Massively Parallel Sequencing of the Filaggrin Gene Reveals an Association Between FLG Loss-of-function Mutations and Leprosy

Wenhao SHI, Zihao MI, Zhenzhen WANG, HuiminZHANG, Na WANG, Zhe WANG, BowenZHANG, QianqianXIA, YueqianYU, GongqiYU, LeleSUN, Xi’anFU, ChuanWANG, HongLIU and FurenZHANG
Shandong Provincial Hospital for Skin Diseases & Shandong Provincial Institute of Dermatology and Venereology, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, Shandong, China

Filaggrin, encoded by the FLG gene, plays a crucial role in the barrier function of epidermis, but the association between FLG loss-of-function mutations and infectious skin diseases has not been systematically studied. FLG coding sequences from 945 patients with leprosy and 916 healthy controls were captured and enriched using an array-based high-throughput system, and subjected to next-generation sequencing. The loss-of-function mutations found were further validated by Sanger sequencing. A total of 21 loss-of-function mutations were found in 945 patients with leprosy, with a carrier rate of 17.53%, while the prevalence of these mutations in 916 healthy controls was 14.77%, which was significantly lower than in patients. Two individual FLG loss-of-function mutations (K4022X and Q1790X) were found to be significantly associated with leprosy. These results suggest a possible role for filaggrin in defending against leprosy pathogens.

Key words: leprosy; filaggrin; FLG; loss-of-function mutation; association study.

Accepted Oct 7, 2020; Epub ahead of print Oct 13, 2020
Acta Derm Venereol 2020; 100: adv00299.

Corr: Zihao MI and Hong LIU, Shandong Provincial Hospital for Skin Diseases & Shandong Provincial Institute of Dermatology and Venereology, Shandong First Medical University & Shandong Academy of Medical Sciences, 27397 Jingshi Lu, Jinan 250022, Shandong Province, China. E-mails: mizihao2015@126.com; hongyue2519@hotmail.com

The FLG gene encodes filaggrin, which is the main constituent of keratohyalin granules and is indispensable for proper functioning of the epidermal barrier. Loss-of-function (LoF) mutations in the FLG gene, which cause pre-stop and frameshift of coding, are the cause of the Mendelian disease ichthyosis vulgaris (IV) (1). A strong association between FLG LoF mutations and atopic dermatitis (AD) has also been shown (2–4). Based on the speculation that an impaired skin barrier may facilitate the invasion of microbes, the association between FLG mutations and skin infection has been studied in patients with AD, and filaggrin deficiency has been shown to increase patients’ susceptibility to bacterial infection (5). Nevertheless, the association between FLG LoF mutations and infectious skin disease has been little studied. In 2017, the low-frequency FLG mutation K4022X was shown to increase susceptibility to leprosy, in a genome-wide analysis of protein-coding variants in a Chinese population (6), highlighting the involvement of FLG in infectious diseases per se.

Leprosy is a chronic infectious disease caused by the pathogen Mycobacterium leprae, which shows strong genetic predisposition. This study sequenced the FLG gene coding sequence in a large cohort of patients with leprosy and matched healthy controls. The results show that the carrier rate of FLG loss-of-function mutations in patients with leprosy was significantly higher than that in healthy controls. Skin barrier deficiency, caused by FLG loss-of-function mutations, may contribute to the pathogenesis of leprosy by increasing the risk of infection through impaired skin. These results suggest a possible role for filaggrin in defending against leprosy pathogens.
MATERIALS AND METHODS

Study subjects and genomic DNA preparation

All patients with leprosy and healthy controls were recruited using uniform criteria from Shandong Provincial Institute of Dermatology and Venereology, as described previously (6). The study was approved by the institutional review board and all subjects gave written informed consent. Leprosy was diagnosed by at least 2 dermatologists based on clinical manifestations and histopathological examination. Blood samples were taken at the time of patient diagnosis. Genomic DNA was extracted using QuickGene DNA Whole Blood Kit L (AADS Biotec, Neyagawa, Japan) following the manufacturer’s instructions.

High-throughput sequencing and Sanger sequencing

A newly reported array-based high-throughput sequencing method was used to sequence the FLG coding sequence (reference gene sequence, NG_016190.1) (16). Briefly, the whole FLG coding sequences of FLG were simultaneously amplified by the Access Array system (Fluidigm, San Francisco, CA, USA) followed by sequencing of pooled amplicons using the Miseq system (Illumina, San Diego, CA, USA). The basic quality of sequencing reads was checked by FastQC v0.11.8, and the raw sequencing reads were trimmed and filtered with Trimmonmatic v0.38 (17). The clean reads obtained were then aligned with hg19 using BWA v0.7.17 (18). Variants were called using GATK (17). The clean reads obtained were then aligned with hg19 using BWA v0.7.17 (18). Variants were called using GATK v4.0.12.0 following GATK Best Practices. --filterExpression "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0” were used for sequencing quality control. The sequencing depth for each fragment was larger than 50x. All identified LoF mutations were further validated by Sanger sequencing. Before Sanger sequencing, FastStart™ High Fidelity PCR System (Roche Applied Science, Mannheim, Germany) was used to amplify target sequences. Sanger sequencing was performed using Applied Biosystems™ 3500xL Dx Genetic Analyzer (Life Technologies, NY, USA). All primers used have been described previously (16).

Statistical analysis

Fisher’s exact test was used to compare the allele frequencies between patients and controls. The odds ratios (ORs) of the significant mutation and the combined FLG LoF mutation genotype were estimated by logistic regression analysis.

RESULTS

All patients and controls were of Chinese Han ethnicity and were enrolled from Shandong province, China. The mean ± standard deviation (SD) age of the 945 patients with leprosy was 63.31 ± 11.48 years (range 23–90 years) and 54.6% (516) of patients were male. The mean ± SD age of the 916 controls was 68.00 ± 11.48 years (range 31–90 years), and 53.60% (491) of controls were male. No sex or age bias was observed between patients and controls.

In total, 21 LoF mutations, including 11 nonsense and 10 frameshift mutations, were discovered and validated in 945 Chinese Han patients with leprosy and 916 healthy controls (Table 1). Among them, 9 mutations had not previously been reported, namely c.12159_12162del:p.Q4053Hfs*18, c.C11227T:p.R3743X, c.G5841A:p.W1947X, c.4774dupT:p.S1592Ffs*3, c.C2185T:p.Q729X, c.G622T:p.E208X, c.428delG:p.S143Tfs*50, c.A418T:p.R140X and c.A388T:p.R130X (Fig. I). The other 12 mutations were previously reported, of which 2 were low-frequency variants (1% < minor allele frequency < 5%) and the remaining 10 were rare variants (minor allele frequency < 1%). The frequencies of these 12 known mutations in the healthy controls showed similar frequencies in public databases (Table SI1).

Table 1. Case-control analysis of 21 FLG mutations

| FLG mutation                        | Cohort | aa | Aa | AA | Total | p-value |
|-------------------------------------|--------|----|----|----|-------|---------|
| c.12159_12162del:p.Q4053Hfs*18     | Leprosy| 0  | 1  | 940| 941   | 1.000   |
| Controls                            | 0      | 915| 915|    |       |         |
| c.A12064T:p.K40222X                | Leprosy| 0  | 9  | 846| 846   | 0.016   |
| Controls                            | 31     | 851| 914|    |       |         |
| c.C11227T:p.R3743X                 | Leprosy| 0  | 2  | 943| 945   | 0.682   |
| Controls                            | 61     | 851| 914|    |       |         |
| c.7945delA:p.S2649Vfs*93           | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 915| 916|    |       |         |
| c.6834_6838del:p.S2279Tfs*39       | Leprosy| 0  | 4  | 941| 945   | 0.375   |
| Controls                            | 0      | 915| 916|    |       |         |
| c.G5841A:p.W1947X                  | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.G5383T:p.E1795X                  | Leprosy| 0  | 2  | 943| 945   | 1.000   |
| Controls                            | 0      | 915| 916|    |       |         |
| c.C5386T:p.Q1790X                  | Leprosy| 0  | 9  | 936| 945   | 0.021   |
| Controls                            | 0      | 915| 916|    |       |         |
| c.4774dupT:p.S1592Ffs*3            | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.4771_4772del:p.G1591Ifs*3        | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.C3909A:p.S1302X                  | Leprosy| 0  | 4  | 941| 945   | 0.542   |
| Controls                            | 0      | 910| 916|    |       |         |
| c.3321delA:p.G1109Efs*12           | Leprosy| 0  | 1  | 908| 945   | 0.102   |
| Controls                            | 36     | 908| 945|    |       |         |
| c.3222_3225del:p.S1074Vfs*46       | Leprosy| 0  | 6  | 939| 945   | 0.288   |
| Controls                            | 0      | 914| 916|    |       |         |
| c.2282_2285del:p.S761Ifs*35        | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.C2185T:p.Q729X                   | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.1510T:p.R501X                    | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.1248dupG:p.S417Vfs*1             | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.G622T:p.E208X                    | Leprosy| 0  | 1  | 944| 945   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.428delG:p.S143Tfs*50             | Leprosy| 0  | 1  | 942| 943   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.A418T:p.R140X                    | Leprosy| 0  | 1  | 942| 943   | 1.000   |
| Controls                            | 0      | 916| 916|    |       |         |
| c.A388T:p.R130X                    | Leprosy| 0  | 2  | 941| 943   | 0.500   |
| Controls                            | 0      | 916| 916|    |       |         |

For FLG LoF mutation combined genotype

| FLG mutation combined genotype | Leprosy | Controls |
|--------------------------------|---------|----------|
| AA                             | 153     | 123      |
| Aa                             | 776     | 786      |
| aa                             | 941     | 914      |

1https://doi.org/10.2340/00015555-3663
The overall carrier rate of 21 mutations in patients with leprosy was 17.53%, which was significantly higher than the 14.77% carrier rate in healthy controls ($p=0.041$, OR 1.31, 95% confidence interval (CI) 1.04–1.66, Table I). Two individual mutations (K4022X and Q1790X) were found to be significantly associated with leprosy ($p<0.05$). Consistent with the results of our previous study, a significant association between mutation c.A12064T:p.K4022X and leprosy was observed (10.10% carrier rate in patients vs 6.89% in controls, $p=0.016$, OR 1.48, 95% CI 1.08–2.04, Table I). The other leprosy-associated rare mutation was c.C5368T:p.Q1790X (0.95% carrier rate in patients vs 0.11% in controls, $p=0.021$, OR 8.80, 95% CI 1.11–69.58, Table I). Among the remaining 19 mutations, 14 showed higher frequency in patients with leprosy than in healthy controls, but with insignificant $p$-values (Fig. 2, Table I). In contrast, only 2 individual mutations showed the same frequency in patients with leprosy and healthy controls, and 3 individual mutations showed insignificant higher frequencies in healthy controls (Fig. 2, Table I).

Regardless of mutation frequency in patients, the distribution of the 21 mutations found across the coding sequence of FLG showed no obvious hot spots. Nevertheless, 17 of 21 mutations (80.95%) were located before the 7th repeat of exon 3 (Fig. 2).

**DISCUSSION**

By employing an array-based high-throughput deep sequencing method, the coding sequence of the filaggrin gene in 945 patients with leprosy and 916 controls was finely mapped. The overall carrier rate of 21 FLG LoF mutations was found to be significantly higher in patients...
with leprosy than in healthy controls. Two individual mutations (K4022X and Q1790X) were shown to be significantly associated with leprosy. It is well known that the FLG LoF mutations, which can cause pre-stop and frameshift of coding leading to a deficiency of filaggrin, can be directly treated as causal variants without further functional analysis (14). Thus, our findings are the first, to our knowledge, to indicate a role of the FLG gene in the pathogenesis of leprosy.

FLG sequencing remains a challenge because exon 3 of FLG is very large (>12 kb) and highly repetitive (2), which makes direct sequencing of FLG using the Sanger sequencing method difficult, costly and time-consuming. Hence, most previous studies focused only on the genotyping of known predominant LoF mutations in patients, and few studies of the sequencing of the entire FLG gene were reported in large cohorts. In recent years, efforts have been made to sequence FLG using target-region enrichment based on the next-generation sequencing (NGS) method. In 2017, a probe capture enrichment-based NGS method (Roche NimbleGen) was first used to sequence FLG in 54 Chinese Han IV patients (13). Another probe capture system-based NGS method (Agilent SureDesign, Agilent Technologies) was also developed and used for FLG sequencing in 741 children with AD (15, 19). In addition to these 2 probe capture-based enrichment methods, a high-throughput PCR enrichment method (Access Array,Fluidigm), followed by NGS, was also developed and used to sequence FLG in 334 Singaporean patients with IV or AD (16). Those studies not only confirmed the known FLG LoF mutations, but also found several novel LoF mutations. Herein, using the high-throughput PCR enrichment method, NGS of the FLG coding sequence was performed in 945 patients with leprosy and 916 healthy controls. In total, we identified 21 FLG LoF mutations and, among them, 9 were newly reported, highlighting again the efficacy of the NGS method in FLG sequencing.

Although M. leprae, the pathogenic agent of leprosy, was discovered more than 100 years ago, the transmission patterns of M. leprae are still not fully understood (20). The most widely accepted transmission pathway is the upper respiratory tract (21), but several studies have reported M. leprae infection following trauma (20), such as tattoos (22–24), injury with a glass bangle (25) and knee injury (26), which suggests the possibility of infection through impaired skin. A previous study found that FLG LoF mutations could increase the risk of recurrent skin bacterial infections in patients with AD (5). Moreover, knockdown of FLG was shown to increase the colonization of Staphylococcus aureus in an epidermal model (27). The current study found that the carrier rate of overall FLG LoF mutations and 2 individual mutations were significantly higher in patients with leprosy than in healthy controls, indicating the possible pathogenic role of the FLG gene in leprosy, which suggested that an impaired skin barrier may facilitate the transmission of M. leprae through direct skin contact.

In addition to barrier function, filaggrin has also been demonstrated to have immune functions. Filaggrin deficiencies were reported to cause immunological dysregulation, such as increased CD11c expression in Langerhans cells, increased systemic Th17 response and decreased expression of IFN-γ in keratinocytes (28–30). Considering that Langerhans cells and keratinocytes play roles in the pathogenesis of leprosy, and that Th17 cells are also an important T cell subset involved in the immune response to leprosy (31, 32), it is possible that immune dysregulation caused by FLG LoF mutations also contribute to the pathogenesis of leprosy. Further research, using animal or human epidermis models, will be needed to confirm this proposal.

Similar to other studies on FLG in AD or IV, our studies have focused only on the variants that cause pre-stop or frameshift of coding, and other types of mutations, such as missense mutations, were not included (13, 15, 16). This was based on the conventional opinion that truncated filaggrin proteins cannot function normally, as with the homozygous R501X, which has been shown to cause the loss of keratohyalin granules in the epidermis granular layer (1). In the current study, all 21 mutations could cause stop-gain and frameshift of coding, generating truncated protein. In particular, 17 of 21 (80.95%) LoF mutations located before the seventh repeat of exon 3 could cause at least a 40% loss in exon 3, with severe effects on filaggrin function. Another limitation of this study is that the p-value (p=0.041) was not corrected for multiple tests, which made it necessary to verify our discovery in other populations and samples with larger sample size.

In summary, a comprehensive analysis of FLG and leprosy using NGS sequencing was performed in 945 patients with leprosy and 916 healthy controls, revealing a significant association between FLG LoF mutations and leprosy. Further investigations will be required to reveal the potential barrier and immunological roles of filaggrin in leprosy pathogenesis.

ACKNOWLEDGEMENTS

The authors thank all the participants involved in this study. We thank XFW Wong and JC Common from the Common of Institute of Medical Biology, A*STAR, Singapore, for their technical support.

This work was funded by the National Key Research and Development Program of China (2016YFE0201500), the Academic Promotion Programme of Shandong First Medical University (2019LJ002, 2019RC007), the National Natural Science Foundation of China (81822038), the Natural Science Foundation of Shandong Province (ZR2018BC020), and the Shandong Province Taishan Scholar Project (tspd20150214, tsqn201812124).

The authors have no conflicts of interest to declare.
REFERENCES

1. 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: 337–342.

2. 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: 441–446.

3. Yu HS, Kang MJ, Jung YH, Kim HY, Seo JH, Kim YJ, et al. Mutations in the filaggrin are predisposing factor in Korean children with atopic dermatitis. Allergy Asthma Immunol Res 2013; 5: 211–215.

4. Zhang H, Guo Y, Wang W, Shi M, Chen X, Yao Z. Mutations in the filaggrin gene in Han Chinese patients with atopic dermatitis. Allergy 2011; 66: 420–427.

5. Cai SC, Chen H, Koh WP, Common JE, van Bever HP, McLean WH, et al. Filaggrin mutations are associated with recurrent skin infection in Singaporean Chinese patients with atopic dermatitis. Br J Dermatol 2012; 166: 200–203.

6. Liu H, Wang Z, Li Y, Yu G, Fu X, Wang C, et al. Genome-wide analysis of protein-coding variants in leprosy. J Invest Dermatol 2017; 137: 2544–2551.

7. Han XY, Silva FJ. On the age of leprosy. PLoS Negl Trop Dis 2014; 8: e2544.

8. Cardoso CC, Pereira AC, de Sales Marques C, Moraes MO. Leprosy susceptibility: genetic variations regulate innate and adaptive immunity, and disease outcome. Future Microbiol 2011; 6: 533–549.

9. Mazini PS, Alves HV, Reis PG, Lopes AP, Sell AM, Santos-Rosa M, et al. Gene association with leprosy: a review of published data. Front Immunol 2015; 6: 658.

10. Sauer ME, Salomao H, Ramos GB, D’Espindula HR, Rodrigues RS, Macedo WC, et al. Genetics of leprosy: expected and unexpected developments and perspectives. Clin Dermatol 2015; 33: 99–107.

11. Wang Z, Sun Y, Fu X, Yu G, Wang C, Bao F, et al. A large-scale genome-wide association and meta-analysis identified four non-synonymous filaggrin gene variants in Chinese Han patients with ichthyosis vulgaris. Br J Dermatol 2017; 177: e202–e203.

12. Zhang FR, Huang W, Chen SM, Sun LD, Liu H, Li Y, et al. Genome wide analysis of protein-coding variants in leprosy. J Invest Dermatol 2018; 138: 1501–1506.

13. Margolis DJ, Mitra N, Gochnauer H, Wubbenhorst B, D’Andrea K, Kraya A, et al. Uncommon filaggrin variants are associated with persistent atopic dermatitis in African Americans. J Invest Dermatol 2013; 131: 1247–1250, 1250 e1241.

14. Irvine AD, McLean WH, Leung DY. Filaggrin mutations are associated with recurrent skin infection in Singaporean Chinese patients with atopic dermatitis. Br J Dermatol 2012; 166: 200–203.

15. Margolis DJ, Mitra N, Gochnauer H, Wubbenhorst B, D’Andrea K, Kraya A, et al. Uncommon filaggrin variants are associated with persistent atopic dermatitis in African Americans. J Invest Dermatol 2013; 131: 1247–1250, 1250 e1241.

16. Wong X, Denil S, Foo JN, Chen H, Tay ASL, Haines RL, et al. Array-based sequencing of filaggrin gene for comprehensive detection of disease-associated variants. J Allergy Clin Immunol 2017; 141: 814–816.

17. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014; 30: 2114–2120.

18. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009; 25: 1754–1760.

19. Margolis DJ, Mitra N, Gochnauer H, Wubbenhorst B, D’Andrea K, Kraya A, et al. Uncommon filaggrin variants are associated with persistent atopic dermatitis in African Americans. J Invest Dermatol 2018; 138: 1501–1506.

20. Bratschi MW, Steinmann P, Wickenden A, Gillis TP. Current knowledge on mycobacterium leprae transmission: a systematic literature review. Lepr Rev 2015; 86: 142–155.

21. Lastoria JC, Abreu MA. Leprosy: review of the epidemiological, clinical, and etiopathogenic aspects – part 1. An Bras Dermatol 2014; 89: 205–218.

22. Ghorpade A. Inoculation (tattoo) leprosy: a report of 31 cases. J Eur Acad Dermatol Venereol 2002; 16: 494–499.

23. Singh G, Tutakne MA, Tiwari VD, Dutta RK. Inoculation leprosy developing after tattooing – a case report. Indian J Lepr 1985; 57: 887–888.

24. Singh RK. Tattoos and paucibacillary leprosy. Travel Med Infect Dis 2009; 7: 325–326.

25. Ghorpade A. Post-traumatic inoculation tuberculoid leprosy after injury with a glass bangle. Lepr Rev 2009; 80: 215–218.

26. Brandsma JW, Yoder L, MacDonald M. Leprosy acquired by inoculation from a knee injury. Lepr Rev 2005; 76: 175–179.

27. van Drongelen V, Haisma EM, Out-Luiting JJ, Nibbering PH, El Ghabbouri A, Reduced filaggrin expression is accompanied by increased Staphylococcus aureus colonization of epidermal skin models. Clin Exp Allergy 2014; 44: 1515–1524.

28. Bonefeld CM, Petersen TH, Bandier J, Agerbeck C, Linneberg A, Ross-Hansen K, et al. Epidermal filaggrin deficiency mediates increased systemic T-helper 17 immune response. Br J Dermatol 2016; 175: 706–712.

29. Dang NN, Pang SG, Song HY, An LG, Ma XL. Filaggrin silencing by shRNA directly impairs the skin barrier function of normal human epidermal keratinocytes and then induces an immune response. Braz J Med Biol Res 2015; 48: 39–45.

30. Leisten S, Oyoshi MK, Galand C, Hornick JL, Gurish MF, Geha RS. Development of skin lesions in filaggrin-deficient mice is dependent on adaptive immunity. J Allergy Clin Immunol 2017; 141: 814–816.

31. de Sousa JR, Sotto MN, Simoes Quaresma JA. Leprosy as a complex infection: breakthrough of the Th1 and Th2 immune paradigm in the immunopathogenesis of the disease. Front Immunol 2017; 8: 1635.

32. Pinheiro RO, Schmitz V, Silva BJA, Dias AA, de Souza BJ, de Mattos Barbosa MG, et al. Innate Immune Responses in Leprosy. Front Immunol 2018; 9: 518.