Data in brief 29 (2020) 105307

Contents lists available at ScienceDirect

Data in brief

journal homepage: www.elsevier.com/locate/dib

Data Article

# Dataset of allele, genotype and haplotype frequencies of four polymorphisms filaggrin gene in Russian patients with atopic dermatitis

Tatyana Belyaeva [a], Irina Ponomarenko [a], Evgeny Reshetnikov [a, *], Alexey Polonikov [b], Inna Aristova [a], Anna Elykova [a], Natalya Rudykh [a], Mikhail Churnosov [a]

[a] Department of Medical Biological Disciplines, Belgorod State University, 308015, Belgorod, Russia
[b] Department of Biology, Medical Genetics and Ecology, Kursk State Medical University, 305041, Kursk, Russia

ARTICLE INFO

*Article history:*
Received 8 January 2020
Received in revised form 11 February 2020
Accepted 13 February 2020
Available online 21 February 2020

*Keywords:*
Single nucleotide polymorphism
Atopic dermatitis
Female
Male
*FLG*

ABSTRACT

Data on the allele, genotype and haplotype frequencies of four single nucleotide polymorphisms (SNPs) (rs3126085, rs12144049, rs471144 and rs4363385) *filaggrin* (*FLG*) gene in Russian patients with atopic dermatitis are presented. Genome-wide association studies identified these SNPs could be significant genetic markers associated with atopic dermatitis. The frequencies of alleles, genotypes and haplotypes of four SNPs were calculated in 3 groups: entire sample, females and males. No significant differences in the allele, genotype and haplotype frequencies between males and females with AD patients were observed.



* Corresponding author.
*E-mail address:* reshetnikov@bsu.edu.ru (E. Reshetnikov).

https://doi.org/10.1016/j.dib.2020.105307

Specifications Table

| | |
|---|---|
| Subject | *Biology* |
| Specific subject area | *Genetics* |
| Type of data | *Table and figure* |
| How data were acquired | *MALDI/TOF mass spectrometry using Sequenom MassARRAY 4.0 platform (Agena Bioscience™)* |
| Data format | *Raw and analyzed data* |
| Parameters for data collection | *Total genomic DNA was isolated from buffy coat using the standard phenol-chloroform method.* |
| Description of data collection | *DNA samples were genotyped using the Sequenom MassARRAY® iPLEX platform, which is based on MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometry* |
| Data source location | *Belgorod, Russia* |
| Data accessibility | *The data is available with this article* |

**Value of the Data**

- The frequencies of alleles, genotypes and haplotypes of rs3126085, rs12144049, rs471144 and rs4363385 of the FLG gene in Russian were not differed between males and females with AD.
- The polymorphisms at the FLG gene may associate with atopic dermatitis.
- The allele, genotype and haplotype frequencies are an important data for understanding the genetic architecture of different populations.
- The data can be used for studying the genetic basis of atopic dermatitis and other skin (i.e. psoriasis) or allergic disease (i.e. asthma) in various populations of the world.

## 1. Data description

The dataset represents the raw data (supplementary Table), frequencies of alleles, genotypes (Table 1) and haplotypes (Fig. 1, Table 2) for four single nucleotide polymorphisms (SNPs) (rs3126085,

**Table 1**
The frequencies of alleles and genotypes for single nucleotide polymorphisms (SNPs) rs3126085, rs12144049, rs471144 and rs4363385 in *FLG* gene in Russian patients with atopic dermatitis.

| SNP genotype or allele | All (n = 350) | | Female (n = 237) | | Male (n = 113) | |
|---|---|---|---|---|---|---|
| | n | frequency | n | frequency | n | frequency |
| rs3126085 | | | | | | |
| AA | 12 | 0.0342 | 9 | 0.0380 | 3 | 0.0265 |
| GA | 77 | 0.2200 | 52 | 0.2194 | 25 | 0.2212 |
| GG | 261 | 0.7457 | 176 | 0.7426 | 85 | 0.7522 |
| A | 101 | 0.1443 | 70 | 0.1477 | 31 | 0.1372 |
| G | 599 | 0.8557 | 404 | 0.8523 | 195 | 0.8628 |
| rs12144049 | | | | | | |
| GG | 26 | 0.0743 | 18 | 0.0759 | 8 | 0.0708 |
| AG | 109 | 0.3114 | 76 | 0.3207 | 33 | 0.2920 |
| AA | 215 | 0.6143 | 143 | 0.6034 | 72 | 0.6372 |
| G | 161 | 0.2300 | 112 | 0.2363 | 49 | 0.2168 |
| A | 539 | 0.7700 | 362 | 0.7637 | 177 | 0.7832 |
| rs471144 | | | | | | |
| TT | 4 | 0.0114 | 4 | 0.0169 | 0 | 0.0000 |
| GT | 42 | 0.1200 | 31 | 0.1308 | 11 | 0.0973 |
| GG | 304 | 0.8686 | 202 | 0.8523 | 102 | 0.9027 |
| T | 50 | 0.0714 | 39 | 0.0823 | 11 | 0.0487 |
| G | 650 | 0.9286 | 435 | 0.9177 | 215 | 0.9513 |
| rs4363385 | | | | | | |
| AA | 66 | 0.1886 | 45 | 0.1899 | 21 | 0.1858 |
| GA | 165 | 0.4714 | 113 | 0.4768 | 52 | 0.4602 |
| GG | 119 | 0.3400 | 79 | 0.3333 | 40 | 0.3540 |
| A | 297 | 0.4243 | 203 | 0.4283 | 94 | 0.4159 |
| G | 403 | 0.5757 | 271 | 0.5717 | 132 | 0.5841 |

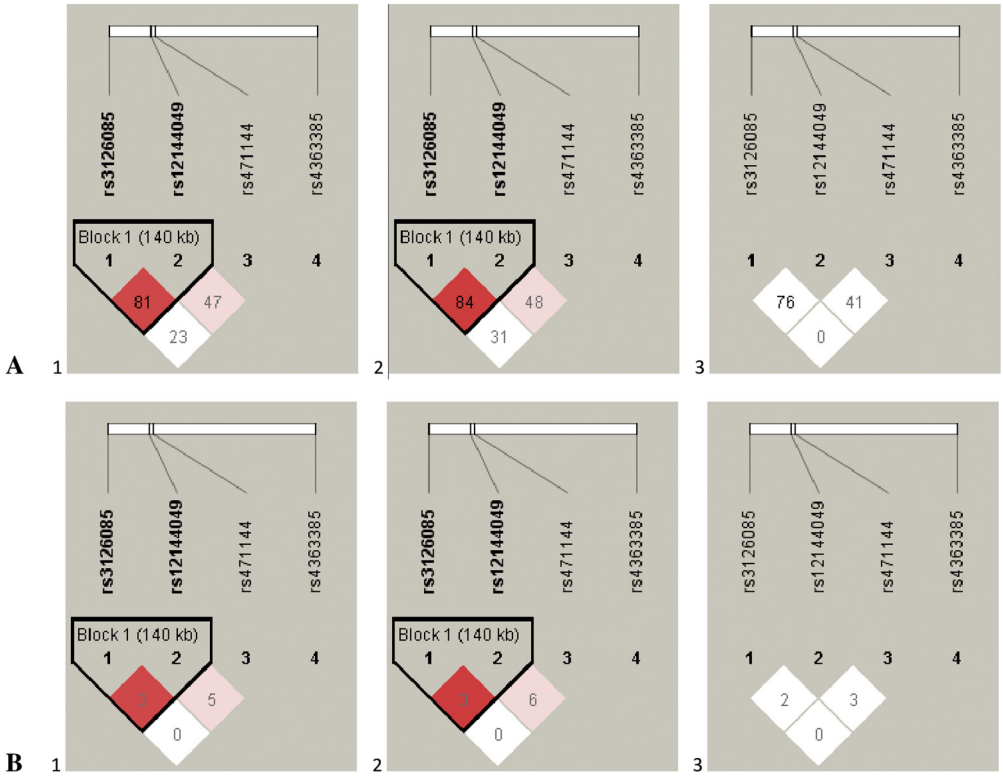


**Fig. 1.** Linkage disequilibrium (LD) between SNPs rs3126085, rs12144049, rs471144, and rs4363385 of the *FLG* gene in Russian patients with atopic dermatitis. LD measures are presented as Lewontin's standardized coefficient D′ (Figure, A) and the square of the correlation Pearson's coefficient ($r^2$) between SNPs (Figure, B). D′ values vary gradually from white color (D′ = 0, no LD between SNPs) to dark red (D′ = 1, SNPs are in complete LD). Figure sections 1, 2 and 3 represent entire sample, females, and males, respectively.

rs12144049, rs471144 and rs4363385) *filaggrin* (*FLG*) gene in Russian patients with atopic dermatitis (AD). These SNPs were associated with AD in previously published genome-wide association studies (GWAS) (Table 3) and also candidate gene studies https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5738205/ [1–5], have significant regulatory potential (Table 4) and influence gene expression level (Table 5). The dataset frequencies of the SNP alleles, genotypes and haplotypes were divided into three groups: entire sample, females and males. The minor allele frequency (MAF) for rs3126085 – 0.1443 (female – 0.1477, male – 0.1372), rs12144049 – 0.2300 (female – 0.2363, male – 0.2168), rs471144 – 0.0714 (female – 0.0823, male – 0.0487) and rs4363385 – 0.4243 (female – 0.4283, male – 0.4159). No significant differences in the allele, genotype and haplotype frequencies were found between males and females with AD patients.

## 2. Experimental design, materials, and methods

### 2.1. Subjects selection

During a period between 2010 and 2016, AD patients were recruited at Dermatovenerologic dispensaries of Belgorod and Kursk regions (Russia). AD was diagnosed by experienced dermatologists according to the UK Diagnostic Criteria [6]. The participants were unrelated Russians born in the

**Table 2**
The frequencies of haplotypes for haploblock of single nucleotide polymorphisms (SNPs) rs3126085 and rs12144049 in *FLG* gene in Russian patients with atopic dermatitis.

| Haplotype (rs3126085 and rs12144049) | All (n = 350), frequency | Female (n = 237), frequency | Male (n = 113), frequency |
|---|---|---|---|
| GA | 0.638 | 0.626 | 0.656 |
| GG | 0.224 | 0.232 | 0.212 |
| AA | 0.138 | 0.142 | 0.132 |

**Table 3**
The literature data about associations of the studied polymorphisms with atopic dermatitis (GWAS data).

| SNP | Position (hg38) | Association (significance) | Reference |
|---|---|---|---|
| rs3126085 | 152,328,341 | OR = 1.22 ($p = 6 \times 10^{-12}$) | [1] |
| rs12144049 | 152,468,434 | OR = 1.53 ($p = 3 \times 10^{-30}$) | [2] |
| | | OR = 1.39 ($p = 1 \times 10^{-16}$) | [3] |
| rs471144 | 152,481,779 | OR = 1.54 ($p = 2 \times 10^{-12}$) | [2] |
| rs4363385 | 153,016,845 | OR = 1.23 ($p = 2 \times 10^{-17}$) | [2] |

Central Russia [7]. The exclusion criteria were as follows: malignant tumors, severe autoimmune diseases, chronic severe diseases of the vital organs (heart, respiratory or renal failure). A total of 350 patients with AD (237 female and 113 male) met these criteria. This work was approved by the Regional Ethics Committee of Belgorod State University and informed consents were obtained from all participants.

### 2.2. DNA analysis

The procedures of whole blood sampling, genomic DNA isolation were described elsewhere [8].

Four SNPs in the *FLG* gene such as rs3126085, rs12144049, rs471144 and rs4363385 were selected for the analysis according to the following criteria [9]: 1) a SNP was reported to be associated with AD risk by genome-wide association, 2) SNP possesses a regulatory potential (regSNP), 3) SNP is associated with changes in gene expression (eSNP), and 4) MAF $\geq$ 5%.

The selected SNPs were found to be associated with the risk of AD, as previously reported by genome-wide association studies (Table 3) and were found to be functionally significant polymorphisms, i.e. they possess significant regulatory potential (Table 4), as determined by the HaploReg online tools, v4.1 update 05.11.2015 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php), and have impact on gene expression level (Table 5), as determined by the GTExportal, (http://www.gtexportal.org).

DNA samples were genotyped using the MALDI-TOF mass spectrometry iPLEX platform (Agena Bioscience Inc, San Diego, CA). To ensure quality control of genotyping blind replicates were included. Laboratory personnel involved in genotyping were completely blinded to patients' information. The repeatability test for 5% of randomly selected samples was performed, yielded 100% reproducibility.

### 2.3. Statistical analysis

Genotypes for the polymorphisms were evaluated regarding their accordance to Hardy-Weinberg equilibrium (HWE) using the chi-square test. Differences in allele, genotype and haplotype frequencies between females and males with AD were assessed by the Kruskall-Wallis test. The linkage disequilibrium (LD) between rs3126085, rs12144049, rs471144 and rs4363385 FLG gene was analyzed using Haploview version 4.2 software (https://www.broadinstitute.org/haploview/haploview). The LD block structure was determined using the Solid Spine of the LD algorithm [10] provided by the Haploview 4.2. The degree of genetic linkage between the 4 SNPs in 3 groups was estimated as Lewontin's coefficient D′ and squared Pearson's correlation coefficient $r^2$. D′ values vary gradually from white color (D′ = 0, no LD between SNPs) to dark red (D′ = 1, SNPs are in complete LD). (Fig. 1).

**Table 4**
Regulatory effects of the 4 SNPs of the FLG gene (HaploReg, v4.1, update 05.11.2015) (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php).

| chr | pos (hg38) | variant | Ref | Alt | AFR | AMR | ASN | EUR | SiPhy | Promoter | Enhancer | DNAse | Proteins | Motifs | NHGRI/EBI | GRASP QTL | Selected eQTL | GENCODE | dbSNP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | freq | freq | freq | freq | cons | histone marks | histone marks | | bound | changed | GWAS hits | hits | hits | genes | func annot |
| 1 | 152328341 | rs3126085 | G | A | 0.53 | 0.36 | 0.59 | 0.15 | | | 7 tissues | | | Foxp3,TEF | 1 hit | 1 hit | 26 hits | FLG-AS1 | intronic |
| 1 | 152468434 | rs12144049 | C | T | 0.67 | 0.8 | 0.76 | 0.74 | | | | | | Irf,Obox6,ZEB1 | | 2 hits | 1 hit | 23kb 3′ of RP1-91G5.3 | |
| 1 | 152481779 | rs471144 | T | G | 0.06 | 0.08 | 0.18 | 0.08 | | LIV, GI | | | | 7 altered motifs | | | 1 hit | 29kb 5′ of LCE5A | |
| 1 | 153016845 | rs4363385 | T | C | 0.78 | 0.54 | 0.66 | 0.59 | | | | | | | | 3 hits | 7 hits | 4.2kb 5′ of SNORA31 | |

**Table 5**
The cis-eQTL values of the 4 SNPs of the FLG gene in skin (according to Genotype-Tissue Expression (GTEx) (http://www.gtexportal.org/)).

| SNP | Gene expression | Reference allele | Alternative allele | Effect Size (β) | *P*-Value | Tissue |
|---|---|---|---|---|---|---|
| rs3126085 | *FLG* | G | A | −0.22 | 0.000000037 | Skin - Sun Exposed (Lower leg) |
| | *RP1-91G5.3* | G | A | 0.38 | 0.00002 | Skin - Sun Exposed (Lower leg) |
| rs12144049 | *CRNN* | C | T | −0.3 | 0.00000000096 | Skin - Sun Exposed (Lower leg) |
| | *CRNN* | C | T | −0.32 | 0.00000003 | Skin - Not Sun Exposed (Suprapubic) |
| rs471144 | *FLG-AS1* | T | G | −0.51 | 0.000023 | Skin - Not Sun Exposed (Suprapubic) |
| rs4363385 | *SPRR1B* | T | C | −0.25 | 0.00000000011 | Skin - Sun Exposed (Lower leg) |
| | *SPRR2D* | T | C | 0.27 | 0.0000000045 | Skin - Sun Exposed (Lower leg) |
| | *LCE3C* | T | C | −0.34 | 0.000000033 | Skin - Sun Exposed (Lower leg) |
| | *LCE3C* | T | C | −0.34 | 0.000000033 | Skin - Not Sun Exposed (Suprapubic) |
| | *SPRR2B* | T | C | −0.24 | 0.00000081 | Skin - Sun Exposed (Lower leg) |
| | *LCE1D* | T | C | −0.27 | 0.0000014 | Skin - Not Sun Exposed (Suprapubic) |
| | *SPRR1B* | T | C | −0.18 | 0.0000025 | Skin - Not Sun Exposed (Suprapubic) |
| | *LCE1D* | T | C | −0.24 | 0.0000046 | Skin - Sun Exposed (Lower leg) |
| | *SPRR2B* | T | C | −0.24 | 0.000014 | Skin - Not Sun Exposed (Suprapubic) |

## Acknowledgements

This is a self-funded work with no external sponsorship.

## Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2020.105307.

## References

[1] L.D. Sun, F.L. Xiao, Y. Li, W.M. Zhou, H.Y. Tang, X.F. Tang, H. Zhang, H. Schaarschmidt, X.B. Zuo, R. Foelster-Holst, et al., Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population, Nat. Genet. 43 (2011) 690–694.

[2] H. Baurecht, M. Hotze, S. Brand, C. Büning, P. Cormican, A. Corvin, D. Ellinghaus, E. Ellinghaus, J. Esparza-Gordillo, R. Fölster-Holst, A. Franke, C. Gieger, N. Hubner, T. Illig, A.D. Irvine, M. Kabesch, Y.A. Lee, W. Lieb, I. Marenholz, W.H. McLean, D.W. Morris, U. Mrowietz, R. Nair, M.M. Nöthen, N. Novak, G.M. O'Regan, , Psoriasis Association Genetics Extension, S. Schreiber, C. Smith, K. Strauch, P.E. Stuart, R. Trembath, L.C. Tsoi, M. Weichenthal, J. Barker, J.T. Elder, S. Weidinger, H.J. Cordell, S.J. Brown, Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms [published correction appears in Am J Hum Genet. 2015 Dec 3;97(6):933], Am. J. Hum. Genet. 96 (2015) 104–120, https://doi.org/10.1016/j.ajhg.2014.12.004.

[3] H. Schaarschmidt, D. Ellinghaus, E. Rodríguez, A. Kretschmer, H. Baurecht, S. Lipinski, U. Meyer-Hoffert, J. Harder, W. Lieb, N. Novak, R. Fölster-Holst, J. Esparza-Gordillo, I. Marenholz, F. Ruschendorf, N. Hubner, E. Reischl, M. Waldenberger, C. Gieger, T. Illig, M. Kabesch, X.J. Zhang, F.L. Xiao, Y.A. Lee, A. Franke, S. Weidinger, A genome-wide association study reveals 2 new susceptibility loci for atopic dermatitis, J. Allergy Clin. Immunol. 136 (2015) 802–806, https://doi.org/10.1016/j.jaci.2015.01.047.

[4] T.M. Belyaeva, Role of interaction of polymorphic loci of the *FLG* gene in the formation of chronic true eczema in women, Res. Results Biomed. 5 (4) (2019) 20–30, https://doi.org/10.18413/2658-6533-2019-5-4-0-1 (In Russian).

[5] C. Shen, L. Liu, Z. Jiang, X. Zheng, L. Meng, X. Yin, J. Gao, Y. Sheng, J. Gao, Y. Li, F. Zhou, F. Xiao, L. Sun, Y. Cui, S. Yang, X. Zuo, X. Zhang, Four genetic variants interact to confer susceptibility to atopic dermatitis in Chinese Han population, Mol. Genet. Genom. 290 (2015) 1493–1498, https://doi.org/10.1007/s00438-015-1014-x.

[6] H.C. Williams, P.G. Burney, A.C. Pembroke, R.J. Hay, The U.K. working party's diagnostic criteria for atopic dermatitis. III. Independent hospital validation, Br. J. Dermatol. 131 (1994) 406–416.

[7] I.N. Sorokina, N.A. Rudykh, I.N. Bezmenova, I.S. Polyakova, Population genetic characteristics and genetic epidemiological research of candidate genes associations with multifactorial diseases, Res. Results Biomed. 4 (4) (2018) 20–30, https://doi.org/10.18413/2313-8955-2018-4-4-0-3 (In Russian).
[8] I. Ponomarenko, E. Reshetnikov, O. Altuchova, A. Polonikov, I. Sorokina, A. Yermachenko, V. Dvornyk, M. Churnosov, Association of genetic polymorphisms with age at menarche in Russian women, Gene 686 (2019) 228–236, https://doi.org/10.1016/j.gene.2018.11.042.
[9] I.V. Ponomarenko, Selection of polymorphic loci for association analysis in genetic-epidemiological studies, Res. Result Med. Pharm. 4 (2) (2018) 40–54, https://doi.org/10.18413/2313-8955-2018-4-2-0-5 (in Russian).
[10] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, Bioinformatics 21 (2005) 263–265.