Staphylococcus prevails in the skin microbiota of long-term immunodeficient mice.

Marc Garcia-Garcerà[1,4] Mireia Coscollà[1,2], Koldo Garcia-Etxebarria[1], Juan Martín-Caballero[3], Fernando González-Candelas[4], Amparo Latorre[4], Francesc Calafell[1]

1 Institut de Biologia Evolutiva CSIC-UPF, Dr.Aiguader 88, 08003 Barcelona, Spain

2 Swiss Tropical and Public Health Institute, Basel, Switzerland

3 Animal Facility, Parc de Recerca Biomèdica de Barcelona, Dr.Aiguader 88, 08003 Barcelona, Spain

4 Joint Research Unit on Genomics and Health, Centre for Public Health Research (CSISP) and Cavanilles Institute for Biodiversity and Evolutionary Biology (University of Valencia). Av. Catalunya, 20 46020. Valencia. Spain.

Corresponding Author

Francesc Calafell

Institute of Evolutionary Biology (UPF-CSIC)

Dr. Aiguader 88

08003 Barcelona (Spain)

Phone: +34933160803

Email: francesc.calafell@upf.edu

Running title: microbiota shifts in immunodeficient mouse skin

SUMMARY

Host-commensal relationships in the skin are a complex system governed by variables related to the host, the bacteria, and the environment. A disruption of this system may lead to new steady states, which, in turn, may lead to disease. We have studied one such disruption by characterizing the skin microbiota in healthy and immunodepressed (ID) mice. A detailed anatomopathological study failed to reveal any difference between the skin of healthy and ID mice. We sequenced the 16S rDNA V1-V2 gene region to saturation in ten healthy and ten ID 8-week old mice, and found than all of the healthy and two of the ID mice had bacterial communities that were similar in composition to that of human skin, although, presumably because of the uniform raising conditions, less interindividual variation was found in mice. However, eight ID mice showed microbiota dominated by *Staphylococcus epidermidis.* Quantitative PCR amplification of 16S rDNA gene and of the *Staphylococcus*-specific TstaG region confirmed the previous results and indicated that the quantitative levels of *Staphylococcus* were similar in both groups while the total number of 16S copies was greater in the healthy mice. Thus, it is possible that, under long-term immunodeficiency, which removes the acquired but not the native immune system, *Staphylococcus epidermidis* may inhibit the growth of other bacteria but does not cause a pathogenic state.

Keywords: skin microbiota / 16s rRNA /immunodeficiency / *Staphylococcus*

INTRODUCTION

The skin, as the most external body barrier, protects the organism against external aggressions and infections. Skin is also the first organ that interacts with the external environment, allowing non-pathogenic microorganisms to coexist and interact mutualistically with the organism(Roth & James 1988). The mammalian skin, as many other ecosystems, consists of a structurally complex surface with multiple niches according to environmental, physical and chemical characteristics. All of its appendages, invaginations, and glands turn the skin into the most complex ecosystem of the human body, which is colonized by a wide range of microorganisms(Grice & Segre 2011; Rosenthal et al. 2011).

Although less than 5% of bacterial species are culturable, molecular approaches allow characterizing the bacterial diversity in a given ecosystem using the 16S ribosomal RNA unit as a phylogenetic marker(Amann et al. 1995). The ability to amplify and sequence the whole range of bacterial 16S rRNA using a set of universal primers plus the revolution of next generation sequencing platforms provide sufficient information to assess and compare bacterial diversity in different skin niches in terms of space, time, and state (Hugenholtz & Pace 1996; Turnbaugh et al. 2007).

Bacterial-host skin interactions range from immune system collaboration to external layer post-processing(Roth & James 1988). The indigenous microbiota may play an important role in skin immunity following two different mechanisms: firstly, resident bacteria may have a main role in the activation of the immune system, enabling

the expansion and maintenance of the CD8+ lymphocyte population in skin, preventing the development of allergic disease, and also stimulating the Toll-like receptors (TLRs) in keratinocytes and dendritic cells that will respond more effectively and efficiently to pathogenic insults(Cogen et al. 2008). Secondly, commensal bacteria may actively inhibit the growth of pathogenic bacteria, by secreting signal molecules that activate the expression of host antimicrobial peptides (AMPs) and recruit AMPs to the skin. The indigenous microbiota may also block pathogen quorum sensing, interfering the progression of the infection (Otto 2009; Mehta et al. 2009).

As in many other systems, complex interactions may lead to complex behaviors that can result pathogenic to the host, under certain host genetic predispositions. Specific microorganism-disease relationships have been suggested for complex disorders such as psoriasis, or atopic dermatitis(Fredricks 2001). But, although some evidence for changes in microbial composition under disease has been shown(Gao et al. 2008), there is no evidence for a disease being caused by such a change. However, given the complex interactions and functions of the skin microbiota in the development of the immune system, the study of the possible trigger effect of the microbiota in complex diseases such as atopic dermatitis or psoriasis is interesting. Nevertheless, disease and healthy states should be compared with caution, as any change in microbial composition may be a consequence, rather than a cause, of the environmental condition change brought on by disease.

The alteration of the immune system homoeostasis can be considered as a change in the state of the whole ecosystem. This environmental change may affect the

fitness of the different bacterial communities, mostly those that interact more intensely with the immune system. The change in fitness will depend on the relationship between the bacterial species and the immune system, and it can lead to a significant reduction, even complete depletion, of the bacterial diversity if the environment change is maintained through time. This implies that, over time, the change in species composition will be more dramatic, thus making it harder to find any posited subtle triggering factor.

Here we have studied skin bacterial diversity after long term immunodeficiency using mice as a model. We have compared ID- to normal healthy- (wild type) skin, and observed a strong reduction in bacterial diversity that was not related to any skin pathogenic state(Ley et al. 2006).

RESULTS

Anatomopathology of ID mouse skin

Healthy (H) and ID mice, as explained in the Methods section, were born and housed together in the same room, with similar interaction protocols to reduce stress levels. Hygienic and feeding procedures were identic for both cohorts. The housing protocol allowed the exchange of microbiota among the cohort, to reduce the possible environmental variability. A careful observation of the skin did not show any skin lesion in either H or ID mice susceptible of cutaneous infection or inflammatory disease. A more detailed anatomopathological observation of the ID-mouse skin did not show neither acanthosis nor hyperkeratosis, features common to ichtyiotic disorders. Other epidermal alterations such as spongiosis or keratinocyte ballooning, associated with acute flares of atopic dermatitis, were also absent. Psoriasiform hyperplasia, characterized by the expansion of the dermal papillae through keratinocyte hyperproliferation was neither present in ID mice. Both healthy- and ID-mouse skin showed a thin flat epidermis, with two or three layers of normal keratinocytes. Stratum Spinosum was not observed in either healhy- or ID-mouse skin. This absence is characteristic of hairy regions where hair, rather than epidermis thickness, acts as the main protector for skin. Since ID mice are albino and healthy C57BL/6J mice are not and have black hair, melanin was only found on the bottom of the hair follicle of healthy mice, as expected. Hair follicles in both skin types were normal, and different only in color. Immunohystochemical analysis showed low, normal levels of CD3+, CD4+/CD8+ and CD56+ cells in healthy uninfected skin. Given that levels of lymphocytes in skin are

quite low they were unable to differentiate between healthy and ID skin, which seemed perfectly normal. At the vascular level, low levels of lymphocyte infiltration were observed in both skin types, making them undistinguishable at this level too. No alteration in the number of eosinophilic/basophilic cells was evident in either of both skin types. In summary, no structural or cytological differences were observed between ID- and healthy-mouse skin that could relate ID mice with any of the most common inflammatory/immune skin disorders even under the immunodepressed condition of the ID mice.

Diversity in mouse skin microbiota

The V1-V2 region of the 16S rDNA gene was amplified from skin samples of 8 week-old mice, an age that may be sufficient for the microbiota to colonize normal skin and reach a stable composition. Amplicons were sequenced using the FLX-Titanium platform, to a depth that reached the plateau of the rarefaction curve. A total of 158,541 sequences was obtained; after applying all filters, 143,908 sequences were used for further analyses, with an average 7,200 sequences per sample (range 3,554 - 15,261).

A total of 13 bacterial phyla was found in healthy mice (Figure 1), with >70% of sequences assigned to Proteobacteria, followed by Firmicutes (5-15%), Bacteroidetes (1-10%) and Actinobacteria (0.5-5%). A total of 167 genera was present in healthy samples, 138 of which were represented by more than three sequences. Only 15 genera were present in all healthy samples; four of these represented ~70% of the sequences: *Acinetobacter* (23%), *Escherichia/Shigella* (20%), *Acidovorax* (13%), and

Diaphorobacter (12%), all of them previously described in mammalian skin(Grice et al. 2008; Grice et al. 2009). Other characteristic genera such as *Corynebacterium*, *Propionibacterium*, *Comamonas, Bacteroides*, and *Staphylococcus* were also present at lower frequencies in all healthy samples (see suppl. table 1 for detailed information on the taxonomic assignment). A slightly higher number of genera was found in other human skin studies by sequencing 16S rDNA at a comparable depth. However, interindividual differences were greater in humans(Grice et al. 2008). In contrast, ID mice skin diversity was severely reduced, with more than 90% of the sequences assigned to *Staphylococcus* spp. in 6 of 10 samples and more than 60% in all but two samples, which showed a microbiota composition similar to that of healthy samples. *Staphylococcus* sequences were in an average frequency of 5.2% in healthy samples, with a range of 0.5%-10%. Bacterial diversity was estimated for each sample with Shannon and Chao1 indices (table 1), and compared between sample type using a two-tailed t-test. ID mice were much less diverse than healthy ones (p-value=0.0074), even though they were raised under the same environmental conditions, including temperature, humidity, food and water access, and, as stated above, their skins were anatomically identical. Sample clustering was tested using correspondence analysis. Samples from the same type clustered together in a 2D plot, with the only exception of the two samples with a healthy-like distribution, as expected (Figure 2).

Validation of Staphylococus *levels*

Preferential amplification of the *Staphylococcus* sequences with the 16S rDNA

primers could partially explain these results. As a control, quantitative PCR (qPCR) was performed with both the 16S rRNA gene, and the *Staphylococcus*-specific *TstaG* region. After normalization, Cq50, defined as the time point where half of the maximum intensity is achieved, was calculated by inference and compared between *TstaG* and 16S rDNA amplifications on each sample group (Suppl. Figure 2). Interestingly, both curves from ID mouse samples nearly overlapped in all but two samples (t-test p-value=0.0594) indicating that, according to diversity results, most bacteria in the ID skin were indeed *Staphylococcus spp.* Moreover, the two outliers that behaved like healthy samples in their 16S rDNA sequences (see above) were also the two samples that amplified independently in qPCR amplifications. In contrast, healthy skin qPCR curves were different for both regions (16S and *TstaG*), and fitted correctly with the diversity results (t-test p-value= 0.00125). Cq50 differences were obtained by substraction and the normal distribution was tested the using Lillefors test (p-value>0.05). Cq50 differences in both ID and healthy mouse skin samples were then compared using an unpaired two-tailed t-test. Significant differences were observed between ID and healthy samples (p-value= 0.0074); that is, the amount of 16S copies present in healthy mouse skin was significantly larger than in ID mice. Hence, the diversity shift observed in the sequencing experiments cannot be attributed to primer bias in the *Staphylococcus* species.

Staphylococcus *diversity in H and ID mice*

We compared the distribution of *Staphylococcaceae* phylotypes between healthy- and ID mice. From the 143,908 sequences used for the study, 64,641 sequences were

assigned to *Staphylococcus* by the RDP database(Maidak et al. 2001; Cole et al. 2007). All these sequences, regardless sample origin, were clustered at 98% identity into 641 phylotypes. Together with the RDP reference sequences, a phylogenetic tree was constructed to assess the relationship between the references and the reads. As expected, all phylotypes were located on the genus *Staphylococcus node*, and were widely distributed among lower levels of the taxa, with no outlier in the remaining genera of the family (Figure 3). One of the main issues was the presence of phylotypes assigned to *Staphylococcus aureus*, which is a biomarker of opportunistic pathogenicity. This species was rare in our samples: only singletons assigned to *S. aureus* were found, representing less than 0.5% of all sequences assigned to the *Staphylococcus* genus. Phylotypes assigned to *S. epidermidis* were much more frequent. Diversity analyses of the phylotypes assigned to *Staphylococcus* did not show significant differences between sample types (t-test p-value=0.21) or within samples of the same type (ANOVA p-value$_{(H)}$=0.09, p-value$_{(ID)}$=0.42) when relative Shannon indexes were compared. However, in most samples, regardless their origin, more than 50% of *Staphylococcus* assigned sequences were classified into one main phylotype. In all but two of the ID samples, *S. epidermidis* was the most prevalent taxon.

DISCUSSION

The work we present focuses on microbial diversity of mammal skin under a highly controlled environment using mice as a model. Given that skin is the most external layer in mammalian bodies, we expect that its microbial composition is highly influenced by the environment, and in consequence, high levels of interindividual variability, depending on the different environments the individual has been in contact with. Previous studies on the microbial diversity of skin, most of them focused in human skin, have reported a high inter- and intra- individual variability (Fierer et al. 2010; Costello et al. 2009). These different results could be due to differences in the sampling method used. While previous studies controlled only a few minimum elements of the complex system (hygiene, antibiotic intake, etc), we have performed a curated method of selection of the sampling region of individuals that were grown in a highly controlled environment, with controlled patterns of interaction and isolation. The high environmental homogeneity of the samples results in a high reproducibility of the variability observed in the same cohort, thus leading to a high statistical confidence. Moreover, the reduction in noise has allowed us to reveal equivalences among samples of both cohorts. Despite the fact that more than 90% of sequences in ID mice were assigned to *Staphylococcus sp.*, it is interesting to note that the remaining sequences in ID samples were assigned to the most prevalent genera of healthy samples, suggesting that the initial colonization was similar for all mice, given their common growth environment. Correspondence Analysis (PCoA) Component 1 (DCA1), which explains

72% of the diversity, links perfectly with the relative abundance of *Staphylococcus spp.* assigned reads. But more interestingly, the association of high values of DCA1 with *Actinobacteria* reads suggests that this phylum is highly associated to the skin, and tends to be more resistant to displacement by *Staphylococcus spp.* But the evolution of ID samples resulted in the complete displacement of *Staphylococcus* species, by the competition and inhibition of the other taxa. On the other hand, healthy mice evolved maintaining the diversity or even acquiring new phylotypes that would fit in the niche.

Despite the actual differences in variability, we have found that healthy-mouse skin has a similar bacterial composition to that found in human skin, according to previous reports. Given the drastic variation in composition depending on the sampled region, our results agree with those reported by Grice et al., who sampled the forearm region (Grice et al. 2008). Although the high levels of Proteobacteria are uncommon in other skin regions, the results by Grice et al., both in humans and mice, were similar to ours. More interestingly, we have found that the skin of all but two of the 8-week old ID mice we have analyzed showed higher 16S Quantification Cycle (Cq) levels in qPCR analyses, which implies that less bacteria were present in ID-mice skin than in that of healthy mice. Although both types of mice harbored similar quantitative levels of *Staphylococcus* according to the qPCR experiments, this genus dominated the microbiota of ID-mouse skin. One of the main concerns was the susceptibility of BALB/c mice to *S. aureus* infections(Köckritz-Blickwede et al. 2008). In agreement with the anotomopathological study showing no visible lesions in the skin of either H or ID mice, low levels of *S.aureus* in the commensal skin microbiota suggests different processes

involving systemic susceptibility to nosocomial infection of pathogenic strains of *S. aureus* and skin microbiota diversity. One possible explanation for the dominance of *S. epidermidis* could be that in ID mice (which have anatomically and immunologically normal skin), *Staphylococcus* inhibits the growth of most other bacteria, a process that may still be under way at 8 weeks, since two ID mice showed normal bacterial profiles.

Staphyloccocus epidermidis, the most abundant species in our samples, is one of the main commensals of skin, being almost ubiquitous in all skin regions(Galdbart et al. 2000; Costello et al. 2009), but is also one of the main nosocomial pathogens(Uçkay et al. 2009). This duality of the relationship of *S.epidermidis* with the host has recently become of great interest (Jean-Baptiste et al. 2011). However, despite the opportunistic pathogenesis, *S.epidermidis* has a mostly benign relationship with the host, by a combination of a low virulence potential (compared to other *Staphylococcus* species) and the ability to evade both innate and acquired host defenses(Otto 2009). Thus, *S.epidermidis* is able to avoid cationic antimicrobial peptides (AMPc), one of the main systems of the host innate defense.The AMPc-mediated signal transduction activates the expression of the *dlt* operon and the *mprF* gene, which leads to decreased atraction of additional AMPc(Peschel et al. 1999; Peschel et al. 2001) by lowering the negative charge of *S. epidermidis* cell surface. Moreover, *S.epidermidis* has been proposed to have a probiotic function in healthy skin by preventing the colonization of other pathogens such as *S.aureus*(Otto et al. 2001). However, according to our results under immunodepression, the antimicrobial activity of *S. epidermidis* could result in the depletion of almost all possible competitors. Cross-inhibition of the quorum sensing

system is one of the main mechanisms that *S. epidermidis* uses to compete against other bacteria. *Staphylococcus* quorum sensing is activated by the recognition of the so-called bacterial pheromones (auto-induced peptides AIPs) by the *agr* system, which in *S.epidermidis* seems to be able to inhibit other *Staphylococcus* species (Otto 2001). Whether this mechanism applies to other bacteria is not known yet. And, even if this is the case, the mechanism used by *Staphylococcus* to inhibit the growth of other taxa remains also unclear. Other mechanisms, such as the production of Epidermicin NI01, have been proved to inhibit the growth of a wide range of gram-positive bacteria, including other *Staphylococci* (Sandiford & Upton 2012). A broader antimicrobial activity has been observed in other taxa such as in *Enterococcus durans.* This bacteria produces a 5KDa bacteriocin that inhibits the growth of a broad range of gram-positive and gram-negative bacteria (Line et al. 2008). But it is unknown if this is the case in *S. epidermidis*, and further research is needed in this field. *S. epidermidis* is also able to activate the innate immune system to reduce the possibility of infection. *S. epidermidis,* through the activation of TLR2, is able to increase the expression of ß-defensins and inhibit the growth of other Gram-positive bacteria(Lai et al. 2010). The interaction between *S.epidermidis* and the innate immune system, together with the lack of an acquired immune system could lead to the inhibition of the growth of other bacteria, resulting in the reduction of diversity that we observe. However, more research is needed and other possible explanations, related or not to the active inhibition of other bacteria by *S. epidermidis,* should also be considered.

From a complex system point of view, skin can be considered, depending on the

source, as a multidimensional system with three main types of variables, host, microbial, and environmental, which interact and lead it to a transient system equilibrium. This equilibrium state will be maintained while all variables are constant, but changes in any of them would change drastically the whole system(Moya et al. 2008; Bäckhed et al. 2005; Ley et al. 2008). Previous studies have shown that a microbial shift may lead to altered inflammatory states and impaired healing during diabetic wound progression(Scharschmidt et al. 2009; Grice et al. 2010). The latter studies exemplify the rupture and recomposition of a system when equilibrium is altered and progresses through a pathogenic state. However, we present here an example of a non-pathogenic system in which an indirect element of the system is altered leading to a shift in microbiota without any pathogenic associated condition. In this case, differences between cohorts are anthropically engineered, eliminating the acquired immune system of one of the populations. The marked reduction of CD3+, CD4+, CD8+ lymphocyte populations, will be translated to a reduction in the adaptive inflammatory response to any pathologic or not pathologic infection. It seems, then, that the only way to maintain *Staphylococcus spp.* below a certain threshold level is through the acquired immune system(M. Li et al. 2007). But, while the acquired host defense against *S.epidermidis* is less well understood, we can speculate, from these results, that the acquired immune system may be the main force maintaining the equilibrium of normal microbiota. Further research is certainly needed to investigate this possibility. Moreover, given that most samples, independently of their origin, presented one main *S.epidermidis* phylotype, this result suggests the existence of a dominant strain that was particularly apt at colonizing

skin, while other phylotypes appeared later and competed with varying degrees of success.

The skin microbiome has been considered as a putative triggering factor for complex diseases such as psoriasis or atopic dermatitis(Gao et al. 2008; Dekio et al. 2007). Common inflammatory diseases present an exacerbation of the activity of the innate immune system activity against the target tissue, in this case the skin, with some commonalities with the innate response to infection(Sun et al. 2006; Lande et al. 2007). However, no important differences have been found in bacterial composition of the skin that could be considered as a triggering factor for such diseases. Previous studies on the involvement of bacteria in the development of complex diseases have been performed once the symptoms had already appeared(Dekio et al. 2007; Gao et al. 2008); even in the case of a real involvement, the triggering factor can be far back in time and location and consequently highly masked by the adaptation of the whole system to the new situation. In consequence, further analyses are needed to establish whether the microbiota triggers the change of the state or the state induces the change in the microbiota. This would be crucial to the understanding of the complex host-microbiota-environment system. Our observation was not associated to any pathological state, suggesting that, in this particular case, staphylococcal dominance is a consequence, and not a cause, of the altered condition, and is not apparently followed by a pathogenic state. The analysis of serial samples may shed light on *Staphylococcus* - host relationships and will result on a better understanding of the complex skin system.

CONCLUSIONS

In this work, under a highly controlled environment and careful sampling scheme, we have been able to observe patterns of bacterial diversity associated to skin of mice which can be replicated among individuals. Given that skin is in constant interaction with the environment and the previous information about the high variability in skin microbiota, the stringency in the sample preparation is crucial to further skin metagenomic studies involving health and pathogenic states.

Our observations suggest that the anatomically normal skin of immunodeficient mice is gradually colonized by *Staphylococcus*. That is, we have been able to describe a dramatic change in the composition of the bacterial community of the skin that is triggered by a remote event and that does not lead to an apparent pathogenic state. This observation may contribute to understanding the host-commensal relationships, and how the disruption of this homoeostasis can be related to skin disease.

EXPERIMENTAL PROCEDURES

Mouse skin sampling

Ten healthy wild type and ten immunodeficient (ID) 8 week-old male mice were euthanized according to a local IRB-board (PRBB, IACU committee) approved protocol, and a region of 3x3cm was excised from the dorsum-lumbar region, using a sterile blade, and frozen in liquid nitrogen to preserve the integrity of the skin. Healthy (H) mice belonged to the C57BL/6J strain, and were provided by Charles River (Wilmington, MA) and ID mice, provided by the same company, belonged to the C.B-17/Icr-PrkdcSCID/IcrIcoCrl strain(Bosma et al. 1983). These mice are homozygous for the severe combined immune deficiency spontaneous mutation *Prkdc*^{scid}, and are characterized by the absence of functional T cells and B cells, lymphopenia, hypogammaglobulinemia, and a normal hematopoietic microenvironment. These strains were initially chosen because they have been reared at our animal facility for a large number of generations and are perfectly adapted to this environment, reducing possible variation introduction from the suppliers(Rodrigue & Lavoie 1996). We consider that the genetic variation among H and SCID mice is low enough to accept C57BL/6J mice as controls for this study. The genetic background of SCID mice is an admixture of Balb/c and C57BL/6J strains, and the genomic differences between these two strains are around 30,000 SNPs, with less of 10% of these SNPs located in coding or regulatory regions(Keane et al. 2011; Mouse Genome Sequencing Consortium et al. 2002).

Moreover, environmental factors such as skin region, supplier, hygienic and

feeding ranges, and social behaviour have been considered as the most important factors introducing variability in skin microbial diversity. Then, to reduce as much as possible the effect of these factors in the study, ID and H parental mice were acquired from the same supplier. The mice used in this study were born and housed in the same room, with identical feeding and hygienic rates, allowing them to freely exchange their microbiota. All the processes involving mice were accredited by AAALAC international.

Skin samples were subsequently split, under freezing conditions to preserve all genetic material, using a 4mm punch blade and stored at -80ºC for subsequent experiments.

Immunohistochemistry

One cylindrical portion from each individual, 4mm wide 5mm deep was stored in neutral-buffered formalin for 24h before embedding in paraffin, after partially removing hair. Hematoxilin-Eosin (H-E) staining was performed to assess possible structural disruption on immunocompromised skin. To assess immune cell infiltration or underlying infection process, CD3, CD4 and CD8 staining with a primary rabbit anti-mouse antibody (1:20 dilution, Dako, Glostrup, Denmark) was performed on paraffin-embedded tissue sections. Biotin-labeled goat anti-rabbit secondary antibody was used (1:500 dilution. Dako), and stained with AEC Substrate kit (Vector Labs, Burlingame, CA). Staining tissue sections were visualized with a Leica AF6000 E Image acquisition station (Leica Microsystems, Netzlar, Germany).

16S rDNA amplification and sequencing

DNA was extracted from a skin portion of comparable weight for each individual, with the DNeasy tissue kit (Qiagen, Valencia, CA) following the modified protocol for Gram positive bacteria and adding a homogenization step with a mechanical homogenizer IKA Ultraturrax (Thermo Scientific, Waltham, MA). Double strain DNA was quantified with Picogreen™ (Thermo Scientific). For each mouse, two replicate 50µL-25cycle PCRs were prepared using fusion primers (AdaptorA-10-nt barcode-8F, AdaptorB-355R) for the Titanium 454 platform (Roche Applied Science. Penzberg. Germany) which amplify the V1-V2 region of the 16S rDNA. Products of both PCR reactions were pooled together, purified by filtration (Macherey-Nagel, Bethlehem, PA) and quantified with Picogreen™(Ley et al. 2005; Stiller et al. 2009). Tagged-PCR products were then pooled together and sequenced with the FLX-Titanium platform following the manufacturer's guidelines for fusion primers with an expected 10,000 reads per sample to saturate the expected diversity according to previous diversity studies in mouse and human skin(Grice et al. 2010; Gotelli & Colwell 2001; Costello et al. 2009). Only reads between 150 and 350 bp, and average quality score ≥ 30, and with assignable or correctable tags were included in the analysis. Filtered reads were separated by barcode using a customized Perl script (M.Garcia-Garcerà, unpublished), clustered using a similarity threshold of 99% with CD-HIT(W. Li & Godzik 2006), and a representative sequence (phylotype) of each cluster was aligned against the RDP database using the *Infernal* (Cole et al. 2009; Nawrocki et al. 2009) alignment software.

16S rRNA diversity and distribution

Sequences were assigned to a specific taxon following the NCBI taxonomy. In case of taxonomic ambiguity in the *Infernal* result, the read was assigned to the lowest common taxonomic level using a modification of the lowest common ancestor (LCA) algorithm that allows to set the deepest taxonomic level accepted(Alstrup et al. 2004). In case that the result was over that level, the read was assigned as “no rank” and removed from subsequent analyses. Unique sequences were also removed from the analysis to avoid missassignment. Only phylotypes with at least 3 reads assigned (which are called singletons) were considered for this study.

Relative diversity and richness were estimated with both Chao1 and Shannon and Simpson indexes using the *vegan* R package(Oksanen et al. 2011). Rarefaction curves were also calculated using the same library to assess the diversity saturation by the number of reads.

Similarity patterns among samples were visualized by correspondence analysis and the difference in bacterial composition present in each mouse group was tested for with a 2-tailed Student's t-test.

Quantitative PCR validation of diversity estimates

To measure the relative abundance of bacteria among samples, 20 ng of DNA were amplified using 1 mM of each 16s rDNA gene primers 63F and 355R(Castillo et al. 2006), 5 µL of FastStart SYBR green master mix (Roche Applied Science) and nuclease-free water to a final volume of 10µL. All reactions were performed three times. Quantitative PCR (qPCR) reactions were performed in a LigthCycler 480 II instrument

under manufacturer's instructions for SYBR green analysis, using the basic relative quantification protocol. The relative abundance of *Staphylococcus* species in all samples was assessed by performing qPCR previously, using the TstaG422-F and TstaG765-R primers, which specifically amplify *Staphylococcus* sequences(Martineau et al. 2001; Morot-Bizot et al. 2004). These primers are specific for a 300bp region of the *tuf* gene for *Staphylococcus* species, and do not amplify this region in other families, allowing us to specifically quantify the relative amount of staphylococci on our samples. The three amplification curves obtained for each sample were averaged. Cq values were calculated for each curve and relative values were calculated and normalized for each sample taking as standard the output of the 16S rDNA gene amplification(Higuchi et al. 1993; Bustin et al. 2009). For each sample, differences between 16S rDNA and TstaG region amplification were calculated inferring the Cq50 value, understood as the PCR cycle that achieved half the maximum fluorescence intensity. The normal distribution of Cq50 was tested using Lilliefors test(Lilliefors 1967), and differences between groups were statistically tested using a two-tailed Student's t-test unpaired for the first comparison. Differences within groups were also tested using ANOVA.

Phylogenetic analysis of *Staphylococcus* reads

16S rDNA sequence reads assigned to the family *Staphylococcaceae* in our samples were further analyzed phylogenetically by comparing them to a fixed reference tree, which was constructed using all complete *Staphylococcaceae* 16S rDNA sequences in the RDP database(Wang et al. 2007; Cole et al. 2009; Cole et al. 2007).

Redundant sequences were removed from the analysis using the greedy incremental clustering algorithm implemented on CD-HIT(W. Li & Godzik 2006). The remaining sequences were aligned using SSU-align(Kolbe 2009), manually adjusted and trimmed with trimAL v. 1.3(Capella-Gutierrez et al. 2009). A phylogenetic tree was constructed from the resulting alignment with RAxML v. 7.2.8(Stamatakis et al. 2005), using the GTR substitution matrix without invariant positions and estimation of the gamma distribution as the best-suited evolutionary model for our dataset as determined by jModelTest v. 0.1.1(Rodríguez et al. 1990; Posada 2008) using a Maximum Likelihood seed tree. Redundant information was removed from the analysis with the option *prune* of ETE(Huerta-Cepas et al. 2010). Phylotypes assigned to the *Staphylococcaceae* family in our samples were realigned to the reference *Staphylococcaceae* alignment with SSU-align to assess differences on the distribution of *Staphylococcus sp.* between both samples. The same diversity measures were applied to the phylotype distribution along the reference tree using the R *vegan* package(Oksanen et al. 2011; Ihaka & Gentleman 1996).

LIST OF ABBREVIATIONS

ID = Immunodeficient

LCA = Lowest Common Ancestor

qPCR = Quantitative PCR

Cq = Quantifiaction cycle

H = Healthy

AMPc = Cationic antimicrobial peptides

COMPETING INTERESTS

The authors declare not to have any financial or non-financial competing interests

AUTHORS' CONTRIBUTIONS

MGG and FC devised the study; JMC provided the mouse samples; MGG and MC extracted and sequenced bacterial DNA; MGG, MC, and KGE performed the statistical analyses. Results were discussed and interpreted by MGG, MC, FGC, AL, and FC. MGG and FC initially wrote the manuscript, which was revised with input from all authors.

ACKNOWLEDGEMENTS

This work was financed by the MICINN (Spanish Ministry of Science and Innovation) grant SAF2010-16240. MGG was supported by a predoctoral fellowship from MICINN. We would like to thank Ferran Palero and Leonor Sanchez-Buso (CSISP) for the critical reading of the manuscript.

REFERENCES

Alstrup, S., Gavoille, C. & Kaplan, H., 2004. Nearest common ancestors: A survey and a new algorithm for a distributed environment. *Theory of Computing Systems*.

Amann, R.I., Ludwig, W. & Schleifer, K.H., 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews*, 59(1), pp.143–169.

Bäckhed, F. et al., 2005. Host-bacterial mutualism in the human intestine. *Science*, 307(5717), pp.1915–1920.

Bosma, G.C., Custer, R.P. & Bosma, M.J., 1983. A severe combined immunodeficiency mutation in the mouse. *Nature*, 301(5900), pp.527–530.

Bustin, S.A. et al., 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, 55(4), pp.611–622.

Capella-Gutierrez, S., Silla-Martinez, J.M. & Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), pp.1972–1973.

Castillo, M. et al., 2006. Quantification of total bacteria, enterobacteria and lactobacilli populations in pig digesta by real-time PCR. *Veterinary microbiology*, 114(1-2), pp.165–170.

Cogen, A., Nizet, V. & Gallo, R., 2008. Skin microbiota: a source of disease or defence? *British Journal of Dermatology*, 158(3), pp.442–455.

Cole, J.R. et al., 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Research*, 35(Database), pp.D169–D172.

Cole, J.R. et al., 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research*, 37(Database), pp.D141–D145.

Costello, E.K. et al., 2009. Bacterial Community Variation in Human Body Habitats Across Space and Time. *Science*, 326(5960), pp.1694–1697.

Dekio, I. et al., 2007. Characterization of skin microbiota in patients with atopic dermatitis and in normal subjects using 16S rRNA gene-based comprehensive analysis. *Journal of Medical Microbiology*, 56(12), pp.1675–1683.

Fierer, N. et al., 2010. From the Cover: Forensic identification using skin bacterial communities. *Proceedings of the National Academy of Sciences*, 107(14), pp.6477–6481.

Fredricks, D.N., 2001. Microbial ecology of human skin in health and disease. *The journal of investigative dermatology Symposium proceedings / the Society for Investigative Dermatology, Inc [and] European Society for Dermatological Research*, 6(3), pp.167–169.

Galdbart, J., Allignet, J. & Tung, H., 2000. Screening for Staphylococcus epidermidis Markers Discriminating between Skin-Flora Strains and Those Responsible for Infections of Joint Prostheses. *Journal of Infectious Diseases*, 182, pp.351–355.

Gao, Z. et al., 2008. Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. *PloS one*,

3(7), p.e2719.

Gotelli, N.J. & Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters*, 4(4), pp.379–391.

Grice, E.A. & Segre, J.A., 2011. The skin microbiome. *Nature reviews Microbiology*, 9(4), pp.244–253.

Grice, E.A. et al., 2008. A diversity profile of the human skin microbiota. *Genome Research*, 18(7), pp.1043–1050.

Grice, E.A. et al., 2010. Longitudinal shift in diabetic wound microbiota correlates with prolonged skin defense response. *Proceedings of the National Academy of Sciences*, 107(33), pp.14799–14804.

Grice, E.A. et al., 2009. Topographical and Temporal Diversity of the Human Skin Microbiome. *Science*, 324(5931), pp.1190–1192.

Higuchi, R. et al., 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Bio/technology (Nature Publishing Company)*, 11(9), pp.1026–1030.

Huerta-Cepas, J., Dopazo, J. & Gabaldón, T., 2010. ETE: a python Environment for Tree Exploration. *BMC bioinformatics*, 11, p.24.

Hugenholtz, P. & Pace, N.R., 1996. Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. *Trends in biotechnology*, 14(6), pp.190–197.

Ikaha, R. & Gentleman, R., 1996. A language for data analysis and graphics. *J Comp Graph Stat*, 5, pp.299–314.

Jean-Baptiste, N. et al., 2011. Coagulase-negative staphylococcal infections in the neonatal intensive care unit. *Infection control and hospital epidemiology : the official journal of the Society of Hospital Epidemiologists of America*, 32(7), pp.679–686.

Keane, T.M. et al., 2011. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*, 477(7364), pp.289–294.

Kolbe, D., 2009. Local RNA structure alignment with incomplete sequence. *Bioinformatics*.

Köckritz-Blickwede, von, M. et al., 2008. Immunological mechanisms underlying the genetic predisposition to severe Staphylococcus aureus infection in the mouse model. *The American journal of pathology*, 173(6), pp.1657–1668.

Lai, Y. et al., 2010. Activation of TLR2 by a small molecule produced by Staphylococcus epidermidis increases antimicrobial defense against bacterial skin infections. *Journal of Investigative Dermatology*, 130(9), pp.2211–2221.

Lande, R. et al., 2007. Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide. *Nature*, 449(7162), pp.564–569.

Ley, R.E. et al., 2008. Evolution of mammals and their gut microbes. *Science*, 320(5883), pp.1647–1651.

Ley, R.E. et al., 2005. Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America*, 102(31), pp.11070–11075.

Ley, R.E., Peterson, D.A. & Gordon, J.I., 2006. Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine. *Cell*, 124(4), pp.837–848.

Li, M. et al., 2007. The antimicrobial peptide-sensing system aps of Staphylococcus aureus. *Molecular microbiology*, 66(5), pp.1136–1147.

Li, W. & Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), pp.1658–1659.

Lilliefors, H., 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*.

Line, J.E. et al., 2008. Isolation and purification of enterocin E-760 with broad antimicrobial activity against gram-positive and gram-negative bacteria. *Antimicrobial agents and chemotherapy*, 52(3), pp.1094–1100.

Maidak, B., Cole, J. & Lilburn, T., 2001. The RDP-II (ribosomal database project). *Nucleic Acids …*.

Martineau, F. et al., 2001. Development of a PCR assay for identification of staphylococci at genus and species levels. *Journal of Clinical Microbiology*, 39(7), pp.2541–2547.

Mehta, P. et al., 2009. Information processing and signal integration in bacterial quorum sensing. *Molecular systems biology*, 5, p.325.

Morot-Bizot, S.C., Talon, R. & Leroy, S., 2004. Development of a multiplex PCR for the identification of Staphylococcus genus and four staphylococcal species isolated from food. *Journal of Applied Microbiology*, 97(5), pp.1087–1094.

Mouse Genome Sequencing Consortium et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), pp.520–562.

Moya, A. et al., 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nature Reviews Genetics*, 9(3), pp.218–229.

Nawrocki, E.P., Kolbe, D.L. & Eddy, S.R., 2009. Infernal 1.0: inference of RNA alignments. *Bioinformatics*, 25(10), pp.1335–1337.

Oksanen, J. et al., 2011. Vegan: community ecology package. *R package version 1.17-10*. Available at: http://CRAN.R-project.org/package=vegan.

Otto, M., 2001. Staphylococcus aureus and Staphylococcus epidermidis peptide pheromones produced by the accessory gene regulator agr system. *Peptides*.

Otto, M., 2009. Staphylococcus epidermidis — the “accidental” pathogen. *Nature reviews Microbiology*, 7(8), pp.555–567.

Otto, M. et al., 2001. Pheromone cross-inhibition between Staphylococcus aureus and Staphylococcus epidermidis. *Infection and immunity*, 69(3), pp.1957–1960.

Peschel, A. et al., 1999. Inactivation of the dlt operon in Staphylococcus aureus confers sensitivity to defensins, protegrins, and other antimicrobial peptides. *The Journal of biological chemistry*, 274(13), pp.8405–8410.

Peschel, A. et al., 2001. Staphylococcus aureus resistance to human defensins and evasion of neutrophil killing via the novel virulence factor MprF is based on modification of membrane lipids with l-lysine. *The Journal of experimental medicine*, 193(9), pp.1067–1076.

Posada, D., 2008. jModelTest: Phylogenetic Model Averaging. *Molecular biology and evolution*, 25(7), pp.1253–1256.

Rodrigue, L. & Lavoie, M.C., 1996. Comparison of the proportions of oral bacterial species in BALB/c mice from different suppliers. *Laboratory animals*, 30(2), pp.108–113.

Rodríguez, F. et al., 1990. The general stochastic model of nucleotide substitution. *Journal of theoretical biology*, 142(4), pp.485–501.

Rosenthal, M. et al., 2011. Skin microbiota: Microbial community structure and its potential association with health and disease. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*, 11(5), pp.839–848.

Roth, R.R. & James, W.D., 1988. Microbial ecology of the skin. *Annual review of microbiology*, 42, pp.441–464.

Sandiford, S. & Upton, M., 2012. Identification, Characterization, and Recombinant Expression of Epidermicin NI01, a Novel Unmodified Bacteriocin Produced by Staphylococcus epidermidis That Displays Potent Activity against Staphylococci. *Antimicrobial agents and chemotherapy*, 56(3), pp.1539–1547.

Scharschmidt, T.C. et al., 2009. Matriptase-Deficient Mice Exhibit Ichthyotic Skin with a Selective Shift in Skin Microbiota. *Journal of Investigative Dermatology*, 129(10), pp.2435–2442.

Stamatakis, A., Ludwig, T. & Meier, H., 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics*, 21(4), pp.456–463.

Stiller, M. et al., 2009. Direct multiplex sequencing (DMPS)--a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Research*, 19(10), pp.1843–1848.

Sun, C. et al., 2006. Peptidoglycan recognition proteins Pglyrp3 and Pglyrp4 are encoded from the epidermal differentiation complex and are candidate genes for the Psors4 locus on chromosome 1q21. *Human genetics*, 119(1-2), pp.113–125.

Turnbaugh, P.J. et al., 2007. The human microbiome project. *Nature*, 449(7164), pp.804–810.

Uçkay, I. et al., 2009. Foreign body infections due to Staphylococcus epidermidis. *Annals of medicine*, 41(2), pp.109–119.

Wang, Q. et al., 2007. Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, 73(16), pp.5261–5267.

FIGURE LEGENDS

Figure 1. Comparison of the relative abundances (x10,000) of bacteria in skin samples.

16S sequences were assigned to the genus level. Only genera with more than three sequences assigned were used for the analysis; then the whole information was clustered to family level. Uncommon/rare families were clustered together into class level. For more detailed information, see the supplementary material.

Figure 2. Correspondence analysis of the bacterial diversity in skin of immunodepressed and healthy mice.

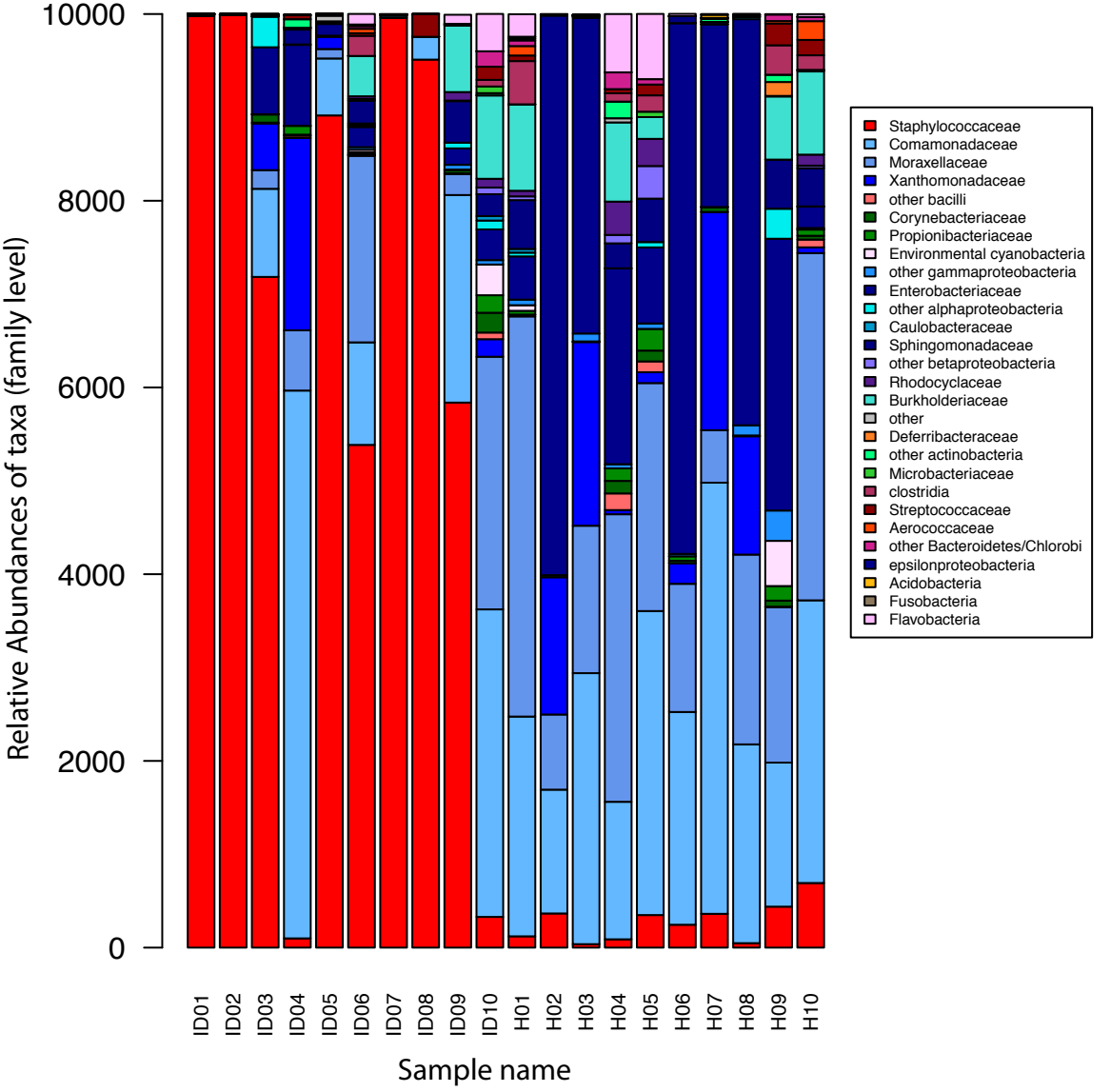
Samples clustered together according to health state using Correspondence Analysis (PCoA) of the unweighted UniFrac distance matrix. Samples are shown by name, and the different taxa are colored by class. Components 1 and 2 together explain > 95% of the variance. DCA1 (72% of explained variability) is associated to the relative abundance of *Staphylococcus* assigned reads, resulting in a marked divergence between the ID samples and the ones that behave as healthy (including ID4 and ID10, which have similar diversity than the healthy ones).

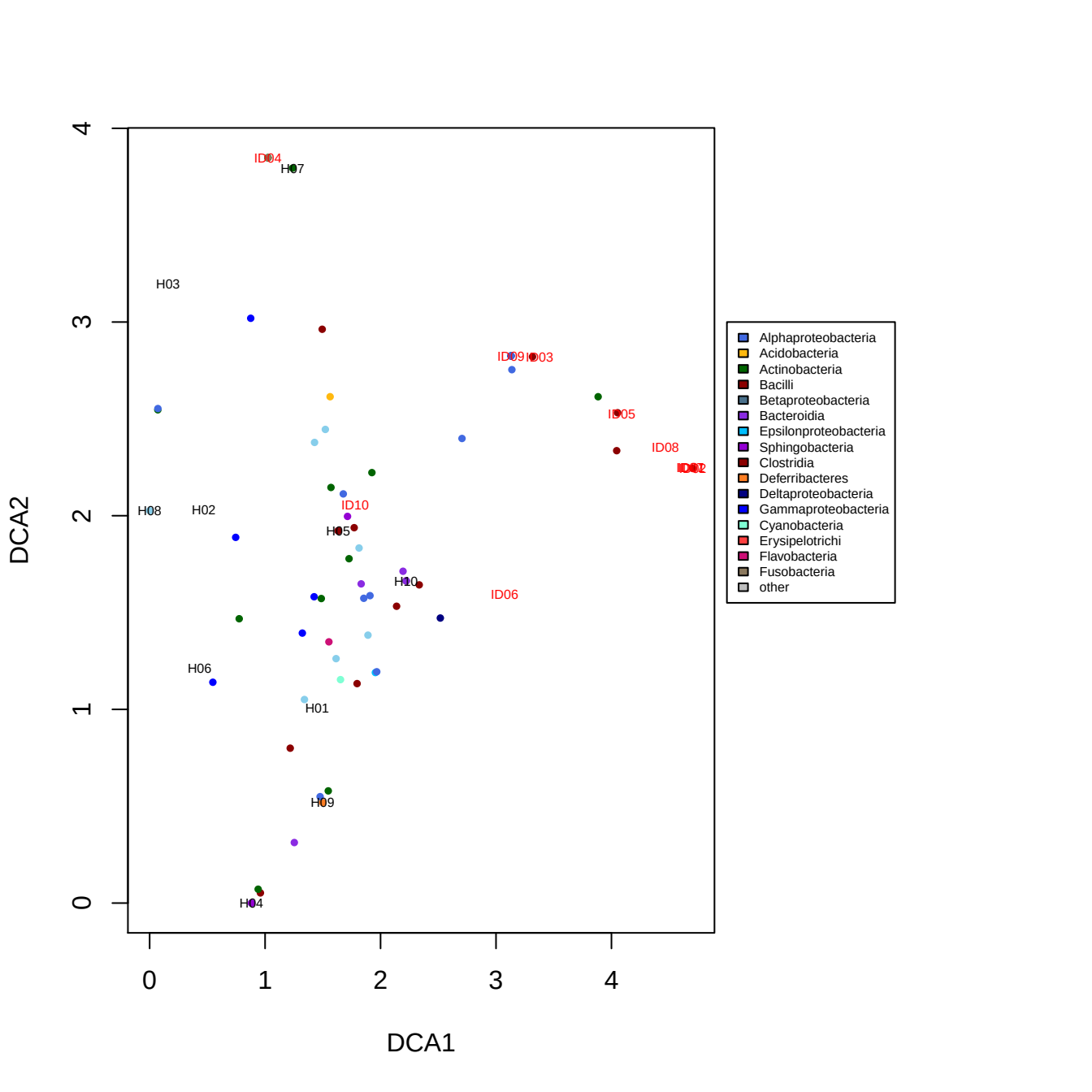
Figure 3. Phylogenetic analysis of the *Staphylococcacae* assigned reads.

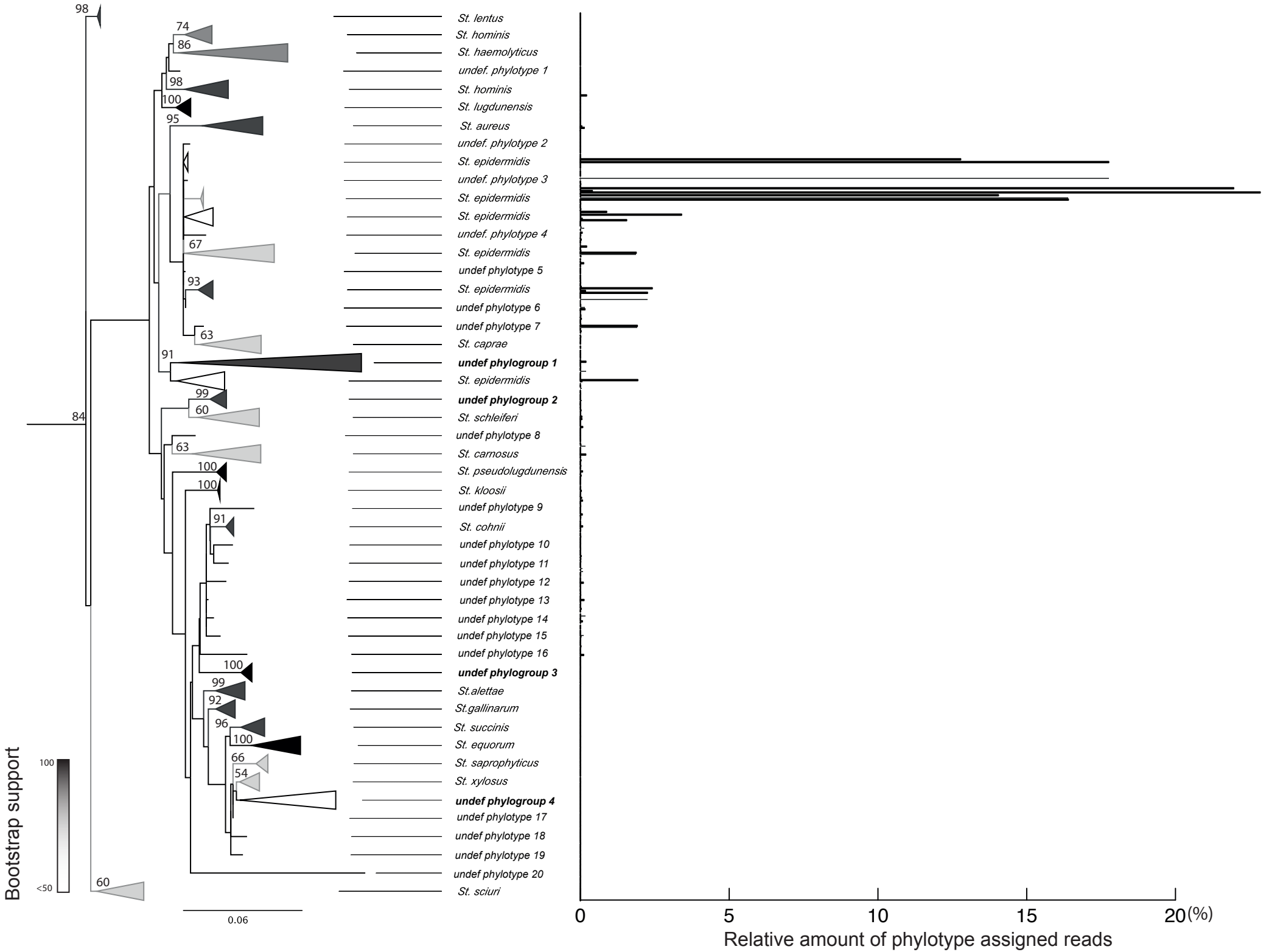
16S sequences belonging to the *Staphylococcaceae* family were obtained from the RDP database to construct a reference.

sample	Shannon	Chao1	SE,Chao1	ACE	SE,ACE
1ID	0.020	11	11.7	14.24	1.513
2ID	0.009	4	1.3	5.11	1.136
3ID	1.078	13	0.7	14.17	1.769
4ID	1.302	21	NaN	23.09	1.975
5ID	0.506	29	10.3	39.96	4.164
6ID	1.521	22	NaN	22.00	1.809
7ID	0.032	7.5	3.7	9.22	1.355
8ID	0.229	3	NaN	NaN	NaN
9ID	1.697	21	NaN	21.00	2.268
10ID	2.167	26	NaN	NaN	NaN
1H	1.832	21	NaN	NaN	NaN
2H	1.206	8	NaN	8.00	0.935
3H	1.434	21.86	1.85	24.28	2.409
4H	2.167	23	NaN	NaN	NaN
5H	2.187	23	NaN	NaN	NaN
6H	1.213	10	NaN	NaN	NaN
7H	1.392	17	NaN	18.56	2.195
8H	1.402	22	0	22.00	2.185
9H	2.248	23	NaN	23.00	2.106
10H	1.827	26	0	26.00	2.148

Table 1. Diversity of microbial families for the 20 samples analyzed. Diversity indeces were calculated for each sample given a family-based abundance table. SE, Standard Error; ACE, Abundance-base Coverage Estimator. Undefined (NaN) values appear when all rare taxa are only assigned as singletons.


10000
8000
6000
4000
2000
0
Relative Abundances of taxa (family level)
ID01
ID02
ID03
ID04
ID05
ID06
ID07
ID08
ID09
ID10
H01
H02
H03
H04
H05
H06
H07
H08
H09
H10
Sample name
Staphylococcaceae
Comamonadaceae
Moraxellaceae
Xanthomonadaceae
other bacilli
Corynebacteriaceae
Propionibacteriaceae
Environmental cyanobacteria
other gammaproteobacteria
Enterobacteriaceae
other alphaproteobacteria
Caulobacteraceae
Sphingomonadaceae
other betaproteobacteria
Rhodocyclaceae
Burkholderiaceae
other
Deferribacteraceae
other actinobacteria
Microbacteriaceae
clostridia
Streptococcaceae
Aerococcaceae
other Bacteroidetes/Chlorobi
epsilonproteobacteria
Acidobacteria
Fusobacteria
Flavobacteria


DCA1
DCA2
0
1
2
3
4
ID04
H07
H03
ID09
ID03
ID05
ID08
ID02
H08
H02
ID10
H05
H10
ID06
H06
H01
H09
H04
Alphaproteobacteria
Acidobacteria
Actinobacteria
Bacilli
Betaproteobacteria
Bacteroidia
Epsilonproteobacteria
Sphingobacteria
Clostridia
Deferribacteres
Deltaproteobacteria
Gammaproteobacteria
Cyanobacteria
Erysipelotrichi
Flavobacteria
Fusobacteria
other


St. lentus
St. hominis
St. haemolyticus
undef. phylotype 1
St. hominis
St. lugdunensis
St. aureus
undef. phylotype 2
St. epidermidis
undef. phylotype 3
St. epidermidis
St. epidermidis
undef. phylotype 4
St. epidermidis
undef phylotype 5
St. epidermidis
undef phylotype 6
undef phylotype 7
St. caprae
undef phylogroup 1
St. epidermidis
undef phylogroup 2
St. schleiferi
undef phylotype 8
St. carnosus
St. pseudolugdunensis
St. kloosii
undef phylotype 9
St. cohnii
undef phylotype 10
undef phylotype 11
undef phylotype 12
undef phylotype 13
undef phylotype 14
undef phylotype 15
undef phylotype 16
undef phylogroup 3
St.alettae
St.gallinarum
St. succinis
St. equorum
St. saprophyticus
St. xylosus
undef phylogroup 4
undef phylotype 17
undef phylotype 18
undef phylotype 19
undef phylotype 20
St. sciuri
98
74
86
98
100
95
67
93
63
91
99
60
84
63
100
100
91
100
99
92
96
100
66
54
60
Bootstrap support
100
<50
0.06
0
5
10
15
20(%)
Relative amount of phylotype assigned reads