes A–C, and Bs of FLG. P008 (A/A), P023 (B/B), P022 (C/C), and P021 (A/Bs) were used as representatives of each genotype.

Finally, we found three major (A–C) and one minor (Bs) genotypes for the FLG gene of 24 Japanese AD patients (Table 1). There are four cases (P001, P005, P022, and P023) for which no heterozygous nucleotide was detected perhaps because DNA sequence of two alleles are identical or one allele is deleted.

To confirm the reliability of these re-constructed FLG sequences, we compared the length of long PCR products which include whole homologous sequence units region (Fig. 3D). We found that the length of long PCR products observed are consistent with the number of homologous sequence units in reconstructed FLG sequences of each genotype. These

Table 1 Genotype and mutation analyses of FLG for 24 Japanese AD patients

Patients	FLG genotype	No. of units per allele	Nonsense mutation
P001	C ^a	12 ^a	—
P002	A/B	11/13	—
P003	A/B	11/13	—
P004	A/B	11/13	—
P005	B ^a	13 ^a	—
P006	A/A	11/11	—
P007	B/B	13/13	—
P008	A/A	11/11	—
P009	B/C	13/12	—
P010	A/B	11/13	—
P011	A/A	11/11	—
P012	B/C	13/12	—
P013	A/C	11/12	—
P014	A/B	11/13	—
P015	A/C	11/12	—
P016	A/B	11/13	8666-8667CC>GA (B)
P017	B/B	13/13	—
P018	A/B	11/13	—
P019	A/Bs	11/11	9887C>A (Bs)
P020	A/B	11/13	8666-8667CC>GA (B)
P021	A/Bs	11/11	9887C>A (Bs)
P022	C ^a	12 ^a	—
P023	B ^a	13 ^a	—
P024	A/B	11/13	—

^a DNA sequence of two FLG alleles are identical or one allele may be deleted.

results indicated that previously reported length polymorphism of FLG were derived from unit number difference among FLG sequences [3].

3.3. Sequence differences among three FLG genotypes

We identified three distinct genotypes (A–C) which represent different number of homologous sequence units (Fig. 4A). The genotype B consists of 13 units including 8A and 6A which are similar to units 8 and 6, respectively. The genotype C consists of 12 units including 8A but one unit is lacked between unit 10 and unit 6A comparing with genotype B. The genotype Bs lacked unit 8A and 9 from genotype B, which would be subtypes of genotype B. Among these genotypes of FLG, we often identified single nucleotide sequence polymorphisms among these three genotypes (Supplementary Table 2). These sequence polymorphisms were valuable to precisely assemble each homologous sequence units, especially when patient's genotypes are heterozygous.

3.4. Two novel nonsense mutations in the FLG of Japanese AD patients

The previous reports showed thirteen different mutations in the FLG for European IV and AD

patients, and two additional nonsense mutations (S2554X and 3323delA) for Japanese IV and AD patients [6–9]. We found two novel nonsense mutations in 24 Japanese AD patients: one mutation 8666-8667CC>GA on unit 9 of allele B in the patient P016 (also patient P020) causes a nonsense mutation S2899X (Fig. 4A and B) and the other mutation 9887C>A on unit 10 of allele Bs in the patient P021 (also patient P019) causes a nonsense mutation S3296X (Fig. 4A and C).

4. Discussion

In this study, we developed a shotgun DNA sequencing method “FLG-shotgun” specialized for determining DNA sequence of the complicated FLG that consists of 11–13 tandem homologous sequence units of 972 or 975 bp. In this method, multiple units of FLG homologous sequence are amplified by PCR using several sets of common primers which are designed at conserved region of multiple units, and DNA sequences of each PCR product were determined. Finally, DNA sequences of FLG in both alleles were re-constructed respectively using DNA sequence differences among each multiple units of homologous sequence. Previous studies used the several primer sets to amplify the specific region of these multiple units [8,9]. This method is one of

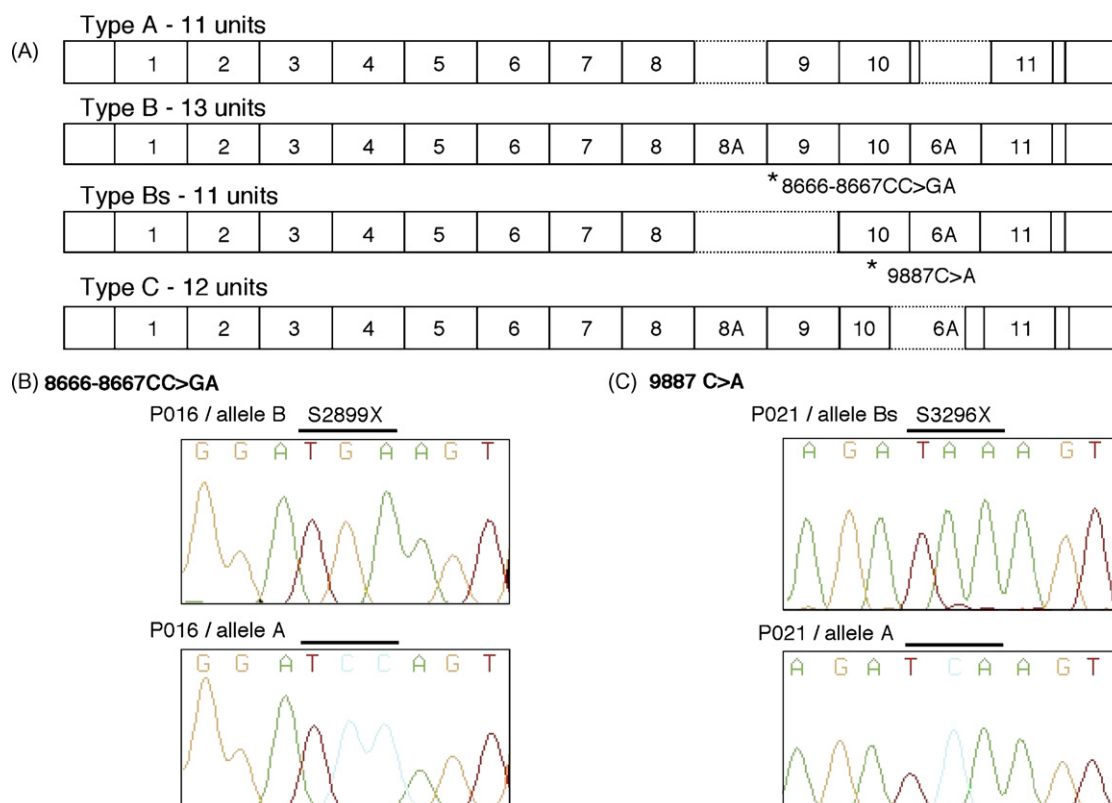


Fig. 4 The genotype of FLG and two novel nonsense mutations. (A) Three major (types A, B, and C) and one minor genotypes (type Bs) of FLG gene in Japanese AD patients. (B) The mutation 8666-8667CC>GA in the allele B of patient P016 as compared to the normal allele A. This mutation causes nonsense codon S2899X. (C) The 9887C>A mutation in the allele Bs of patient P021 as compared to the normal allele A. This mutation causes nonsense codon S3296X.

efficient method to determine the FLG DNA sequence, however, it is difficult to distinguish the DNA sequence of two alleles. Furthermore, there is a little possibility to miss the detection of the novel FLG genotype which is not amplified using current primer sets. Actually, number of these multiple units of homologous sequence is different in each individual [3,6]. Therefore, novel DNA sequencing method for FLG is worthwhile to know the detailed FLG DNA sequence information of each genotype.

As a result of FLG-shotgun of 24 Japanese AD patients, we identified three major genotypes (A–C) regarding difference in the number of homologous sequence unit. Previous report indicated that number of FLG homologous sequence unit varies from 10 to 12 [3]. We observed such variation but it ranged from 11 to 13 in Japanese AD patients, suggesting a significant difference in human races. This variation was also linked often with small nucleotide sequence differences within homologous sequence units.

We found two novel nonsense mutations (S2899X and S3296X) in Japanese AD patients. Previously, the most of nonsense mutations of FLG were identified

from IV patients, and then were confirmed in AD patients. In this study, we directly found these two mutations in AD patients. Currently, there is no information whether all FLG mutations responsible for AD are common with IV. Further studies would be needed to confirm the association between IV and two novel FLG mutations that we found.

We could not find any other mutations that were reported for European and Japanese AD patients. Other study also reported no European type nonsense mutations in Japanese AD patients [6,9]. Therefore, the types of nonsense mutation would be different between European and Japanese populations. The nonsense mutations produce a truncated shorter profilaggrin which may still be processed to filaggrins. If this was the case, the reduced amounts of filaggrin would be insufficient to form strong skin barrier and in turn it would directly affect the patient phenotype. Further study will be necessary to confirm the relation between FLG gene mutations and production level of filaggrin in the AD patients. Moreover, in the EDC locus, there are six other FLG-related genes and these genes are possible to associate with the pathogenesis of AD. Therefore, the FLG-shotgun method would also be

applicable to the association study of those genes in EDC to further elucidate the genetic basis of AD.

Acknowledgements

The authors thank to Dr. Kazunori Shibuya and Dr. Akiharu Kubo for useful discussion, and thank to M. Furuhashi, T. Adachi, and S.K. Ishikawa for their excellent technical assistance. This work was supported by a fund from Keio Medical Association, and by Health and Labour Sciences Research Grants for Research on Allergic Disease and Immunology from the Ministry of Health, Labour, and Welfare of Japan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jdermsci.2008.02.009](https://doi.org/10.1016/j.jdermsci.2008.02.009).

Conflict of interest

None.

References

- [1] Listwan P, Rothnagel JA. Keratin bundling proteins. *Methods Cell Biol* 2004;78:817–27.
- [2] Candi E, Schmidt R, Melino G. The cornified envelope: a model of cell death in the skin. *Nat Rev Mol Cell Biol* 2005;6:328–40.
- [3] Gan SQ, McBride OW, Idler WW, Markova N, Steinert PM. Organization, structure, and polymorphisms of the human profilaggrin gene. *Biochemistry* 1990;29:9432–40.
- [4] Judge MR, McLean WHI, Munro CS. Disorders of keratinisation. *Rook's textbook of dermatology*, vol. 2. 2004. pp. 34.54–34.56.
- [5] Sybert VP, Dale BA, Holbrook KA. Ichthyosis vulgaris: identification of a defect in synthesis of filaggrin correlated with an absence of keratohyaline granules. *J Invest Dermatol* 1985;84:191–4.
- [6] Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet* 2006;38:337–42.
- [7] Sandilands A, O'Regan GM, Liao H, Zhao Y, Terron-Kwiatkowski A, Watson RM, et al. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. *J Invest Dermatol* 2006;126:1770–5.
- [8] Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, et al. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 2007;39:650–4.
- [9] Nomura T, Sandilands A, Akiyama M, Liao H, Evans AT, Sakai K, et al. Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. *J Allergy Clin Immunol* 2007;119:434–40.
- [10] Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, et al. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 2006;38:441–6.
- [11] Marenholz I, Nickel R, Rüschendorf F, Schulz F, Esparza-Gordillo J, Kerscher T, et al. Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J Allergy Clin Immunol* 2006;118:866–71.
- [12] Stemmler S, Parwez Q, Petrasch-Parwez E, Epplen JT, Hoffman S. Two common loss-of-function mutations within the filaggrin gene predispose for early onset of atopic dermatitis. *J Invest Dermatol* 2007;127:722–4.
- [13] Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP, et al. Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *J Invest Dermatol* 2007;127:564–7.
- [14] Mischke D, Korge BP, 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* 1996;106:989–92.
- [15] Presland RB, Haydock PV, Fleckman P, Nirunskis W, Dale BA. Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. *J Biol Chem* 1992;267:23772–81.
- [16] Markova NG, Marekov N, Chipev CC, Gan SQ, Idler WW, Steinert PM. Profilaggrin is a major epidermal calcium-binding protein. *Mol Cell Biol* 1993;13:613–25.

Available online at www.sciencedirect.com

