Improving diagnostic yield for filaggrin; hidden mutations in the Dutch population

Michel van Geel^{1,2}, Carolien N. van Oosterhoud², Stijn A.I. Ghesquiere², Roel R.H.F.J. Brandts², Peter M. Steijlen¹
Departments of 1) Dermatology and 2) Clinical Genetics, Maastricht University Medical Centre, Maastricht, The Netherlands

INTRODUCTION

Filaggrin (filament-aggregating protein) has an important role in the terminal differentiation of the skin’s epidermis, in forming a barrier against allergens and preventing trans epidermal water loss. Recessive loss-of-function mutations are causally linked to ichthyosis vulgaris (IV) and heterozygous carriers have a predisposition for developing atopic dermatitis, allergic rhinitis and asthma. The profilaggrin gene (*FLG*) on chromosome 1q21, has a complex structure consisting of 3 exons, of which 2 are coding and the last very large exon encompassing 8 to 10 repeated units approximately 1 kilobase in size. Until now 85 unique *FLG* mutations are published in literature mainly identified in the European and Asian populations. Molecular diagnostics with conventional Sanger sequencing for IV has been hampered by the notoriously difficult to analyse *FLG* gene, caused by its homologous and polymorphic repeated units. By implementation of single molecule molecular inversion probes (smMIPs) and next generation sequencing (NGS), an alternative screening strategy for analysis of the entire coding region of the *FLG* gene becomes feasible. Genetic analysis of the whole gene instead of screening for only population-specific mutations, would improve diagnostic yield by scrutinizing also for rare family-specific mutations or specific mutations in ethnicities not previously studied.

MATERIAL & METHODS

The easy to implement, affordable smMIP-NGS strategy (Figure 1) was employed to screen for the whole *FLG* gene. Since exclusion of NGS-duplicate-reads is possible with the single molecule (sm) tag, mutation-percentages can be related and assigned to polymorphic duplicated filaggrin-repeat-unit 8 and 10 if a long distance specific PCR-fragment sizing experiment is included. The filtering of relevant variants was done with an in-house developed software tool and displayed with IGV (Figure 2).

Figure 1 Schematic overview of smMIP-NGS method. Pool of smMIPs is hybridised (blue and red sequence) to the target genomic DNA, extended (orange) with DNA polymerase and subsequently ligated to form circular DNA fragments. The linearized DNA (genomic and untargeted probes) is digested with exonuclease. Circularised library DNA is captured and PCR amplified with sequence-patient specific index primers and pooled for NGS sequencing. Data is collected, in silico gathered per individual through the specific index, duplicates removed according to their specific tag (yellow) and aligned to the reference genome sequence and filtered for pathogenic variants.

RESULTS

In a cohort of previously screened Dutch patients (N=70) for only the population-specific mutations, retrospectively the whole *FLG* gene was analysed. Since all known mutations result in premature protein termination, focus of attention was on identifying nonsense and small insertion or deletion mutations. In several (8/70) of the screened patients additional novel truncating mutations were identified (table 1), elucidating their previously unexplained (more severe) clinical presentations.

CONCLUSIONS

This study emphasises the need for screening the entire *FLG* gene for mutations, to improve the diagnostic yield in IV and identify hidden variants in the homologous repeated units of the gene. Herein, the smMIP-NGS method proves to be a reliable straightforward strategy to boost clinical diagnostics for IV and opens possibilities to facilitate patient stratification in large cohort studies.

Figure 1 General overview of smMIP-NGS strategy for analysis of the whole coding sequence of the *FLG* gene. Genomic structure of the *FLG* gene (top). Enlargement of part of *FLG* to indicate the double tiling (on forward and reverse strand) smMIP design on coding *FLG* sequence (middle). Snapshot of IGV sequence alignment (with removal of duplicates) to reference sequence with a heterozygous c.2282_2285del mutation (bottom).

Table 1 Results of patients tested for the most frequent *FLG* mutations in the Western European population with Sanger sequencing and for the whole coding region of *FLG* with smMIP-NGS analysis.

patient ID	gender	mutation Sanger	smMIPs	phenotype
1	F	c.1501C>T ht	na	eczema, ichthyosis vulgaris
2	F		n/n	ichthyosis vulgaris
3	M	c.2282_2285del4, c.7339C>T	na	ichthyosis vulgaris
4	M	c.2282_2285del4 hm	na	ichthyosis vulgaris
5	F		n/n	severe eczema
6	M	c.2282_2285del4, c.7339C>T	na	ichthyosis vulgaris
7	F		c.1063C>T p.(Gln355*) ht	ichthyosis vulgaris
8	F	c.2282_2285del4 ht	na	dry skin, pruritis, palmar hyperlinearity
9	F	c.1501C>T, c.2282_2285del4	na	ichthyosis vulgaris, constitutional eczema
10	M	c.2282_2285del4, c.7339C>T	na	ichthyosis
11	M		n/n	ichthyosis vulgaris
12	M		n/n	congenital ichthyosis
13	F		n/n	eczema, atopic, ichthyosis vulgaris?
14a	M	c.2282_2285del4 ht	na	ichthyosis
14b	F		na	eczema, bronchitis
15	M		n/n	ichthyosis
16a	M		c.9595C>T p.(Gln3199*) ht	eczema, ichthyosis vulgaris
16b	M		c.9595C>T p.(Gln3199*) ht	eczema, ichthyosis vulgaris
17	M	c.2282_2285del4 ht	na	dry skin, ichthyosis vulgaris
18	M		c.7487del ht	generalised eczema
19	F	c.1501C>T, c.2282_2285del4	na	ichthyosis vulgaris
20	M		na	ichthyosis
21	M		na	severe eczema, atopic constitution
22	F	c.2282_2285del4 ht	c.2282_2285del ht	severe eczema
23	M		n/n	severe atopic eczema, food allergies, asthma
24	F		n/n	ichthyosis vulgaris
25	F		n/n	eczema
26	M		n/n	eczema, ichthyosis
27a	M		n/n	ichthyosis
27b	M		n/n	ichthyosis
28	M		n/n	pruritis
29	M		n/n	ichthyosis vulgaris
30	F		n/n	dry skin
31	M	DSG1 mutation	n/n	ichthyosis vulgaris?
32	M	c.2282_2285del4, c.7339C>T	na	ichthyosis
33	F		na	collodion baby, ichthyosis
34	M	NIPAL4 mutation	na	ichthyosis, erythroderma, palmoplantar keratoderma striata
35	M	c.2282_2285del4 ht	c.2282_2285del + c.4042C>T (p.(Gln1348*)) ht	atopic eczema, ichthyosis
36	M	c.1501C>T ht	c.1501C>T ht	dry skin
37	F	c.2282_2285del4 hm	na	ichthyosis vulgaris
38	M	c.1501C>T, c.2282_2285del4	na	severe constitutional eczema
39	M		n/n	ichthyosis vulgaris
40	M		na	severe ichthyosis
41	M		c.6950_6957del, c.7945del	ichthyosis vulgaris
42	F	c.1501C>T, c.2282_2285del4	na	atopic eczema, pityriasiform squamae
43	M		n/n	dry skin, erythematous plaques
44	M	c.2282_2285del4 hm	na	severe eczema, ichthyosis
45	M	c.2282_2285del4 ht, +STS mutation	na	ichthyosis, atopic eczema, food allergy
46	M		n/n	ichthyosis vulgaris
47	M	c.2282_2285del4 ht, +STS mutation	c.2282_2285del4 ht	ichthyosis
48	F	c.1501C>T, c.9740C>A	na	ichthyosis, food allergies, asthma
49	M	c.9740C>A ht	c.9740C>A ht	ichthyosis
50	M	STS mutation	na	dry ichthyosiform skin
51	F	c.1501C>T, c.2282_2285del4	na	dry ichthyosiform skin
52	F		c.3551C>A (p.(Ser1184*)) + c.63dup	ichthyosis vulgaris
53	M	c.2282_2285del4 hm	na	ichthyosis, keratosis pilaris, eczema
54	F	c.2282_2285del4 hm	na	ichthyosis vulgaris
55	M	c.1501C>T hm	na	dry skin, atopic dermatitis, pruritis, allergic rhinitis, asthma
56	M		n/n	ichthyosis vulgaris
57	M	STS mutation	na	ichthyosis
58	F		n/n	ichthyosis vulgaris
59	M	STS mutation	na	ichthyosis
60	M		n/n	ichthyosis vulgaris
61	M		n/n	pruritis
62	M		n/n	ichthyosis
63	M	STS mutation	n/n	ichthyosis
64	M	c.1501C>T hm	c.1501C>T hm + c.4720_4721del ht	eczema, dry skin
65	F		c.3757G>T p.(Gly1253*) ht	ichthyosis vulgaris
66	M		n/n	atopic dermatitis, food allergies
67	F		n/n	ichthyosis?

n/n: no mutation detected, na: not analysed, ID a and b: one family