ORIGINAL ARTICLE

Optimization of long-range PCR protocol to prepare filaggrin exon 3 libraries for PacBio long-read sequencing

Chiara Mareso[1] · Elena Albion[1] · William Cozza[1] · Benedetta Tanzi[2] · Stefano Cecchin[2] · Paolo Gisondi[3] · Sandro Michelini[4] · Francesco Bellinato[3] · Serena Michelini[5] · Silvia Michelini[6] · Matteo Bertelli[1,2] · Giuseppe Marceddu[1]

Abstract

Background The filaggrin (FLG) protein, encoded by the *FLG* gene, is an intermediate filament-associated protein that plays a crucial role in the terminal stages of human epidermal differentiation. Loss-of-function mutations in the *FLG* exon 3 have been associated with skin diseases. The identification of causative mutations is challenging, due to the high sequence homology within its exon 3 (12,753 bp), which includes 10 to 12 filaggrin tandem repeats. With this study we aimed to obtain the whole FLG exon 3 sequence through PacBio technology, once 13-kb amplicons have been generated.

Methods and results For the preparation of SMRTbell libraries to be sequenced using PacBio technology, we focused on optimizing a 2-step long-range PCR protocol to generate 13-kb amplicons covering the whole *FLG* exon 3 sequence. The performance of three long-range DNA polymerases was assessed in an attempt to improve the PCR conditions required for the enzymes to function properly. We focused on optimization of the input template DNA concentration and thermocycling parameters to correctly amplify the entire *FLG* exon 3 sequence, minimizing non-specific amplification.

Conclusions Taken together, our findings suggested that the PrimeSTAR protocol is suitable for producing the amplicons of the 13-kb FLG whole exon 3 to prepare SMRTbell libraries. We suggest that sequencing the generated amplicons may be useful for identifying LoF variants that are causative of the patients' disorders.

Keywords Long-range PCR · *FLG* · PacBio · SMRTbell library · Protocol optimization · NGS

Introduction

The human filaggrin gene (*FLG*) is located on chromosome 1q21.3 in the epidermal differentiation complex (EDC) [1] and encodes profilaggrin, an insoluble highly phosphorylated ~ 400 kDa precursor polyprotein. When keratinocytes of granular layers turn into corneocytes, profilaggrin is dephosphorylated and processed by serine proteases, to yield multiple 37 kDa-filaggrin monomers [2]. Filaggrin is an intermediate filament-associated protein [3] that binds and organizes keratin intermediate filaments [4]. It promotes cornified-envelope formation to ensure appropriate barrier function [5], and once degraded it provides free amino acids [6], forming the natural moisturizing factor (NMF) [7], which helps maintaining epidermal hydration, and urocanic acid, which plays a role in protecting against UV radiation and in modulating immune functions [8, 9]. The human *FLG* gene has an unusually large exon 3, which involves a series of related but nonidentical 972 bp filaggrin tandem repeats, flanked by partial, imperfect filaggrin repeats [10]. The size of *FLG* gene exon 3 varies from 12.7 to 14.7 kb in the population, depending on repeat number, which ranges from 10 to 12 [11, 12]. It encodes most of the N-terminal domain and complete filaggrin monomers, each of 324 amino acids [1] (Fig. 1). Loss-of-function (LoF) mutations in this exon have been associated with two common inherited keratinization disorders, ichthyosis vulgaris (OMIM 146,700) and atopic dermatitis (ATOD2) (OMIM 605,803), and with a possible predisposition to food allergies, asthma and allergic rhinitis [13, 14].

Chiara Mareso and Elena Albion have contributed equally to this work.

✉ William Cozza
   william.cozza@assomagi.org;
   magiadmin@magigroup.community

Extended author information available on the last page of the article

Fig. 1 Scheme of FLG gene structure and profilaggrin organization, generated in PowerPoint Version 2021
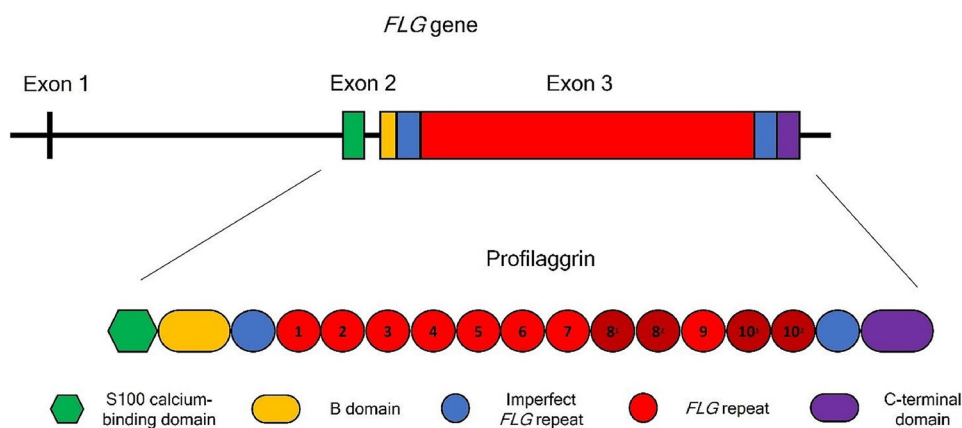
Due to high sequence homology between tandem repeated regions of *FLG* exon 3 (see Supplemental Fig. S1), the identification of LoF variants in this exon is challenging, as short-sequencing reads map with low quality [15]. To date, most studies aimed at FLG genotyping have screened for common LoF mutations but have avoided sequencing the entire FLG coding region, which may result in underreporting the genetic contribution. The FLG exon 3 sequencing through NGS would result too complicated to be performed, due to the limits of NGS technology to analyze low complexity regions with high homology [10]. The NGS limitations in the analysis of genomic regions are not only limited to the FLG exon 3 sequence. As a matter of fact, it is well documented that NGS sequencing technology presents limitations in the mapping of different genomic regions, despite the use of sophisticated bioinformatics tools [16]. These regions include those which present base modifications in repetitive regions, as well as regions with structural variations, very high GC content and sequences with multiple homologous elements [17, 18].

To overcome limitations of "short-read sequencing" applications, Third-Generation sequencing (TGS), like PacBio methodologies, were developed through time, as they permit long-reads production with an average length of more than 10 kb. "Long-reads sequencing" is more representative of chromosomes, and consequently produce more contiguous genome reconstruction. They also permit to identify even the more complex variations and structural changes. In addition, the absence of PCR amplification in TGS approaches allows less biases and more homogeneous genome coverage. TGS library construction is simplified and less time-consuming if compared to the short-reads methods. [19–21]. In the FLG exon 3 specific case, the sequencing through PacBio technology enables highly accurate (99.8%) long high-fidelity (HiFi) reads [22] spanning the whole exon sequence, and can possibly identify even less common and less studied LoF variants.

In addition, TGS methods are shown to be more advantageous for long-reads in the sequence alignment phase, if compared to NGS methodologies. It is well known that, NGS technologies present strong limitations in sequencing long-reads of DNA sequences, due to the high number of biases introduced in the alignment phase. The sequenced strand of DNA pass through the fragmentation and amplification phases which lead to random clones that must be assembled into single continuous sequences, leading to high bias rates. In addition, NGS methodologies fail to generate a sufficient overlap between the DNA fragments. In contrast, TGS methodologies reduce the possibility of these biases: they reduce amplification bias, and generate DNA fragments which present reasonable lengths, that permit better overlapping for better sequence assembly [23].

With regard to the bioinformatic analysis of the sequencing data obtained, we referred to our internally developed bioinformatic pipeline "PacMAGI" [24]. The workflow of such bioinformatic pipeline presents a quality control performed by using FastQC and MultiQC, an alignment phase with reference genotype, through the use of Pbmm2, and the generation of mapping statistics through Picard and Qualimap tool. "PacMAGI" calling phase includes variants calling through Longshot, SVs calling through PBSVs. "PacMAGI" variants are finally annotated, filtered and interpreted using VarSome.

Long-range PCR (LR-PCR) is the most common approach to enriching long DNA fragments, and when combined with sequencing it can be used to screen for genetic variants [25]. We succeeded in finding suitable conditions for amplifying *FLG* exon 3 and for developing a LR-PCR protocol specifically designed for the preparation of SMRTbell libraries to be sequenced using PacBio technology.

Results

SMRTbell library preparation starts with two PCR-rounds. The first-round PCR aims to amplify the chromosomal segment of interest (on chromosome 1q21.3) using target-specific primers tailed to Forward/Reverse universal sequences, which allow annealing of Forward/Reverse Barcoded Universal Primers (BUP) provided by PacBio during the second-round of PCR, resulting in amplicon barcoding.

Optimization of first-round PCR protocol

Phusion Hot Start II High-Fidelity DNA Polymerase, KAPA HiFi HotStart DNA Polymerase, and PrimeSTAR GXL DNA Polymerase were selected on the basis of their features as candidate enzymes to amplify the ~ 13 kb DNA fragment of interest, which spans the whole *FLG* exon 3 sequence. As suggested by PacBio guidelines, we started performing our assays using Phusion Hot Start II DNA Polymerase, which is an extremely processive Pyrococcus-like polymerase, having an error rate of 4.4×10^{-7}. This enzyme is able to amplify long amplicons, such as 20 kb λ DNA, and was therefore ideal for amplifying the ~ 13 kb DNA fragment of interest. Since Phusion works better at high denaturation and annealing temperatures, because of the high salt concentration in the reaction buffer, denaturation temperature was set at 98 °C for a short time (30 s), as recommended. Annealing temperature was set at 65 °C, as determined using the suggested online tool 'Tm calculator' provided by Thermofisher (www.thermofisher. com/tmcalculator). For the second attempt at first-round PCR optimization, we chose KAPA HiFi HotStart DNA Polymerase, which is also widely used in next-generation sequencing (NGS) library preparation. It belongs to the DNA polymerase B-family, and is engineered to improve yield, speed, sensitivity and DNA affinity without requiring accessory proteins or DNA binding domains. KAPA is a high-fidelity polymerase having an error rate of 1 error per 3.6×10^{6} nucleotides incorporated, and is capable of amplifying DNA targets up to 15 kb or even 20 kb in the case of less complex targets; it was therefore chosen to amplify the ~ 13 kb target region of interest. KAPA HiFi HotStart ReadyMix has a higher salt concentration than conventional PCR ready-mixes, which affects primer annealing. The annealing temperature was therefore set at 64 °C, 5 °C higher than primer melting temperature (59 °C). For the two-assay series performed with Phusion and KAPA, extension time was calculated considering 1 min/kb and set at 13 min and 20 s. In the third series of assays, PrimeSTAR GXL DNA Polymerase was

selected to amplify *FLG* gene exon 3. A study comparing six DNA polymerase performances in LR-PCR found that PrimeSTAR GXL Polymerase was the best, by virtue of its ability to amplify large amplicons, even of different sizes and melting temperatures, under identical PCR conditions [26]. It is a high-fidelity polymerase capable of amplifying DNA fragments, even ≥ 30 kb in length. In the case of 10 to 30 kb amplicons, the manufacturer's guidelines suggest a two-step PCR, which only consists of a denaturing (98 °C for 10 s) and an extension step (68 °C for 10 min), skipping the annealing phase. To produce sufficient amplicons for downstream applications, all the prepared reaction mixes were allowed 100 ng of template DNA. Reaction mixtures and thermocycling conditions adopted for the first-round PCR are listed in Table 1.

All the PCR products were loaded and run on 1 × agarose gel. The gel electrophoresis results (Fig. 2) showed that all three polymerases seemed to amplify *FLG* gene exon 3, since the corresponding bands on the agarose gel indicated amplification of a ~ 13 kb PCR product. To further confirm the results of the previous experiment, nested PCRs were performed starting from template DNA and on an aliquot of sample DNA, used as positive control. The resulting amplicons were run on 2% agarose gel, shown in Fig. 3. This test confirmed that the three previous assays effectively and correctly amplified *FLG* gene exon 3, since all gel bands proved amplification of the expected 473 bp in-target region of interest. Moreover, the intensity of the sample bands was higher than that of the gel band of the control DNA, indicating that the amplified sample target regions were at higher concentrations, since they had already been amplified in the previous LR-PCR. However, as shown in Fig. 2, the two ~ 13 kb gel bands of the PCR products amplified by Phusion and KAPA were very weak compared to that resulting from the amplification performed by PrimeSTAR, which produced amplicons at higher concentrations. We therefore chose the optimized PrimeSTAR PCR protocol (Table 1) to perform the second-round PCR for the preparation of *FLG* exon 3 SMRTbell libraries.

Optimization of second-round PCR protocol

For the first attempt to optimize the second-round PCR, we used the same thermocycling conditions as for the previously optimized first-round PCR protocol. The gel electrophoresis results shown in Fig. 4a show a DNA smear pattern, indicating non-specific amplification, possibly due to the large number of cycles. Cycle number was therefore lowered from 25 to 10 (Figs. 4b lane 3), and from 10 to 7 (Figs. 4b lane 5) in a subsequent assay. The gel electrophoresis results (Fig. 4b) showed that the number of non-specific products decreased with decreasing cycle number. Due to the significantly reduced non-specific amplification resulting

Table 1 (a) The optimized protocol of the first-round LR-PCR using three different DNA polymerases. The optimized PrimeSTAR* PCR protocol was the first-round PCR protocol chosen for the preparation of FLG exon 3 SMRTbell libraries. (b) The optimized protocol of the second-round LR-PCR using PrimeSTAR* PCR protocol

LR-PCR		
a	First-round	
Enzyme	Reaction mixture	PCR conditions
Phusion	100 ng template DNA, 1.25 µl 10 µM primers, 12.5 µl 2× Phusion HS II HF Master Mix, and distilled water to 25 µl	98 °C 30 s 30 cycles 98 °C 10 s 65 °C 15 s 72 °C 13 min and 20 s 72 °C for 7 min Hold at 4 °C
KAPA	100 ng template DNA, 0.75 µl 10 µM primers, 12.5 µl 2×KAPA HiFi HotStart ReadyMix, and distilled water to 25 µl	95 °C 3 min 30 cycles 98 °C 20 s 64 °C 15 s 72 °C 13 min and 20 s 72 °C 13 min and 20 s Hold at 4 °C
PrimeSTAR*	100 ng template DNA, 0.5 µl 10 µM primers, 5 µl 5×PrimeSTAR GXL Buffer, 2 µl dNTP, 0.5 µl PrimeSTAR® GXL DNA polymerase, and distilled water to 25 µl	25 cycles 98 °C 10 s 68 °C 10 min Hold at 4 °C
b	Second-round	
Enzyme	Reaction mixture	PCR conditions
PrimeSTAR	3 µl first-round PCR DNA, 1 µl Barcoded Universal Primers, 5 µl 5×PrimeSTAR GXL Buffer, 2 µl dNTP, 0.5 µl PrimeSTAR® GXL DNA polymerase, and distilled water to 25 µl	7 cycles 98 °C 10 s 68 °C 10 min Hold at 4 °C

from the second assay, the cycle number was also kept at 7 in the subsequent test. The same test was performed starting with 3 µl of first-round PCR amplicons, to determine the appropriate amount of first-round PCR product to use and to increase final amplicon concentration. The resulting second-round PCR products were run on 1×agarose gel. The gel image (Fig. 4c) showed very feeble gel bands related to nonspecific products close to the one of interest, which on the other hand was quite intense, indicating a high target amplicon concentration. The PCR product aliquot was purified with 0.45X AMPure PB beads, and DNA concentration was measured by Qubit dsDNA BR Assay Kit, performed with a Qubit Fluorometer. For SMRTbell library preparation 3000 ng of total input DNA is required, based on the amplicons size (~ 13 kb). Thus, samples were pooled in equimolar amounts, to obtain an equal representation of each amplicon in the data.

Discussion

Several assays were performed to optimize the 2-step long-range PCR protocol for the preparation of *FLG* exon 3 SMRTbell libraries. Generating long amplicons in order to have the benefits of long-read PacBio sequencing for analyzing *FLG* exon 3 was challenging, due among other things to factors related to the target chromosomal region that includes GC content (55.3%) and highly homologous sequences. Suitable template DNA concentration and thermocycling conditions were determined according to amplicon length, primer melting temperature and reaction buffer composition. We assessed the performance of three long-range DNA polymerases, advertised as able to generate amplicons up to 20 kb or more. Compared to KAPA and Phusion, PrimeSTAR generated greater amounts of target amplicons, minimizing nonspecific product amplification.

Fig. 2 Gel electrophoresis of first-round PCR products, amplified using different polymerases. Lane M, DNA marker; 1, Phusion Hot Start II High-Fidelity DNA Polymerase; 3, KAPA HiFi HotStart DNA Polymerase; 5, PrimeSTAR GXL Polymerase; 2, 4, 6 were negative controls for the corresponding polymerases. The full-length gel is presented in Supplementary Fig. S2. The original unedited gel image is shown in the supplementary dataset, and the image was not joined from different parts of the gel
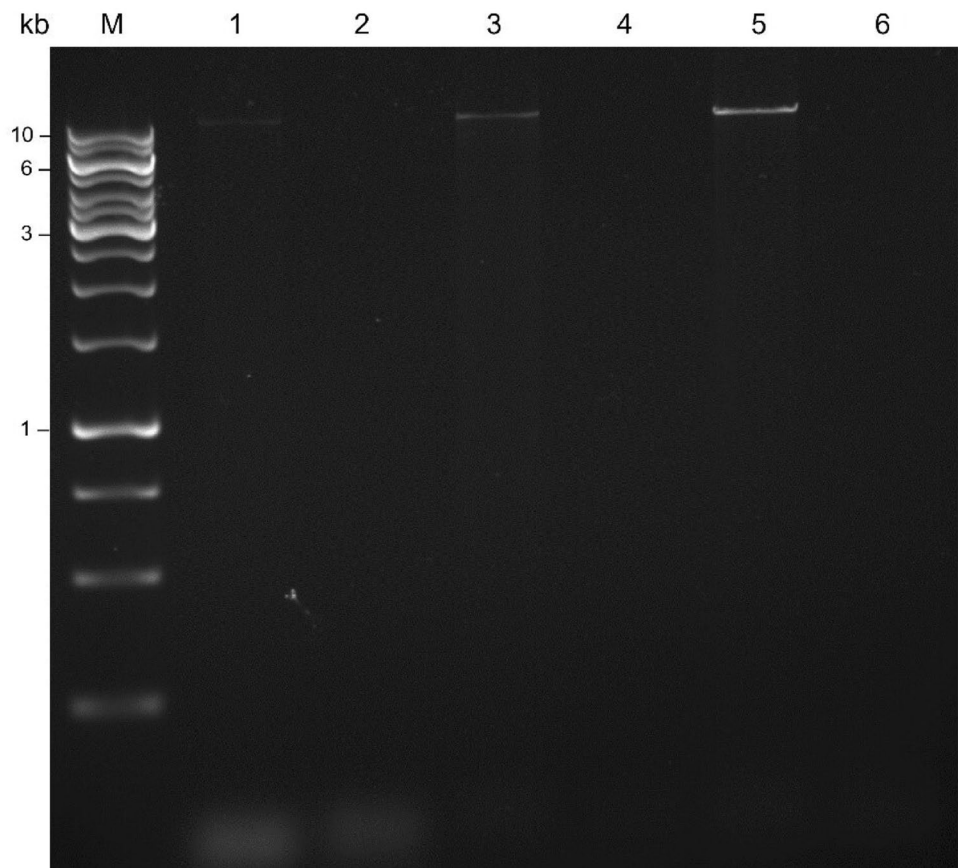
Fig. 3 Gel electrophoresis of nested PCR products, performed on PCR products resulting from the previous tests and using different polymerases. Lane M, DNA marker; 1, Phusion Hot Start II High-Fidelity DNA Polymerase; 2, KAPA HiFi HotStart DNA Polymerase; 3, PrimeSTAR GXL Polymerase; 4, positive control; 5, negative control (5). The full-length gel is presented in Supplementary Fig. S3. Original unedited gel image is shown in the supplementary dataset and image was not joined from different parts of the gel
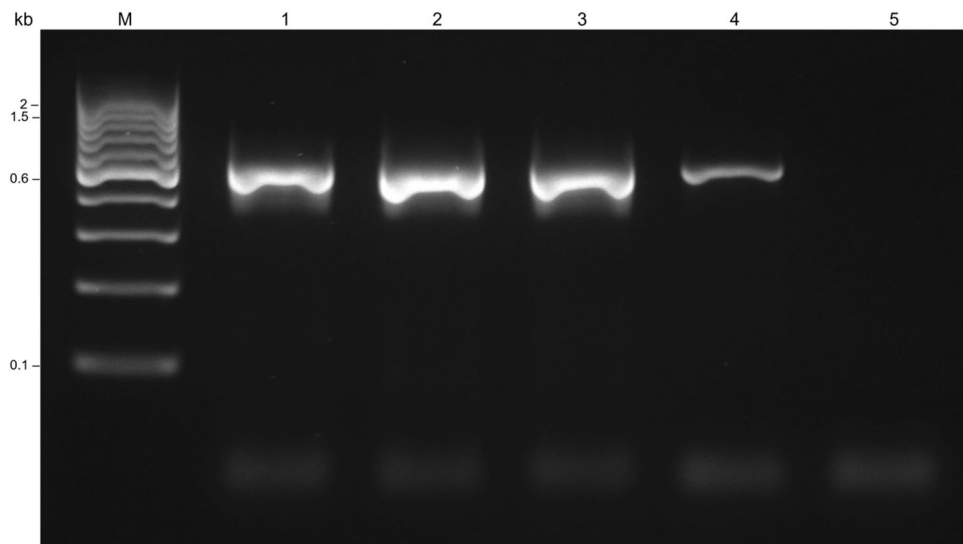
Unlike the other two polymerases, a two-step PCR protocol was adopted for PrimeSTAR. It did not include an annealing phase, but only denaturing (98 °C for 10 s) and extension (68 °C for 10 min) steps. The better performance of Prime-STAR is probably explained by the fact that this enzyme was developed specifically to amplify products ≥ 30 kb in length, even if they are rich in GC-content [26]. The results

showed that PrimeSTAR was suitable for producing *FLG* exon 3 amplicons for the preparation of SMRTbell libraries.

After a series of attempts to optimize the two-step LR-PCR protocol, we succeeded in amplifying the 13-kb genomic fragment spanning the entire sequence of *FLG* exon 3. This protocol was used to prepare SMRTbell libraries to sequence *FLG* exon 3 in 63 Italian patients with
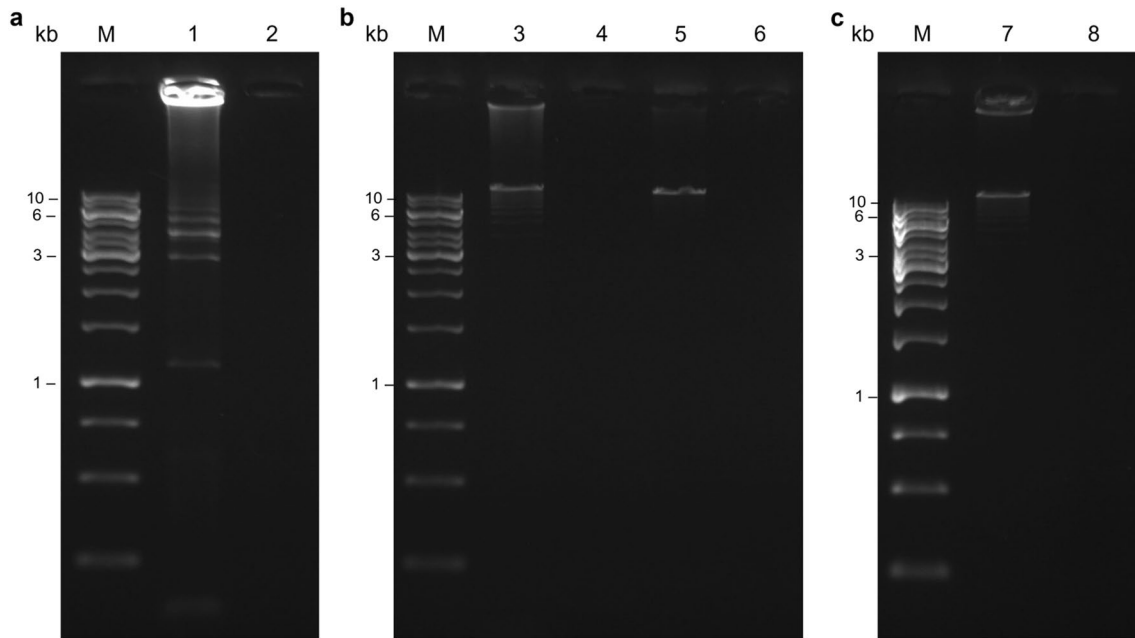
Fig. 4 Gel electrophoresis of second-round PCR products amplified using PrimeSTAR GXL Polymerase, performed following the previously optimized LR-PCR protocol (**a**), and subsequently optimizing cycle number to 10 and 7 (**b**). Test perfomed using 3 µl of first-round PCR amplicons and 7 PCR cycles (**c**). Lane M, DNA marker; 1, 3, 5, 7 s-round PCR products; 2, 4, 6, 8 corresponding negative controls. The full-length gel is presented in Supplementary Fig. S4. Original unedited gel image is shown in the supplementary dataset and image was not joined from different parts of the gel

keratinization disorders. The quality of the results was determined, to verify the accuracy of protocol described above, which are reported in supplementary material [Supplementary FigS5; Supplementary material 2_MultiQC report for 63 patients].

Indeed, sequencing analysis of the entire *FLG* exon 3 at single nucleotide resolution would also enable screening for less common LoF variants, which may be more frequent in the Italian population. No treatment has yet been found for ichthyosis vulgaris and atopic dermatitis. There is only palliative care, like topical medications to ease symptoms including itch and pain, and to prevent these conditions from worsening. However, more comprehensive genetic testing using new generation technologies could offer patients the chance to be candidates for future therapeutic options.

Methods

The SMRTbell library to sequence *FLG* gene exon 3 was prepared following the PacBio protocol guidelines Procedure & Checklist—Preparing SMRTbell® Libraries using PacBio® Barcoded Universal Primers for Multiplexing Amplicons—Part Number 101–791-800 Version 02 (April 2020) (https://www.pacb.com/wp-content/uploads/Procedure-Checklist-Preparing-SMRTbell-Libraries-using-PacBio-Barcoded-Universal-Primers-for-Multiplexing-Amplicons.pdf). All assays were performed using the same genomic DNA sample, derived from transformed reference cell lines provided by Coriell Institute for Medical Research as template. For the purpose of this work, we purchased the NA20509 cell line (INTERNATIONAL HAP-MAP PROJECT—TOSCANI IN ITALIA (TUSCANS IN ITALY)) from Coriell Cell Repositories (http://ccr.coriell.org/). Reactions were performed using a SureCycler 8800 Thermal Cycler (Agilent). The first step of the SMRTbell library preparation workflow consists of two PCR-rounds, which require optimization according to the target region features.

For the first-round PCR, forward target-specific primer (5'-AGGGTTATTTTGAGCTCTTTGTGAA-3') and reverse target-specific primer (5'-TGGGACAGTGATTATGTTGGAGA-3') tailed to F/R universal sequences were designed to amplify the ~ 13 kb DNA fragment spanning the entire *FLG* exon 3 sequence. Primer design was performed using the freely available program Primer3web (https://primer3.ut.ee) and, to avoid mispriming, the primer couple was tested in silico using NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast) and the UCSC tool In-Silico PCR (https://genome.ucsc.edu/cgi-bin/hgPcr). Three long-range DNA polymerases were selected to perform the LR-PCR, including Phusion Hot Start II DNA Polymerase (Thermo Scientific, Vilnius, Lithuania), KAPA HiFi HotStart DNA Polymerase (KAPA

Biosystems, Cape Town, South Africa), and PrimeSTAR GXL DNA Polymerase (TaKaRa Bio, Shiga, Japan). For each selected polymerase, assays were optimized according to the manufacturers' guidelines and set up as shown in Table 1. To assess the results of LR-PCR amplification, PCR products were run on 1% agarose gel and visualized by staining with GelRed Nucleic Acid Gel Stain (Biotium, Hayward, CA). Image scanning was performed with GeneSys software version 1.5.4 (https://www.syngene.com/software/genesys-rapid-gel-image-capture/). To confirm that the amplicons were derived from *FLG* exon 3, nested PCRs were performed using in-target forward primer (5'-AGTGATAGTGAGGGACATTCAGAG-3') and in-target reverse primer (5'-CCGTCTCCTGATTGTTTGTCCT-3') to amplify a 473 bp DNA fragment inside the previously amplified target region. These reactions containing 1 µl of first-round PCR products, 1 µl of the two F/R primers (10 µM), 10 µl AmpliTaq Gold FastPCR Master Mix (Thermo Scientific, Vilnius, Lithuania) and distilled water up to 20 µl were heated to 95 °C for 10 min, followed by 30 cycles at 96 °C for 3 s, 59 °C for 3 s, and 68 °C for 10 s, and a final extension at 72 °C for 10 s.

The first attempt to optimize the second-round PCR protocol was performed mixing 1 µl of first-round PCR product, 1 µl BUP, 5 µl 5 × PrimeSTAR GXL Buffer, 2 µl dNTP, 0.5 µl PrimeSTAR® GXL DNA polymerase and distilled water up to 25 µl, using the same cycling conditions as for the previously optimized first-round PCR protocol (Table 1). Subsequently, second-round PCR protocol was optimized by decreasing cycle number to seven and increasing the amount of template DNA to 3 µl. Second-round PCR products were purified using 0.45X AMPure PB beads (Pacific Biosciences) and DNA concentrations were read on a Qubit 2.0 Fluorometer using the Qubit dsDNA Broad Range Assay Kit (Invitrogen Life Technologies, USA).

Declarations

References

1. 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 (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 38(3):337–342. https://doi.org/10.1038/ng1743

2. Sandilands A, Sutherland C, Irvine AD, McLean WH (2009) Filaggrin in the frontline: role in skin barrier function and disease. J Cell Sci 122(Pt 9):1285–1294. https://doi.org/10.1242/jcs.033969

3. Steinert PM, Cantieri JS, Teller DC, Lonsdale-Eccles JD, Dale BA (1981) Characterization of a class of cationic proteins that specifically interact with intermediate filaments. Proc Natl Acad Sci USA 78(7):4097–4101. https://doi.org/10.1073/pnas.78.7.4097. PMID:6170061;PMCID:PMC319733

4. Jensen JM, Proksch E (2009) The skin's barrier. G Ital Dermatol Venereol 144(6):689–700

5. Furue M (2020) Regulation of filaggrin, loricrin, and involucrin by IL-4, IL-13, IL-17A, IL-22, AHR, and NRF2: pathogenic implications in atopic dermatitis. Int J Mol Sci 21(15):5382. https://doi.org/10.3390/ijms21155382

6. Henry J, Toulza E, Hsu CY, Pellerin L, Balica S, Mazereeuw-Hautier J, Paul C, Serre G, Jonca N, Simon M (2012) Update on the epidermal differentiation complex. Front Biosci (Landmark Ed) 17(4):1517–1532. https://doi.org/10.2741/4001

7. Brettmann EA, de Guzman Strong C (2018) Recent evolution of the human skin barrier. Exp Dermatol. 27(8):859–866. https://doi.org/10.1111/exd.13689

8. Irvine AD, McLean WH, Leung DY (2011) Filaggrin mutations associated with skin and allergic diseases. N Engl J Med 365(14):1315–1327. https://doi.org/10.1056/NEJMra1011040

9. Gibbs NK, Norval M (2011) Urocanic acid in the skin: a mixed blessing? J Invest Dermatol 131(1):14–17. https://doi.org/10.1038/jid.2010.276

10. Margolis DJ, Mitra N, Wubbenhorst B, Nathanson KL (2020) Filaggrin sequencing and bioinformatics tools. Arch Dermatol Res. 312(2):155–158. https://doi.org/10.1007/s00403-019-01956-3

11. Gan SQ, McBride OW, Idler WW, Markova N, Steinert PM (1990) Organization, structure, and polymorphisms of the human pro-filaggrin gene. Biochemistry 29(40):9432–9440. https://doi.org/10.1021/bi00492a018.Erratum.In:Biochemistry.1991Jun11;30(23):5814

12. 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 (2007) Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. Nat Genet 39(5):650–654. https://doi.org/10.1038/ng2020

13. Brown SJ, McLean WH (2012) One remarkable molecule: filaggrin. J Invest Dermatol. 132(3 Pt 2):751–62. https://doi.org/10.1038/jid.2011.393

14. Hoober JK, Eggink LL (2022) The discovery and function of filaggrin. Int J Mol Sci 23(3):1455. https://doi.org/10.3390/ijms23031455

15. Mandelker D, Schmidt RJ, Ankala A, McDonald Gibson K, Bowser M, Sharma H, Duffy E, Hegde M, Santani A, Lebo M, Funke B (2016) Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing. Genet Med 18(12):1282–1289. https://doi.org/10.1038/gim.2016.58

16. Quick J, Loman N, Duraffour S et al (2016) Real-time, portable genome sequencing for Ebola surveillance. Nature 530:228–232. https://doi.org/10.1038/nature16996

17. Mantere T, Kersten S, Hoischen A (2019) Long-Read Sequencing Emerging in Medical Genetics. Front Genet 7(10):426. https://doi.org/10.3389/fgene.2019.00426

18. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouil Q (2020) Opportunities and challenges in long-read sequencing data analysis. Genome Biol 21(1):30. https://doi.org/10.1186/s13059-020-1935-5

19. Athanasopoulou K, Boti MA, Adamopoulos PG, Skourou PC, Scorilas A (2021) Third-generation sequencing: the spearhead towards the radical transformation of modern genomics. Life (Basel) 12(1):30. https://doi.org/10.3390/life12010030

20. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C (2018) The third revolution in sequencing technology. Trends Genet 34(9):666–681. https://doi.org/10.1016/j.tig.2018.05.008

21. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17(6):333–351. https://doi.org/10.1038/nrg.2016.49

22. Wenger AM, Peluso P, Rowell WJ, Chang PC, Hall RJ, Concepcion GT, Ebler J, Fungtammasan A, Kolesnikov A, Olson ND, Töpfer A, Alonge M, Mahmoud M, Qian Y, Chin CS, Phillippy AM, Schatz MC, Myers G, DePristo MA, Ruan J, Marschall T, Sedlazeck FJ, Zook JM, Li H, Koren S, Carroll A, Rank DR, Hunkapiller MW (2019) Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol. 37(10):1155–1162. https://doi.org/10.1038/s41587-019-0217-9

23. Mobley, I. (2022, November 13). Long-read sequencing vs short-read sequencing. Front Line Genomics. https://frontlinegenomics.com/long-read-sequencing-vs-short-read-sequencing

24. Sorrentino E, Albion E, Modena C, Daja M, Cecchin S, Paolacci S, Miertus J, Bertelli M, Maltese PE, Chiurazzi P, Stuppia L, Colombo L, Marceddu G (2022) PacMAGI: A pipeline including accurate indel detection for the analysis of PacBio sequencing data applied to RPE65. Gene. 20(832):146554. https://doi.org/10.1016/j.gene.2022.146554

25. Jia H, Guo Y, Zhao W, Wang K (2014) Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer. Sci Rep 18(4):5737. https://doi.org/10.1038/srep05737

26. Rhoads A, Au KF (2015) PacBio sequencing and its applications. Genom Proteom Bioinform 13(5):278–89. https://doi.org/10.1016/j.gpb.2015.08.002

Authors and Affiliations

Chiara Mareso[1] · Elena Albion[1] · William Cozza[1] ⓘ · Benedetta Tanzi[2] · Stefano Cecchin[2] · Paolo Gisondi[3] · Sandro Michelini[4] · Francesco Bellinato[3] · Serena Michelini[5] · Silvia Michelini[6] · Matteo Bertelli[1,2] · Giuseppe Marceddu[1]

Chiara Mareso
chiara.mareso@assomagi.org

Elena Albion
elena.albion@assomagi.org

Benedetta Tanzi
benedetta.tanzi@assomagi.org

Stefano Cecchin
laboratorio@assomagi.org

Paolo Gisondi
paolo.gisondi@univr.it

Sandro Michelini
sandro.michelini@aslroma6.it

Francesco Bellinato
francesco.bellinato@univr.it

Serena Michelini
serenamichelini@gmail.com

Silvia Michelini
silviamichelini1992@gmail.com

Matteo Bertelli
matteo.bertelli@assomagi.org

Giuseppe Marceddu
giuseppe.marceddu@assomagi.org

[1]   Diagnostics Unit, MAGI EUREGIO, 39100 Bolzano, Italy

2   Diagnostics Unit, MAGI'S LAB, 38068 Rovereto, Italy

3   Section of Dermatology and Venereology, Department of Medicine, University of Verona, 37129 Verona, Italy

4   Vascular Diagnostics and Rehabilitation Service, Marino Hospital, ASL Roma 6, 00047 Marino, Italy

5   Unit of Physical Medicine, "Sapienza" University of Rome, 00189 Rome, Italy

6   Unit of Neurosurgery, University of "Tor Vergata", 00133 Rome, Italy