Late cornified envelope proteins are predominantly expressed in the skin and other cornified epithelia. On the basis of sequence similarity, this 18-member homologous gene family has been subdivided into six groups. The LCE3 proteins have been the focus of dermatological research because the combined deletion of LCE3B and LCE3C genes (LCE3B/C-del) is a risk factor for psoriasis. We previously reported that LCE3B/C-del increases the expression of the LCE3A gene and that LCE3 proteins exert antibacterial activity. In this study, we analyzed the antimicrobial properties of other family members and the role of LCE3B/C-del in the modulation of microbiota composition of the skin and oral cavity. Differences in killing efficiency and specificity between the late cornified envelope proteins and their target microbes were found, and the amino acid content rather than the order of the well-conserved central domain of the LCE3A protein was found responsible for its antibacterial activity. In vivo, LCE3B/C-del correlated with a higher beta-diversity in the skin and oral microbiota. From these results, we conclude that all late cornified envelope proteins possess antimicrobial activity. Tissue-specific and genotype-dependent antimicrobial protein profiles impact skin and oral microbiota composition, which could direct toward LCE3B/C-del–associated dysbiosis and a possible role for microbiota in the pathophysiology of psoriasis.

INTRODUCTION

The skin protects against dehydration and infection owing to the barrier function of the epidermis. In the last living epithelial cell layer, the stratum granulosum, structural proteins such as FLG, hornerin, and the family of late cornified envelope (LCE) proteins are expressed. For long, these were thought to be exclusively contributing to the formation of corneocytes and the physical barrier of the skin. Recent studies however uncovered their antimicrobial activity (Gerstel et al., 2018; Hansmann et al., 2015; Miajlovic et al., 2010; Niehues et al., 2017). Hornerin and LCEs are classified as cationic intrinsically disordered antimicrobial peptides (AMPs), in which high content of disorder-promoting amino acids define the antimicrobial effect of these proteins (Latendorf et al., 2019). These latest findings highlight the function and importance of previously thought to be exclusive structural epidermal proteins in skin barrier homeostasis.

The cutaneous microbiome consists of bacteria, fungi/yeasts, and viruses that contribute to skin health, as exemplified by the ample evidence of (inflammatory) cutaneous diseases associated with microbial dysbiosis (Fyhrquist et al., 2019; Holmes, 2013; Kong et al., 2012; Langan et al., 2018; Tomida et al., 2013). The human skin microbiota composition can be affected by multiple endogenous factors, such as AMP expression profiles, (microbial) nutrient abundance or scarcity, or other niche-specific abiotic characteristics. The first is highlighted by the near absence of skin infections in psoriasis lesions characterized by high levels of AMPs, in contrast to prevalent skin infections in atopic dermatitis where AMP levels are relatively low (Büchau and Gallo, 2007; de Jongh et al., 2005; de Koning et al., 2011; Harder et al., 2010). The latter is illustrated by the cutaneous microbiome of patients with ichthyosis vulgaris. In this study, deficiency of FLG protein, due to loss-of-function sequence variations in the FLG gene, relates to a lower abundance of the so-called Gram-positive anaerobic cocci, probably owing
to the lack of bacterial nutrients normally provided by FLG breakdown products (Zeeuwen et al., 2017a). Less Gram-positive anaerobic cocci are also found in patients with atopic dermatitis, where FLG levels are also lower either owing to the genetic predisposition (FLG sequence variations) or the local inflammatory milieu (Fyhrquist et al., 2019). Current knowledge about the interaction between the cutaneous microbiome and its host is growing and contributes to the notion that skin health relies on the interaction with (commensal) cutaneous microbiota.

One family of epidermal AMPs, the LCE proteins, is of particular interest because certain genetic variants in this family are presumed to be predisposing factors to the chronic inflammatory skin disease psoriasis. High levels of AMPs in psoriasis might be considered beneficial, that is, by reducing the risk of skin infections. However, AMPs such as hBD2, also exert anti-inflammatory and proinflammatory activities that may contribute to the inflammatory processes in psoriasis (Hollox et al., 2008; Jansen et al., 2009). The deletion of LCE3B and LCE3C genes (LCE3B/C-del), as we have previously reported, is a commonly replicated risk factor for developing psoriasis (de Cid et al., 2009; Zhang et al., 2009). LCE genes are expressed in the epidermis and the oral epithelium and were assumed to encode structural proteins with a role in physical epithelial barrier formation (Bergboer et al., 2011; de Koning et al., 2012). We have shown that despite the high sequence similarity of the LCE family members, their baseline expression in the skin as well as their expression regulation in disease are different (Niehues et al., 2016). These data suggested a nonredundant role of the different LCE proteins. Our hypothesis-driven functional studies have revealed that LCE3B/C-del causes an upregulation of the flanking LCE3A gene and that the loss of these two LCE proteins does not have an obvious effect on skin barrier function. No genotype-dependent effect was observed for the inside-out or outside-in physical skin barrier function in vitro in three-dimensional epidermal equivalents derived from LCE3B/C-del and wild-type keratinocytes. Remarkably, further analysis revealed a completely other function of the LCE3 proteins, namely the antimicrobial activity against a variety of bacterial taxa at low micromolar concentrations (Archer et al., 2017; Niehues et al., 2017).

In this study, we aimed to investigate the antimicrobial properties of members of all six LCE subgroups on a broad panel of microbes. We investigated the mechanism of bacterial killing by LCE proteins and the downstream consequence of high LCE3A expression due to the LCE3B/C-del genotype on the cutaneous and oral microbiota of healthy individuals.

RESULTS

Members of all LCE subgroups have antibacterial activity

In a previous study, we have shown that a subgroup of the LCE family, the LCE3 proteins, exerts antibacterial activity (Niehues et al., 2017). In this study, we analyzed other LCE protein subgroups for their antimicrobial activity. First, we performed a high-throughput screening assay with full-length synthetic LCE proteins with one member of each LCE subgroup (LCE1 to LCE6) to test the antibacterial activity against a panel of skin-associated commensal and pathogenic bacteria. For this, an assay was developed to determine whether 10 μM LCE protein is sufficient to kill all bacteria (10^4–10^5 colony-forming units [CFUs/ml] present in a low salt buffer within 2 hours. Complete killing would result in the lack of bacterial growth after addition of a growth medium (optical density [OD] < 20%) to the assay plate. Importantly, the presence of bacterial growth (OD >20%) does not imply a lack of antimicrobial efficacy, considering that incomplete killing will ultimately result in bacterial growth in this assay by the remainder of live bacteria. All LCE proteins were able to completely kill at least one or more strains tested (Figure 1a). The most broad-spectrum antibacterial activity was demonstrated for LCE6A, being active against all strains tested, whereas LCE1F only showed the complete killing of Corynebacterium aurimucosum. This bacterium appeared to be sensitive to all other tested LCE proteins. Next, we analyzed the effect of a panel of LCE proteins on inhibition of two skin pathogenic bacteria, Staphylococcus aureus and Pseudomonas aeruginosa, using the gold standard CFU counting by serial dilution and plate cultures. Results indicate a sensitivity of S. aureus against most LCE proteins at a concentration of 10 μM (Figure 1b). P. aeruginosa was affected more strongly: all LCE proteins, except for LCE1F and LCE6A, killed this bacterium even at a concentration of 3 μM (Figure 1c). Remarkably, LCE3A showed the highest killing efficiency against P. aeruginosa with a concentration of 1 μM resulting in over 99% killing.

Amino acid content of the central LCE protein domain determines antibacterial activity

We then focused on the correlation of the antibacterial activity of LCE proteins to their well-conserved amino acid sequence domains (Figure 2a) (Henry et al., 2012; Latendorf et al., 2019). We compared the antibacterial activity of the central domain of LCE3A in its original arrangement and in three alternative forms with the same amino acids but in random (scrambled) order (Figure 2b). Antibacterial activity against P. Aeruginosa was found to be independent of the amino acid order because all forms induced similar antibacterial strength.

LCE3 proteins disrupt the integrity of the bacterial cell envelope

To gain more insight into the possible target of LCE-mediated killing, we exposed a Gram-negative (P. aeruginosa) and a Gram-positive (Staphylococcus epidermidis) bacterium to LCE3A, LCE3B, and LCE3C proteins and analyzed the cell morphology before and after exposure by transmission electron microscopy (Figure 3). The blebbing observed after LCE exposure hints toward an LCE3-evoked disturbance of the integrity and stability of the cell envelope, eventually causing extrusion of cell content and cell lysis.

LCE proteins are fungicidal against yeasts

Because not only bacteria but also fungi are part of the cutaneous microbiome, these could be targeted by LCE proteins. To investigate such potential, the activity of different LCE proteins (one for each LCE group) at various concentrations against the fungus Candida albicans was analyzed. LCE2D, LCE4A, and LCE6A reduced C. albicans CFU at a concentration of 3 μM. For LCE3A and LCE5A, fungicidal effects were observed at 1 μM (Figure 4). LCE1F did not affect C. albicans viability at all concentrations tested.
Antimicrobial Activity of LCE Proteins

Antimicrobial efficacy of LCE3A on the skin microbiome
LCE proteins should not solely be effective in a direct in vitro assay when using one specific Microbe but should also effectively kill microbes in a complete microbiome sample, such as the in vivo setting. To analyze such potential, lower back skin swabs of three healthy volunteers were incubated with 10 μM LCE3A protein for 2 hours, followed by blood agar plate culturing (Figure 5). LCE3A protein clearly reduced the total number of CFU, varying from 10- to 100-fold reduction up to complete growth inhibition in individual cases.

Microbiome diversity is different in LCE3B/C-del individuals
The in vitro antimicrobial function of LCE proteins led us to question the biological relevance of this function in vivo. We hypothesized that local, tissue-specific LCE expression and antimicrobial activity thereof contributes to microbial diversity by specifically targeting the growth and survival of microbes at these niches. We took advantage of the naturally occurring LCE3B/C-del genotype in the population, which is known to specifically upregulate LCE3A (Niehues et al., 2017). The antimicrobial activity of LCE3A (as shown in Figures 1–5) may therefore finetune the skin microbiota composition in vivo. We analyzed skin and saliva microbiota in healthy LCE3B/C wild-type and healthy homozygous LCE3B/C-del individuals (Schirmer et al., 2016) (Supplementary Table S1) by 16S ribosomal RNA (rRNA) marker gene Illumina sequencing (Supplementary Table S2). From a biological point of view, the average microbiota profiles were according to expectations, with niche-specific dominant genera characteristic for skin and saliva, most notably Cutibacterium (50.4%), Corynebacterium (7.7%), and Staphylococcus (5.6%) for skin (Supplementary Excel file S1a) and Streptococcus (20.9%), Prevotella (17.5%), and Veillonella (10.6%) for saliva (Supplementary Excel file S1b). In an analysis of the LCE3B/C-del sample contrast, for both microbial niches, we did not find specific bacterial taxa that strongly correlated with the LCE3B/C-del genotype of individuals, nor did we find a clear clustering of samples on this genotype on the basis of their microbiota profiles (Supplementary Figures S1 and S2). Nonetheless, the most prominent differences between the LCE3B/C genotype for skin were in the genus Micrococcus (from 2.2% in LCE3B/C wild-type to 2.8% in LCE3B/C-del on average, P = 0.017) and Cutibacterium (55.1–48.0%, P = 0.080) and for saliva, with a slight decrease in Neisseria (3.9–2.7%, P = 0.052) and Haemophilus (3.1–3.0%, P = 0.074), but these did not survive multiple testing.

No difference in alpha-diversity was observed in terms of phylogenetic distance whole tree and Shannon index for the underlying numbers. Interestingly, on the basis of the UniFrac beta-diversity distance metric, we found that for the saliva niche, the distances between samples within the LCE3B/C-del group were significantly larger than between samples within the LCE3B/C wild-type group. This effect was also significant for skin, albeit to a lesser degree (see Figure 6a and b, which shows the UniFrac distances between samples visualized in principal component analysis space; see Figure 6d and e, where these data are expressed in box plots, with P = 0.014 for skin and P = 0.0003 for saliva). We performed a similar analysis on gut samples available for all volunteers in this study (Figure 6c). In this study, we expected not to find any difference between genotypes because LCE proteins are not expressed in the gut. UniFrac is not compatible with metagenomics data, so we applied the
**Figure 2.** Antibacterial activity of the central LCE3A domain against *Pseudomonas aeruginosa.* (a) Full-length LCE proteins were aligned (https://www.ncbi.nlm.nih.gov/tools/cobalt/). Red indicates the highly conserved positions, and blue indicates lower conservation. In the central domains, positively charged amino acids are indicated in green, hydrophobic amino acids are indicated in yellow, and cysteine residues are indicated in purple. (b) The original central domain of LCE3A as well as three scrambled forms were tested for their antibacterial activity at 10 μM. Graphs show CFU count after 2 hours after incubation, followed by overnight growth on blood agar. A total of $1 \times 10^2$ CFU/ml is the detection limit, therefore values $<1 \times 10^2$ CFU/ml are not depicted. Assays were performed in triplicate. Statistical significance for control versus different dosages, for each LCE protein, was tested by nonparametric one-sided Mann–Whitney U ($^{*}P = 0.05$). CFU, colony-forming unit; LCE, late cornified envelope.

**DISCUSSION**

Since the discovery of the LCE gene family (Marshall et al., 2001), it took almost 10 years before the LCE3B/C-del was identified as a risk factor for developing psoriasis (de Cid et al., 2009), and another decade later, we provided
evidence regarding the biological function of the LCE3 proteins as antimicrobials (Niehues et al., 2017). However, the function and the exact antimicrobial mechanism of LCE3 and/or other LCE family members in tissues where these proteins are expressed and their role in psoriasis remain to be explored.

In this study, we report the specificity of antibacterial activity for members of all the six LCE subgroups against a panel of human host-associated bacteria. Despite the high similarity between the LCE proteins at the amino acid sequence level (Figure 2a), antibacterial activity against P. aeruginosa shows remarkable differences in efficacy, indicating that small amino acid variations in LCE proteins are responsible for specific antibacterial activity toward different species. We showed that the central protein domain of LCE3A was equally effective as the full-length protein, irrespective of the amino acid order. This could be explained by the fact that LCE proteins are small; are very cationic, especially in their central domain (Supplementary Table S3); and have an extremely high cysteine content, properties also described for other AMPs that target the bacterial membrane (Harder et al., 2013; Jiang et al., 2008).

We presume that the antimicrobial action of LCE proteins takes place in the stratum corneum. The fact that the stratum corneum is not an aqueous solution but a rather dry layer of lipids and protein makes any prediction about relevant LCE concentrations or even the ionic strength of this (micro)milieu very difficult, if not impossible. In vivo, LCE proteins may be lower than the LCE proteins used in our in vitro assays, but their continuous presence on the skin could therefore still be equally effective. Therefore, we performed an additional experiment in which we tested the antimicrobial activity of LCE3A against S. aureus in a concentration and time series. The resulting data show that a longer incubation time with a low concentration of LCE3A has the same effect as that of a short incubation time, with a relatively high concentration of the antibiotic (Supplementary Figure S3). Moreover, we have previously shown that in a direct experimental comparison of LCE3A and the well-described hBD2 protein, LCE3A has stronger antibacterial activity (Niehues et al., 2017). According to the standards in this research field, the LCE

Figure 4. Antifungal activity of LCE protein against Candida albicans. C. albicans was incubated at concentrations of 1, 3, or 10 μM with full-length LCE proteins of all LCE subgroups (n = 3). Graphs show the CFU count 24 hours before incubation with LCE, followed by overnight growth on blood agar. A total of 1 × 10² CFU/ml is the detection limit, therefore values <1 × 10² CFU/ml are not depicted. Statistical significance for control versus different concentrations for each LCE protein was tested by nonparametric one-sided Mann–Whitney U (*P = 0.05). CFU, colony-forming unit; LCE, late cornified envelope.

Figure 5. Growth of bacteria of an in vivo microbiome swab incubated with LCE3A protein. Full-microbiome swabs of three volunteers were taken from the lower back skin and incubated with or without 10 μM full-length LCE3A protein in sodium phosphate buffer (10 mM, pH 7.4) for 2 hours. Blood agar solid culture media (on Petri dishes) show the CFUs after overnight aerobic culture of these samples in a dilution series (clockwise starting at 12-o’clock). CFU/ml of all bacteria growing on these plates were calculated for all samples as presented at the bottom. CFU, colony-forming unit.
proteins are considered antimicrobial because the LD90 (lethal dose that kills 90% of the inoculum) is found between 0.1 and 10 μM. Considering the LD99 (lethal dose that kills 99% of the inoculum), most of the tested LCE protein (parts) would reach this limit. Although LCE1 appeared less antimicrobial in vitro, the net activity in vivo could possibly be equal to the more potent LCEs, considering their higher expression levels in healthy skin (Bergboer et al., 2011). We hypothesize that only when needed, for example, after skin barrier disruption and at risk for infection, LCE3 proteins are upregulated.

Furthermore, we showed that LCE3A protein can effectively kill microbes isolated from cutaneous ex vivo microbiome samples (Figure 5). The LCE3B/C-del has a dual effect on the LCE expression repertoire, as we have previously shown. In the homozygous state, the expression of both LCE3B and LCE3C will obviously be completely lacking, but the LCE3A gene is expressed at significantly higher levels in the normal skin of these individuals (Niehues et al., 2017). This different epidermal host defense repertoire could result in altered microbiota composition of the skin and oral cavity. We previously speculated that this might provide an evolutionary benefit and hence favor the spread of LCE3B/C-del in the populations outside Africa (allele frequency is 60–70% [Bassaganyas et al., 2013; de Cid et al., 2009]). An increased risk to develop psoriasis could be regarded as the evolutionary cost of having a more favorable antimicrobial response in the epidermis and oral cavity. Our sequencing data revealed that LCE3B/C-del correlated with a higher beta-diversity in the skin and oral microbiota. Low alpha- and beta-diversity is often seen in lesional atopic dermatitis skin, where decreased expression of AMPs is associated with microbiological dysbiosis (de Jongh et al., 2005; Fyhrquist et al., 2019; Kong et al., 2012). However, microbiota studies on the psoriatic lesion skin show conflicting results regarding alpha- and beta-diversity (Yerushalmi et al., 2019). The most
significant difference in within-group beta-diversity between LCE3B/C genotypes was found in the saliva, speculatively related to the higher constitutive LCE3 expression levels as known for stratified oral epithelia when compared with that of the normal skin (Niehues et al., 2016). In other words, LCE3B/C-del samples are on average less similar to each other than wild-type (shown by larger distances between samples), and this effect is much stronger in the oral niche than in the skin. Consequently, the oral microbiome of patients with psoriasis would be interesting to analyze in more depth for LCE3B/C-del genotype-dependent alterations and also on the functional level. However, it should be noted that transcriptome/proteome data for the individuals in this study is lacking, and therefore a direct correlation between microbiota data and LCE expression levels cannot be made.

LCE3B/C-del has been associated with psoriasis and other related autoimmune disorders. The involvement of other LCE family members is reported in several inflammatory diseases, as illustrated in Supplementary Figure S4. Because these associations are often based only on genetic studies, extensive functional studies are necessary to confirm causal relationships between the presence or absence of specific LCE proteins and disease mechanisms and the potential role of microbiome dysbiosis herein.

We conclude that all LCE family members can be regarded as natural antimicrobials and that the absence of two genes (LCE3B/C-del) causes dysbiotic changes in the oral and cutaneous microbiota composition. These differences in diversity may be attributed to increased LCE3A expression in LCE3B/C-del individuals. Because LCE3B/C-del is a genetic risk factor for the development of psoriasis, the relationship between skin and oral microbiota and disease pathogenesis would be a valid option for further studies in genetically defined groups of patients and healthy controls.

MATERIALS AND METHODS

Synthetic LCE proteins

Full-length LCE3A and its central domain were synthesized by solid-phase peptide synthesis, purified to 85–90% purity by reverse-phase HPLC and characterized by electrospray mass spectroscopy (Peptide Suzhou, China). Scrambled forms of the LCE3A central protein domain were designed using an online randomization tool (http://www.cellbiol.com/scripts/randomizer/dna_protein_sequence_rando mizer.php) and were synthesized and purified to >98% purity by HPLC (Proteogenix, Schiltigheim, France). Sequences of scrambled protein are presented in Figure 2b.

Microbial strains and growth conditions

For details of the used bacterial and fungal strains and their growth conditions, see Supplementary Materials and Methods.

Antimicrobial assay

Bacteria were harvested from exponential growth liquid cultures by centrifugation (2,100g, 5 minutes) and resuspended in sodium phosphate buffer (10 mM, pH 7.4) in a concentration of 10^4–10^5 CFU/ml. Bacterial suspensions were exposed to LCE peptides in a volume of 100 µl for 2 hours at 37 °C. After this preincubation period with LCE peptides, the antimicrobial effects were determined through OD measurement or CFUs counting.

As a first antibacterial high-throughput screening method, 100 µl of brain heart infusion was added to the LCE-treated suspensions. Owing to the salt sensitivity of LCE peptides, the antibacterial process was stopped by the administration of brain heart infusion. The bacteria were then grown overnight at 37 °C, and the OD at 490 nm of LCE-treated samples was compared with that of non-LCE-treated controls. Growth ≤ 20% was determined as representing effective LCE antibacterial activity.

In the standard antibacterial activity assay (CFU counting), the LCE-treated suspensions were serially diluted in PBS, plated on Columbia agar with 5% sheep blood, and incubated overnight at 37 °C.

Antifungal assay

C. albicans colonies were brought into sodium phosphate buffer (10 mM, pH 7.4). The OD of this suspension was determined using a spectrophotometer at 530 nm. The suspension was diluted to a concentration of ~10^5 CFU/ml and exposed to LCE peptides in a volume of 100 µl for 24 hours at 35 °C. The LCE-treated suspensions were serially diluted in PBS, plated on Columbia agar with 5% sheep blood, and incubated overnight at 35 °C for CFU counting.

Microbiome sampling and sample preparation

Healthy individuals from the 500 functional genomics project cohort (Schirmer et al., 2016) who are homozygous for LCE3B/C-del (n = 51) or wild-type (n = 25) were invited to donate saliva and skin swabs of the lower back. Furthermore, healthy volunteers (n = 3) were swabbed from their lower back for ex vivo testing of LCE antibacterial activity on whole microorganisms. Study procedures, sample collection, and microbial genomic DNA isolation are described in the Supplementary Materials and Methods. In advance, medical ethical committee (Commissie Mensgebonden Onderzoek Arnhem-Nijmegen) approval and individual written informed consent were obtained. The study was conducted according to the Declaration of Helsinki principles.

16S rRNA gene library preparation

To generate PCR amplicon libraries for skin and oral microbiota sequencing, sample-specific amplicons were generated in house by PCR with primers targeting the V3–V4 hypervariable region of the 16S small-subunit rRNA gene. Microbial genomic DNA (10–25 ng) was used as a template using 357F and 802RV2 primers appended with Illumina adaptor sequences (Illumina, San Diego, CA) (Zeeuwen et al., 2017b).

16S rRNA gene sequencing

Illumina 16S libraries (Illumina) were barcoded by an additional PCR (12 cycles), multiplexed, and sequenced at BaseClear BV (Leiden, The Netherlands) on an Illumina MiSeq paired-end 300 cycles system (Illumina). For further details, see Supplementary Materials and Methods.

16S rRNA marker gene sequencing data analysis

For 16S rRNA marker gene sequencing analysis, see Supplementary Materials and Methods and Supplementary Excel files S1 and S2.

Comparison of skin and saliva microbiota composition with gut microbiota data

For a comparison between (genus level) microbiota composition by 16S rRNA sequencing for skin and saliva samples and with gut microbiota samples by metagenomics from the 500 functional genomics cohort (Schirmer et al., 2016), see Supplementary Materials and Methods and Supplementary Excel files S1 and S2.
H Niehues et al.
Antimicrobial Activity of LCE Proteins

Statistics and multivariate analyses
For details of the statistical analysis of sequencing data between LCE3B/C-del genotypes, see Supplementary Materials and Methods.

Data availability statement
The authors confirm that all data supporting the findings of this study are available within the article and its Supplementary Materials. Datasets related to this article can be found at the European Nucleotide Archive database (Paksershet et al., 2014) under study accession number PRJEB46654 (http://www.ebi.ac.uk/ena). The sequencing data are available in fastq-format, including corresponding metadata for each sample.

ORCiDs
Hanna Niehues: http://orcid.org/0000-0002-6954-6955
Danique A. van der Krieken: http://orcid.org/0000-0002-6795-7881
Thomas H.A. Ederveen: http://orcid.org/0000-0003-0068-1275
Jos P.H. Smits: http://orcid.org/0000-0003-0915-8624
Patrick A.M. Jansen: http://orcid.org/0000-0002-7894-2140
Laura van Niftrik: http://orcid.org/0000-0002-7011-7234
Rob Mesman: http://orcid.org/0000-0003-0877-3676
Mihai G. Netea: http://orcid.org/0000-0003-2421-6052
Jos P.H. Smits: http://orcid.org/0000-0003-0915-8624
Joost Schalkwijk: http://orcid.org/0000-0002-1308-1319
Ellen H. van den Bogaard: http://orcid.org/0000-0003-4846-0287
Patrick L.J.M. Zeeuwen: http://orcid.org/0000-0002-6878-2438

CONFLICT OF INTEREST
The authors state no conflict of interest.

ACKNOWLEDGMENTS
The authors thank all volunteers that have participated in this study. This study is funded by a TOP grant from ZonMw (91211052) and by a grant from the LEO Foundation (LF18068).

AUTHOR CONTRIBUTIONS
Conceptualization: HN, JS, EHVDB, PLJMZ; Data Curation: HN, DAVIDK, THAE, PAMJ, LVN, RM, MGN, JPJS, JS, EHVDB, PLJMZ; Formal Analysis: HN, DAVIDK, THAE, JPJS, PAMJ, LVN, RM, MG; Funding Acquisition: JS, EHVDB, PLJMZ; Investigation: HN, DAVIDK, THAE, PAMJ, LVN, RM, MGN, JPJS, JS, EHVDB, PLJMZ; Methodology: HN, EHVDB, PLJMZ, DAVIDK, THAE, JPJS, PAMJ; Project Administration: HN, DAVIDK, THAE, PAMJ, LVN, RM, MGN, JPJS, JS, EHVDB, PLJMZ; Resources: HN, DAVIDK, THAE, PAMJ, LVN, RM, MG, JPJS, JS, EHVDB, PLJMZ; Software: THAE; Supervision: EHVDB, PLJMZ; Validation: HN, DAVIDK, THAE, PAMJ, LVN, RM, MGN, JPJS, JS, EHVDB, PLJMZ; Visualization: HN, DAVIDK, THAE, PAMJ, LVN, RM, MGN, JPJS, JS, EHVDB, PLJMZ; Writing: Original Draft Preparation: HN, DAVIDK, THAE, JPJS; Writing - Review and Editing: EHVDB, PLJMZ, DAVIDK, THAE, PAMJ, LVN, RM, MG, JPJS, JS

Disclaimer
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2021.11.036.

REFERENCES
Niehues H, van Vlijmen-Willems IM, Bergboer JG, Kersten FF, Narita M, Hendriks WJ, et al. Late cornified envelope (LCE) proteins: distinct expression patterns of LCE2 and LCE3 members suggest nonredundant


SUPPLEMENTARY MATERIALS AND METHODS

Microbial strains and growth conditions
Staphylococcus aureus (ATCC 29213), S. epidermidis (ATCC 12228), Pseudomonas aeruginosa (ATCC 27853), Acinetobacter baumannii (ATCC 345059), Streptococcus pyogenes (ATCC 12344), Corynebacterium aurimucosum (clinical isolate), S. capitis (clinical isolate), and S. hominis (clinical isolate) were grown on Columbia agar with 5% sheep blood (BD, Franklin Lakes, NJ). From this plate, one colony was transferred to brain heart infusion medium (Mediaproducts BV, Groningen, The Netherlands) and incubated overnight at 37 °C. The next day, the bacterial suspensions were diluted 100 times in brain heart infusion and incubated for another 2.5 hours to achieve exponential growth of the bacteria.

Subcultures of Candida albicans (clinical isolates) were grown on Columbia agar with 5% sheep blood and cultured overnight at 35 °C, followed by a second subculture on Columbia agar incubated overnight at 35 °C. From this second subculture, colonies were used in antifungal assays.

Transmission electron microscopy
Approximately 10^5 colony-forming units/ml bacteria were incubated in sodium phosphate buffer (10 mM, pH 7.4) with 10 µM late cornified envelope peptide for 1 hour. Thereafter, 2.5 µl of this reaction was loaded on a 100 mesh copper TEM grid containing a carbon-coated Formvar support film. Grids were incubated for 5 minutes at room temperature before blotting away the excess liquid and then air dried for about 24 hours. Samples were imaged in a Jeol 1010 transmission electron microscope, operating at 60 kV.

Microbiome sampling
All suitable volunteers in this study were selected according to the inclusion/exclusion criteria as approved by a protocol from the National Institutes of Health (Bethesda, MD) Human Microbiome Project (https://www.hmpdacc.org/) and as previously published by our group (Zeeuwen et al., 2017). In advance, medical ethical committee (Commissie Mensgebonden Onderzoek Arnhem-Nijmegen) approval and individual written informed consent were obtained. The study was conducted according to the Declaration of Helsinki principles.

Skin samples were collected from the lower back and obtained by swabbing 4 cm² skin area using Sterile Catch-All Sample Collection Swabs (Epicientre Technologies, San Diego, CA) soaked in sterile SCF-1 solution (50 mM Tris buffer [pH 8], 1 mM EDTA, and 0.5% Tween-20) as previously described (Zeeuwen et al., 2017). As negative controls, we took two mock swabs, which were only exposed to ambient air. DNA was extracted from the swabs using the Mobio Ultracean Microbial DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, CA) with modifications as described previously (Zeeuwen et al., 2012). DNA samples were stored at −20 °C until further processing.

Subjects had to spit roughly 5 ml saliva into a 50 ml tube, which was centrifuged at 2,600g for 15 minutes to get a pellet of buccal cells. The supernatant was transferred to two 2 ml Eppendorf tubes that were subsequently centrifuged at 10,000g for 15 minutes to collect bacteria. The supernatant was discarded until there was 150 µl of liquid left above both bacterial pellets. Both tubes were pooled into one tube, which was supplemented with 300 µl Mo Bio Bead solution, and DNA extraction was performed as described earlier.

Study procedures
1. Subjects should not eat and drink 1 hour before saliva collection.
2. Subjects should avoid the use of body lotion on the lower back skin.
3. Subjects should avoid the use of tanning beds for 48 hours before sampling visit.
4. Subjects do not shower on the sampling day (they can wash their hair but avoid leaking soap to the rest of their body and do not rub up against their lower back skin [avoid soap as much as possible]).
5. Subjects do not scrub the lower back skin with a towel, they should dab a little if necessary.
6. Subjects should avoid the use of body lotion on the lower back on the sampling day.
7. Subjects should refrain from using mouthwash on the day of saliva collection.
8. Subjects should not eat and drink 1 hour before saliva collection.

16S ribosomal RNA gene sequencing
Illumina 16S libraries (Illumina, San Diego, CA) were barcoded by an additional PCR (12 cycles), multiplexed, and sequenced at BaseClear BV (Leiden, The Netherlands) on an Illumina MiSeq paired-end 300 cycles system (Illumina). The sequencing run was analyzed with the Illumina CASAVA pipeline (version 1.8.3, Illumina) with demultiplexing on the basis of the sample-specific barcodes. The raw sequencing data produced were processed by removing the sequence reads of too low quality (only passing filter reads were selected) and discarding reads containing adaptor sequences or failing PhiX Control with an in-house filtering protocol. A quality assessment on the remaining reads was performed using the FASTQC quality control tool, version 0.10.0. Multiplexed FASTQ files as provided by BaseClear were first used to generate Illumina paired-end sequence pseudoreads by PEAR (Zhang et al., 2014) using the default settings.

16S ribosomal RNA marker gene sequencing data analysis
For 16S ribosomal RNA marker gene sequencing analysis, a customized Python workflow based on Quantitative Insights Into Microbial Ecology (version 1.8) (Caporaso et al., 2010b) was adopted (http://qiime.org). Reads were filtered for chimeric sequences using the UCHIME algorithm, version 4 (Edgar, 2010). The operational taxonomic units (OTUs) clustering (by open reference OTU picking), taxonomic assignment, and reference alignment were done with the pick_open_reference_otus.py workflow script of Quantitative Insights Into Microbial Ecology, using UCLUST (USEARCH) (Edgar, 2010) as the clustering method (97% identity) and Greengenes, version 13.8, as a 16S reference database (McDonald et al., 2012). OTUs consisting of only a single sequence were removed. Sequences that could not be aligned by PyNAST (Caporaso et al., 2010a) against the 16S reference alignment were removed. Hierarchical clustering of samples was performed using unweighted pair group method with arithmetic mean, with weighted UniFrac as its
distance measure as implemented in Quantitative Insights Into Microbial Ecology 1.8. Alpha-diversity metrics (PD whole tree, Chao1, Observed Species, and Shannon) were calculated by bootstrapping 19,330 reads per sample for skin samples or 26,122 reads per sample for saliva samples and by taking the average over 10 trials. Figures resulting from Quantitative Insights Into Microbial Ecology clustering analyses were generated using the interactive tree of life tool (Letunic and Bork, 2011). Note that owing to technical limitations in the resolution of 16S ribosomal RNA marker gene sequencing, OTU calling on the level of species should be interpreted with caution. The 16S read, OTU, and alpha-sequencing, OTU calling on the level of species should be interpreted with caution. The 16S read, OTU, and alpha-diversity statistics can be found in Supplementary Excel file S2b, and c and see Supplementary Excel file S1a and b for the 16S genus level compositional matrices for skin and saliva, respectively.

Comparison of skin and saliva microbiota composition with gut microbiota data of 500 functional genomics cohort
For a comparison between (genus level) microbiota composition by 16S ribosomal RNA gene for skin and saliva samples and with gut microbiota samples by metagenomics, the Bray–Curtis dissimilarity metric was applied. This direct comparison was possible because the exact same volunteers were available for both the 500 functional genomics study and this study (see Supplementary Excel file S2a for the overlap between samples). In short, log_{10}-transformed genus level microbial relative abundances were used to calculate the Bray–Curtis distance between samples of the same dataset using the SciPy spatial distance Python module (http://docs.scipy.org). For details about the 500 functional genomics cohort and its applied methodology, we refer to Schirmer et al. (2016). Please see Supplementary Excel file S1c for the metagenomics genus level compositional matrix for the gut (fecal) samples according to 500 functional genomics (i.e., for the sample overlap in volunteers).

Statistics and multivariate analyses
For the microbiota data in this manuscript, statistical significance between late cornified envelope contrasts with regard to taxonomy abundances or distance measures was tested by a nonparametric Mann–Whitney U. Principal component analysis as well as multivariate redundancy analysis were done using Canoco 5.04 (ter Braak and Smilauer, 2012) using default settings of the analysis type Unconstrained or Constrained, respectively. The principal component analysis was corrected for age and gender, and log-transformation was set to 100. Relative abundance values for taxa were used as response data (log-transformation was set to 100, and data were centered, as default by Canoco), and for redundancy analysis, the sample classes as explanatory variables (i.e., late cornified envelope status). Redundancy analysis calculates P-values by randomly permuting the sample classes for 1,000 times. For the colony-forming unit count data of the antimicrobial assays, statistical analysis was performed using nonparametric one-sided Mann–Whitney U by comparing different dosages of one late cornified envelope protein with those of its control.

SUPPLEMENTARY REFERENCES
Supplementary Figure S1. Skin microbiota clustering of LCE3B/C wild-type and LCE3B/C-del volunteers. Each leaf of the tree represents a single skin sample of a volunteer (LCE3B/C wild type encoded as LCE2 and homozygous LCE3B/C-del as LCE0). Samples were clustered on the basis of beta-diversity (between-sample distance) using weighted UniFrac as a distance measure and hierarchical UPGMA as a clustering method. Horizontal bars show the relative abundance microbiota composition on the genus level, whereas reads that could not be classified up to this level are in white. This figure was generated with the iTOL (Letunic and Bork, 2011). HVxxx = HV000; iTOL, interactive tree of life; LB, lower back (skin); LCE, late cornified envelope; LCE3B/C-del, deletion of LCE3C and LCE3C genes; UPGMA, unweighted pair group method with arithmetic mean.
Supplementary Figure S2. Saliva microbiota clustering of LCE3B/C wild-type and LCE3B/C-del volunteers. Each leaf of the tree represents a single saliva sample of a volunteer (LCE3B/C wild type encoded as LCE2 and homozygous LCE3B/C-del as LCE0). Samples were clustered on the basis of beta-diversity (between-sample distance) using weighted UniFrac as a distance measure and hierarchical UPGMA as a clustering method. Horizontal bars show the relative abundance microbiota composition on the genus level, whereas reads that could not be classified up to this level are in white. This figure was generated with the iTOL (Letunic and Bork, 2011). iTOL, interactive tree of life; LCE, late cornified envelope; LCE3B/C-del, deletion of LCE3C and LCE3C genes; OR, oral; UPGMA, unweighted pair group method with arithmetic mean.
Supplementary Figure S3. Time-course antibacterial activity of LCE proteins against *Staphylococcus aureus*. Bacteria were incubated with 1, 3, or 10 μM LCE3A protein for different time periods. Bacterial survival was analyzed by counting CFU after overnight incubation on blood agar. A total of $1 \times 10^2$ CFU/ml is the detection limit, therefore values $<1 \times 10^2$ CFU/ml are not depicted. The assay was performed in triplicate. Statistical significance for control versus different dosages, for each LCE protein, was tested by nonparametric one-sided Man–Whitney U ($^*P = 0.05$). CFU, colony-forming unit; LCE, late cornified envelope.

Supplementary Figure S4. Relationship of LCE proteins with biological functions and human diseases. The TenWise KMAP platform captures validated scientific knowledge about human genes, micro-organisms, metabolites, pathways, phenotypes, diseases, drug compounds, and research workflows in over 200 million biological relations. For the generation of the figure, an API key was used to get Python-scripted access to KMAP for capturing relations of the six LCE human gene groups with diseases and functions. The thickness of the arrows between concepts indicates the amount of available literature describing this relation. By clicking on the gray circles, one is redirected to the PubMed overview page displaying the underlying publications that describe this relation. For downloading the interactive figure, visit https://www.tenwise.nl/wp/wp-content/uploads/2021/01/LCE_290121_final.pdf. API, application programming interface; KMAP, Karnaugh map; LCE, late cornified envelope.
### Supplementary Table S1. Cohort of Healthy Individuals Genotyped for the LCE3B/C Deletion

<table>
<thead>
<tr>
<th>Cohort Details</th>
<th>LCE3B/C Wild-Type</th>
<th>LCE3B/C-del (Homozygous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number individuals</td>
<td>25</td>
<td>51</td>
</tr>
<tr>
<td>Male/female distribution</td>
<td>10:15</td>
<td>25:26</td>
</tr>
<tr>
<td>Mean age</td>
<td>31.6 ± 14.8</td>
<td>32.4 ± 15.0</td>
</tr>
</tbody>
</table>

Abbreviation: LCE3B/C-del, deletion of LCE3C and LCE3C genes.

These numbers are retrieved from Supplementary Excel file S2a. Healthy individuals were part of the 500 Functional Genomics project cohort (Schirmer et al., 2016).

### Supplementary Table S2. Skin and Saliva 16S rRNA Marker Gene Illumina Sequencing Data

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Number of Sequence Reads (Min)</th>
<th>Number of Sequence Reads (Max)</th>
<th>Mean Sequence Reads/Sample</th>
<th>Mean OTU/Sample</th>
<th>OTU (%) Assigned at Genus Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>20,547</td>
<td>53,071</td>
<td>39,059 ± 5,847</td>
<td>841 ± 266</td>
<td>96.3 ± 2.8</td>
</tr>
<tr>
<td>Saliva</td>
<td>26,948</td>
<td>1,37,253</td>
<td>45,093 ± 17,390</td>
<td>1,107 ± 335</td>
<td>91.2 ± 5.3</td>
</tr>
</tbody>
</table>

Abbreviations: max, maximum; min, minimum; OTU, operational taxonomic unit.

These numbers are retrieved from Supplementary Excel File S2b and c.

### Supplementary Table S3. Mw and IEP

<table>
<thead>
<tr>
<th>LCE Full-Length Protein</th>
<th>Amino Acid Count</th>
<th>Mw (g/mol)</th>
<th>IEP</th>
<th>LCE Central Domain</th>
<th>Amino Acid Count</th>
<th>Mw (g/mol)</th>
<th>IEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCE1F</td>
<td>118</td>
<td>11,654</td>
<td>8.83</td>
<td>LCE1F</td>
<td>21</td>
<td>2,615</td>
<td>11.93</td>
</tr>
<tr>
<td>LCE2D</td>
<td>110</td>
<td>11,180</td>
<td>8.53</td>
<td>LCE2D</td>
<td>21</td>
<td>2,642</td>
<td>10.39</td>
</tr>
<tr>
<td>LCE3A</td>
<td>89</td>
<td>9,146</td>
<td>8.81</td>
<td>LCE3A</td>
<td>21</td>
<td>2,517</td>
<td>9.60</td>
</tr>
<tr>
<td>LCE3B</td>
<td>95</td>
<td>9,812</td>
<td>8.57</td>
<td>LCE3B</td>
<td>21</td>
<td>2,517</td>
<td>9.60</td>
</tr>
<tr>
<td>LCE3C</td>
<td>94</td>
<td>9,729</td>
<td>8.73</td>
<td>LCE3C</td>
<td>21</td>
<td>2,636</td>
<td>11.40</td>
</tr>
<tr>
<td>LCE4A</td>
<td>99</td>
<td>9,980</td>
<td>8.64</td>
<td>LCE4A</td>
<td>21</td>
<td>2,523</td>
<td>9.38</td>
</tr>
<tr>
<td>LCE5A</td>
<td>118</td>
<td>11,795</td>
<td>8.75</td>
<td>LCE5A</td>
<td>21</td>
<td>2,495</td>
<td>11.40</td>
</tr>
<tr>
<td>LCE6A</td>
<td>80</td>
<td>9,022</td>
<td>9.14</td>
<td>LCE6A</td>
<td>21</td>
<td>2,547</td>
<td>11.45</td>
</tr>
</tbody>
</table>

Abbreviations: IEP, isoelectric point; LCE, late cornified envelope; Mw, molecular weight.

Mw and IEP of LCE proteins and their corresponding central domains are calculated with the Expacy Compute IEP/Mw webtool (https://web.expasy.org/compute_pi/).