

ARTICLE

<https://doi.org/10.1038/s41467-019-12253-y>

OPEN

# Microbe-host interplay in atopic dermatitis and psoriasis

Nanna Fyhrquist  et al.<sup>#</sup>

Despite recent advances in understanding microbial diversity in skin homeostasis, the relevance of microbial dysbiosis in inflammatory disease is poorly understood. Here we perform a comparative analysis of skin microbial communities coupled to global patterns of cutaneous gene expression in patients with atopic dermatitis or psoriasis. The skin microbiota is analysed by 16S amplicon or whole genome sequencing and the skin transcriptome by microarrays, followed by integration of the data layers. We find that atopic dermatitis and psoriasis can be classified by distinct microbes, which differ from healthy volunteers microbiome composition. Atopic dermatitis is dominated by a single microbe (*Staphylococcus aureus*), and associated with a disease relevant host transcriptomic signature enriched for skin barrier function, tryptophan metabolism and immune activation. In contrast, psoriasis is characterized by co-occurring communities of microbes with weak associations with disease related gene expression. Our work provides a basis for biomarker discovery and targeted therapies in skin dysbiosis.

---

Correspondence and requests for materials should be addressed to H.A. (email: [harri.alenius@ki.se](mailto:harri.alenius@ki.se)). <sup>#</sup>A full list of authors and their affiliations appears at the end of the paper.

Interactions between commensal or pathogenic microbes and the hosts they colonize are central to the maintenance of homeostasis and the initiation of disease<sup>1</sup>. This rapidly advancing field is now starting to bear the fruits of interdisciplinary efforts but our understanding of microbe–host interactions is still limited<sup>2</sup>.

Skin represents a primary tissue interface, where accessibility to both microbiome species and underlying tissue provides a unique opportunity for studies into host–microbiome interactions<sup>3</sup> in the initiation and maintenance of atopic/allergic or autoimmune-type inflammation<sup>4,5</sup>. Recent advances in analyzing microbial gene sequences in healthy skin has provided a comprehensive understanding of the classes of microbes and their diversity occupying distinct topographical niches<sup>6</sup>. Pioneering studies have expanded into shotgun sequencing approaches to further probe the taxonomic diversity and biogeography of microbes in small cohorts of healthy individuals<sup>7</sup>. Commensal skin microbes control adaptive skin immune homeostasis through interaction with specific subsets of antigen-presenting dendritic cells (DCs) and effector T-cell populations<sup>8,9</sup>, supporting their own survival, and protecting against the overgrowth of pathogens. Insights into the association of distinct microbial classes with inflammatory skin disease and their impact on the host genome are just emerging. In atopic dermatitis (AD), as a model of atopic/allergic inflammatory disease, commensal skin microbes are associated with disease flares<sup>10</sup>. AD-related dysbiosis is frequently characterized by the colonization by *Staphylococcus aureus*, and simultaneous loss of other, potentially beneficial species. *S. aureus* colonizes skin effectively, and expresses several virulence factors<sup>11</sup> with proven roles in the pathogenesis of AD by studying effects in cell<sup>12</sup> and animal models<sup>13,14</sup>. *S. aureus* is associated with severity of the disease, and may be the result of a combination of detrimental effects from *S. aureus* on the one hand, and the loss of beneficial effects from other members of the skin microbiota on the other hand. Therefore, the use of antimicrobial therapeutics that target *S. aureus* may not be the optimal choice as they may also wipe out beneficial species or strains, and break mutualistic interactions between the skin and its microbiota. Psoriasis (PSO) appears to induce physiological changes at the lesion site, selecting for a specific microbiota. The psoriasis (PSO)-associated skin microbiome displays disease-specific features that might have a diagnostic value<sup>15–17</sup>, but whether the differential microbiota has pathophysiologic significance remains undetermined.

Here, we present a large-scale, comprehensive analysis of the microbiome and microbiome-associated host transcriptome in skin of healthy volunteers (HV), AD, and PSO patients, revealing two distinct patterns of host–microbe interactions in chronic skin inflammation. We report a significant increase in the abundance of *S. aureus* and loss of anaerobic species in AD, while PSO is characterized by the co-occurrence of multiple organisms, including *Corynebacterium* and *Finnegoldia* species. In AD lesions, which are abundantly colonized by *S. aureus*, toxin production and metabolic reprogramming are highly over-represented microbial functions. The host, in turn, responds to the changes in the microbiota via altered expression of genes related to barrier function, metabolic reprogramming, antimicrobial defense mechanisms and T helper type 2 (T<sub>H</sub>2) signaling. In PSO, our results suggest that members of *Corynebacterium* may play a regulatory role, which is reduced in disease.

## Results

**The skin microbiotas in AD and PSO are highly distinct.** Skin swabs were collected using a standardized protocol (Supplementary Fig. 1a–c). A total of 3.36 million 16S rRNA gene reads were analyzed using QIIME v1.8.0<sup>18</sup> resulting in the identification

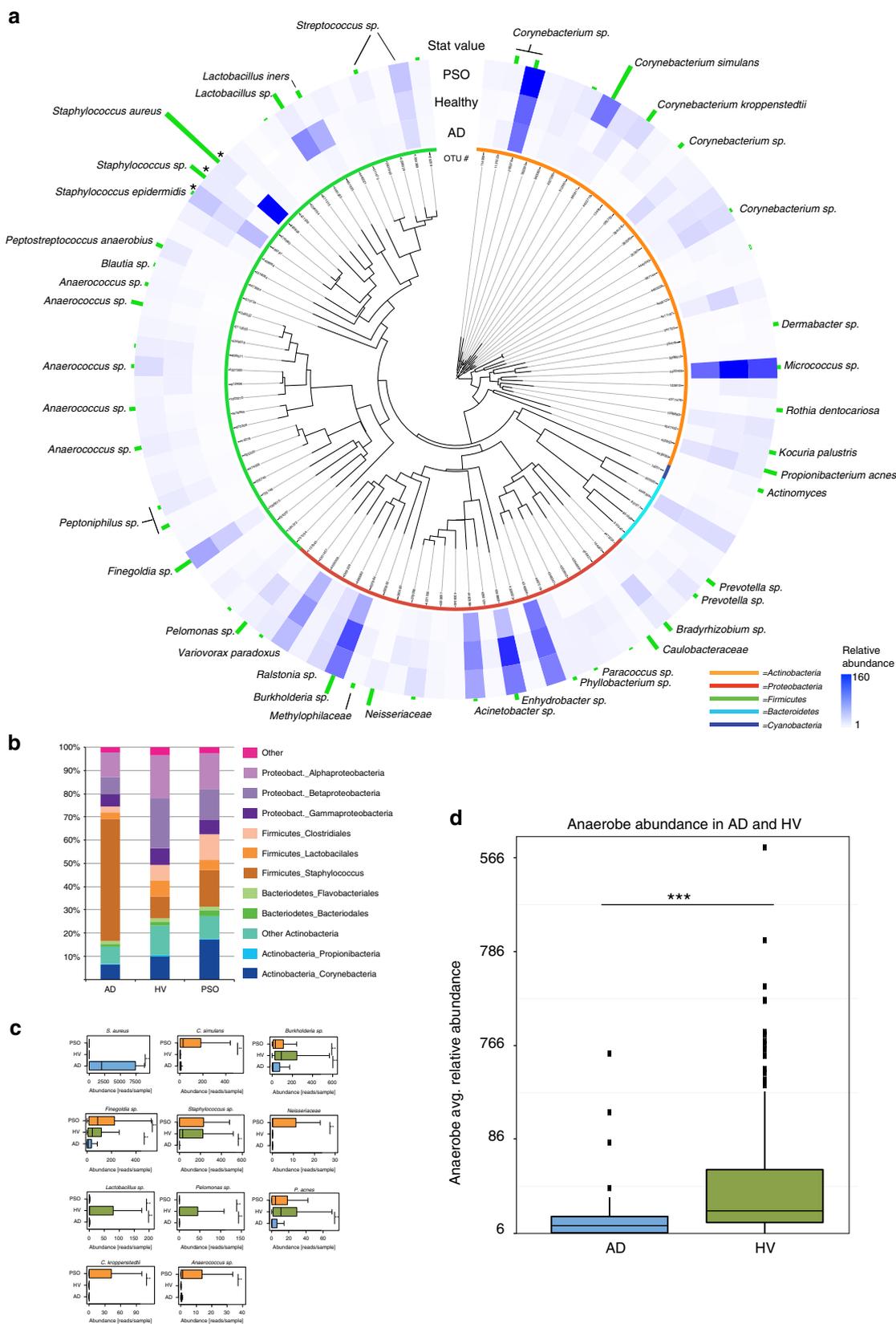
of 17,725 operational taxonomic units (OTUs) at 99.3% identity level. After removal of rare OTUs, 3342 remained. Blasting sequences against the Greengenes 16S rRNA gene database revealed healthy skin microbiomes consistent with previous reports<sup>6</sup>. An overall analysis indicated clear differences between AD and HV (Fig. 1a, b, Supplementary Fig. 2), confirmed by nonmetric multidimensional scaling (NMDS) analysis of the 95 most abundant OTUs (Supplementary Fig. 3a, Supplementary Table 1). AD also showed a reduction in diversity (Supplementary Fig. 3b). A total of 51 highly abundant OTUs showed significant differences between HV, AD and PSO (Fig. 1a).

We additionally carried out a careful analysis of the influence of confounding factors, including age, anatomical location, gender and clinical center (Supplementary Table 2, Supplementary Fig. 3c). After correction for the confounding effects, 8 out of 17 confounded OTUs from AD and 5 of 13 OTUs from PSO retained significance. For example, *C. kroppenstedtii* was associated with age, *S. aureus* with anatomical location and *Lactobacillus* sp. with gender, but they all remained significantly associated with disease after correction ( $p < 0.01$ ). Associations were tested using the Kruskal–Wallis test for body site and institution, the Mann–Whitney *U* test for gender, and Spearman correlation for age.

The remaining 11 most significant OTUs are shown in Fig. 1c. The most significant result is an increase in the abundance of *S. aureus* in AD, associated with a significantly lower abundance of OTUs representing strictly anaerobic bacteria in AD (Fig. 1d). The loss of anaerobes in AD is not driven by *S. aureus*, indicated by repeated analysis of samples devoid of *S. aureus* (Supplementary Fig. 2d). Significant changes in PSO compared to HV include increases in the abundance of *C. simulans* and *C. kroppenstedtii* as well as *Finnegoldia* and *Neisseriaceae* species. *Lactobacilli*, *Burkholderia* spp. and *P. acnes* were lower in abundance in both AD and PSO compared to healthy skin.

**Microbiota-based classification of AD and PSO.** We next asked whether the skin microbiome discriminates inflammatory skin pathologies. A supervised learning pipeline was employed to construct classifiers to explore the key sets of microbial taxa. We identified 26 microbes discriminating AD vs. HV cohorts with an area under the curve (AUC) of 0.94 (class errors HV = 0.03, AD = 0.27, Fig. 2a). The most discriminative taxa were of the genus *Staphylococcus*, including *S. aureus* ( $Z = 14.0$ ), *S. epidermidis* ( $Z = 5.8$ ), *Staphylococcus* spp. ( $Z = -6.9$ ) and *Burkholderia* spp. ( $Z = -7.5$ ). An NMDS analysis highlights the role of *S. aureus* in differentiating AD (red dots) from HV (blue dots; Supplementary Fig. 4a). We identified 24 microbes that differentiate PSO vs. HV with an AUC of 0.85 (class errors HV = 0.09, PSO = 0.33, Fig. 2b). The top discriminating microbes were *C. simulans* ( $Z = 15.5$ ), *Neisseriaceae* g. spp. ( $Z = 6.9$ ), *C. kroppenstedtii* ( $Z = 5.5$ ), *Lactobacillus* spp. ( $Z = -8.4$ ) and *Lactobacillus iners* ( $Z = -3.5$ ). An NMDS analysis demonstrated a clear separation boundary between HV (green dots) and PSO (green dots; Supplementary Fig. 4b). We identified 15 microbes that differentiate between AD and PSO with AUC of 0.86 (class errors AD = 0.29, PSO = 0.13, Supplementary Fig. 4c). The top microbes were *S. aureus* ( $Z = 14.0$ ), *S. epidermidis* ( $Z = 3.9$ ), *Finnegoldia* spp. ( $Z = -6.5$ ), *C. simulans* ( $Z = -5.0$ ), and *C. kroppenstedtii* ( $Z = -4.4$ ). An NMDS analysis highlights the importance of *S. aureus* in AD (Supplementary Fig. 4d).

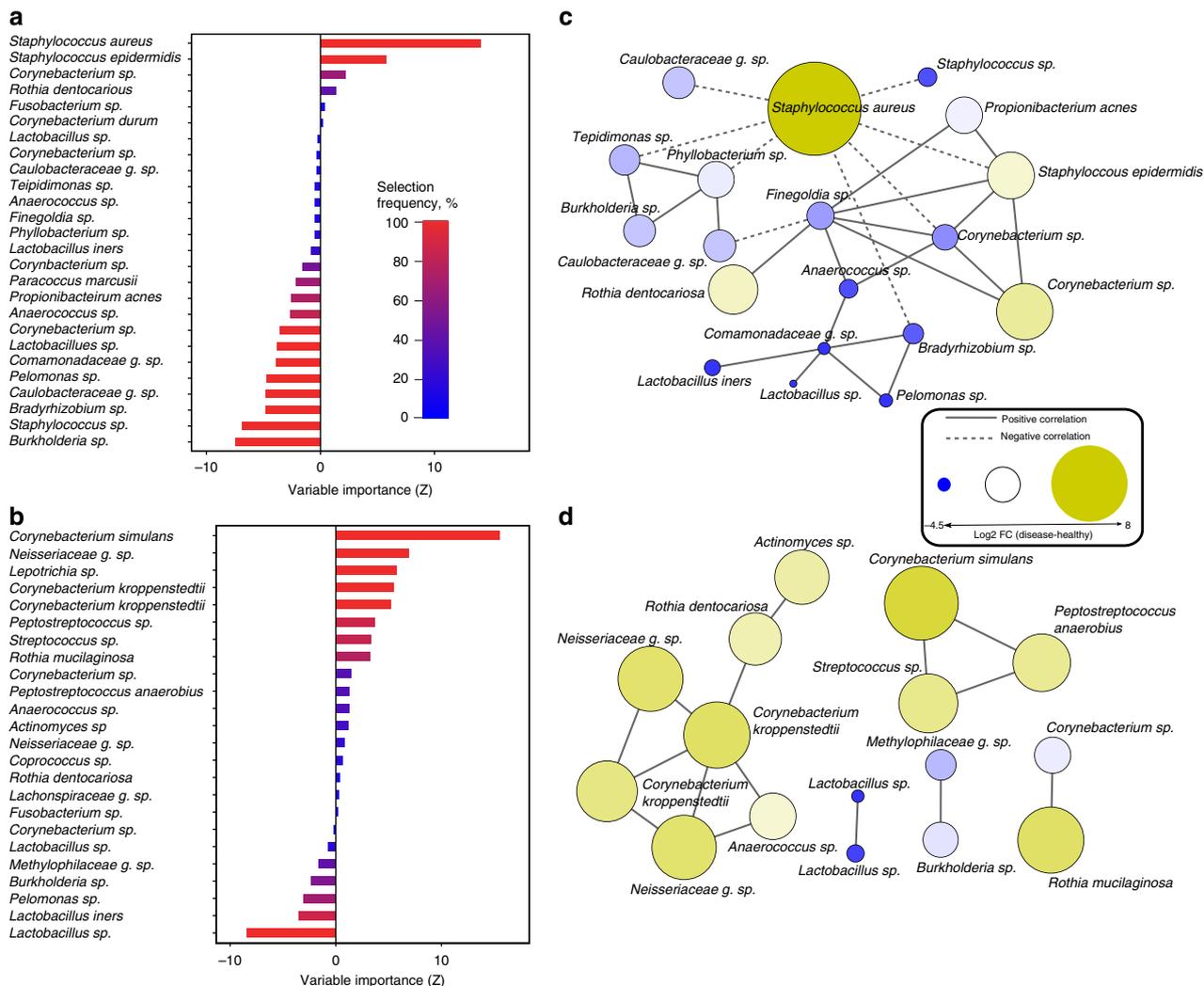
**Communities of microbes associate with AD or PSO.** To understand the interactions between communities of microbes under different disease states, we used network principles to express co-occurrence relationships. We found distinct



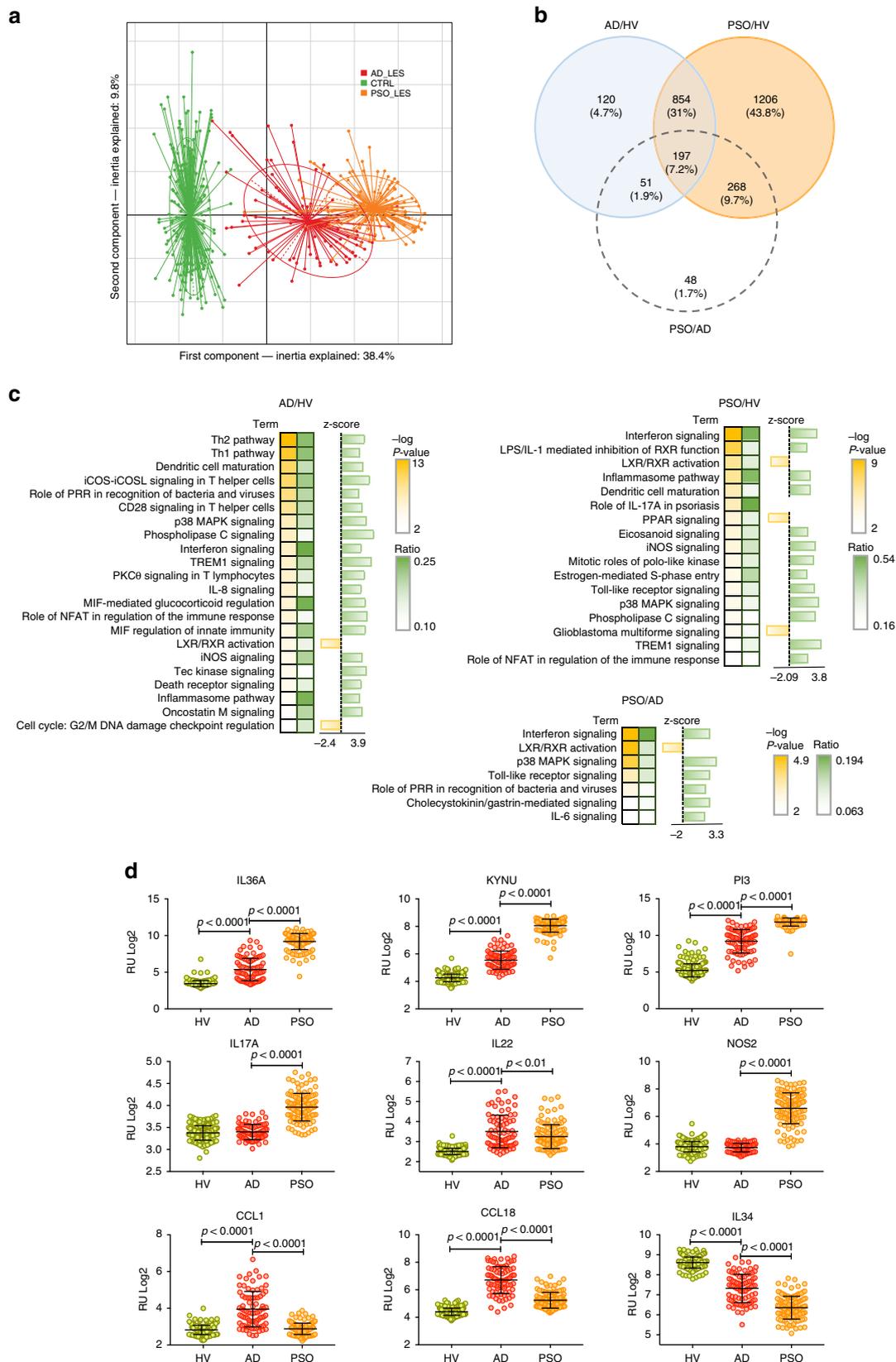
differences between the community structures of microbes associated with AD and PSO. For microbes associated with AD, SparCC<sup>19</sup> correlation between taxa resulted in 19 species with a correlation above a threshold of 0.2 ( $p < 0.05$ , Fig. 2c). *S. aureus* negatively correlated with species including *Corynebacterium*

sp., *S. epidermidis*, *Tepidimonas* spp. and *Phyllobacterium* spp. Of the 24 microbes identified as important for PSO classification (Fig. 2b), 16 species showed significant SparCC correlation (SparCC > 0.2,  $p < 0.05$ , Fig. 2d). The most discriminant taxa *C. simulans* and *C. kroppenstedtii* displayed positive correlations

**Fig. 1** Characterization of the skin microbiome in AD and PSO. The results show a typical range of skin microbiomes in HV ( $n = 115$ ) and significant changes in AD ( $n = 82$ ) and PSO ( $n = 119$ ). **a** An evolutionary tree based on 16S rRNA gene sequences, abundance and statistical significance of the 95 most abundant OTUs. The blue color intensity in the heat map shows the relative abundance of each OTU. The three *Staphylococcus* OTUs (indicated by asterisks) were calculated using a wider scale, due to their relatively high abundance in certain samples. The length of the green vertical bars indicate nonparametric statistical score (Kruskal–Wallis test, FDR,  $p < 0.05$ ). The color bars by each OTU number indicates bacterial phylum (green: *Firmicutes*; red: *Proteobacteria*; cyan: *Bacteroidetes*; blue: *Cyanobacteria*; orange: *Actinobacteria*). **b** The most abundant bacterial groups depicted for HV, AD and PSO. **c** Statistical analysis (Mann–Whitney  $U$  test (FDR,  $p < 0.05$ )) of the 11 OTUs showing the most significant changes in AD and/or PSO vs. HV after correction for confounding factors. The values on the  $x$  axis are in number of reads, out of a total of 8495 reads/sample. The asterisks indicate statistically significant differences and correspond to  $p < 0.01$  (\*\*) and  $p < 0.001$  (\*\*\*). The center line in the boxplots corresponds to the median, the bounding box is the interquartile range (IQR) and the whiskers are defined as 1.5 times IQR. **d** Statistical analysis (Mann–Whitney  $U$  test,  $p < 0.05$ ) of the relative abundance of OTUs representing strictly anaerobic bacteria in AD lesions and HV. The  $x$ -axis units and the elements of the boxplots are as in (c). The source data files used to generate the present figure are available from the NCBI Sequence Read Archive under accession PRJNA554499



**Fig. 2** Microbiota-based classification of AD and PSO. **a** Variable importance for the best set of discriminatory AD ( $n = 82$ ) taxa identified through Random Forest feature selection and classification (RF) analysis. Bars are colored by selection frequency (Red = OTU selected in all folds, blue = OTU selected in one fold). **b** Variable importance for the best set of discriminatory PSO ( $n = 119$ ) taxa identified by RF analysis. **c** Co-occurrence network of AD-associated OTUs selected by RF selection. Pairwise correlations were calculated over AD samples using SparCC and OTUs with a mean  $Z > 0.2$  are shown. A connection represents a correlation  $> 0.2$  and significant  $p < 0.05$  correlation. Solid and dashed lines respectively represent positive and negative correlations. The size and color of each node is proportional to the log<sub>2</sub> fold change between healthy and disease. **d** Correlation network of PSO-associated OTUs selected by RF. The source data files used to generate the present figure are available from the NCBI Sequence Read Archive under accession PRJNA554499



with *Streptococcus* spp., *P. anaerobius* and *Anaerococcus* spp., *Neisseriaceae* g. spp., and *Rothia dentocariosa* respectively. Comparison across microbial interactions in PSO indicate that rather than a single species dominating the microbial landscape (as in AD), we observe multiple species associated with this disease type.

**Overlaps and distinctions of AD and PSO transcriptomes.** The transcriptomes of full-thickness skin biopsies were defined based on a stringent  $10^{-5}$  FDR level and a fold change (FCH) of 1.5, or greater. A principal component analysis (PCA) revealed a clear separation between HV, AD, and PSO (Fig. 3a). Differential gene expression analysis identified 1232 genes differing between AD

**Fig. 3** The AD and PSO skin transcriptomes. **a** Projection of AD (red,  $n = 82$ ), PSO (orange,  $n = 119$ ) and HV (green,  $n = 115$ ) transcriptome profiles in the subspace spanned by the two first components of the principal component analysis (PCA) performed on the 1000 most variant genes. **b** Venn diagram of the differentially expressed genes in the AD vs. HV, PSO vs. HV and PSO vs. AD contrasts (identified using the *Limma* linear model and empirical Bayes method, cut-off:  $\log_2 FC > 0.58$  and  $FDR p < 1 \times 10^{-5}$ , corrected using the Benjamini–Hochberg method). **c** IPA canonic pathway analysis of significantly enriched functions in AD and PSO gene signatures and the PSO vs. AD contrast. **d** Statistical analysis (unpaired *t* test) of selected top up- and downregulated genes. The center line in the dot plots corresponds to the mean, and the error bars to the standard deviation. The source data files used to generate the present figure are available from EBI ArrayExpress under accession E-MTAB-8149

and HV, and 2525 between PSO and HV, and 1051 genes shared between AD and PSO (Fig. 3b).

For functional insight into the disease-specific genes, we performed Ingenuity Pathway Analysis (IPA), revealing significant overrepresentation of  $T_H2$  and  $T_H1$  signaling, dendritic cell maturation and iCOS-iCOSL signaling in helper T cells in AD, and Interferon signaling, LPS-IL-1-mediated inhibition of RXR function, the Inflammasome pathway and Th17 signaling in PSO. The levels of Interferon signaling and p38 MAPK signaling distinguished PSO from AD (Fig. 3c).

*TNF* and *IFNG* were identified as upstream regulators in both AD and PSO, whereas *IL4* and *IL13* emerged as unique features in AD, underpinning the central role of  $T_H2$ -associated signaling in AD (Supplementary Fig. 5b). Gene Ontology (GO) enrichment analysis highlighted chemotaxis, inflammatory response, and extracellular matrix organization in AD, and leukocyte activation in both AD and PSO (Supplementary Fig. 6). Among top upregulated genes in AD and PSO we observed inflammatory mediators (*S100* proteins, defensins, matrix metalloproteinases, IL-1 family cytokines), T helper-related genes (*CCL1*, *CCL18*, *IL17A*, *IL22*, *PI3/Elafin*), barrier genes (*KRT16*, *SERPINB4*, *KLK9*, *FLG2*, *LCE5A*, *CLDN8*), and genes involved in tryptophan (trp) metabolism (*KYNU*). Top downregulated genes included *IL34*, anti-inflammatory *IL37*, and *NOS2* (Fig. 3d, Supplementary Fig. 7).

**The *S. aureus*-induced host gene signature in AD.** Taking advantage of our large microbiome and transcriptome datasets, we addressed the interplay between host and microbes. We stratified AD samples into “high” and “low” groups, based on microbial abundance, including top ( $n = 27$ ) and bottom ( $n = 25$ ) tertiles in the analysis. The low abundance samples were devoid of *S. aureus*, whereas high abundance samples exhibited in all cases a high abundance of *S. aureus* (87–99%).

First, we explored the microbial gene repertoire of *S. aureus* high and low groups. Using PiCRUST that predicts functional content based on the 16S rRNA gene marker, we observed the enrichment of bacterial toxins, the two-component system and glycolysis in the *S. aureus* high group compared with the low group (Supplementary Fig. 8a–b). The majority of the functions were contributed by *S. aureus* OTUs (Supplementary Fig. 8c). To validate OTU-based predictions, we performed whole genome sequencing (WGS) on a limited sample set, showing close agreement between the two independent sequencing methods (Supplementary Fig. 9). WGS data confirmed the enrichment of major bacterial toxins, hld and plc (alpha- and delta-toxin, respectively), and genes associated with galactose metabolism, the phosphotransferase system, the two-component system, and glycolysis and gluconeogenesis, in the *S. aureus* high samples (Supplementary Fig. 8c).

Next, we investigated gene expression profiles in the underlying skin. Comparison of the transcriptomes between *S. aureus* high and low samples revealed a set of 256 significant genes ( $FDR < 0.05$ ,  $FCH \geq 1.5$ ) (Supplementary Fig. 10a, Supplementary Table 3). To explore whether the *S. aureus*-regulated genes were relevant to global features of AD pathophysiology, we created a

co-expression network based on pairwise Pearson correlation ( $r > 0.7$ ), using all significant AD-associated genes. Network community detection<sup>20</sup> identified ten distinct modules, enriched for disease-relevant functions (Fig. 4a, Supplementary Fig. 10c). Projecting *S. aureus*-regulated genes onto the AD network revealed significant enrichment in genes that mapped to modules M1 and M5 (hypergeometric test,  $FDR < 0.05$ ), associated with keratinocyte differentiation, and extracellular matrix organization, respectively (Fig. 4b).

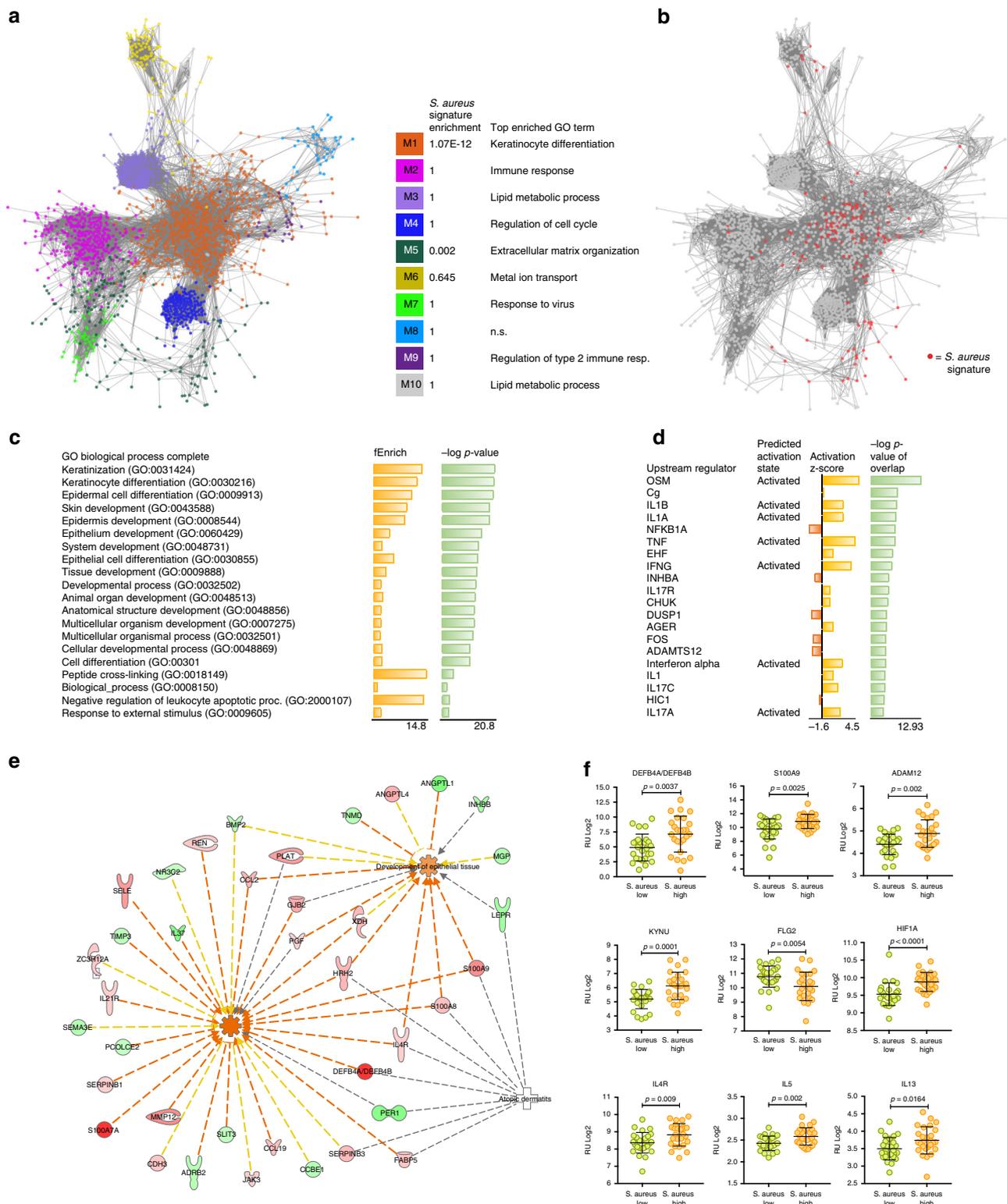
Functional analysis of the *S. aureus*-regulated genes using GO and IPA revealed the enrichment of keratinization and skin development (Fig. 4c), and  $T_H17$  signaling and tryptophan (trp) degradation (Supplementary Fig. 11a), respectively. IPA predicted *IL1B*, *TNF* and *IFNG* as top upstream regulators (Fig. 4d) and leukocyte migration and development of epithelial tissue as downstream effects (Fig. 4e). While upregulated genes were enriched for inflammatory signaling (Supplementary Fig. 11b), downregulated genes were over-represented for mainly skin development (Supplementary Fig. 11c). Top genes included skin barrier and antimicrobial factors (*S100A7*, *DEFB4A/B*, *S100A9*, *MMP12*, *FLG2*, *CLDN8*, *ADAM12*), components of trp metabolism (*KYNU*, *TDO2*, *KMO*), immune activation (*IL1B*, *CCL2*, *CCL19*), and  $T_H2$  signaling (*IL4R*, *IL5*, *IL13*, *PI3*, *TNFRSF4*, *CCR4*). Moreover, *HIF1A* and its targets *HK2* and *PFKP* were among the significantly regulated genes (Fig. 4e, f, Supplementary Fig. 10b, d–e).

Since the kynurenine pathway was enriched in the *S. aureus* “high” samples, we performed transcriptomic reconstruction of trp breakdown, indicating the accumulation of 3-hydroxyanthralic acid (3-HAA) in the skin (Supplementary Fig. 12a). Further, we investigated to what extent AD-associated strains of *S. aureus* may depend on trp, and isolated 32 *S. aureus* strains from AD patients with moderate to severe manifestation of the disease. We found that 66% of the isolated strains grew independent of trp (Supplementary Fig. 12b, Supplementary Table 4), and this observation was further supported by our WGS data (73% trp biosynthesis-related genes in the *S. aureus* high samples) (Supplementary Fig. 12c).

Finally, to explore the impact of *S. aureus* on gene targets identified by this study, organotypic human epidermal equivalents were topically exposed to  $10^6$  CFU *S. aureus* for 24 h, followed by measurement of gene expression by qPCR (Supplementary Fig. 13a–c). *S. aureus* significantly induced the expression of *DEFB4*, *PI3*, *IL4R* and *S100A9* (Supplementary Fig. 13d).

### Integrative analysis of the PSO microbiome and transcriptome.

Next, we combined the PSO microbiome with associated skin transcriptomes. Testing the top discriminating microbes in PSO in terms of their ability to partition the skin transcriptome did not yield significant results. Therefore, we constructed a co-expression network based on PSO differentially expressed genes, and partitioned it into 12 modules as described above<sup>20</sup> (Supplementary Fig. 14a). We identified associations between the top 25 most differentially abundant taxa between PSO and HV, and module eigengenes derived from the PSO co-expression



network, using linear models implemented in the MaAsLin package<sup>21</sup>. A linear model controlling for body site, institution, age and gender effects was fit for each microbe–module eigen-gene pair. Overall, six associations were identified (FDR < 0.20) (Supplementary Fig. 14b). Negative associations were identified between *Corynebacterium* spp. and three co-expression modules; M2, M3 and M10. M3 was strongly associated with the cell cycle, whereas M2 and M10 were enriched for inflammatory pathways including interferon signaling, and T<sub>H</sub>1 and T<sub>H</sub>2 activation

(Supplementary Fig. 14c). These results suggest that *Corynebacterium* spp. may play a regulatory role which is reduced in disease however. Overall, our results indicate that microbe–host associations in PSO are considerably less well defined than observed in AD.

**Microbiome associates with clinical severity in AD.** To investigate the effect of microbiota and gene expression on clinical

**Fig. 4** The “*S. aureus* signature” and functional associations. **a** Using AD-associated genes (AD patients,  $n = 82$ ) identified at FDR level  $10^{-5}$ , we created an AD gene co-expression network, which was partitioned into modules by network community detection. Top enriched GO terms are indicated for each network module. **b** Differential analysis between *S. aureus* “high” ( $n = 27$ ) and “low” ( $n = 25$ ) samples revealed 256 differentially expressed genes (FDR < 0.05, FCH  $\geq 1.5$ ). Hypergeometric tests revealed significant enrichment in *S. aureus*-associated genes (colored red) that mapped to modules M1 and M5. Gene annotations and Ingenuity Pathway analysis identified **c** enriched gene ontology terms in the *S. aureus* signature, and **d** predicted upstream regulators, respectively. **e** Molecular networks generated between top functions and associated genes. The red color of the gene symbols indicates upregulated genes, green indicates downregulated genes. Red colored edges indicate predicted activation, yellow edges indicate inconsistent findings, and gray edges lack a predicted effect. **f** Statistical analysis (unpaired *t* test) of RNA expression levels of selected genes in *S. aureus* “high” and “low” abundance samples. The center line in the dot plots corresponds to the mean, and the error bars correspond to the standard deviation. The source data files used to generate the present figure are available from the NCBI Sequence Read Archive under accession PRJNA554499, and from EBI ArrayExpress under accession E-MTAB-8149

severity, we performed feature selection coupled with multivariate regression. Genes and microbes were ranked according to their correlation with SCORAD or PASI indices and then regression models were trained onto the top  $N$  features ( $N = 5\text{--}50$  in increments of five), and the best performing model was selected. Prediction of SCORAD from microbial abundance was optimal with the top 35 species with Random Forest regression (Supplementary Fig. 15a). The accuracy was relatively modest (MAE = 12.35, correlation true vs. predicted = 0.55) suggesting that variability in microbial abundance and clinical severity are likely explained by other factors. Amongst the most highly correlated species were *Tepidimonas* spp., *Propionibacterium acnes*, and *S. aureus* (Supplementary Fig. 15b). Most of the selected species were negatively correlated with the exception of *S. aureus*, indicating that microbial diversity may be inversely associated with clinical severity in AD. Diversity correlation with SCORAD revealed a weak but significant correlation ( $\text{cor} = -0.27$ ,  $p = 0.01$ ).

Transcriptomic data outperformed microbial abundance for predicting SCORAD. The top model identified 15 genes using linear regression as the best set of predictive genes (Supplementary Fig. 15c), (MAE = 9.84, correlation true vs. predicted = 0.66). Amongst the top genes were *SEMA3D*, *IGSF10*, *LGR5* and *CASP10* (Supplementary Fig. 15d). No associations between clinical severity and microbial abundance in PSO were identified.

Finally, oozing and crusting in the AD lesions indicative of infection, correlated significantly with the abundance of *S. aureus*, supporting the link between colonization by *S. aureus* and clinical severity in AD (Supplementary Fig. 16).

## Discussion

Microbe–host interactions may be central to skin homeostasis, and dysbiosis may drive disease<sup>10,15</sup>. However, to what extent the skin microbiota may associate with the host skin phenotype and underlying mechanisms, remains elusive. To achieve a better understanding of the dialogue between the skin and the skin microbiome, we used skin as a model and AD and PSO as proxies for  $T_H2$ -associated atopic and  $T_H17$ -associated autoimmune inflammation, respectively, giving us the opportunity to compare two types of inflammation and the associated microbiomes and transcriptomes. Here we present the findings of a large cohort ( $n = 316$ ) combining the analysis of skin microbiomes and associated transcriptomes, identifying unique gene profiles that characterize healthy vs. inflamed skin. We show that while AD is dominated by one single microbial species, multiple species associate with PSO, and the abundance of the dominating species in AD, *S. aureus*, correlates with disease relevant gene expression (overview of the main findings in Fig. 5).

Analysis of the two patient groups shows distinct differences between microbiomes. The most significant change, the increase in *S. aureus* abundance in AD, is not present in all samples. A

number of AD lesions contain little or no *S. aureus*, representing potentially different endotypes of the disease which deserve further investigation. The changes associated with psoriasis are more complex than for AD and involve many different bacteria, including *Corynebacteria* and *Finnegoldia*. Through exploring classification models as a way to discriminate disease state, several of the OTUs were also strong predictors of disease status. *Lactobacillus* appeared to be consistently depleted in both diseases, whereas disease-specific species such as *S. aureus* in AD and *C. simulans* and *C. kroppenstedtii* in PSO were important for prediction in disease. Moreover, we identified two distinct community patterns by means of co-occurrence analysis. In AD, it was clear that *S. aureus* dominated the microbial landscape and negatively correlated with several skin commensals, such as *S. epidermidis* and *Corynebacterium* spp, and thus may be associated with the depletion of potentially regulatory or protective microbes. Importantly, recent studies have shown that *S. epidermidis* may specifically limit the growth of *S. aureus*<sup>22</sup>, and disease severity is inversely correlated with the abundance of *S. epidermidis* relative to *S. aureus*<sup>14</sup>. In contrast to the dynamics revealed in AD, the PSO-associated interaction network suggested that multiple, co-occurring species, is a more representative model. Previous studies have shown that while single taxa are unable to discriminate between PSO and healthy skin, the combined relative abundances of selected genera (*Corynebacterium*, *Streptococcus* and *Staphylococcus*) attain significance across the groups<sup>15</sup>. In our hands, the same genera display clear patterns of co-occurrence, increasing in relative abundance in PSO lesions compared with healthy skin.

We observed a remarkable loss of strictly anaerobic bacteria in AD, indicating a switch in the skin microbiome from anaerobic to aerobic metabolism. Healthy skin is normally  $O_2$  deprived<sup>23</sup>, but a dry flaky skin and an impaired epidermal barrier function—characteristics seen in AD<sup>24</sup>, may increase oxygenation and select for the low abundance of strictly anaerobic bacteria such as *Lactobacillus* spp or *Finnegoldia* spp. At anaerobic conditions bacteria are fermenting organic matter, e.g. in skin the amino acid serine originating from filaggrin degradation, forming in particular lactic acid, propionic acid and other short chain fatty acids (SCFA). These metabolites lower the skin pH to  $\text{pH} < 5.5$ , maintaining a protective acidity of the skin. Potassium lactate is also one of the most relevant “natural moisturizing factors” of healthy skin. Furthermore, gram-positive anaerobe cocci such as *Finnegoldia*, *Anaerococcus*, and *Peptoniphilus*, stimulate rapid induction of antimicrobial peptides response in human keratinocytes, which could be an important signaling mechanism to the keratinocytes when the skin is injured<sup>25</sup>. In the complete or partial absence of these organisms, danger signaling in keratinocytes and other barrier functions could be impaired, potentially favoring colonization by *S. aureus*.

We identified AD and PSO transcriptomes which overlap substantially with previously published studies<sup>26–29</sup>. While



biofilms<sup>45</sup>, triggering hypoxia inducible factor (HIF) transcription factors, which drive the metabolic changes and shape the immunological response. In the *S. aureus* high samples, we observed significant upregulation of *HIF1A*, and its target genes *HK2*, *PFKP* and *IL1B*. *HK2* and *PFKP* are key regulatory enzymes in glucose metabolism. While glycolysis in the host is necessary to combat infection, a recent study shows that staphylococcal glycolysis is equally necessary to cause infection<sup>44</sup>. In line, we observed increased levels of glucose and galactose metabolism and glucose transport, in the *S. aureus* high associated microbiome, suggesting enhanced ability to compete with the skin for limited oxygen and glucose. Staphylococci also express hypoxia activated genes<sup>46</sup>, some of them which are part of the two-component system, regulating metabolic activity in anaerobic conditions<sup>44</sup>. Accordingly, the two-component system was significantly enriched in the *S. aureus* high samples, including genes which facilitate anaerobic respiration and biofilm formation<sup>47,48</sup>.

The essential amino acid trp and its catabolites are emerging as important mediators of host–microbe interactions, with multiple effects on host physiology on the one hand, and consequences for the survival of the microorganisms, on the other hand. Recent studies have shown attenuated trp metabolic pathways in the skin microbiota of AD patients, and protective functions of microbial trp catabolites in AD<sup>49</sup>. In turn, the host may use trp “depletion”, or “starvation” as an antimicrobial strategy<sup>50</sup>. We generated a transcriptomic reconstruction of trp breakdown in atopic inflammation (Supplementary Fig. 13a), and detailed analysis of the pathway indicated significant upregulation of *TDO2*, *KMO* and *KYNU*, while alternative endpoints of the pathway were not regulated. Hence, the result suggests accumulation of 3-hydroxyanthranilic acid (3-HAA), and the pathway being locked into the generation of this metabolite. Interestingly, 3-HAA is considered an inflammatory mediator, and has been shown to interact directly with neuronal cells<sup>51</sup>. The depletion of key nutrients by the host is thought to be a mode of host defense during bacterial colonization, and certain microbes, including *staphylococci*, are sensitive to the depletion of trp by the host<sup>50</sup>. While investigating whether such mechanisms are at play in AD during *S. aureus* colonization, we found that the majority of *S. aureus* strains in AD patients were independent of supplied trp.

Taken together, our findings in AD patients imply that *S. aureus* may actively modulate the skin barrier function, while triggering an inflammatory response, including Th2 type signaling, potentially interfering with the regulation of *S. aureus* colonization and perhaps contributing to the exacerbation of clinical symptoms. Moreover, dense colonization by *S. aureus* may cause local hypoxia, leading to metabolic reprogramming of the host and its microbiota, and amplification of the inflammatory response. Our findings also suggest that *S. aureus* promotes trp degradation and the accumulation of proinflammatory metabolites in the host, which may help the host to maintain a defensive immune barrier, but may also feed into AD pathology. Finally, AD patients may be preferentially colonized by *S. aureus* strains that are capable of trp biosynthesis, providing them with a colonization advantage in the atopic environment. However, the defense mechanism of limiting trp may also result in adverse effects, i.e. repressing colonization by benign commensal bacteria, creating more dysbiosis.

Unlike in AD, where one species, *S. aureus*, was identified as the dominant microbe, we observed that multiple species demonstrated increased abundance in PSO. To establish a relationship between host transcriptomic profiles and microbial abundances, we screened PSO-associated microbes against modules of co-expressed genes, but detected only weak associations between potential pathogens and the expression of host transcripts. Several modules were found to be associated with

microbial abundance, four of which were associated with one particular taxon corresponding to *Corynebacterium* spp. This species was negatively associated with a module enriched for “interferon signaling”, suggesting a potential protective, or regulatory role of pathways relating to psoriatic inflammation. This is of particular importance, since interferon signaling represents a key event during the initiation of psoriatic inflammation<sup>52</sup>. However, since the identified correlations with bacteria were weak, further studies are warranted for unraveling the potential role of fungi or viruses in psoriatic inflammation.

Finally, we applied regression algorithms to highlight potential links between transcriptomic and microbial biomarkers with disease severity. In AD a set of 35 microbes was identified and linked to clinical severity. Of these species, all, except *S. aureus*, were negatively correlated with disease severity. The association between the abundance of *S. aureus* and severity has been observed in previous studies<sup>10,53</sup>, and specific strains of *S. aureus* contribute to a varying degree to severity, adding to the complexity of AD disease<sup>14</sup>. We lack *S. aureus* strain information in this study, but identify genes which appear to be driven by *S. aureus*. More than half of these genes are relevant to AD pathophysiology, overlapping with the AD transcriptome, but only a handful of these genes are associated with disease severity, suggesting that the *S. aureus* signature is driven mainly by *S. aureus*, not by severity. Lastly, we conclude that the host transcriptome predicted patient SCORAD to greater accuracy than the microbiome, suggesting that gene expression is a better predictor of clinical severity.

This study represents a large cohort with a rich dataset giving an opportunity for further detailed analysis of specific microbe–host interaction. For additional insight, full genetic diversity of the microorganisms need to be explored, adding information concerning strains and the functional potential, as well as information related to yet unexplored microorganisms including fungi, viruses and archaea. Principles of microbe–host interactions will need to be added to the concept of homeostasis and disease, and only an integrated approach will be able to follow the complex ecology of human health, paving the way for medical interventions that aim at preserving health-associated homeostasis between humans and their microbiota.

## Methods

**Subject recruitment and sampling.** Adult patients (18–83 years) with moderate-to-severe chronic AD (SCORAD score > 25,  $n = 91$ ) and plaque-type PSO (PASI score > 7,  $n = 134$ ) as well as healthy volunteers ( $n = 126$ ) were recruited for the study (Supplementary Fig. 1). Microbiome samples and skin biopsies were obtained from areas with active disease in the upper back or posterior thigh in AD patients, and from the upper and lower back in PSO patients. Healthy volunteers were sampled in corresponding skin areas.

Each subject underwent a physical examination by a dermatologist and the medical history was recorded. The diagnoses were made by a dermatologist based on clinical presentation, personal history, laboratory findings and the criteria of Hanifin and Rajka<sup>54</sup>. The exclusion criteria included concomitant autoimmune diseases (e.g. rheumatoid arthritis, diabetes, alopecia areata, etc.), the use of systemic antibiotics within 2 weeks and systemic immunosuppressive therapy or phototherapy or systemic biologic agents within the previous 12 weeks prior to screening. Before skin sampling, the biopsy sites were left untreated for at least 2 weeks and cleansing with only the non-antibacterial Dove soap was allowed and washing was avoided for 24 h prior to sampling. The patients or healthy volunteers who did not match these clinical exclusion criteria were removed from the study. The following biological samples were then obtained and submitted to analysis: (1) microbiome samples from upper/lower back, posterior thigh or buttocks (PSO, AD, HV) with no prior cleaning or preparation of the skin surface using sterile gloves to prevent cross-contamination were obtained placing a sterile ring (2.5 cm diameter) onto the appropriate skin area, 1.5 ml phosphate-buffered saline (PBS) was supplemented into the ring and the area sampled scraping a glass rod in a circular motion ten times to the left and to the right. Subsequently, the microbiome-enriched PBS was harvested and stored. In addition, mock samples containing only PBS were collected at each sampling time in order to assess contamination. (2) 6 mm punch biopsies from skin at the “microbiome” sites were taken in local anesthesia. Subsequently, samples were stored in RNAlater (Sigma-Aldrich) and

subjected to further analyses (Supplementary Fig. S1a). The study was approved by the appropriate local Institutional Review Boards (University of Helsinki, Dnro 91/13/03/00/2011; Heinrich Heine University Düsseldorf, 3647/2011; King's College London, 11/H0802/6) and all subjects provided written informed consent before participation.

Upper back and thigh posterior were chosen to represent skin areas with AD lesions, while upper back and lower back were chosen for PSO lesions, at the same time taking care in minimizing harm due to the surgical procedure of obtaining skin biopsies (Supplementary Fig. 1a). Only samples from active skin disease, i.e. lesional skin, were included in the study, and normal skin from HV was used as a control. In HV, samples were taken from sites corresponding to the sampling sites in AD and PSO. Identical SOPs were used in all three clinical centers and quality and homogeneity was assessed and confirmed in a pilot study prior to starting the major sample collection. Final analysis included 82 AD, 119 PSO and 115 HV samples integrating the microbiome with the transcriptome.

**16S rRNA gene sequencing and analyses.** DNA was extracted from the clinical swab and mock samples using Qiagen's Pathogen Lysis Tubes and the QIAamp UCP Pathogen Mini Kit (Cat.No: 19092) according to the manufacturer's instructions. In brief, sample pellets were resuspended in 500  $\mu$ l Buffer ATL and vortexed for 10 min at maximum speed using Pathogen Lysis Tubes containing glass beads. The samples were transferred to fresh Beckman tubes and incubated in 16.5 mg/ml lysozyme (Sigma) for 30 min at 37 °C. Fifty microliters proteinase K was added and the samples were then incubated for 10 min at 56 °C. Addition of 250  $\mu$ l of Buffer APL2 was followed by incubation at 70 °C for 10 min. Ten microliters RNA-grade glycogen (20 mg/ml, Thermo Scientific) were added to maximize DNA recovery. Ethanol was added to a final concentration of 25%. DNA was extracted and washed using spin columns, and subsequently eluted in 50  $\mu$ l of Buffer AVE.

For 16S rRNA gene amplification and preparation for sequencing, 2.5  $\mu$ l template were amplified in RT-PCR GradeWater (Life technologies), 3% DMSO, with 1 $\times$  PCR HF buffer using Phusion Hot start II DNA polymerase, 200  $\mu$ M dNTPs (all Thermo Scientific), and 500 nM custom primers (Eurofins MWG Operon). One universal forward primer (341f 5'-CCTACGGGNGGCWGCAG with adaptor B, Lib-L) was paired with one of 104 barcoded reverse primers (805r 5'-GACTACHVGGGTATCTAATCC with adaptor A, Lib-L) (Supplementary Table 5). Each barcode consisted of seven nucleotides, contained no homopolymers, and a pair of barcodes differed in at least two positions. Each PCR was run in triplicates and the PCR products from each sample were pooled. A negative control PCR reaction lacking template was included for all primer pairs in each run. The PCR was run for 30 cycles. The PCR products were purified from the reaction using Dynabeads® MyOne™ Carboxylic Acid (Life Technologies, Cat.No: 35401) and TruSeq precipitation buffer (16% PEG-6000, 1.5 M NaCl) on the Magnatrix 1200 (LBH Advanced Bioservices AB, Sweden). The purity of the amplicons was visualized on the Agilent 2100 BioAnalyzer using High sensitivity DNA chips and reagents (Agilent Technologies, Cat.No: 5067-4626) according to the manufacturer's instructions. DNA concentrations were measured by real-time PCR (KAPA Library Quantification Kits For Roche 454 GS Titanium platform, Cat. No: KK4821 and BioRad CFX96 Touch™ Real-Time PCR Detection System; C1000 Thermal cycler) according to the manufacturer's instructions with samples diluted 1:500, 1:1000, and 1:2000 in 10 mM Tris-HCl, pH 8.0. Extension time was 90 s. Finally, the samples were adjusted to  $1.0 \times 10^8$  DNA molecules for each sample before pooling 50–60 samples per 454 sequencing run.

For 454 amplicon sequencing, emulsionPCR was performed on the amplicon library using a large volume emPCR (Lib-L, v2 reagent kit) according to the manufacturer's amplicon protocols and pyrosequenced (one way read direction) on a Genome Sequencer FLX-Titanium instrument (Roche/454 Life Sciences) at Science For Life Laboratory (SciLifeLab), Stockholm. Each library was sequenced in both regions of a two region gasketed 70  $\times$  75 mm Titanium PicoTiterPlate, and base calling was performed with the on-instrument amplicon filter settings. Samples containing only water were sequenced in order to assess contamination during the sequencing process.

For demultiplexing and preprocessing of 454 reads, all sequence reads were assigned to their samples using the unique sample barcodes. Raw sequence reads were analyzed with AmpliconNoise version 1.25 to remove 454 sequencing and PCR artifacts and PerseusD from the same program package to remove PCR chimaeras, using default parameter values. The output from each sample was further processed in QIIME (Quantitative Insights Into Microbial Ecology) version 1.8.0 if the number of processed high-quality reads exceeded 3000 per sample. Otherwise, the sample was resequenced.

For OTU clustering and taxonomy assignment, the preprocessed dataset comprised of a total of 3,357,091 high-quality reads. The following analysis steps were performed using QIIME version 1.8.0. OTUs were picked at 99.3% identity using the `pick_open_reference_otus.py` command and `uclust 1.2.22q`. Taxonomy was assigned using `blast-2.2.22`. The reference data files used for both OTU clustering and taxonomy assignment were downloaded from the Greengenes Database Consortium<sup>55</sup>. As AmpliconNoise did not perform very well in identifying chimeric sequences in our dataset, ChimeraSlayer<sup>56</sup> was applied here within the QIIME pipeline and identified chimeric sequences were removed from the OTU table and the phylogenetic tree.

**Analysis of 16S rRNA gene data.** The protocol for 16S rRNA gene data production used long amplicons and deep 454 sequencing in order to make it possible to exhaustively distinguish as many different OTUs as possible and to discover close to all relevant bacteria in the skin. The noninvasive sampling protocol was designed to specifically sample bacteria on the surface of the skin as the total amount of bacteria in the deeper skin layers is expected to be low. 16S rRNA gene PCR was carried out directly from the DNA preparation, without another amplification step to avoid amplification-induced bias to the data. The analysis of the 16S rRNA gene data was carried out according to standard protocols and relevant controls and low-quality data was removed using AmpliconNoise, which is the gold standard in the field. The same, easily reproduced sampling protocol was used in all cases and the DNA preparations, PCRs, sequencing and analysis were carried out in one facility using standard methods.

Three samples of poor quality were removed from the OTU table (1 PSO and 2 HV, Supplementary Fig. 1b). Abundances were normalized using the Trimmed Mean of M-values method (TMM), implemented in the edgeR Bioconductor package. Good's coverage and Shannon diversity index was calculated using QIIME version 1.8.0.

To preserve statistical power, only OTUs present in more than 25% of all samples were analyzed. Differentially abundant OTUs were identified from the filtered set by comparing the abundance distribution of each OTU across the three clinical groups (HV vs. PSO vs. AD) with the Kruskal–Wallis test (FDR,  $p < 0.05$ ). Microbe–disease-specific associations were detected testing for differences in the abundance distribution of each differentially abundant OTU between the diseased (PSO and AD) and HV groups with the Mann–Whitney  $U$  test (FDR,  $p < 0.05$ ).

For analysis of confounding variables, associations between OTU abundance and nonclinical factors (age, gender, body site and institution) were tested for within the HV group using: (1) Kruskal–Wallis test for body site and institution; (2) Mann–Whitney  $U$  test for gender and (3) Spearman correlation for age. FDR correction was used to correct for multiple testing. Additionally, the same association tests were carried out within the PSO and AD groups to search for disease-factor interactions. Then, a linear model was fitted for each OTU to account for the discovered nonclinical factor associations in the microbe–disease association test. The general linear model equation is:  $OTU\_abundance = clinical\_group + nonclinical\_factors$  and it was fitted using the `lm` function of R.  $p$  values from the test for the disease were FDR-corrected for multiple hypothesis testing.

Bacterial aerobes vs. strict anaerobes were determined according to Bergey's Manual of Systematic Bacteriology<sup>57</sup>. Anaerobe abundance comparison between AD and HV was performed using a Mann–Whitney  $U$  test. Additionally, to account for the systematic increase of *S. aureus* in AD, the same test was repeated only with samples in which the aforementioned bacterium was not detected.

For visualization of results, the phylogenetic tree in Fig. 1a was calculated using ClustalW2 and it was visualized, along with associated data using interactive Tree of Life, iTOL<sup>58</sup> and Adobe Illustrator. NMDS analysis was performed using phyloseq. Boxplots were generated in R version 3.1.0 with the `ggplot2` package version 1.0.1, with whiskers extending to max 1.5 IQR. Outliers were removed.

**Analysis of 16S rRNA gene sequencing negative controls.** For OTU clustering and taxonomy assignment, OTU picking was performed independently from the volunteer samples. The `pick_closed_reference_otus.py` pipeline from QIIME version 1.9.1 was run on the samples, with the same parameters and the same version of the Greengenes 16S database used to process the volunteer samples. The procedure yielded a total of 425 identified OTUs in 40 samples.

To reduce false positives in the contamination analysis, OTUs were discarded if they were present in less than five samples or if the raw abundance of the OTU in any sample was less than 100 reads. Filtering criteria were selected to preserve both low abundant bacteria whose presence is supported by many negative controls and bacteria with enough read coverage to be trustworthy. The filtering procedure reduced the total number of detected OTUs in the mock samples from 425 to 73.

OTU ids obtained from the negative controls were compared against the list of differentially abundant OTU ids from the volunteer samples. None of the OTUs from the negative controls matched the OTUs identified in the volunteer samples.

**Generation and analysis of metagenomics shotgun data.** Libraries for shotgun sequencing were prepared from the same DNA preparations that were used for 16S rRNA gene PCR. The libraries were prepared using the ThruPLEX DNA-seq kit (Rubicon Genomics, Cat No. R400406) and sequenced using the Illumina HiSeq technology. The resulting reads were quality controlled and human sequences were removed. Primary analysis was carried out using MetaPhlan2 and humann2.

**16S rRNA gene sequence-based metagenomic feature inference.** Metagenome inference from 16S rRNA gene sequences was carried out using PICRUSt<sup>59</sup>. Briefly, gene content of individual OTUs was inferred using the Greengenes v.13.5 database and KEGG Orthologs (KOs) were extracted on multiple hierarchical levels (KEGG tiers 1–4). Relevant microbial pathways were extracted from the tier 1 categories “metabolism” and “signaling and cellular processes”. Contribution of each OTU to the respective functional terms was quantified using the “`metagenome_contributions.py`” subroutine, where the relative contribution of OTUs was compared

between *S. aureus* high and *S. aureus* low groups using Student's *t* test followed by Benjamini–Hochberg multiple testing correction. Visualization was carried out using STAMP software<sup>60</sup> and R v.3.5.1.

**Generation of microarray transcriptional profiles.** The skin biopsy samples were stored in RNAlater and total RNA was extracted from the tissue samples using the RNeasy Fibrous Tissue Mini kit (Qiagen). Tissue samples were homogenized using the FastPrep-24 instrument (Nordic Biolabs AB), and RNA was extracted according to the manufacturer's instructions. The yield and purity of RNA in the samples were controlled using a Nanodrop spectrophotometer and Qubit fluorometer to verify absence of inhibitors (R260/280: 2.1; R260/230 nm: 1.3). RNA integrity was quantified by electrophoresis and performed using Agilent dedicated Lab-on-chip (RNA6000 Nano and Pico kits). RNA Integrity numbers and 28S/18S ratio averages were respectively 8.6 and 2.

One hundred nanograms of total RNA was amplified according to Affymetrix protocols (Affymetrix® GeneChip® Whole Transcript (WT) Expression Arrays). Based on expertise of Institut Curie genomic platform, MAQC A RNA samples (Universal RNA, Stratagene, P/N: 740000) were implemented to series of RNA amplification in order to monitor target preparation. In practice, series of 47 RNA (from healthy volunteers, and patients) and 1 Universal RNA were amplified, monitored and labeled. During synthesis steps, purified molecules were quantified using a multichannel Nanodrop (ND8000, Thermo) to normalize amount of molecules used for DNA synthesis (10 µg) and hybridization (5.5 µg). Molecules were also controlled on high throughput electrophoresis (QIAxcel DNA, Qiagen) in order to monitor the size of complementary RNAs (average: 500nt), and fragmented DNA (average: 50nt), to ensure quality of targets and hybridization of microarrays. Series of 96 targets were hybridized onto Affymetrix Gene ST 2.1 96 array plates, including in total two Universal RNA, using an Affymetrix Genetitan MC system. Quality of raw data and normalized data was monitored to control dynamics of the measurements, across series of synthesis, and series of hybridization using bacterial spike in controls added to total RNA, and using Universal RNA.

An automated quality control pipeline based on the arrayQualityMetrics method was used to capture quality failures in microarray data. A total of 12 samples were removed from the dataset. Data were then normalized using the Robust Multi-array Average (RMA) approach implemented in the affy Bioconductor package<sup>61</sup>.

Technical batch effects originating from sample preparation for microarray analysis were removed (SVA-package with Combat function). For identification of differentially expressed genes, a linear model (R package *limma*) was fitted to the data (using age, gender, anatomical location and clinical center as covariates), and pairwise comparisons were done using the empirical Bayes method. Transcriptomes were defined based on a fold change of 1.5 or greater and a Benjamini–Hochberg adjusted *p* value less than 0.05. Functional enrichment analyses were performed using web-based tools (<http://amp.pharm.mssm.edu/Enrichr/> and the PANTHER Classification System version 13.1) and the Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, [www.qiagen.com/ingenuity](http://www.qiagen.com/ingenuity)).

High RNA and microarray quality was assured by standard methods including measurement of RNA quantity and degradation, the use of control probes and QC and preprocessing methods provided the Bioconductor repository of R libraries. The high quality of samples is reflected by a very low rate of exclusion due QC in S16 (3/331 subjects) or Affymetrix (12/331 subjects) analyses. Overall the data were robust and reliable.

**3D human epidermal equivalents and bacterial stimulations.** Human epidermal equivalents (HEEs) were generated by seeding 300,000 human primary keratinocytes (obtained from abdominal plastic skin surgery) on polycarbonate cell culture inserts, followed by 11 days of culture (3 days submerged and 8 days air-exposed) (Supplementary Fig. 13). The last day of culture 10<sup>6</sup> CFU *S. aureus* was topically applied on the HEEs for 24 h, followed by analysis of morphology, histology (protein expression) and qPCR analysis (gene expression).

For bacterial cultures, *S. aureus* (ATCC 29213) was obtained from the Department of Medical Microbiology of the Radboudumc. Bacteria were inoculated on Columbia agar with 5% sheep blood (Dickinson and Company (BD), Sparks, MD) overnight (*S. aureus*) or for 4 days (*F. magna*) at 37 °C. One single colony of each plate was picked and cultured in Brain Heart Infusion medium (Mediaproduits BV, Groningen, The Netherlands). Bacteria were collected by centrifugation, washed two times with PBS and finally resuspended in PBS resulting in bacterial concentrations of 10<sup>7</sup> CFU/ml. To determine the amount of bacteria that was brought on the keratinocyte cultures, bacterial suspensions were serially diluted in steps of 5. Ten microliters of each dilution was placed on sheep blood agar plates and incubated overnight or for 4 days at 37 °C in aerobic (*S. aureus*) conditions, respectively. Visible colonies on the plate were counted for each dilution. The number of CFU was calculated: counted CFU × dilution factor.

For the development of 3D HEE, primary human keratinocytes were obtained from abdominal plastic skin surgery, and isolated and expanded according to the Rheinwald–Green protocol<sup>62</sup> and stored in liquid nitrogen. For the generation of HEE's 300,000 primary human keratinocytes were seeded onto 0.4 µm pore size transwell filters (Thincerts, Greiner Bio-One). Prior to seeding, these transwells were incubated with 100 µg/ml rat-tail collagen (BD Biosciences) diluted in sterile

cold PBS at 4 °C for 1 h. Afterwards, excessive coating was carefully aspirated and filters were washed once with cold, sterile PBS. First, the primary keratinocytes were cultured in a submerged manner in proliferation medium (CnT-prime, CELLnTEC) for 2 days to form a homogeneously distributed monolayer. One day later, medium was changed to differentiation medium (60% 3D barrier, CELLnTEC complemented with 40% Dulbecco's modified Eagle medium, Sigma). On the third day the transwell inserts were lifted to the air–liquid interface to induce differentiation and stratification of the epidermis. From now on cells were cultured in differentiation medium until the end of the culture and they were refreshed with differentiation medium every other day.

On day 7 of air–liquid interface culture 10<sup>6</sup> CFU *S. aureus* in PBS (cultured as described above) was applied topically on the HEEs. For control HEEs only PBS was applied. After 24 h HEEs were harvested for qPCR analysis.

RNA isolation, cDNA synthesis and qPCR analysis was performed as described earlier<sup>63</sup>. All primers were designed and used as described previously<sup>64</sup> (Supplementary Table 5). Target gene expression was normalized to the expression of the house-keeping gene human acidic ribosomal phosphoprotein P0 (*RPLP0*). The  $\Delta\Delta C_t$  method was used to calculate relative mRNA expression levels<sup>65</sup>.

For morphological analysis, HEEs were fixed in 4% formalin solution for 4 h and embedded in paraffin. 6 µm Six-micrometer sections were stained with hematoxylin and eosin (H&E, Sigma-Aldrich) or DAPI stain solution (DAPI Fluoromount-G, Southern Biotech) as described previously<sup>66</sup>.

**Tryptophan dependence assay.** The growth of *S. aureus* strains isolated from moderate-to-severe atopic dermatitis patients was analyzed in tryptophan-containing vs. tryptophan-depleted culture media. The tryptophan-depleted, conditioned medium was derived from IFN-γ activated, IDO-positive human glioblastoma cells 86HG39 that efficiently inhibit the growth of tryptophan-auxotrophic *S. aureus* strains<sup>67</sup>. 86HG39 cells were cultured in Iscove's modified Dulbecco's medium (IMDM) (Gibco, Grand Island, USA), supplemented with 5% heat-inactivated fetal calf serum (FCS) in culture flasks (Costar, Cambridge, USA) in a humidified incubator (37 °C, 10% CO<sub>2</sub>). 1 × 10<sup>6</sup> cells per cell culture flask were stimulated with 1000U/ml IFN-γ (R&D Systems, Minneapolis, USA) for 74 h and the conditioned medium was harvested. As controls, medium from unstimulated 86HG39 cells or tryptophan-free RPMI 1640 Medium (Gibco, Grand Island, USA) (with or without tryptophan supplementation) were used. For the tryptophan dependence assay, a 24 h old colony was picked, resuspended in PBS (Gibco, Grand Island, USA) and serially diluted. The conditioned medium as well as control media (200 µl) were inoculated with 10 µl of the bacterial dilution containing 10–100 colony forming units (cfu). Cultures were incubated overnight. Bacterial growth was monitored by measuring the optical density of resuspended cultures at 620 nm. Tryptophan dependency was validated by growth restoration in tryptophan-containing cultures.

**Classification and feature selection.** In order to identify the best set of representative microbiota for disease class, we used a supervised learning approach based upon Random Forest classification models. The pipeline was used to train models for three contrasts (Healthy vs. AD, Healthy vs. PSO and PSO vs. AD) with the aim of identifying the features that best discriminate disease groups. To identify the most stable predictors of disease, we first removed under-represented species by filtering out those that were present in less than 15 samples for each contrast. Next, to select the most discriminant OTU features, Random Forest feature selection implemented in the R package Boruta was used under a tenfold cross-validation framework with ten randomized repeats. Selected features were ranked by Boruta according to the variable importance score *Z*, which represents the average loss in accuracy after permutation of attribute values across samples. OTUs with a mean *Z* score greater than 0.2 were considered for further analysis. The fold change between healthy and disease patients was calculated to identify depletion or increased abundance in disease. After selection of features, the classification model was built within the R package randomForest. We evaluated the performance of the model using the R package ROCR and report the mean AUC across all folds<sup>68</sup>. NMDS with Bray–Curtis distance using the selected feature sets with the function metaMDS from the R package Vegan was used for visualization purposes only. Selection frequency was calculated as the percentage of times a variable was selected by Boruta across all folds over all randomizations.

**Microbial co-occurrence networks.** We constructed interaction networks on all OTUs present in at least 5% of samples resulting in a core microbiome of 569 OTUs. We employed SparCC<sup>19</sup> on the raw OTU abundance, which calculates a corrected correlation coefficient designed specifically for assessing the correlative relationships between taxa in microbiome studies. The statistical significance of correlations was evaluated against an empirical null distribution obtained with 100 bootstrap iterations (*p* < 0.05 and SparCC > 0.2). Network visualizations were generated in Cytoscape.

**Dimensionality reduction of transcriptomics data.** To select genes for integration, differential analysis between healthy and diseased samples was performed using the limma package reducing the total number of host transcripts under consideration from 32,633 to 16,716 AD and 22,433 PSO-associated transcripts

(FDR < 0.05). Co-expression networks were constructed using the pairwise Pearson correlations between differentially expressed genes. Edges were determined by a hard threshold of  $r < 0.7$  and the main component of the network was retained for further analysis. This reduced the overall transcriptome to a core disease-associated network of 1833 AD and 2653 PSO-associated genes. To partition the network into modules of genes that display similar expression profile, the Louvain method for community detection was employed. Co-expression networks were visualized in Cytoscape.

**Identification of host–microbial interactions.** For each module identified through network community detection, feature reduction was performed by calculation of the module eigengene which is equivalent to the first principal component of module gene expression<sup>69</sup>. To test for associations between microbial taxa and module eigengenes, we used the MaAsLin (<https://huttenhower.sph.harvard.edu/maaslin>) package<sup>21</sup> which fits linear models between arcsine square root transformed OTU relative abundances and metadata after removal of statistical outliers. We assessed the relationship between the top 25 most significant disease-associated microbes and module eigengene pairs while correcting for potential confounding factors using the model:  $\text{microbe} \sim \text{module} + \text{anatomical location} + \text{institution} + \text{age} + \text{gender}$ .  $p$  values were corrected using the Benjamini–Hochberg method and microbe–module associations with FDR < 0.20 were considered as significant. In AD, the top most abundant microbe (*S. aureus*) was used for stratification of the patient cohort, and differential gene expression was compared between samples exhibiting the top and bottom tertiles of *S. aureus* abundance. Network modules enriched for *S. aureus*-associated genes were identified using hypergeometric tests followed by Benjamini–Hochberg multiple testing correction. For modules with significant microbial association, we evaluated the overrepresentation of pathways using Ingenuity pathway analysis.

**Microbiome, transcriptome and disease severity associations.** Regression algorithms were employed for the task of identifying gene and taxonomical markers that can predict the severity of disease indices. For transcriptomic data, one score is computed for each gene, quantifying its degree of correlation with severity. The genes are then ranked in descending order according to their scores and a series of regression models are trained onto the top  $N$  genes ( $N = 5–50$  in increments of five). The trained model can then be used to predict the severity of new samples. To ensure that the constructed regression model can generalize well to unseen samples, a leave-one-out cross-validation scheme is applied to identify performance. Briefly, one sample is made blind to the process of calculation of correlation and training of regression model, and is used to estimate the accuracy of the trained model. In leave-one-out cross-validation, each sample is held out exactly once and the prediction errors are averaged over all samples. The optimal number of genes is then chosen as the set that corresponds to the lowest mean absolute prediction error. The same procedure was applied to OTU abundance data to identify a signature that correlates with severity of disease. Six multivariate regression methods were applied including linear regression, SVM, 5-NN, M5P, RF and Cubist and the model with the lowest mean absolute error was chosen.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data underlying Figs. 1–4, Supplementary Figs. 1–12, 14–16 and Supplementary Tables 1–3 are available from the NCBI Sequence Read Archive under accession PRJNA554499 and from EBI Array Express under accession E-MTAB-8149. Data underlying Supplementary Fig. 13 is provided as a Source Data file. All other data are available from the corresponding author upon reasonable requests.

Received: 17 May 2018 Accepted: 27 August 2019

Published online: 16 October 2019

## References

- Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**, 811–818 (2007).
- Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.* **13**, 260–270 (2012).
- Grice, E. A. & Segre, J. A. The skin microbiome. *Nat. Rev. Microbiol.* **9**, 244–253 (2011).
- Bieber, T. Atopic dermatitis. *N. Engl. J. Med.* **358**, 1483–1494 (2008).
- Nestle, F. O., Kaplan, D. H. & Barker, J. Psoriasis. *N. Engl. J. Med.* **361**, 496–509 (2009).
- Grice, E. A. et al. Topographical and temporal diversity of the human skin microbiome. *Science* **324**, 1190–1192 (2009).
- Oh, J. et al. Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**, 59–64 (2014).
- Naik, S. et al. Commensal-dendritic-cell interaction specifies a unique protective skin immune signature. *Nature* **520**, 104–108 (2015).
- Linehan, J. L. et al. Non-classical immunity controls microbiota impact on skin immunity and tissue repair. *Cell* **172**, 784–796.e718 (2018).
- Kong, H. H. et al. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res.* **22**, 850–859 (2012).
- Geoghegan, J. A., Irvine, A. D. & Foster, T. J. Staphylococcus aureus and atopic dermatitis: a complex and evolving relationship. *Trends Microbiol.* **26**, 484–497 (2018).
- Vu, A. T. et al. Staphylococcus aureus membrane and diacylated lipopeptide induce thymic stromal lymphopoietin in keratinocytes through the Toll-like receptor 2-Toll-like receptor 6 pathway. *J. Allergy Clin. Immunol.* **126**, 985–993 (2010). 993 e981–983.
- Nakamura, Y. et al. Staphylococcus delta-toxin induces allergic skin disease by activating mast cells. *Nature* **503**, 397–401 (2013).
- Byrd, A. L. et al. Staphylococcus aureus and Staphylococcus epidermidis strain diversity underlying pediatric atopic dermatitis. *Sci. Transl. Med.* **9**, eaal4651 (2017).
- Alekseyenko, A. V. et al. Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome* **1**, 31 (2013).
- Gao, Z., Tseng, C. H., Strober, B. E., Pei, Z. & Blaser, M. J. Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. *PLoS ONE* **3**, e2719 (2008).
- Statnikov, A. et al. Microbiomic signatures of psoriasis: feasibility and methodology comparison. *Sci. Rep.* **3**, 2620 (2013).
- Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
- Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **8**, e1002687 (2012).
- Newman, M. E. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **103**, 8577–8582 (2006).
- Morgan, X. C. et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* **13**, R79 (2012).
- Nakatsuji, T. et al. Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis. *Sci. Transl. Med.* **9**, eaah4680 (2017).
- Tholen, S. et al. Skin barrier defects caused by keratinocyte-specific deletion of ADAM17 or EGFR are based on highly similar proteome and degradome alterations. *J. Proteome Res.* **15**, 1402–1417 (2016).
- Weidinger, S. & Novak, N. Atopic dermatitis. *Lancet* **387**, 1109–1122 (2016).
- Zeeuwen, P. L. et al. Gram-positive anaerobe cocci are underrepresented in the microbiome of filaggrin-deficient human skin. *J. Allergy Clin. Immunol.* **139**, 1368–1371 (2017).
- Esaki, H. et al. Identification of novel immune and barrier genes in atopic dermatitis by means of laser capture microdissection. *J. Allergy Clin. Immunol.* **135**, 153–163 (2015).
- Inkeles, M. S. et al. Comparison of molecular signatures from multiple skin diseases identifies mechanisms of immunopathogenesis. *J. Invest. Dermatol.* **135**, 151–159 (2015).
- Quaranta, M. et al. Intraindividual genome expression analysis reveals a specific molecular signature of psoriasis and eczema. *Sci. Transl. Med.* **6**, 244ra290 (2014).
- Suarez-Farinas, M. et al. Alopecia areata profiling shows TH1, TH2, and IL-23 cytokine activation without parallel TH17/TH22 skewing. *J. Allergy Clin. Immunol.* **136**, 1277–1287 (2015).
- Kobayashi, S. D., Malachowa, N. & DeLeo, F. R. Pathogenesis of Staphylococcus aureus abscesses. *Am. J. Pathol.* **185**, 1518–1527 (2015).
- Craven, R. R. et al. Staphylococcus aureus alpha-hemolysin activates the NLRP3-inflammasome in human and mouse monocytic cells. *PLoS ONE* **4**, e7446 (2009).
- Rohrl, J., Yang, D., Oppenheim, J. J. & Hehlgans, T. Human beta-defensin 2 and 3 and their mouse orthologs induce chemotaxis through interaction with CCR2. *J. Immunol.* **184**, 6688–6694 (2010).
- Glaser, R. et al. The antimicrobial protein psoriasis (S100A7) is upregulated in atopic dermatitis and after experimental skin barrier disruption. *J. Invest. Dermatol.* **129**, 641–649 (2009).
- Abtin, A. et al. The antimicrobial heterodimer S100A8/S100A9 (calprotectin) is upregulated by bacterial flagellin in human epidermal keratinocytes. *J. Invest. Dermatol.* **130**, 2423–2430 (2010).
- Wang, S. et al. S100A8/A9 in Inflammation. *Front. Immunol.* **9**, 1298 (2018).
- De Benedetto, A. et al. Tight junction defects in patients with atopic dermatitis. *J. Allergy Clin. Immunol.* **127**, 773–786 (2011). e771–777.
- Mansouri, Y. & Guttman-Yassky, E. Immune pathways in atopic dermatitis, and definition of biomarkers through broad and targeted therapeutics. *J. Clin. Med.* **4**, 858–873 (2015).

38. Mukhopadhyay, S. et al. Matrix metalloproteinase-12 is a therapeutic target for asthma in children and young adults. *J. Allergy Clin. Immunol.* **126**, 70–76 (2010). e16.
39. von Hoven, G. et al. Dissecting the role of ADAM10 as a mediator of Staphylococcus aureus alpha-toxin action. *Biochem. J.* **473**, 1929–1940 (2016).
40. Garlanda, C., Dinarello, C. A. & Mantovani, A. The interleukin-1 family: back to the future. *Immunity* **39**, 1003–1018 (2013).
41. Giustizieri, M. L. et al. Keratinocytes from patients with atopic dermatitis and psoriasis show a distinct chemokine production profile in response to T cell-derived cytokines. *J. Allergy Clin. Immunol.* **107**, 871–877 (2001).
42. Pedersen, M. B., Skov, L., Menne, T., Johansen, J. D. & Olsen, J. Gene expression time course in the human skin during elicitation of allergic contact dermatitis. *J. Invest. Dermatol.* **127**, 2585–2595 (2007).
43. Brunner, P. M. et al. A mild topical steroid leads to progressive anti-inflammatory effects in the skin of patients with moderate-to-severe atopic dermatitis. *J. Allergy Clin. Immunol.* **138**, 169–178 (2016).
44. Wickersham, M. et al. Metabolic stress drives keratinocyte defenses against Staphylococcus aureus infection. *Cell Rep.* **18**, 2742–2751 (2017).
45. Lone, A. G. et al. Colonization of epidermal tissue by Staphylococcus aureus produces localized hypoxia and stimulates secretion of antioxidant and caspase-14 proteins. *Infect. Immun.* **83**, 3026–3034 (2015).
46. Wilde, A. D. et al. Bacterial hypoxic responses revealed as critical determinants of the host-pathogen outcome by TnSeq analysis of Staphylococcus aureus invasive infection. *PLoS Pathog.* **11**, e1005341 (2015).
47. Niemann, V. et al. The NreA protein functions as a nitrate receptor in the staphylococcal nitrate regulation system. *J. Mol. Biol.* **426**, 1539–1553 (2014).
48. Le, K. Y. & Otto, M. Quorum-sensing regulation in staphylococci—an overview. *Front. Microbiol.* **6**, 1174 (2015).
49. Yu, J. et al. A tryptophan metabolite of the skin microbiota attenuates inflammation in patients with atopic dermatitis through the aryl hydrocarbon receptor. *J. Allergy Clin. Immunol.* **143**, 2108–2119 (2018).
50. MacKenzie, C. R., Heseler, K., Muller, A. & Daubener, W. Role of indoleamine 2,3-dioxygenase in antimicrobial defence and immuno-regulation: tryptophan depletion versus production of toxic kynurenines. *Curr. Drug Metab.* **8**, 237–244 (2007).
51. Bohar, Z., Toldi, J., Fulop, F. & Vecsei, L. Changing the face of kynurenines and neurotoxicity: therapeutic considerations. *Int. J. Mol. Sci.* **16**, 9772–9793 (2015).
52. Reizis, B., Colonna, M., Trinchieri, G., Barrat, F. & Gilliet, M. Plasmacytoid dendritic cells: one-trick ponies or workhorses of the immune system? *Nat. Rev. Immunol.* **11**, 558–565 (2011).
53. Zollner, T. M. et al. Colonization with superantigen-producing Staphylococcus aureus is associated with increased severity of atopic dermatitis. *Clin. Exp. Allergy* **30**, 994–1000 (2000).
54. Rothe, M. J. & Grant-Kels, J. M. Diagnostic criteria for atopic dermatitis. *Lancet* **348**, 769–770 (1996).
55. McDonald, D. et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).
56. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
57. Ruan, J. [Bergey's Manual of Systematic Bacteriology (second edition) Volume 5 and the study of Actinomycetes systematic in China]. *Wei Sheng Wu Xue Bao* **53**, 521–530 (2013).
58. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).
59. Langille, M. G. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).
60. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124 (2014).
61. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
62. Rheinwald, J. G. & Green, H. Formation of a keratinizing epithelium in culture by a cloned cell line derived from a teratoma. *Cell* **6**, 317–330 (1975).
63. Bergboer, J. G. et al. Psoriasis risk genes of the late cornified envelope-3 group are distinctly expressed compared with genes of other LCE groups. *Am. J. Pathol.* **178**, 1470–1477 (2011).
64. Zeeuwen, P. L. et al. Genetically programmed differences in epidermal host defense between psoriasis and atopic dermatitis patients. *PLoS ONE* **3**, e2301 (2008).
65. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> Method. *Methods* **25**, 402–408 (2001).
66. Smits, J. P. H. et al. Immortalized N/TERT keratinocytes as an alternative cell source in 3D human epidermal models. *Sci. Rep.* **7**, 11838 (2017).
67. Muller, A., Heseler, K., Schmidt, S. K., Spekker, K., Mackenzie, C. R. & Daubener, W. The missing link between indoleamine 2,3-dioxygenase mediated antibacterial and immunoregulatory effects. *J. Cell Mol. Med.* **13**, 1125–1135 (2009).
68. Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. ROCr: visualizing classifier performance in R. *Bioinformatics* **21**, 3940–3941 (2005).
69. Langfelder, P. & Horvath, S. Eigengene networks for studying the relationships between co-expression modules. *BMC Syst. Biol.* **1**, 54 (2007).

## Acknowledgements

The authors want to thank Prof. A. Andersson (KTH Royal Institute of Technology and Science for Life Laboratory (SciLifeLab)) for providing the 16S primer designs and acknowledge support from SciLifeLab, the national infrastructure SNISS, and Uppmax for assistance in massive parallel sequencing and computational infrastructure. Technical support by D. Lundin, BILS (Bioinformatics Infrastructure for Life Sciences) and T. Keyvanfar is also acknowledged. The research has received funding from the FP7/2007–2013 (Grant 261366). The study was partially funded by the Knut and Alice Wallenberg Foundation, the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation (guysbrc-2012-1) Trust, and Dunhill Medical Trust, the Association pour la Recherche contre le Cancer (ARC), the European Research Council (Grant IT-DC 281987), Institute National de la Santé et de la Recherche Médicale (BIO2012-02 and BIO2014-08), INCA (2011-1-PL BIO-12-IC-1), Fondation ARSEP (R12023JJ), ANR (ANR-13-BSV1-0024-02, ANR-10-IDEX-0001-02 PSL\*, ANR-11-LABX-0043 and ZonMw MKMD grant 114021503) and BIOMAP IMI2 (Grant 821511). Open access funding provided by Karolinska Institute.

## Author contributions

Conceptualization, H.A., B.A., J.B., B.H., J.K., A.L., F.L.-S., F.N. A.R., J.-M.S., V.S., S.T.; Resources, J.B., B.H., A.L., A.R., A.M.B., M.F., K.J., E.L., S.M., S.S., I.T.; Methodology, J.K., T.S., S.P.-N., B.A.; Investigation, S.P.-N., H.D., T.S.; Formal analysis, M.B., M.J., S.P.-N., V.S., S.T., M.B.-S., M.J., G.M., L.Y., D.G., N.F., H.A., D.P., P.O., E.v.d.B., P.Z., G.R., J.S., T.E., W.D., S.K., H.S., H.N.; Data curation, P.H., G.J.-C., V.S.; Visualization, B.A., H.A., N.F., M.J., G.M., S.P.-N., S.T., F.N., V.S.; Supervision, B.A., F.N., S.T., V.S., J.-M.S.; Writing—original draft, H.A., B.A., N.F., B.H., M.J., A.L., G.M., F.N., V.S., S.T.; Writing—review and editing, H.A., B.A., B.H., N.F., G.M., S.T., M.F., R.S.G., P.K., S.K., J.K., A.L., F.L.-S., D.P., S.P.-N., J.-M.S., U.G., B.H., A.R., H.S.; Equal contribution: S.T., A.L., V.S., F.N., B.H., B.A., H.A.; Project administration, H.A., N.F.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-12253-y>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer Review Information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Nanna Fyhrquist<sup>1,2,26</sup>, Gareth Muirhead<sup>3,4,26</sup>, Stefanie Prast-Nielsen<sup>5,26</sup>, Marine Jeanmougin<sup>6,7,8,9,26</sup>, Peter Olah<sup>10,11</sup>, Tiina Skoog<sup>12</sup>, Gerome Jules-Clement<sup>6,7,8,9</sup>, Micha Feld<sup>10</sup>, Mauricio Barrientos-Somarribas<sup>13</sup>, Hanna Sinkko<sup>1,2</sup>, Ellen H. van den Bogaard<sup>14</sup>, Patrick L.J.M. Zeeuwen<sup>14</sup>, Gijs Rikken<sup>14</sup>, Joost Schalkwijk<sup>14</sup>, Hanna Niehues<sup>14</sup>, Walter Däubener<sup>15</sup>, Silvia Kathrin Eller<sup>15</sup>, Helen Alexander<sup>16</sup>, Davide Pennino<sup>4</sup>, Sari Suomela<sup>17</sup>, Ioannis Tassis<sup>17</sup>, Emilia Lybeck<sup>17</sup>, Anna M. Baran<sup>10</sup>, Hamid Darban<sup>13</sup>, Roopesh Singh Gangwar<sup>18</sup>, Ulrich Gerstel<sup>19</sup>, Katharina Jahn<sup>10</sup>, Piia Karisola<sup>2</sup>, Lee Yan<sup>3</sup>, Britta Hansmann<sup>19</sup>, Shintaro Katayama<sup>12</sup>, Stephan Meller<sup>10</sup>, Max Bylesjö<sup>20</sup>, Philippe Hupé<sup>6,7,8,21</sup>, Francesca Levi-Schaffer<sup>18</sup>, Dario Greco<sup>22,23,24</sup>, Annamari Ranki<sup>17</sup>, Jens M. Schröder<sup>19</sup>, Jonathan Barker<sup>16</sup>, Juha Kere<sup>12,25</sup>, Sophia Tsoka<sup>3</sup>, Antti Lauerma<sup>17</sup>, Vassili Soumelis<sup>6,9</sup>, Frank O. Nestle<sup>4</sup>, Bernhard Homey<sup>10</sup>, Björn Andersson<sup>13</sup> & Harri Alenius<sup>1,2</sup>

<sup>1</sup>Institute of Environmental Medicine, Karolinska Institutet, Stockholm 17177, Sweden. <sup>2</sup>Department of Bacteriology and Immunology, Medicum, University of Helsinki, Helsinki 00014, Finland. <sup>3</sup>Department of Informatics, Faculty of Natural and Mathematical Sciences, King's College London, London WC2R 2LS, UK. <sup>4</sup>Cutaneous Medicine Unit, St. John's Institute of Dermatology and Biomedical Research Centre, Faculty of Life Sciences and Medicine, King's College London, London SE1 9RT, UK. <sup>5</sup>Centre for Translational Microbiome Research (CTMR), Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm 17177, Sweden. <sup>6</sup>Institut Curie, 26 rue d'Ulm, 75248 Paris, France. <sup>7</sup>INSERM, U900, 75248 Paris, France. <sup>8</sup>Mines ParisTech, 77300 Fontainebleau, France. <sup>9</sup>INSERM, U932, 75248 Paris, France. <sup>10</sup>Department of Dermatology, University Hospital Duesseldorf, Duesseldorf 40225, Germany. <sup>11</sup>Department of Dermatology, Venereology and Oncodermatology, University of Pécs, Pécs 7632, Hungary. <sup>12</sup>Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm 17177, Sweden. <sup>13</sup>Department of Cell and Molecular Biology, Science for Life Laboratory, Karolinska Institutet, Stockholm 17177, Sweden. <sup>14</sup>Department of Dermatology, Radboud University Medical Center, Radboud Institute for Molecular Life Sciences, Nijmegen 6525, The Netherlands. <sup>15</sup>Institute for Medical Microbiology and Hospital Hygiene, Heinrich Heine University Duesseldorf, Duesseldorf 40225, Germany. <sup>16</sup>St John's Institute of Dermatology, Division of Genetics and Molecular Medicine, Faculty of Life Sciences and Medicine, Kings College London, London SE1 9RT, UK. <sup>17</sup>Department of Dermatology, Allergology and Venereology, University of Helsinki and Helsinki University Hospital, Inflammation Centre, Helsinki 00250, Finland. <sup>18</sup>Pharmacology Unit, School of Pharmacy, The Institute for Drug Research, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem 91120, Israel. <sup>19</sup>Department of Dermatology, University Hospital Schleswig-Holstein, Kiel 24105, Germany. <sup>20</sup>Fios Genomics, Edinburgh EH9 3JL, UK. <sup>21</sup>CNRS, UMR144, 75248 Paris, France. <sup>22</sup>Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33520, Finland. <sup>23</sup>Institute of Biomedical Technology, University of Tampere, Tampere 33520, Finland. <sup>24</sup>Institute of Biotechnology, University of Helsinki, Helsinki 00014, Finland. <sup>25</sup>School of Basic and Medical Biosciences, King's College London, London SE1 9RT, UK. <sup>26</sup>These authors contributed equally: Nanna Fyhrquist, Gareth Muirhead, Stefanie Prast-Nielsen, Marine Jeanmougin.