HHS Public Access

Author manuscript

Exp Dermatol. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:

Exp Dermatol. 2018 September ; 27(9): 989–992. doi:10.1111/exd.13691.

Tiled array-based sequencing identifies enrichment of loss-of-function variants in the highly homologous *filaggrin* gene in African American children with severe atopic dermatitis

Mary Elizabeth Mathyer, BA[1,2,3], **Ashley M. Quiggle, PhD**[1,2,3], **X.F. Colin C. Wong, BSc**[4], **Simon L.I.J. Denil, PhD**[4], **Monique G. Kumar, MD**[1], **Heather M. Ciliberto, MD**[1], **Susan J. Bayliss, MD**[1], **John E. Common, PhD**[4], and **Cristina de Guzman Strong, PhD**[1,2,3]

[1]Division of Dermatology, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA

[2]Center for Pharmacogenomics, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA

[3]Center for the Study of Itch, Washington University School of Medicine, St. Louis, MO, USA

[4]Institute of Medical Biology, *A*STAR*, Singapore 138648, Singapore

Abstract

Filaggrin (FLG) loss of function (LOF) variants are a major risk factor for the common inflammatory skin disease, atopic dermatitis (AD), and are often population-specific. African American (AA) children are disproportionately affected with AD, often later developing asthma and/or allergic rhinitis and comprise an atopy health disparity group for which the role of *FLG* LOF is not well known. Discovery of *FLG* LOF using exome sequencing is challenging given the known difficulties for accurate short-read alignment to *FLG*'s high homology repeat variation. Here we employed an array-based sequencing approach to tile across each *FLG* repeat and discover *FLG* LOF in a well-characterized cohort of AA children with moderate to severe AD. Five *FLG* LOF were identified in 23% of our cohort. Two novel *FLG* LOF singletons, c.488delG and p.S3101*, were discovered as well as p.R501*, p.R826*, and p.S3316* previously reported for AD. p.S3316* (rs149484917) is likely an African ancestral *FLG* LOF, reported in African individuals in ExAC (Exome Aggregation Consortium), Exome Variant Server (ESP), and 4 African 1000G population databases (ESN, MSL, ASW, and ACB). The proportion of *FLG* LOF (11.5%) among the total *FLG* alleles in our cohort was significantly higher in comparisons to *FLG* LOF reported for African individuals in ExAC (2.5%; $P = 4.3\times10^{-4}$) and ESP (1.7%; $P = 3.5\times10^{-5}$) suggesting a disease-enrichment effect for *FLG* LOF. Our results demonstrate the utility of array-based sequencing in discovering *FLG* LOF, including novel and population-specific, that are of higher prevalence in our AA severe AD group than previously reported.

Corresponding Author: Cristina de Guzman Strong, PhD, Division of Dermatology, Center for Pharmacogenomics, Center for the Study of Itch, Department of Medicine, Washington University School of Medicine, 4523 Scott Avenue, Campus Box 8123, St. Louis, MO 63110, USA, Tel: (314) 362-7695. Fax: (314) 362-8159, cristinastrong@wustl.edu.

Keywords

Atopic dermatitis; African American; Filaggrin; Health disparity

INTRODUCTION

Atopic dermatitis (AD) is a common and chronic inflammatory skin disease with dry skin and pruritus [1]. AD greatly impacts 15–20% of the pediatric population and 1–3% of adults worldwide [2]. As the incidence of AD worldwide has almost tripled in the past few decades [3], the prevalence also continues to increase in young children and in low-income countries [2]. In the U.S. alone, up to 16% of African American children are disproportionately affected with AD in comparison to 9.8% reported for Caucasian American children and for which the risk factors are not clear [4]. Loss-of-function (LOF) variants in the skin barrier gene filaggrin (*FLG*), initially discovered as causative in ichthyosis vulgaris (dry, scaly skin) in Europeans [5], are a major risk factor for AD and for other associative atopic diseases involved in the atopic march including asthma, allergic rhinitis, and food allergies [1]. *FLG* is primarily expressed in differentiated epidermal keratinocytes with *FLG* LOF variants resulting in filaggrin haploinsufficiency associated with skin barrier defects. Copy number variations (CNV) in the number of intragenic filaggrin monomer repeats comprise 3 different *FLG* alleles (10-, 11-, or 12-repeats) [6]. The *FLG* 10-repeat allele harbors 10 repeats whereas the 11-repeat allele contains either a duplication of the 8^{th} repeat or the 10^{th} repeat and the 12-repeat allele possess both the 8^{th} and 10^{th} duplicated repeats. *FLG* CNV has also been linked to disease risk with the addition of each *FLG* repeat reducing the odds of developing AD by 0.88 per repeat [1, 7, 8]. To date, 110 population-specific *FLG* LOF variants have been reported in AD (Supplemental Table 1). Specifically, multiple and rare *FLG* LOF have been observed in Asian populations in contrast to more common *FLG* LOF observed in European populations [9]. Together, the data suggests recent and parallel emergences of these variants in each population that are also not well understood.

The determination of the prevalence of pathogenic *FLG* LOF variants in African Americans (AA) with AD has been a decade-long, active area of investigation. A previous study in search of four common European *FLG* LOF (R501*, 2282del4, R2447*, and S3247*) found 5.8% of AA AD with *FLG* LOF [10]. A higher prevalence of the same European *FLG* LOF (22%) was identified among 11 AA children with both AD and ichthyosis vulgaris despite only 11% of *FLG* open reading frame being Sanger-sequenced [11]. Together, these studies identified lower percentages of known pathogenic *FLG* LOF in AA children in contrast to the ~50% frequency reported in moderate-severe AD European cohorts [1] but were limited in detecting additional *FLG* LOF.

Discovery of *FLG* LOF using exome and whole genome sequencing poses its challenges for the following reasons. *FLG* is one of the known high homology genes to determine variants owing to the difficulties in the accurate alignments of short read sequencing data for *FLG [12]*. Furthermore, careful consideration of the appropriate reference allele for FLG for either 10-, 11-, or 12-repeats is needed to properly determine variants, namely in the duplicated 8^{th} and 10^{th} repeats found in either of the 11-repeat and in both for the 12-repeat

alleles. The conventional human reference allele for *FLG*, ENSG00000143631, to which most LOF mutations are mapped, is the 10-repeat allele. A recent exome sequencing study in African American children with AD with high read depth (185X) reported a 6.3% proportion of *FLG* LOF alleles, including 6 newly reported and 3 previously described for AD to suggest multiple and perhaps rare *FLG* LOF in this ethnic group [13] (Supplemental Table 1). In support of this finding, the allele frequency for the total number of *FLG* LOF is 2.5% for African individuals in the Exome Aggregation Consortium (ExAC) (whereby each LOF is <1%) and 1.7% in the Exome Variant Server (ESP) for African Americans with no documented knowledges of AD in these datasets [14] (Supplementary Tables 2, 3). Moreover, the prevalence of *FLG* LOF is higher in African individuals compared to those of European and South and Central American but lower than South and East Asian populations (1000 Genomes Phase 3) (Figure 2 of Eaaswarkhanth *et al. [15]*). Together, the population genetics data further supports the global existence of *FLG* LOF variants in individuals of African ancestry that warrants further investigation with respect to AD disease.

We previously identified that fewer filaggrin monomer repeats correlate with more severe AD in our AA pediatric patients whose AD was well-characterized according to the widely accepted AD criteria by the United Kingdom Working Party [UKWP] [7]. We sought to discover *FLG* LOF variants in our AA AD patients (n=39) using our array-based targeted sequencing approach for *FLG* with alignment to the human 12-repeat allele [16] in contrast to the human reference 10-repeat allele [ExAC and EPS]. Our method enables concurrent discoveries for both intragenic *FLG* CNV and *FLG* variant detection that were validated by Sanger sequencing and long-range PCR, respectively, as previously described [16].

METHODS

Patients

African American children with atopic dermatitis according to the UK Working Party criteria for AD [17] were recruited to the study as previously described [7]. The study was approved by the Washington University in St. Louis Institutional Review Board and conforms to the US Federal Policy for the Protection of Human Subjects. All persons gave informed consent prior to his or her inclusion in the study. Disease severity was previously assessed using SCORAD [18] with moderate AD (>25) and severe AD (>50).

FLG Array-based Sequencing

Targeted sequencing for *FLG* was performed as previously described [16]. Briefly, a set of 48 amplicons that overlaps and tiles across the entire *FLG* coding gene from each patient was generated using the Fluidigm Access Array 48/48 chip and sequenced using Illumina MiSeq 2×250 bp reads mapped to the *FLG* 12-repeat allele with at least 100X coverage and confirmed by Sanger sequencing.

RESULTS

FLG LOF variants were identified in 23% of our AA AD patients (9 out of 39 recruited; eligible age range 3 months to 18 years). A total of five *FLG* LOF were identified in 9 heterozygous AA AD patients (c.488delG, p.R501*, p.R826*, p.S3101*, and p.S3316*;

Table 1). c.488delG and p.S3101* are novel *FLG* LOF as they have not been reported in dbSNP, ExAC, and ESP variant databases. p.S3316* (also known as rs149484917) was found in 3 unrelated patients and had been previously reported in a separate AA AD group [13]. The higher proportion of *FLG* LOF in our AA AD group (11.5%) compared to the 2.5% African ExAC frequency indicated an enrichment for *FLG* LOF for AA AD and was statistically significant (p=4.3E-04, Fisher's Exact Test). Similarly, the significance was further demonstrated in comparison to the 1.7% *FLG* LOF in ESP for African Americans (p=3.5E-05, Fisher's Exact Test).

All patients with *FLG* LOF (age≥1 year) reported having at least one additional allergic disease: asthma, allergic rhinitis, or food allergies and either xerosis (n=8) or ichthyosis vulgaris (n=1, c.488delG) and a family history of atopy in at least one first-degree relative. Seven of the nine *FLG* LOF patients (77%) exhibited severe AD (SCORAD >50). Patients with *FLG* LOF exhibited a trend toward a greater percentage of the body affected with the disease (SCORAD component A mean, 56) compared to non-*FLG* LOF patients (mean, 40) but this was not significant (p=0.13, two-tailed t-test, data not shown) and warrants further investigation to test this emerging hypothesis.

We sought to further investigate the distribution and frequency of rs149484917 that results in *FLG* LOF S3316* using ExAC, ESP, and 1000G population datasets given the concurrent findings for AA AD in our cohort and a previous study [13]. rs149484917 has been reported primarily in African individuals with 96% and 100% of the minor allele reported in Africans in ExAC and ESP, respectively. A closer investigation into the 1000G populations revealed that rs149484917 was not found in any European, American, East Asian, or South Asian populations. However, of the seven 1000G African (AFR) ancestral populations, rs149484917 was reported as a less common variant in 4 AFR populations; ESN (Esan in Nigeria) (Minor allele frequency, [MAF], 0.02), MSL (Mende in Sierra Leone) (0.012), ASW (African ancestry in Southwest US) (0.016), and ACB (African Caribbean in Barbados) (0.010). Taken together, the population genetics data identifies rs149484917 (S3316*) as a population-specific *FLG* LOF unique to several populations of African ancestry.

We next sought to determine the genetic background from which our newly discovered *FLG* LOF mutations arose in the context of the 10-, 11-, or 12-repeat *FLG* alleles [1]. The 10-repeat allele is the major allele in Africans (73%) and is interpreted to be ancestral [15]. c.488delG was found in the incomplete, N-terminal repeat (repeat 0) [7] and inferred to be on the 10-repeat allele as the patient was homozygous for the 10-repeat allele (Table 1). The other novel *FLG* LOF mutation (p.S3101*) was identified on the 9th repeat and also on the 10-repeat allele. Thus, the existences of c.488delG, p.S3101*, and p.S3316* on the 10-repeat allele suggest that these variants most likely arose independently and after the separation of the 11- and 12- expanded repeat, derived alleles.

DISCUSSION

In summary, our targeted sequencing strategy to identify *FLG* LOF variants in a well-characterized UKWP-defined AD group of African American children identifies a

significant enrichment of pathogenic *FLG* LOF variant alleles (11.5% versus 2.5% [Africans; ExAC] and also 1.7% [African American; ESP]). We discovered 2 pathogenic *FLG* LOF not previously reported for AD (c.488delG and p.S3101*). The overlap of our p.S3316* findings in our cohort, in an independent AA AD study [13], and in African individuals (ExAC, ESP, and 1000G) supports p.S3316* as a population-specific *FLG* LOF variant for African ancestry. Our discovery of five pathogenic *FLG* LOF variants (including 2 newly discovered) in African American pediatric patients with severe AD has brought forth a hypothesis for the wider research community to test, *FLG* LOF variants among the African American population are enriched in AD, and thus justifies more precision medicine efforts for this health disparity group.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the patients and their families for their participation; Emily Beck, MD, Emily Gurnee, MD, and Colleen Cotton, MD for patient recruitment; Kara Gulewicz, MD for patient recruitment and study design, and Erin Brettmann, PhD for critical reading of the manuscript. The study was supported by Washington University Faculty Diversity Scholar Award and in part by NIAMS (R00AR055948, R01AR065523) to CdGS; T32GM007067 to MEM; A*STAR SPF genetic orphan diseases (IAF SPF 2012/005) to XFCCW, and A*STAR SPF grants for basic and translational research (IAF SPF 2013/004; IAF SPF2013/005) to SLIJD and JEC. This publication is solely the responsibility of the authors and does not necessarily represent the official views of NIAMS or IAF.

REFERENCES

[1]. Brown SJ, McLean WH, One remarkable molecule: filaggrin, J Invest Dermatol 132(3 Pt 2) (2012) 751–62. [PubMed: 22158554]

[2]. Nutten S, Atopic dermatitis: global epidemiology and risk factors, Ann Nutr Metab 66 Suppl 1 (2015) 8–16. [PubMed: 25925336]

[3]. Mallol J, Crane J, von Mutius E, Odhiambo J, Keil U, Stewart A, Group IPTS, The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three: a global synthesis, Allergol Immunopathol (Madr) 41(2) (2013) 73–85. [PubMed: 22771150]

[4]. Shaw TE, Currie GP, Koudelka CW, Simpson EL, Eczema prevalence in the United States: data from the 2003 National Survey of Children's Health, J Invest Dermatol 131(1) (2011) 67–73. [PubMed: 20739951]

[5]. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, Liao H, Evans AT, Goudie DR, Lewis-Jones S, Arseculeratne G, Munro CS, Sergeant A, O'Regan G, Bale SJ, Compton JG, DiGiovanna JJ, Presland RB, Fleckman P, McLean WH, Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris, Nat Genet 38(3) (2006) 337–42. [PubMed: 16444271]

[6]. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, Carrick T, Evans AT, Liao H, Zhao Y, Campbell LE, Schmuth M, Gruber R, Janecke AR, Elias PM, van Steensel MA, Nagtzaam I, van Geel M, Steijlen PM, Munro CS, Bradley DG, Palmer CN, Smith FJ, McLean WH, Irvine AD, Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema, Nat Genet 39(5) (2007) 650–4. [PubMed: 17417636]

[7]. Quiggle AM, Goodwin ZA, Marfatia TR, Kumar MG, Ciliberto H, Bayliss SJ, de Guzman Strong C, Low filaggrin monomer repeats in African American pediatric patients with moderate to severe atopic dermatitis, JAMA Dermatol 151(5) (2015) 557–9. [PubMed: 25564772]

[8]. Brown SJ, Kroboth K, Sandilands A, Campbell LE, Pohler E, Kezic S, Cordell HJ, McLean WH, Irvine AD, Intragenic copy number variation within filaggrin contributes to the risk of atopic

dermatitis with a dose-dependent effect, J Invest Dermatol 132(1) (2012) 98–104. [PubMed: 22071473]

[9]. Irvine AD, McLean WH, Leung DY, Filaggrin mutations associated with skin and allergic diseases, N Engl J Med 365(14) (2011) 1315–27. [PubMed: 21991953]

[10]. Margolis DJ, Apter AJ, Gupta J, Hoffstad O, Papadopoulos M, Campbell LE, Sandilands A, McLean WH, Rebbeck TR, Mitra N, The persistence of atopic dermatitis and filaggrin (FLG) mutations in a US longitudinal cohort, J Allergy Clin Immunol 130(4) (2012) 912–7. [PubMed: 22951058]

[11]. Polcari I, Becker L, Stein SL, Smith MS, Paller AS, Filaggrin gene mutations in African Americans with both ichthyosis vulgaris and atopic dermatitis, Pediatr Dermatol 31(4) (2014) 489–92. [PubMed: 24920311]

[12]. Mandelker D, Schmidt RJ, Ankala A, McDonald Gibson K, Bowser M, Sharma H, Duffy E, Hegde M, Santani A, Lebo M, Funke B, Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing, Genet Med 18(12) (2016) 1282–1289. [PubMed: 27228465]

[13]. Margolis DJ, Mitra N, Gochnauer H, Wubbenhorst B, D'Andrea K, Kraya A, Hoffstad O, Gupta J, Kim B, Yan A, Chiesa Fuxench Z, Nathanson KL, Uncommon filaggrin variants are associated with persistent atopic dermatitis in African Americans, J Invest Dermatol (2018).

[14]. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Exome Aggregation C, Analysis of protein-coding genetic variation in 60,706 humans, Nature 536(7616) (2016) 285–91. [PubMed: 27535533]

[15]. Eaaswarkhanth M, Xu D, Flanagan C, Rzhetskaya M, Hayes MG, Blekhman R, Jablonski NG, Gokcumen O, Atopic Dermatitis Susceptibility Variants in Filaggrin Hitchhike Hornerin Selective Sweep, Genome Biol Evol 8(10) (2016) 3240–3255. [PubMed: 27678121]

[16]. Wong X, Denil S, Foo JN, Chen H, Tay ASL, Haines RL, Tang MBY, McLean WHI, Sandilands A, Smith FJD, Lane EB, Liu J, Common JEA, Array-based sequencing of filaggrin gene for comprehensive detection of disease-associated variants, J Allergy Clin Immunol 141(2) (2018) 814–816. [PubMed: 29056476]

[17]. Williams HC, Burney PG, Hay RJ, Archer CB, Shipley MJ, Hunter JJ, Bingham EA, Finlay AY, Pembroke AC, Graham-Brown RA, et al., The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis, Br J Dermatol 131(3) (1994) 383–96. [PubMed: 7918015]

[18]. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis, Dermatology 186(1) (1993) 23–31. [PubMed: 8435513]

Table 1:

FLG LOF variants and *FLG* repeat alleles in African American pediatric patients with atopic dermatitis (n=39) reported with respect to human reference *FLG* 10-repeat allele.

***FLG* LOF Variant**	**hg19 coordinate/ RSID (if known)**	**Location of variant to *FLG* repeat**	***FLG* repeat allele 1**	***FLG* repeat allele 2**	**Total # *FLG* repeats**	**Age**	**SCORAD**	**Asthma**	**Allergic rhinitis**	**Food Allergies**	**Ichthyosis Vulgaris (IV)/ Xerosis (X)**
c.488delG¶	chr1:152,286,874	0	10	10	20	7	73	Yes	Yes	Yes[1,2,3]	IV
p.R501*	rs61816761	1	11	11	22	11	83	Yes	No	None	X
p.R501*	rs61816761	1	10	11	21	13	38	Yes	Yes	None	X
p.R826*	rs115746363	2	10	10	20	5	51	Yes	No	Yes[1]	X
p.R826*	rs115746363	2	10	10	20	1	65	Yes	No	None	X
p.S3101*¶	chr1:152,278,060	9	10	10	20	10	41	Yes	Yes	None	X
p.S3316*	rs149484917	9	10	11	21	4	50	Yes	No	Yes[4]	X
p.S3316*	rs149484917	9	10	11	21	2	94	No	No	Yes[1]	X
p.S3316*	rs149484917	9	10	10	20	0.5	85	No	No	None	X

¶ denotes novel *FLG* LOF

[1] Peanut

[2] Shellfish

[3] Seafood

[4] Eggs

Reports of asthma, allergic rhinitis, and food allergies are in response to our study questionnaire and obtained during medical exam history.