Deciphering the Role of Skin Surface Microbiome in Skin Health: An Integrative Multiomics Approach Reveals Three Distinct Metabolite–Microbe Clusters

Pierre-François Roux¹, Thierry Oddos¹ and Georgios Stamatas¹

The advent of 16S RNA profiling and shotgun metagenomics has enabled a holistic approach to the study of the skin microbiome composition. Despite the interesting findings in this rapidly developing scientific area, the big question remains: What role does the microbiome play in skin physiology? To begin answering this question, we employed an integrative methodology for microbiome and metabolome analysis of skin surface samples collected from the volar forearm of healthy infants aged 3–6-months. Whereas the infant skin metabolome was dominated by amino acids, lipids, and xenobiotics, the primary phyla of the microbiome were Firmicutes, Actinobacteria, and Proteobacteria. Zooming in on the species level revealed a large contribution of commensals belonging to the Cutibacterium and Staphylococcus genera, including Cutibacterium acnes, Staphylococcus epidermidis, and S. aureus. This heterogeneity was further highlighted when combining the microbiome with metabolome data. Integrative analyses delineated the coexistence of three distinct metabolite–microbe clusters: one dominated by Cutibacterium linked to hydrophobic elements of the skin barrier, one associating Staphylococcus genus with amino acids relevant to the water holding capacity and pH regulation of the skin surface, and one characterized by Streptococcus and independent of any particular metabolomic profile.

INTRODUCTION

Skin is the body’s first line of defense against infections and environmental stressors. It acts as a major physical and immunological protective barrier but also plays a critical role in temperature regulation, water holding, vitamin D production, and sensory perception. The outermost surface of the skin consists of a lipid- and protein-laden cornified layer dotted with hair follicles and eccrine glands that secrete lipids, antimicrobial peptides, enzymes, salts, etc. It harbors microbial communities living in a range of physiologically and anatomically distinct niches. Overall, this constitutes a highly heterogeneous and complex system.

The skin surface is colonized immediately after parturition and is dynamically evolving during the first years of life. Although the long-term impact of delivery mode on the future composition of the skin microbiome remains unclear, it appears that the skin surface of infants born through cesarean section is predominantly colonized by commensal skin bacteria (Streptococcus, Staphylococcus, Propionibacterium), whereas the skin surface of vaginally delivered newborns is mostly colonized by microorganisms common to the female urogenital tract (Lactobacillus, Prevotella, Candida) (Chu et al., 2017; Dominguez-Bello et al., 2016, 2010; Ward et al., 2018). In the first weeks of life, microbial communities start developing site specificity (depending on dry, moist, or lipid-rich niches) while increasing in diversity (Bouslimani et al., 2015; Grice et al., 2009; Mukherjee et al., 2016). At puberty, the stimulation of sebaceous gland secretion by hormones markedly shifts the physicochemical properties of the skin surface and favors the development of lipophilic taxa (Corinebacterium and Propionibacterium) (Mukherjee et al., 2016). During adulthood and in the absence of shifts in external factors, the individual skin microbiome remains relatively stable (Oh et al., 2016), despite the large interindividual variability (Bouslimani et al., 2015), suggesting that mutualistic and commensal interactions exist among microbes and between microbes and host, even for bacterial species often considered as opportunistic pathogens. Under healthy skin conditions, most of the microbes living on the skin behave as commensal or mutualistic organisms. Such microbes inhibit the spread of opportunistic parasites employing various mechanisms, including the stimulation of secretion of innate immunity factors secretion (e.g., IL-1α) (Naik et al., 2012) or antimicrobial peptides by keratinocytes (Nakatsuji et al., 2017). Moreover, commensal microbes contribute to the education of the immune system and to healthy skin barrier homeostasis. In case of skin barrier breach or immunosuppression, these carefully balanced relationships may transition from commensalism to pathogenicity, a transition referred to as dysbiosis (Chen et al., 2018), enabling the overgrowth of pathogenic species, common in skin conditions such as acne (Agak et al., 2014; Bamard...
et al., 2016; Fitz-Gibbon et al., 2013), psoriasis (Gao et al., 2008), ulcer (van Rensburg et al., 2015), and atopic dermatitis (Byrd et al., 2017).

Since the early 1950s, studies involving microbial cultures were undertaken, aiming to understand the role of the skin microbiome in physiology and disease (Kong and Segre, 2012; Roth and James, 1988). The systematic survey of the human microbiome has gained significant momentum over the past decade with the advent of 16S RNA profiling and shotgun metagenomic approaches coupled with second-generation sequencing technologies. Such methods enable the identification of potential causal relationships between microbial communities and clinical outcomes (Schmidt et al., 2018). Studies focusing on the role of individual species in skin physiology have followed a reductionistic approach. More recently, the metabolome has emerged as the Rosetta stone warranting the understanding of the molecular bases of microbial influence on host physiology through the production, modification, or degradation of bioactive metabolites (Shaffer et al., 2017) in diseases ranging from obesity (Maruvada et al., 2017), depression (Valles-Colomer et al., 2019), autism (Sharon et al., 2019), inflammatory bowel disease (Lavelle and Sokol, 2020), diabetes (Liu et al., 2020), neurological conditions (Hertel et al., 2019), as well as heart conditions (Liu et al., 2019; Vojinovic et al., 2019). Despite being successful in identifying metabolic pathways and bacterial targets to improve health, these more holistic, integrative approaches have so far been limited to the study of the gut microbiome (Chen et al., 2019).

Following these examples and applying the approaches used in the study of the gut microbiome to the skin, we aim to understand the role of the skin microbiome in skin physiology. This study constitutes an integrative analysis of skin microbiome and metabolome, examining the case of healthy infants.

RESULTS
To characterize the infant skin metabolomic profile and microbiome composition, we analyzed skin swab and tape samples collected on the dorsal forearm of 16 healthy subjects (9 females, 7 males, aged 118 ± 29 days) (Figure 1 and Supplementary Table S1, see Materials and Methods for an overview of inclusion and exclusion criteria). Skin surface pH and skin surface hydration (SSH) values were also recorded. Matched swab samples (left and right arms) were subjected to 16S ribosomal RNA sequencing followed by profiling of microbial community taxonomic composition defining amplicon sequence variants (ASVs). Skin tapes were analyzed using a combination of ultrahigh-performance liquid chromatography-tandem mass spectrometry and gas chromatography-tandem mass spectrometry. The profiling was carried out using sensitive, high-resolution mass spectrometers in nontargeted mode, capturing a large number of known and uncharacterized metabolites. In addition, parents were asked to provide information on the delivery mode of the infant’s birth.

Overview of the healthy skin surface microbiome and metabolome
The composition and heterogeneity of the skin microbiome and metabolome in this cohort were analyzed first by estimating the relative contribution of each metabolic pathway and bacterial taxon and then grouped into superpathways and phyla, respectively. Overall, from the metabolome perspective, the leading superpathways were amino acids (28.2% of total metabolites), lipids (17.6%), and xenobiotics.
and from the microbiome perspective, the leading phyla were Firmicutes (68.9%), Proteobacteria (15.2%), and Actinobacteria (13.6%) (Figure 2).

The core metabolome present in all the samples at ≥1.4% relative abundance consisted of 24 metabolites and contained fatty acid derivatives (2-hydroxyarachidate, eicosanoylsphingosine, phytosphingosine), amino acids and derivatives (asparagine, hydroxyproline, methionine, N-acetylglucine, dimethylaminoethanol), nucleosides (N6-carbamoylthreonyladenosine), carboxylic acids (1-methyl-4-imidazoleacetic acid), as well as uncharacterized compounds in even proportion across all subjects (Supplementary Figure S1a). Lowering the prevalence threshold to ≥8 samples while increasing the abundance threshold to ≥3% revealed that amino acids (N-acethylthreonine, phenylalanine, arginine, histidine, gamma-glutamylhistidine, gamma-glutamylleucine, etc.) were largely contributing to the core metabolome, together with tricarboxylic acid cycle and (an)aerobic cellular respiration byproducts (alpha-ketoglutarate, pyruvate, lactate), alpha-tocopherol, and lactose (Supplementary Figure S1b). When focusing only on metabolites that were on average contributing the most to the overall skin metabolome without putting any restriction in terms of prevalence, we found that among the most abundant compounds, a significant proportion belonged to xenobiotics (salicylate, propyl 4-hydroxybenzoate, 4-acetamidophenol, triethanolamine, bicine, dexamethasone), likely originating from skincare routines (Supplementary Figure S1c).

The core skin microbiome consisted of 14 genera present in ≥8 samples at ≥1% relative abundance and was dominated by Streptococcus (52.8%), Cutibacterium (11.8%), and Staphylococcus (8.1%) (Supplementary Figure S1d). This overall contribution of major genera was highly heterogeneous across samples: for example, the microbiome from sample 1101 was dominated by Cutibacterium (≥75% of the core microbiome), whereas the one from sample 1111 by Moraxella (≥50% of the core microbiome).

The skin surface metabolome shapes bacterial communities and impacts microbiome diversity

To visualize in a graph the relationship between demographics (age and sex), mode of delivery, skin physicochemical properties, and microbial richness, we employed factor analysis of mixed data, a principal component method specifically designed to explore data from both continuous and categorical variables (Figure 3a). This analysis revealed an association between skin pH, microbiome diversity (Chao1), and SSH. Looking at individual pairwise correlations, we confirmed a weak positive correlation between SSH and microbial richness (Figure 3b). Whereas birth mode appeared to be associated with skin surface pH and SSH, no association was detected with skin microbial diversity in this...
Furthermore, we found no correlation between metabolite diversity (Shannon index) and pH or SSH (Supplementary Figure S3). To explore the association between the skin microenvironment of the individual subject (skin pH and SSH) and bacterial communities, we computed the pairwise Pearson’s correlation coefficient between bacterial genus abundance and skin pH and bacterial genus abundance and SSH. Bacterial genera are color coded according to the phylum they belong to. A more positive correlation with SSH reflects an association between the phylum and a relatively better-hydrated environment, and the opposite holds for a negative correlation. A more positive correlation with pH reflects an association between the phylum and a relatively alkali environment, and the opposite holds for a negative correlation.

**Figure 3. Skin surface microbiome and metabolome correlate with pH and hydration.** (a) Biplot for a FAMD. Variables indicated with an outlined triangle are well-projected in the reduced dimensional plan ($\cos^2 > 0.5$). (b) Dotplot depicting the correlation between SSH and Chao1 alpha diversity index for ASV. (c) Dotplots depicting the Pearson’s correlation coefficient between bacterial genus abundance and skin pH and bacterial genus abundance and SSH. Bacterial genera are color coded according to the phylum they belong to. A more positive correlation with SSH reflects an association between the phylum and a relatively better-hydrated environment, and the opposite holds for a negative correlation. A more positive correlation with pH reflects an association between the phylum and a relatively alkali environment, and the opposite holds for a negative correlation. (d) Dotplots depicting the Pearson’s correlation coefficient between metabolic pathways weight and skin pH and metabolic pathways weight and SSH. Metabolic pathways are color coded according to the superpathway they belong to. A more positive correlation with SSH reflects an association between the species and a relatively better-hydrated environment, and the opposite holds for a negative correlation. A more positive correlation with pH reflects an association between the species and a relatively alkali environment, and the opposite holds for a negative correlation. ASV, amplicon sequence variant; FAMD, factor analysis of mixed data; SSH, skin surface hydration.
Figure 4. Skin surface microbiome and metabolome are highly entangled. Heatmaps (right) and correlation circles (left) depicting canonical correlations—as defined with regularized generalized canonical correlation analysis—between (a) bacteria phyla and metabolic superpathways, (b) bacteria genera and metabolic pathways, and (c) ASVs and metabolites. For (b) and (c), only correlations above R^2 = 0.5 are shown. ASV, amplicon sequence variant; MUFA, monounsaturated fatty acid; SFA, saturated fatty acid.
Figure 5. Unsupervised multiblock sPLS analysis on metabolome and microbiome data highlights three distinct clusters. (a) Biclustering of metabolome and microbiome data (row Z-score) with the k-means clustering results overplotted for both individuals and variables and delineating three metabolite–microbe clusters. (b) Sample plot from the metabolome perspective. (c) Sample plot from the microbiome perspective. (d) Boxplots showing the distribution for pH, SSH, and Chao1 microbiome diversity in the three metabolite–microbe clusters. (e) Contingency heatmaps showing the association between the three metabolite–microbe clusters and birth delivery mode. (f) Heatmap depicting top-correlated predicted microbial metabolic activities (Pearson’s correlation > 0.8) with Cutibacterium, Staphylococcus, and Streptococcus abundance. NS, not significant; sPLS, sparse partial least square; SSH, skin surface hydration; TCA, tricarboxylic acid.
correlation coefficient between skin pH and bacterial genera abundance and between SSH and bacterial genera abundance. By combining in a single graph the coefficient values of the two correlations (genera abundance—pH vs. genera abundance—SSH), we were able to determine the affinity of each genus for distinct skin niches in terms of acidity and moisturization (Figures 3c and Supplementary Figure S2d).

Whereas Pseudomonas, Ruminococcus, Atopeobium, Schaealia, and Lactobacillus were present on subjects with relatively acidic and hydrated skin, Cutibacterium was more abundant on subjects with relatively basic and slightly dry skin, and Moraxella, Agrobacterium, and Acinetobacter were more abundant in those with slightly acid and slightly dry skin. This analysis also revealed that the genera inside a given phylum were settling in heterogeneous niches, hence the significance to study microbiome at the finest possible grain.

We then applied the same method to explore the potential associations between the skin microenvironment and the measured metabolite concentrations (Figures 3d and Supplementary Figure S2e). As expected, amino acids and tricarboxylic acid— and urea cycle—derived metabolites were mostly associated with more acidic and more hydrated skin. We also noticed a broad distribution of lipid-related metabolites across niches, reflecting the broad spectra of chemical properties of these metabolite classes. Whereas long-chain unsaturated fatty acids tended to be associated with slightly more acidic and drier skin, phospholipids were found at higher concentrations in relatively basic and more hydrated sites. Relatively basic and drier niches were more enriched in ceramides.

Skin microbiome aggregates around three distinct communities characterized by their metabolite microenvironment

To investigate the complex relationships between the skin microbiome and metabolome, we applied a regularized Canonical Correlation Analysis at different taxonomic levels: (i) bacterial phyla versus metabolic superpathways, (ii) bacterial genera versus metabolic pathways, and (iii) bacterial species versus metabolites. At the higher taxonomic level, this analysis revealed a strong positive correlation between the abundance of xenobiotics, cofactors, and vitamins and the relative abundance of Actinobacteria as well as a strong anticorrelation between the metabolic superpathways mentioned earlier and Firmicutes (Figure 4a). Zooming in on the genus and metabolic pathway levels (Figure 4b and Supplementary Figure S4a and b) revealed three major clusters: (i) the first one built on the association between Cutibacterium, Acinetobacter, and Corynebacterium in a niche enriched in fatty acid (free fatty acids, monounsaturated fatty acids, and saturated fatty acids), benzoate, tocopherol, and dihydroceramides (Supplementary Figure S4c); (ii) the second one associating Dermacoccus, Agrobacterium, Moraxella, Schaealia, Clostridium, and Staphylococcus with sugars (fructose, mannose), amino acids (leucine, isoleucine), peptides, and vitamin B6 (Supplementary Figure S4d); and (iii) the last one dominated by Streptococcus in a niche independent of any particular correlation with the metabolic pathways mentioned earlier (Figure 4b).

The composition of these three communities was characterized in more detail by examining the microbiome and metabolome data at the species and individual metabolite levels (Figure 4c).

To validate the existence of the three clusters, we applied multomic sparse partial least square unsupervised analysis, integrating microbiome genera abundance data together with metabolome abundance data (Figure 5). Retaining 15 variables in each -omic bloc was sufficient to properly discriminate three clusters of metabolomic and microbe variables (Figure 5a—c and Supplementary Figure S5). The first group of samples (violet cluster) was characterized by an association between fatty acid metabolites and ceramides with Cutibacterium, Actinobacterium, and Bergeyella and is less rich from the microbiome perspective (Figure 5d). The second group of samples (turquoise cluster) was driven by the association between Streptococcus, Porphyrinoma, Propionibacterium, Dermacoccus, and Trueperella in a niche mostly independent of the presence of fatty acids, ceramides, sugars, and pyrimidine and is richer from the microbiome perspective (Figure 5d). The third group (green) was characterized by a richer microbiome associating Schaealia, Corynebacterium, Atopeobium, Lactobacillus, Clostridium, Escherichia, and Staphylococcus with an environment rich in lysine, sugar, and tricarboxylic acid byproducts. Subjects born vaginally tended to host more frequently the first and the third cluster (Figure 5e). We did not find any further association with other parameters.

Finally, we correlated the abundance of Cutibacterium, Streptococcus, and Staphylococcus with the predicted microbial metabolic pathway potential (Figure 5f and Supplementary Figure S6). Streptococcus contributed mainly to acetylene degradation, galactose degradation, and nucleosides catabolism. Interestingly, the abundance of Staphylococcus was highly correlated with the predicted metabolic activities involved in amino acid degradation (L-arginine and L-glutamate catabolism), whereas the abundance of Cutibacterium was highly correlated with oxides and pyruvate metabolism. Both oxides and pyruvate are important intermediates from the serine synthesis pathway, which is used in combination with palmitoyl-CoA during sphingosine synthesis, the precursor of ceramides.

DISCUSSION

Since the late 19th century, the presence of microbes has been associated with diseases. However, mostly through a better understanding of the gastrointestinal tract, we have come to realize that there are commensal and mutualistic species living inside and on us. The particular anatomic location and function of the skin as the interface between the organism and the environment, where microbes are ubiquitous, makes it suitable for microbial colonization. We now understand the skin microbiome as an integral part of the organism’s interface with the environment, which among others, restrains potential colonization by opportunistic pathogens. However, the actual mechanisms of microbe–host interactions and the role of the microbiome in skin physiology remain obscure.

As is the case for the whole human organism, skin is undergoing dramatic changes after birth. At parturition, the newborn starts its journey experiencing a drastic change from a constant-temperature, wet, and sheltered environment to a...
Distinct Skin Metabolite–Microbe Clusters

C. acnes is a major skin commensal and is the dominating species of the pilosebaceous gland, accounting for up to 90% of the total microbiome in sebum-rich sites such as the scalp or the face (Grice et al., 2009). Although accumulating evidence shows its role in enhancing sebaceous gland lipogenesis and triglycerides synthesis in vitro and in vivo (Iinuma et al., 2009), its interplay with stratum corneum lipid metabolism remains elusive. Our data showed that C. acnes had a greater affinity for lipid-rich skin surface and accumulated at sites with greater amounts of fatty acids (2-hydroxyystearate, 2-hydroxypalmitate, myristoleate, arachidate, palmitoleate), cholesterol, and ceramides (N-palmitoyl-sphinganine, N-palmitoyl-sphingosine, N-2-hydroxypalmitoyl-sphingosine, N-stearoyl-D-sphingosine, N-arachidoyl-D-sphingosine). Whether organized into broad bilayers in the intercorneocyte spaces or covalently bound to the corneocyte envelope in the stratum corneum, lipids are essential constituents of the human epidermis, supporting skin barrier function, cell signaling, and antimicrobial defense (van Smeden et al., 2014). Considering the importance of lipids in skin barrier function as well as the role of C. acnes in acne vulgaris, these results are of utmost relevance.

Sta. aureus is known to be involved in the pathology of atopic dermatitis (Leyden et al., 1974). In fact, Sta. aureus typically dominates the microbiome composition on atopic lesions and is responsible for the observed decline in the overall microbiome diversity (Kong et al., 2012). This species relies on the branched-chain amino acids (isoleucine, leucine, valine) for the synthesis of proteins and membrane branched-chain fatty acids. These amino acids are therefore crucial for the species metabolism, adaptation, and virulence (Kaiser et al., 2018).

Untargeted metabolomic profiling of skin samples remains challenging and expensive and compelled us to focus on a restricted number of samples. To overcome this limitation, we designed relevant processing and analytical strategies, mostly relying on the nonparametric statistical framework. Furthermore, we examined the datasets at the pathway/genealogy and genus/phylum levels, avoiding potential problems arising from simple pairwise correlations computed on isolated and sparse variables.

The microbial biomass of the skin is drastically lower than that in the gut, leading to a high host-to-microbe DNA content ratio in the collected samples. To cope with the host DNA contamination that represents up to 90% in skin swab samples (Bjerre et al., 2019), shotgun metagenomic sequencing should be performed at appropriate depth—which remains costly to date—to provide useful data on the microbial community. Despite being less resolutive from the phylogenetic standpoint, 16S ribosomal RNA sequencing offers a more affordable alternative, enabling accurate characterization of the microbial community at the phylum and genus level and avoiding issues of host DNA contamination. We therefore opted for 16S ribosomal RNA profiling in this study.

Metabolic analysis of skin surface samples has been used in previous works to supplement skin microbiome data (Bouslimani et al., 2019, 2015). In this work, we aimed to integrate the statistical analysis of the two -omics datasets to unlock insights focusing on the cross-talk between metabolome and microbiome on healthy infant skin.
This study provides insights on the interplay between metabolome and microbiome on healthy skin and opens new directions in research focusing on the association of ceramides, fatty acids, and Cutibacterium sp. but also relating to longitudinal studies focusing on the evolution of the diverse metabolite–microbe clusters as characterized in this report. It remains to be seen whether the three skin microbial communities identified in this work persist in puberty and adulthood and whether they are predictive of—or preclusive to—pathophysiological outcomes later in life.

MATERIALS AND METHODS

Clinical study, measurements, and sample collection
A single-center, randomized, evaluator-blind, 5-week trial (NCT03457857) was conducted to assess the effects of two skincare regimens on the cutaneous microbiome, metabolome, and skin physiology of healthy infants aged between 3 and 6 months in general good health on the basis of medical history and without any skin conditions or family history of known allergies. In this report, we used only the baseline data to assess the cross-talk between the microbiome, metabolome, and skin physiology. An institutional review board (IntegReview, Austin, TX) approved the study, and parents/legally authorized representatives (LARs) of study participants provided written informed consent. Parents/LARs of prospective participants were screened for eligibility criteria using an institutional review board–approved screener. Parents/LARs were required to be aged at least 18 years. Participant eligibility was assessed at an initial screening visit by the primary investigator, and the study physician confirmed the eligibility of each participant before enrollment. All eligible study participants entered a 7-day washout period, during which parents/LARs were instructed to use a marketed gentle baby cleanser (Johnson’s Head-To-Toe Wash & Shampoo: Johnson & Johnson Consumer, Skillman, NJ) in place of their infant’s normal body cleanser, at least three times during the week, and to refrain from the use of any type of moisturizer or lotion. They were also instructed not to bathe or cleanse the children for at least 12 hours before the scheduled visit. Sample collection from the left or right dorsal forearm was determined by randomization, with one arm used for skin swabs for microbiome analysis and skin tape left or right dorsal forearm was determined by randomization, with

Microbiome profiling
To profile skin microbiota, sequencing was conducted by RTL Genomics. Briefly, DNA was extracted using MagAttract PowerSoil DNA Isolation (Qiagen, Hilden, Germany) on the KingFisher 96-well extraction robot (Thermo Fisher Scientific, Waltham, MA) following the manufacturer’s instructions. Sample amplification for sequencing was conducted using primers encompassing variable regions 1 through 3 (28 forward: GAGTTTGATCNTGGCTCAG, 519 reverse: GTNTTACNCGCGCGKGTG) of the 16S ribosomal RNA gene as previously described (Phillips et al., 2017). Sequencing was conducted on the Illumina MiSeq platform (Illumina, San Diego, CA) according to the manufacturer instructions and targeting a minimum depth of 10,000 taxonomically classified reads per sample, a threshold defined by rarefaction analysis (Supplementary Figure S7). Raw paired-end sequencing reads were first merged using custom R script, and PCR primers were removed from the obtained sequences. These sequences were further quality trimmed, filtered and denoised using DADA2 framework (Callahan et al., 2016) to infer ASVs. Among the 1,647,259 read pairs generated, 1,071,553 were kept. Taxonomy was assigned using the HiMAP National Center for Biotechnology Information–derived database (Segota and Long, 2019). ASV abundance matrix, sample metadata, and taxonomy were finally stored as a phyloseq object (McMurdie and Holmes, 2013). ASVs detected in less than two samples were excluded from the analysis. Microbiome analysis of the negative control (blank swab) showed no relevant contamination (only one ASV detected from the blank swab, among the 1,328 detected in the global experimental design). The ASV count matrix was finally used as an input for PICRUSt2 to define microbiome metabolic pathway potential (Douglas et al., 2019).

Metabolomics
Untargeted metabolomic analysis of skin tape samples was performed by Metabolon (Durham, NC), as previously described (Evans et al., 2009). Individual compounds were identified by comparing the mass spectroscopy data of the samples with a library of >4,500 authenticated purified standards. The library includes information on the retention-time index, mass-to-charge ratio, and chromatographic data (including tandem mass spectrometry data) on all molecular entries. In a given sample, the peak intensities corresponding to each metabolite were normalized to the total intensity count for that sample.

Statistical analyses
All statistical analyses were performed in R, version 4.0.0, and rely on the packages mixOmics (Rohart et al., 2017), FactorMineR (Lé et al., 2008), vegan, and phyloseq (McMurdie and Holmes, 2013).

References

Distinct Skin Metabolite–Microbe Clusters

Factorial analysis of mixed data was applied on a matrix containing pH, SSH, microbiome Chao1 index, as well as sex and mode of birth information for each sample. Regularized Canonical Correlation Analysis was performed on the combination of the metabolomic abundance matrix and the microbiome relative abundance matrix after regularization through Ridge regression ($\ell_2$ penalties) of parameters $\lambda_1$ and $\lambda_2$ using a leave-one-out cross-validation procedure. To define metabolite–microbe clusters, a block sparse Partial Least Square analysis was applied on the combination of the metabolomic abundance matrix (pathway level) and the microbiome relative abundance matrix (genera level) after fine tuning the numbers of dimensions and variables to select using a k-fold cross-validation procedure. The samples and the selected variables were then clustered using k-means biclustering. The optimal number of sample clusters was defined using the gap statistic. When relevant, group comparisons were performed using nonparametric Kruskal–Wallis and Wilcoxon–Mann–Whitney rank-sum tests, and a $p$-value threshold cutoff at 0.05 was considered. Group associations with descriptive parameters were evaluated using the Chi-square test. Correlations were evaluated using Pearson’s correlation and Spearman’s rank method.

Ethical statements

An institutional review board (IntegReview) approved the study (NCT03457857), and parents/LARs of study participants provided written informed consent.

Data availability statement

Raw data are available at the Sequence Read Archive (Bioproject: PRJNA707369). Processed data are available in Supplementary Table S2 (metabolome) and Supplementary Table S3 (microbiome).

ORCIDs

Pierre-François Roux: http://orcid.org/0000-0002-0695-4206
Thierry Oddos: http://orcid.org/0000-0002-6777-5324
Georgios Stamatas: http://orcid.org/0000-0003-4544-7597

CONFLICT OF INTEREST

All authors are employees of Johnson & Johnson Santé Beauté France.

ACKNOWLEDGMENTS

We thank the volunteers who participated in this study as well as Kimberly Capone, Lorena Telofski, Janeta Nikolovski, and Diana Frisca for managing the clinical study.

AUTHOR CONTRIBUTIONS

Conceptualization: TO, GS; Data Curation: PFR; Formal Analysis: PFR; Funding Acquisition: TO, GS; Investigation: TO, GS, PFR; Methodology: GS, TO; Data Curation: PFR; Formal Analysis: PFR; Writing - Original Draft Preparation: TO, GS, PFR; Writing - Review and Editing: TO, GS, PFR

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2021.07.159.

REFERENCES


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/
Supplementary Figure S1. Overview of the healthy surface skin microbiome and metabolome. Barplots depicting the weight of (a) core metabolites with RA > 1.4% in 16 samples, (b) core metabolites with RA > 3% in eight samples, (c) top 20 contributing metabolites, and (d) core microbial genus with RA > 1% in eight samples. The bars on the left of each graph show the average distribution across samples. RA, relative abundance.
Supplementary Figure S2. Skin surface microbiome and metabolome are highly entangled. (a–c) Boxplots highlighting the relationships between birth delivery mode and (a) Chao1 diversity, (b) pH, and (c) SSH. (d) Dotplots depicting the correlation between SSH (green) and pH (red) and Pseudomonas, Granulicatella, and Cutibacterium abundance. The red and green lines correspond to the linear regression for pH (red) and SSH (green). (e) Dotplots depicting the correlation between SSH (green) and pH (red) and urea cycle–related metabolites, ceramides, and long-chain PUFA. The red and green lines correspond to the linear regression for pH (red) and SSH (green). PUFA, polyunsaturated fatty acid; SSH, surface skin hydration.

Supplementary Figure S3. Metabolomic diversity versus SSH and pH. Scatter plots showing the distribution of metabolomic Shannon diversity index as a function of the SSH (left) and pH (right). RA, relative abundance; SSH, skin surface hydration.
Supplementary Figure S4. Top correlated metabolites from the lipid category for *Cutibacterium* and from the amino acid category for *Staphylococcus*. (a–b) Hierarchical biclustering based on pairwise (a) Pearson’s correlation or (b) Spearman’s correlation between microbial and metabolite abundances summarized at the genus and the pathway level. Red boxes delineate the outstanding associations that are conserved across methods. (c–d) Dotplots showing the top correlated metabolites with (c) *Cutibacterium* RA and (d) *Staphylococcus* RA. RA, relative abundance.
Supplementary Figure S5. Unsupervised multiblock sparse partial least square analysis on metabolome and microbiome data highlights three distinct clusters. Hierarchical biclustering of metabolome and microbiome data (row Z-score) with the k-means clustering results from Figure 3a overlapped for both individuals and variables. ASV, amplicon sequence variant.
Supplementary Figure S6. Sparse partial least sPLS-DA on PICRUSt2 metabolic pathway potential data. Biclustering of PICRUSt2 metabolic pathway potential data (row Z-score) with the k-means clustering results overplotted for individuals. The shortlist of MetaCyc biological pathways was selected using an sPLS-DA. sPLS-DA, sparse partial least square discriminant analysis.

Supplementary Figure S7. Rarefaction analysis on microbiome data. Rarefaction curves showing the number of expected ASVs detected as a function of the subsampled reads. ASV, amplicon sequence variant; SSH, skin surface hydration.