Supporting Information

Bouslimani et al. 10.1073/pnas.1424409112

SI Materials and Methods

Molecular Networking. The molecular network was constructed from LC-MS/MS data collected from person 1 and person 2 samples (Datasets S2 and S3), as well as from the beauty products used by these volunteers (Dataset S4). LC-MS/MS data from chemical standards, cosmetic ingredients, bacteria and yeast strains, cultured human basal and differentiated keratinocytes (Datasets S5–S7), and human skin tissues (Dataset S8) were also included in this network. LC-MS/MS data were converted to the *.mzXML* open file format in Compass data analysis software (Bruker Daltonics). This text-based format is used to access MS, scan number, and all MS/MS spectra for each MS run. All *.mzXML* data were used in the analysis (1). To reduce the redundancy of many spectra from identical molecules, MSCluster was used to group MS/MS spectra that originated from identical compounds together (allowing for a precursor *m/z* tolerance of 0.05 Da). These clusters were then represented with consensus spectra, constructed by averaging the peak intensities of the constituent spectra. This set of consensus spectra was used to construct the molecular network.

For every possible pair of MS/MS spectra, spectrum–spectrum similarities were determined using a spectral alignment algorithm calculating a modified cosine. Cosine similarity scores range from 0 to 1, where 1 indicates perfectly matched MS/MS spectra. A cosine threshold of 0.65 was used to construct this molecular network, and spectral pairs matched with a cosine below this value were discarded. Only MS/MS spectral pairs with at least four matched fragment ions were included in this molecular network, and each MS/MS spectrum was only allowed to connect to its top 10 scoring matches and only if the match was in the top 10 scoring matches for both spectra, thus resulting in a maximum of 10 connections per node. The algorithm parameters included a mass tolerance of 0.05 Da for parent mass and 0.05 Da for mass fragments. The molecular network resulting from retaining only connections (i.e., edges) between MS/MS spectra as described above was imported in Cytoscape 2.8.1 for visualization (2). The molecular network was displayed in Cytoscape as nodes and edges (3), where each node represents a specific *m/z* with its corresponding MS/MS consensus spectrum. Network organization was performed using the FM3 layout plugin (ref. 4 and apps.cytoscape.org/apps/fm3). MS/MS spectra corresponding to the background (Blank for swabs samples, growth media, and solvents) that were not connected to a sample node were removed. MS/MS data from MassBank (5) and from METLIN (6) in .mgf format were compared with the MS/MS consensus spectra from the male and female samples. Putative identifications were imported into the network and displayed in Cytoscape as attributes.

All of the MS/MS data used will be available as a public dataset on the MassIVE data repository. Additionally, all identified compounds will be uploaded to the Global Natural Product Social Molecular Networking (GnPS) knowledge base to facilitate rapid future compound identifications for follow-up work.

LTQ-FT-ICR Verification of Heme and HNPs. The verification of the Heme and HNPs by MS/MS and MS/MS/MS was done by LTQ-FT–ion cyclotron (ICR). LTQ analysis was performed on a Finnigan LTQ-FT mass spectrometer (Thermo Electron) coupled to an Advion NanoMate 100 (Advion), using a solvent mixture of 50:50 methanol/water with 1% formic acid. Two swab samples from the female neck area and female groin area were analyzed to confirm the identification of two predicted molecules using MS/MS and MS/MS/MS fragmentation: the Heme in the groin area (Fig. 6*B*) and HNPs (HNP-1, 2, and 3) in the neck area (Fig. 6*C*). All molecules were detected in the MALDI data. Chemical standards of HNP-1 and HNP-3 (AnaSpec, Inc.) were also submitted to tandem MS, and MS/MS (MS^2) spectra were compared with those of HNP-1 and HNP-3 in the neck swab sample. Full-mass spectra were acquired from both neck swab samples and standards in IT ion-positive mode, in the mass range *m/z* 200–2,000. IT-MS settings were as follows: 3 microscans and total acquisition time of 2 min. MS spectra from the groin sample were acquired in FT-MS mode in the mass range *m/z* 200–2,000 with the following settings: 1 microscan and total acquisition time of 5 min. For a direct infusion experiment, with an internal standard, the mass accuracy will be less than 1 ppm. MS/MS fragmentation was performed on manually selected parent ions: *m/z* 616, $[M+H]^+$ ion of the Heme detected in the groin sample, and *m/z* 861, 843, and 872 corresponding to $[M+4H]^{4+}$ ions of HNP-1, HNP-2, and HNP-3, respectively, detected in the neck sample. Parent ions *m/z* 861 and 872 corresponding to $[M+4H]^{4+}$ of HNP-1 and HNP-3 standards were also selected for MS^2. MS^2 was performed in IT-positive mode for *m/z* 861, 843, and 872 in swab samples and standards and in FT-positive mode for *m/z* 616. MS^2 settings were as follows: 5 *m/z* isolation window (IW) for *m/z* 616 and IW 5, 3, and 4 *m/z* for *m/z* 861, 843, and 872, respectively. For standards, the following settings were used: IW 8 *m/z* for *m/z* 861 (HNP-1) and IW 4 *m/z* for *m/z* 872 (HNP-3); 3 microscans in IT and 1 microscan in FT; 2 min of total acquisition time; and 35% normalized collision energy. Product ion *m/z* 557 (from MS^2 of *m/z* 616) was manually isolated within a window of 6 *m/z* and submitted to MS/MS fragmentation (MS^3). MS^3 was performed in IT mode with the following settings: 3 microscans and total acquisition time of 2 min. Acquired mass spectra were analyzed using Xcalibur software (Thermo Scientific).

MS Signal Processing.

MALDI-TOF. The collected MALDI-TOF spectra were converted into the *.mzXML* format with the CompassXport routine (Bruker Daltonics) and imported into the MATLAB software (The Mathworks Inc.) by using the mzxmlread routine from the MATLAB Bioinformatics toolbox. First, each spectrum was normalized by using the Total Ion Count method by dividing each spectral intensity by the summary intensity of the spectrum. The spectra without peaks were reduced in intensity using a heuristic method to compensate for increase in intensity of such spectra after the TIC normalization compared with normal spectra showing a set of peaks. Then, each spectrum was subjected to a two-stage baseline correction by applying two times the msbackadj routine from the MATLAB Bioinformatics toolbox with the following parameters: WindowSize, 200; RegressionMethod, pchip; StepSize, 200; SmoothMethod, none; QuantileValue, 0.05. The mass shift in one of the MALDI plates compared with the other plates was compensated by shifting *m/z* values with data-estimated shifts. The *m/z* shifts used for this were at most 0.08, 0.3, and 0.3 for metabolic, peptidomics, and proteomic spectra, respectively. Finally, all spectra for the same molecular class (metabolites, peptides, proteins) were interpolated onto the same grid of *m/z* values.

LC-MS. All LC-MS datasets were converted into the mzML format by using Compass data analysis software (Bruker Daltonics). For LC-MS feature extraction, we considered datasets from 36 sampling spots (half from male, half from female); the spots were manually selected to cover all anatomical regions. Then, we

applied the OpenMS software to automatically select features for each of these LC-MS datasets and to store selected features in the featureXML format. The FeatureFinderCentroided algorithm was used; for the parameters (Dataset S10). Each detected feature represents a set of *m/z* RT regions corresponding to isotopes of a molecule. The *m/z* and RT widths of these regions are automatically detected based on the data. Then, all detected features were merged into one list. For this analysis, items in the feature list were compared one by one, and if two features had an intersection area of more than 10%, only the feature with maximal intensity was included in the final list. Overall, 51,801 LC-MS features were considered in the final list.

Human Skin Cells and Bacterial Strain Cultures. Neonatal human epidermal keratinocytes (nHEKs) (Life Technologies) and differentiated nHEKs were cultured in Epilife media (Life Technologies) (Dataset S6), supplemented with epidermal-defined growth supplement (Life Technologies) and antibiotic/antimycotic (Life Technologies). nHEKs were differentiated by removing growth supplements and elevating calcium to 1 mM in the culture medium for 4 d. We detached 10^6 basal or differentiated cells using trypsin solution (Life Technologies) and washed them two times with 200 μL PBS (Life Technologies) to remove media traces. Then, 100 μL ethanol (50% solution in water) was added to cell pellets to mimic metabolite extraction from human swab samples. Culture media blanks (controls) were diluted in 50% ethanol in water for analysis.

Commensal bacteria (from the following genera: *Propionibacterium*, *Corynebacterium*, *Staphylococcus*, *Acinetobacter*, *Bacillus*, *Streptococcus*, *Lactobacillus*, *Pseudomonas*, and *Escherichia*), all frequently found on human skin, as well as yeasts (from the following genera: *Malassezia* and *Candida*), were cultured in Petri dishes using appropriate media and growing conditions (Dataset S7). Cell biomass collected from each culture was first washed in 200 μL water, and then 100 μL of 50% (vol/vol) ethanol in water was added. To capture the background, agar pieces cut from blanks containing growth media (controls) were treated following the same protocol. We analyzed 50% ethanol extracts from human cells, bacteria, yeast cultures, and blanks using the same UPLC-MS/MS method described above.

For the study of products of triolein hydrolysis, *P. acnes* was cultured in Tryptic Soy Agar (TSA) media supplemented or not with 1% of triolein (Sigma-Aldrich). *P. acnes* pellets from both culture conditions as well as agar pieces cut from blanks (TSA and TSA supplemented with 1% of triolein) were first washed in 200 μL of water and then extracted in 100 μL of chloroform/methanol (50:50). Extractions were then analyzed using the same UPLC-MS/MS method.

Preparation of Human Skin Tissues. All experiments that involve human skin tissues were performed according to the Institutional Review Board approved by the University of California, San Diego (Project 140144). Normal skin tissues from the back, the scalp, and the neck were obtained from excess healthy tissue removed during sterile surgical procedures for excision of non-melanoma skin cancers (Dataset S8). One normal skin tissue from the back was obtained from the skin bank. The donors included three men 64, 35, and 37 y old, for the scalp, neck, and back skin, respectively. One normal back skin tissue was obtained from a man donor 66 y old, via the New York Firefighters Skin Bank (Dataset S8).

Sterile biopsies (1 cm) were cut from the dermis to the epidermis and were excised from the center of the excess normal skin samples immediately after surgery. Epidermis skin tissue (15 cm) was obtained from the skin bank. We used 0.4 g of each skin tissue for solvent extraction, and 1 mL of solvent was added to 1 g of tissue. Solvent extractions were performed on homogenized tissues, using ethanol (50% solution in water), to mimic metabolite extraction from human swab samples, and using Ethyl acetate/MeOH (1:1). Extractions were performed on ice (4 °C) for 4 h, and then samples were centrifuged and the supernatants were analyzed using the same UPLC-QTOF MS/MS method described above (*LTQ-FT-ICR Verification of Heme and HNPs*).

Shannon Analysis. The diversity analysis was performed separately for 16S rRNA amplicon and LC-MS data and separately for the female and male individuals. For each sample, we calculated the value of the Shannon diversity index for intensities of all features (taxa corresponding to genera for 16S rRNA amplicon and MS features for LC-MS). Then, the values of the Shannon diversity indices for all samples were visualized on the 3D models with the color map used for visualization of other data (blue corresponding to the minimal value and red corresponding to the maximal value of the Shannon index for the individual).

1. Frank AM, et al. (2008) Clustering millions of tandem mass spectra. *J Proteome Res* 7(1): 113–122.
2. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* 27(3): 431–432.
3. Watrous J, et al. (2012) Mass spectral molecular networking of living microbial colonies. *Proc Natl Acad Sci USA* 109(26):E1743–E1752.
4. Hachul S, Junger M (2004) Drawing large graphs with a potential-field-based multilevel algorithm. *Graph Drawing* 3383:285–295.
5. Horai H, et al. (2010) MassBank: A public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* 45(7):703–714.
6. Smith CA, et al. (2005) METLIN: A metabolite mass spectral database. *Ther Drug Monit* 27(6):747–751.

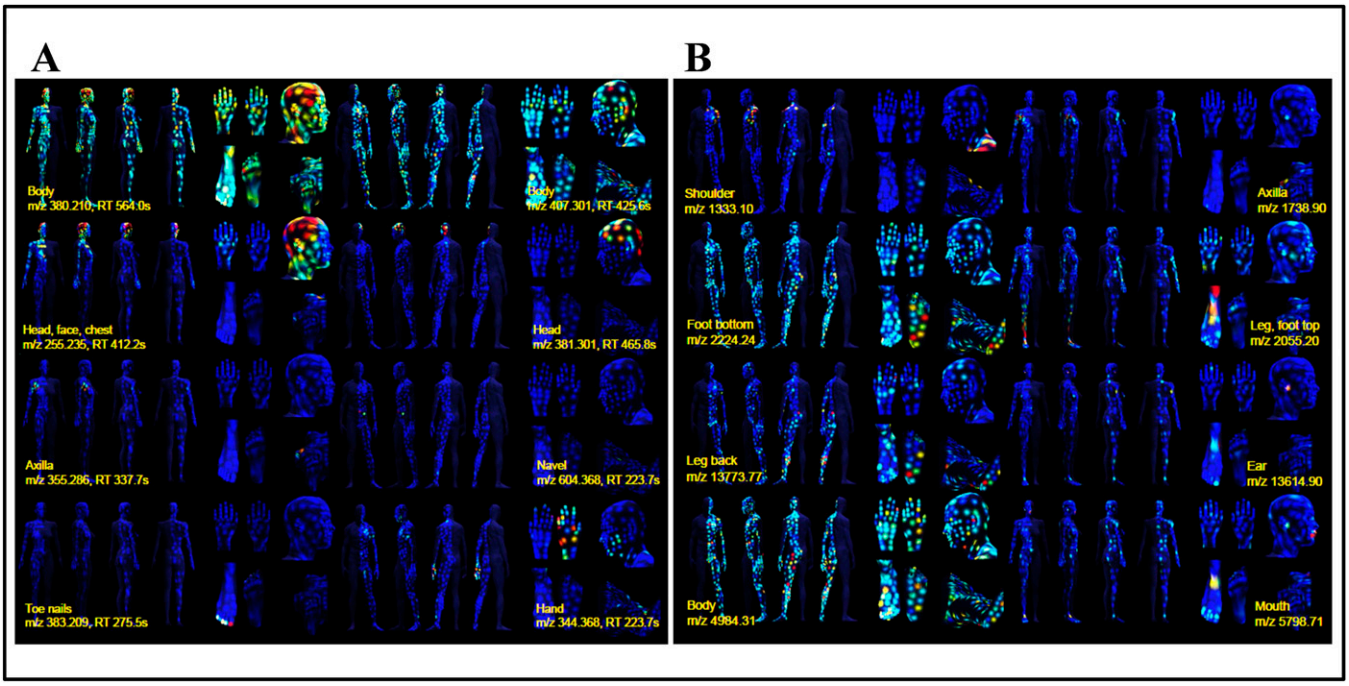


Fig. S1. Representative topographical maps constructed from (*A*) UPLC-QTOF data and (*B*) MALDI data, showing specific and distinct body distributions of molecular features in person 1 and person 2. *m/z*, mass-to-charge ratio; RT, retention time.

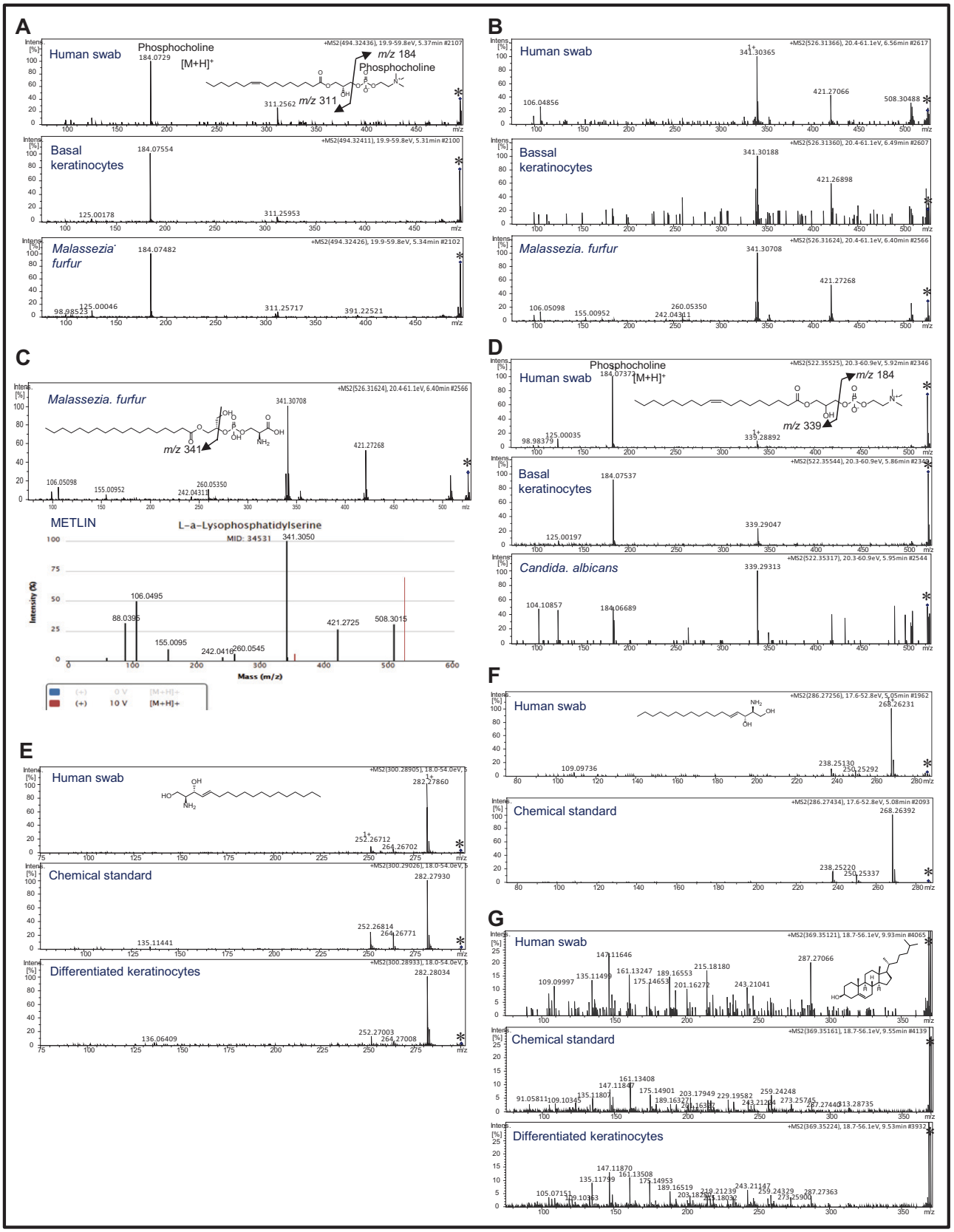


Fig. S2. ESI-QTOF MS/MS fragmentation of PCs, PSs, sphingosines, and cholesterol. (*A*) Lyso-PC (16:1) *m/z* 494, RT 5.4 min in the human swab sample, basal keratinocytes, *M. furfur* ATCC 163486 strain. The loss of the phosphocholine group (*m/z* 184, $[M+H]^+$) induces the release of fatty acid chain (*m/z* 311). Identification based on MS/MS fragmentation and METLIN database interrogation. It is important to note that we drew a specific location of the double bond

Legend continued on following page

and the double bond is drawn as cis, the location of the double bond, or the cis/trans isomer is not known. If a sufficient sample was available, ozonolysis could be used to determine the location of the double bond. What we have here is a match to a family member of Lyso-PC 16:1. (*B*) PS 18:0/0:0 *m/z* 526, RT 6.6 min in human swab sample, differentiated keratinocytes, and *M. furfur* strain. Identification based on METLIN database (Fig. S2*C*). (*C*) A database match of PS 18:0/0:0 *m/z* 526 from *M. furfur*, based on METLIN database. ESI-QTOF MS/MS of *m/z* 526 (RT 6.4) shows that loss of the phosphoserine group releases the fatty acid chain (*m/z* 341). MS/MS of *m/z* 526 in *M. furfur* is identical to PS 18:0/0:0 from the METLIN database. At this level of metabolomic matching, the fatty acid side chain can also be on the other hydroxyl group. (*D*) 1-Oleoylglycerophosphocholine (Oleoyl-PC) *m/z* 522, RT 5.9 min detected in the human swab sample, basal keratinocytes, and *C. albicans*. The loss of the phosphocholine group (*m/z* 184, $[M+H]^+$) induces the release of fatty acid chain (*m/z* 339). It is important to note that we drew a specific location of the double bond and the double bond is drawn as cis, the location of the double bond, or the cis/trans isomer is not known although the RTs match. If a sufficient sample was available, ozonolysis could be used to determine the location of the double bond. This is a match to a molecular family member. (*E*) D-*erythro*-sphingosine (18:1) *m/z* 300, RT 5.3 min in the human swab sample, chemical standard, and differentiated keratinocytes. It is important to note that although the RTs are the same as the standard, the location of the double bond, stereochemistry, or the cis/trans isomer is not known. If a sufficient sample was available, ozonolysis could be used to determine the location of the double bond. What we have here is a match to a molecular family member. (*F*) D-*erythro*-sphingosine (17:1) *m/z* 286, RT 5.1 min in the human swab sample and in the chemical standard. It is important to note that although the RTs are the same as the standard, the location of the double bond, stereochemistry, or the cis/trans isomer is not known. If a sufficient sample was available, ozonolysis could be used to determine the location of the double bond. What we have here is a match to a molecular family member. (*G*) ESI-QTOF MS/MS fragmentation of *m/z* 369 $[M+H-H_20]^+$, RT 9.9 min in the human swab sample, cholesterol standard, and differentiated keratinocytes. It is important to note that although the RTs are similar to the standard, this did not ID the sterol with 100% certainty. MS also cannot determine the stereochemistry. What we have here is a match to a sterol molecular family member. These are the inherent limitations of ID by MS. An asterisk represents a precursor ion.

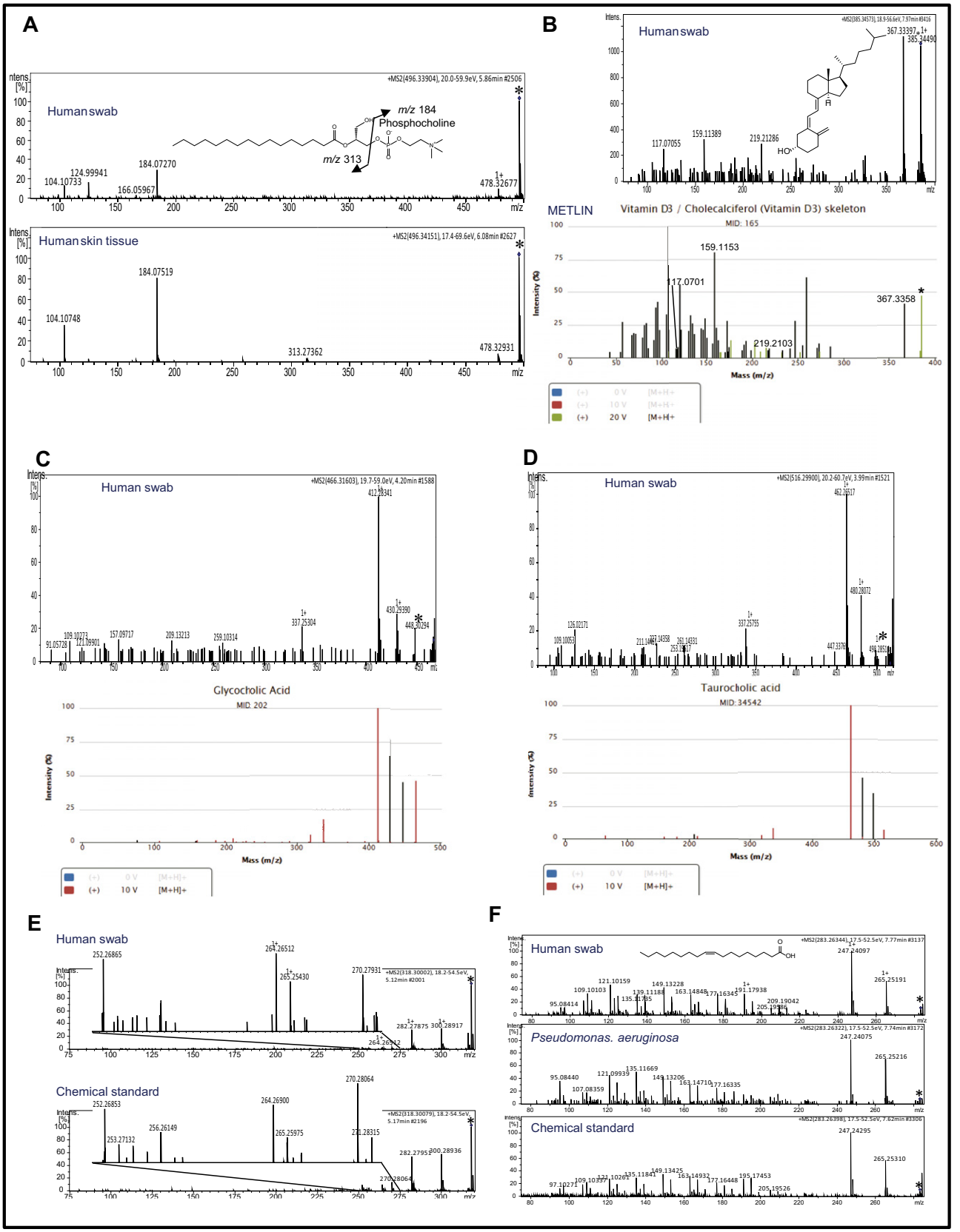


Fig. S3. ESI-QTOF MS/MS ESI-QTOF MS/MS fragmentation of PC 16:0, vitamin D3, Cholic acids, phytosphingosine, and oleic acid. (*A*) ESI-QTOF MS/MS 2-Palmitoyl-sn-glycero-3-phosphocholine (PC 16:0) *m/z* 496, RT 6 min in the human swab sample and human skin tissue. The loss of the phosphocholine group (*m/z* 184, $[M+H]^+$) induces the release of the fatty acid chain (*m/z* 313). Identification is based on MS/MS fragmentation and METLIN database interrogation.

Legend continued on following page

(*B*) Database match to vitamin D3 (7-dehydrocholesterol) based on the METLIN database. ESI-QTOF MS/MS of *m/z* 385, RT 8.0 in the swab sample (*Top*), from the METLIN database (*Down*). Here we have a family that has similar fragmentation patterns to vitamin D3. Therefore, we call this the vitamin D3 molecular family. The identification is not 100% certain. This is the limitation of MS-based identification. MS cannot determine stereochemistry. (*C*) Putative match to glycocholic acid based on METLIN database. MS/MS of *m/z* 466 $[M+H]^+$, RT 4.2 min in the swab sample (*Top*), from the METLIN database (*Down*). This is a match to the glycocholic acid molecular family. (*D*) Database match to taurocholic acid based on METLIN database. Shown is ESI-QTOF MS/MS of *m/z* 516 $[M+H]^+$, RT 4.0 min in the swab sample (*Top*), from the METLIN database (*Down*). This is a match to the taurocholate molecular family. (*E*) ESI-QTOF MS/MS fragmentation and RT of phytosphingosine *m/z* 318, RT 5.2 min detected in the human swab sample (*Top*) and chemical standard (*Down*). (*F*) ESI-QTOF MS/MS fragmentation of cis-oleic acid *m/z* 283 $[M+H]^+$, RT 7.8 min in the swab sample (*Top*), *Pseudomonas aeruginosa* strain (*Middle*), and chemical standard (*Down*). We cannot differentiate between cis/trans nor is the location of the double bond identified. If a sufficient sample was available, ozonolysis could be used to determine the location of the double bond. These are the inherent limitations of MS-based ID. An asterisk represents a precursor ion.

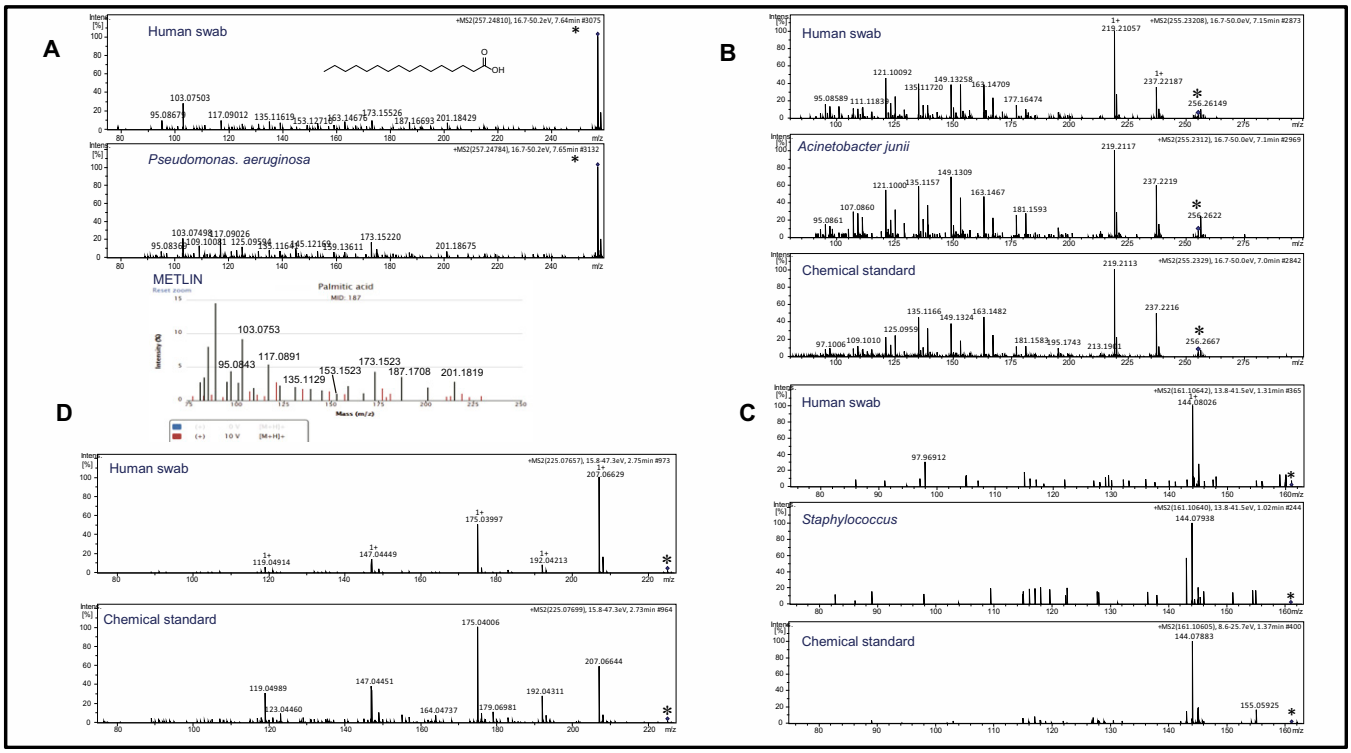


Fig. S4. ESI-QTOF MS/MS fragmentation of palmitic acid, palmitoleic acid, tryptamine, and sinapic acid. (*A*) Palmitic acid match based on METLIN database. Shown is the ESI-QTOF MS/MS of *m/z* 257 $[M+H]^+$, RT 7.6 min in the swab sample (*Top*), *P. aeruginosa* strain (*Middle*), and from the METLIN database (*Down*). Because there is no double bond, this match is good for a saturated fatty acid. (*B*) ESI-QTOF MS/MS of *m/z* 255 $[M+H]^+$, RT 7.1 min in the human swab sample (person 1, head), *Acinetobacter junii* ATCC 17908, and palmitoleic acid standard. (*C*) ESI-QTOF MS/MS of tryptamine *m/z* 161 $[M+H]^+$, RT 1.3 min in the swab sample, *Staphylococcus* culture, and chemical standard. This is a match to the tryptamine family. (*D*) ESI-QTOF MS/MS of sinapic acid *m/z* 225, RT 2.7 min in the human swab sample and in chemical standard. This is a match to the sinapic acid family. An asterisk represents a precursor ion.

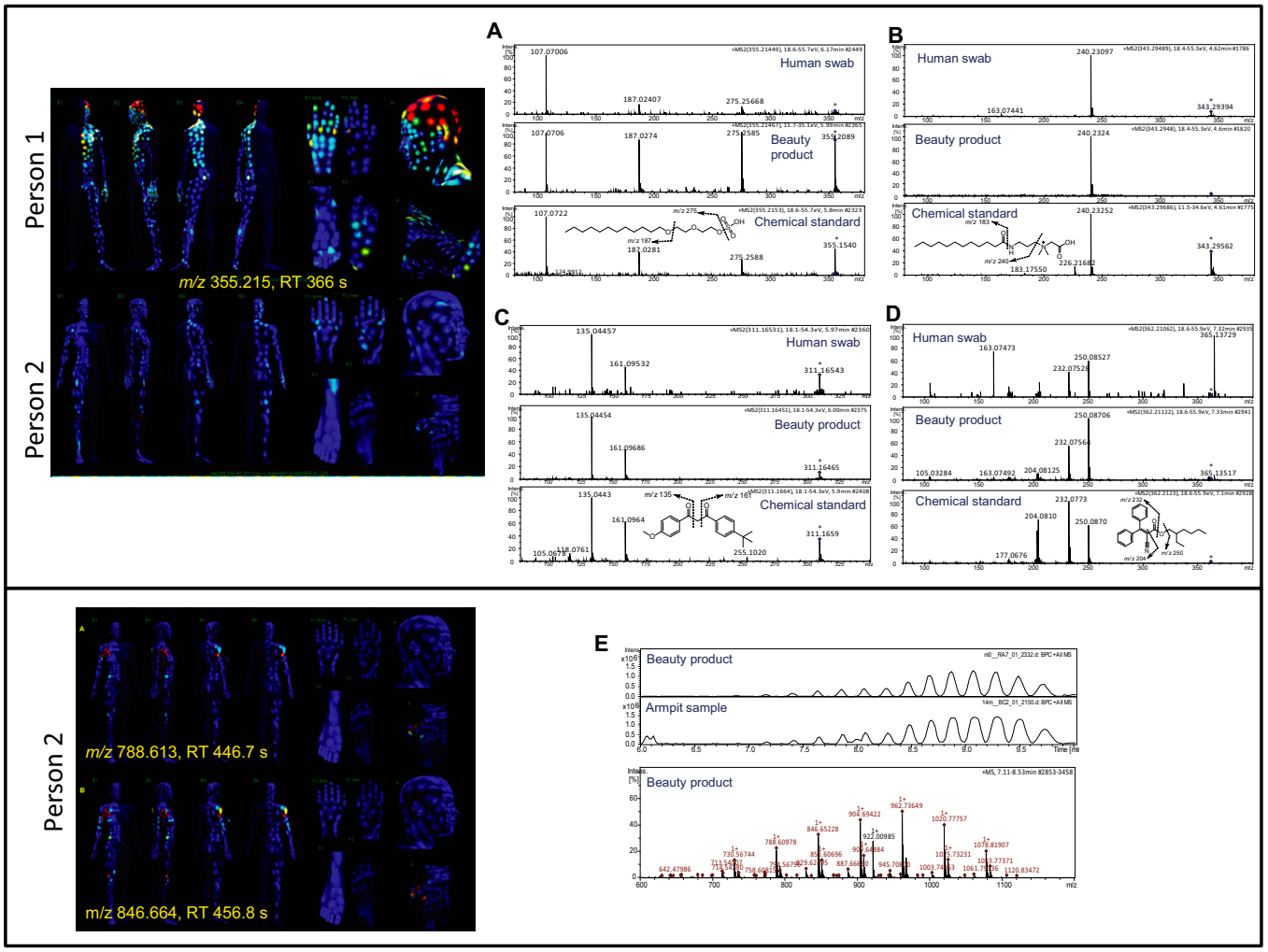


Fig. S5. ESI-QTOF MS/MS identification of beauty products and representative maps generated from UPLC-QTOF data. (*Upper Left*) Representative maps generated from UPLC-QTOF data, showing person 1 and person 2 distribution of *m/z* 355. The surfactant SLES was mainly localized on person 1's head. (*Upper Right*) ESI-QTOF MS/MS fragmentation of beauty products: sodium lauryl ether sulfate, cocoamidopropylbetaine, avobenzone, and octocrylene. (*A*) ESI-QTOF MS/MS fragmentation of C12-2EO SLES where *m/z* 355 $[M+H]^+$, RT 6.2 represents C12 alkyl chain SLES and contains 2 ethoxylated units (2EO), without sodium counterion (*m/z* 355), detected in the human swab sample (*Top*), beauty product (man's shampoo, *Middle*), and cosmetic ingredient SLES (*Down*). The main fragments correspond to desulfated C12-2EO (*m/z* 275) and fatty acid chain (*m/z* 187). It is the most common monomer of the SLES composition, containing 2 or 3 ethoxylated units. (*B*) ESI-QTOF MS/MS fragmentation of *m/z* 343 $[M+H]^+$, RT 4.6 detected in the human swab sample (*Top*), beauty product (person 2's shampoo, *Middle*), and cosmetic ingredient cocoamidopropylbetaine (C12-CAPB, *Down*). Dominant fragmentation product (*m/z* 240) corresponding to the cleavage of the bond between the methylene group and the quaternary nitrogen atom releases the fragment $[M+H–(N(CH_3)_2CH_2COO]^+$. (*C*) ESI-QTOF MS/MS fragmentation of avobenzone *m/z* 311 $[M+H]^+$, RT 6.0 min detected in the human swab sample (*Top*), beauty product (sunscreen lotion, *Middle*), and cosmetic ingredient avobenzone (*Down*). The alpha-cleavage of the keto form of avobenzone releases the aromatic groups *m/z* 135 and *m/z* 161. (*D*) ESI-QTOF MS/MS fragmentation of octocrylene *m/z* 362 $[M+H]^+$, RT 7.3 min detected in the human swab sample (*Top*), beauty product (sunscreen lotion, *Middle*), and cosmetic ingredient octocrylene (*Down*). Mainly, fragmentation corresponding to the protonated acid occurs at *m/z* 250 and acylium ion at *m/z* 232, which then eliminates CO to yield *m/z* 204. (*Lower Left*) Representative maps generated from UPLC-QTOF data, highlighting an exclusive person 2 armpit distribution of *m/z* 788.613, RT 446.7 s and *m/z* 846.664, RT 456.8 s. (*Lower Right*) Polymeric polypropylen glycol $(CH_2CH_2CH_2O)_n$ identification. (*E*, *Top*) Total ion chromatogram (RT range, 6–10 min) of beauty product (person 2's deodorant). (*E*, *Middle*) Armpit woman sample, highlighting polymer profile. A difference of 58 Da is observed in the average ESI-QTOF mass spectrum (RT range, 7.11–8.53 min), from the beauty product sample, characteristic of the polypropylene glycol subunit (*E*, *Down*). An asterisk represents a precursor ion.

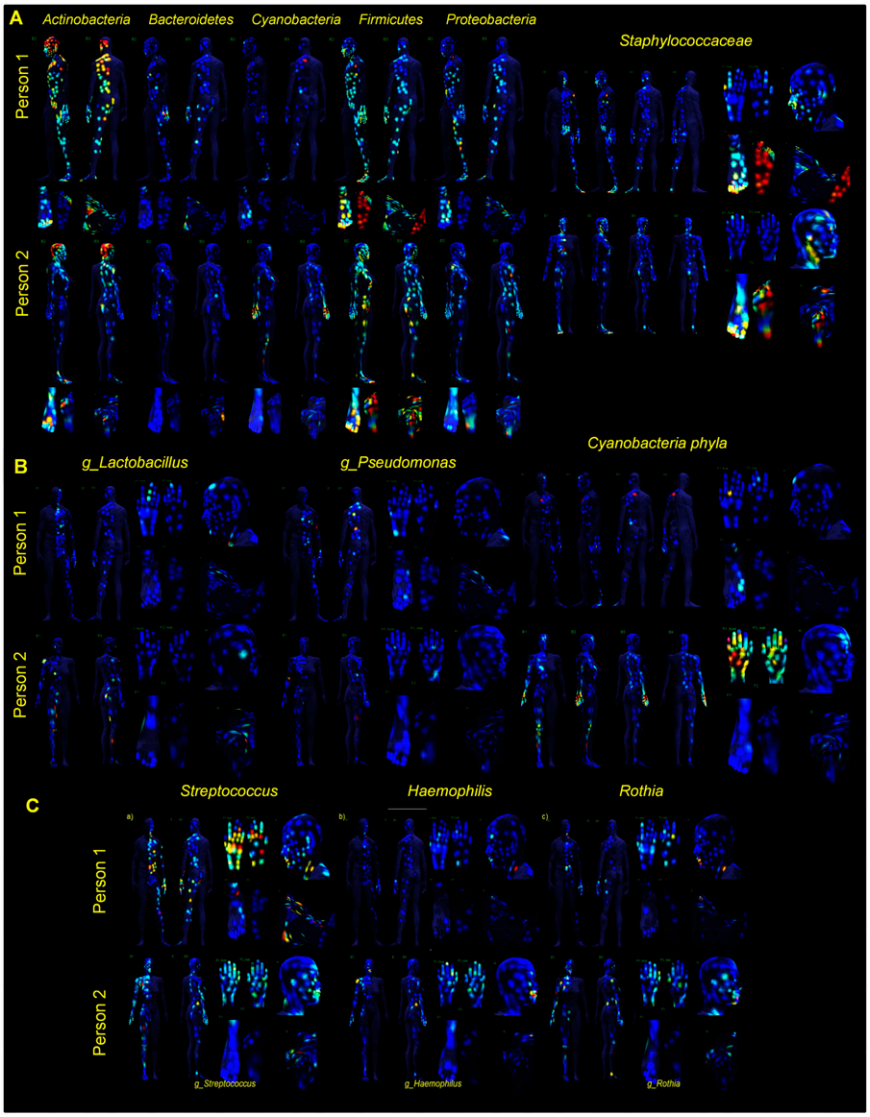


Fig. S6. The topographical maps of representative bacteria detected on the skin surface, at each body location, for person 1 and person 2. The color images, as heat maps, show red as the highest percentage of that phylum that was detected and blue as the lowest percentage; all other colors are between. No spot indicates that the specific phylum was not detected on that location. (*A*) Maps of representative microbial phyla detected on the skin surface, including *Actinobacteria*, *Firmicutes*, *Proteobacteria*, *Cyanobacteria*, *Bacteroidetes* (*Left*), and *Staphylococcaceae* family (*Right*). *Staphylococcaceae* was found in moist areas, such as the foot of both volunteers. (*B*) Person 1 and person 2 body distribution of genera *Lactobacillus and Pseudomonas* (*Left*) and chloroplast class from *Cyanobacteria* phyla (*Right*). High abundance of *Cyanobacteria* in person 2's hands, knees, and lower leg. (*C*) Person 1's and person 2's body distribution of *Streptococcus*, *Haemophilis*, *and Rothia*. They are present on the skin of both individuals, and they were present in larger quantities around the mouth of the male.

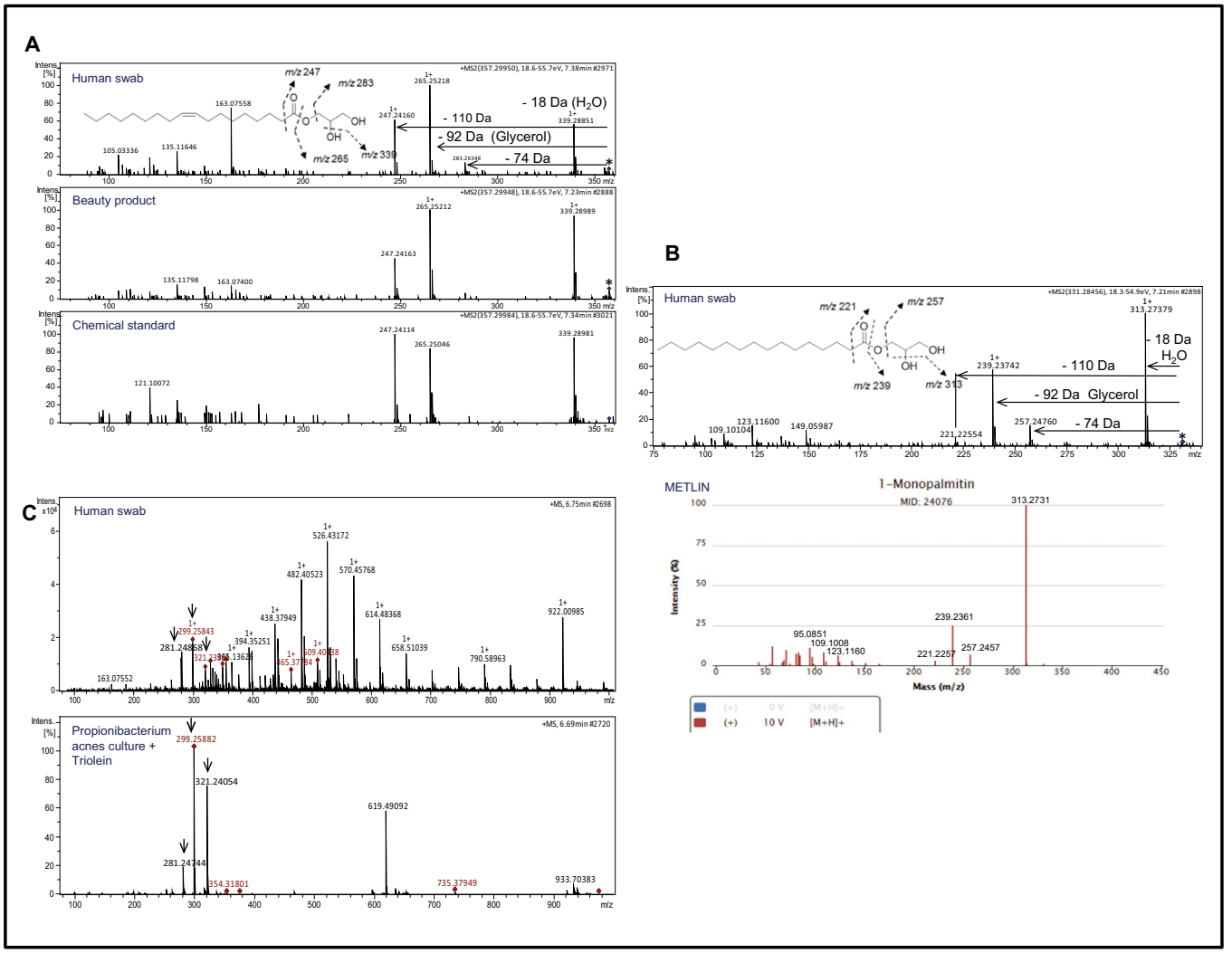


Fig. S7. ESI-QTOF MS/MS fragmentation of (*A* and *B*) monoacylglycerol compounds and (*C*) ESI-QTOF MS spectrum from the human swab sample and the *P. acnes* + triolein sample. (*A*) ESI-QTOF MS/MS fragmentation of monoolein *m/z* 357, RT 7.4 min in a swab sample and beauty product (person 2's shampoo) and chemical standard. Shown are three main fragments ions: *m/z* 283, *m/z* 265 released after the loss of the glycerol group (92 Da), and *m/z* 247 that corresponds to the fatty acid chain. MS/MS fragmentation of *m/z* 357 matches one of the monoolein in the METLIN database. MS/MS cannot differentiate the position of the double bond or the cis versus trans isomer. If a sufficient sample was available, ozonolysis could be used to determine the location of the double bond. Further, the position of the fatty acid on the glycerol backbone is not known with 100% certainty. This is an inherent limitation with ID by MS. (*B*) ESI-QTOF MS/MS fragmentation of 1-Monopalmitin in human swabs and from the METLIN database. One cannot differentiate the location of the fatty acid. (*C*) ESI-QTOF-MS spectrum from the human swab sample and *P. acnes* + triolein sample. Both spectra exhibit $[M+H]^+$ ion of *m/z* 299 as well as $[M+H\text{-}H_20]^+$ (*m/z* 281) and $[M+Na]^+$ (*m/z* 321) (red arrows) showing the presence of the oxidized oleic acid in both the swab sample and *P. acnes* culture containing triolein. An asterisk represents a precursor ion.

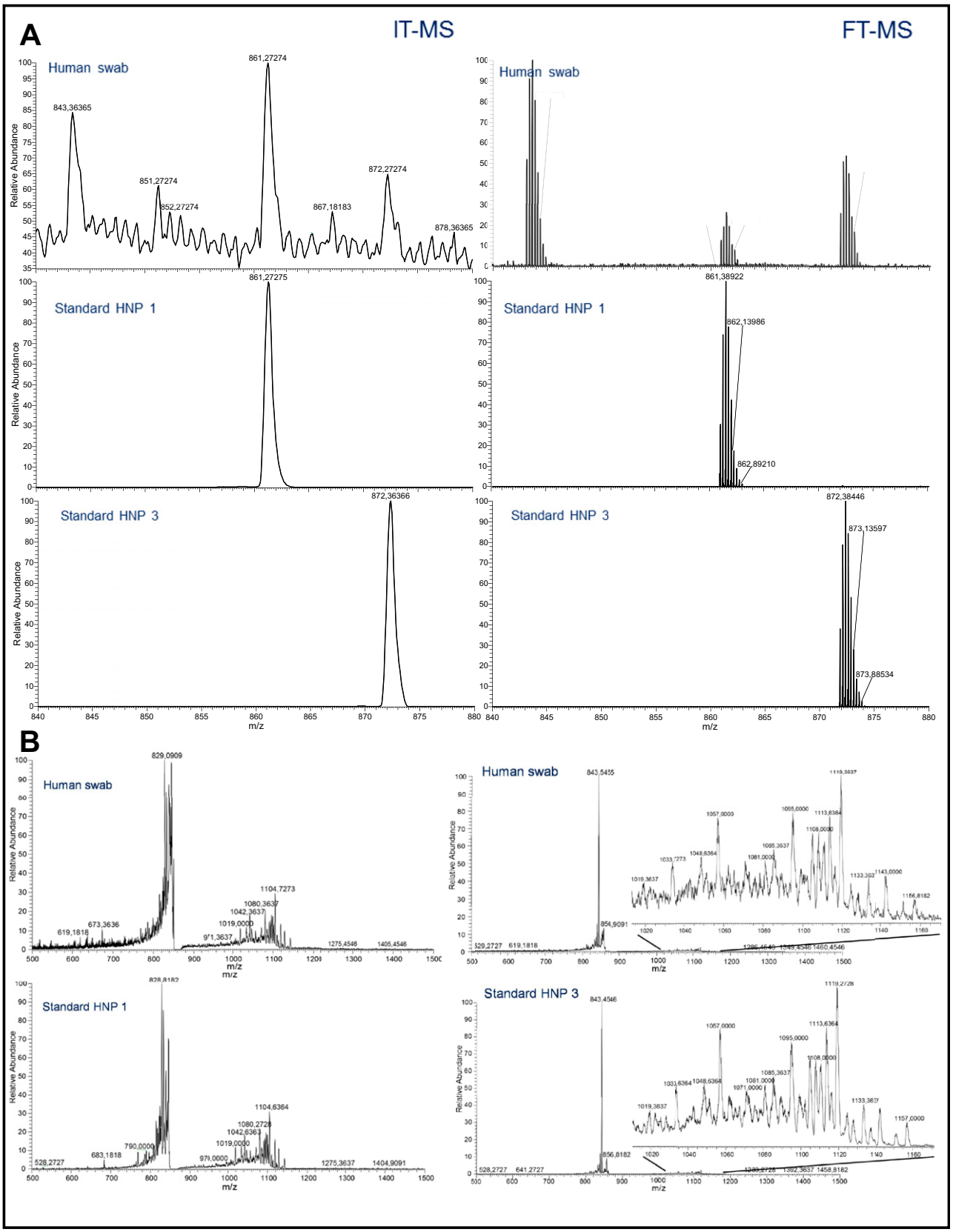


Fig. S8. Identification of HNP in human swab samples. (*A*) IT-MS (*Left*) and FT-MS (*Right*) spectra collected from the neck swab sample, standard HNP-1 *m/z* 861 $[M+4H]^{4+}$, and standard HNP-3 *m/z* 872 $[M+4H]^{4+}$. (*B*, *Left*) ESI-MS/MS fragmentation of *m/z* 861 $[M+4H]^{4+}$ in the swab sample and standard HNP-1 and (*B*, *Right*) ESI-MS/MS fragmentation of *m/z* 872 $[M+4H]^{4+}$ in the swab sample and standard HNP-3.

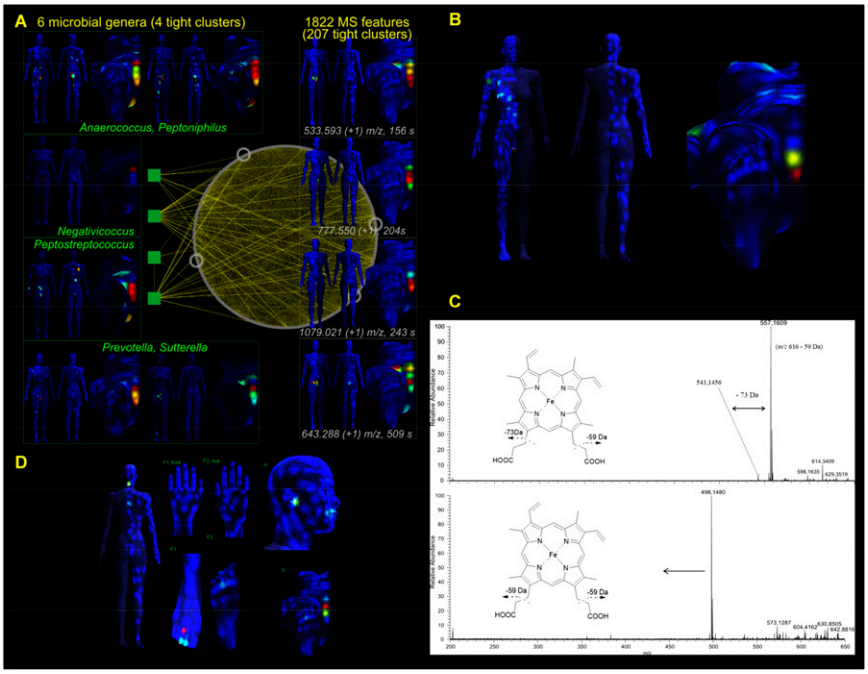
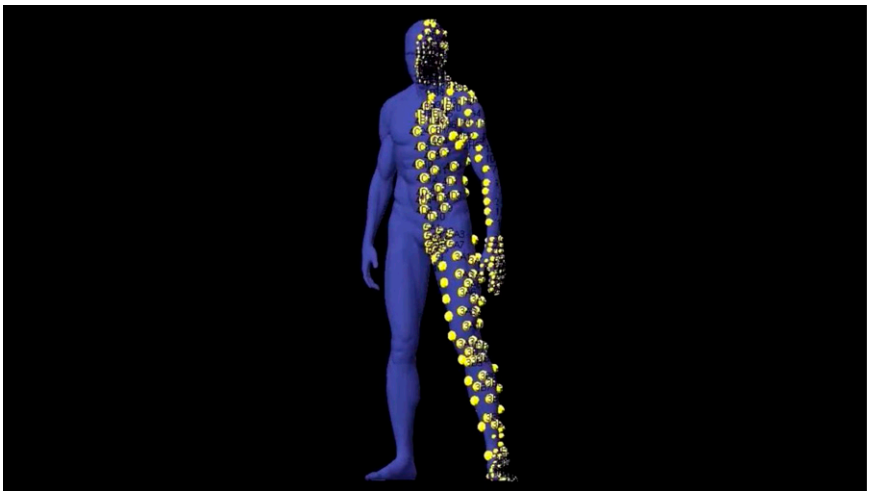
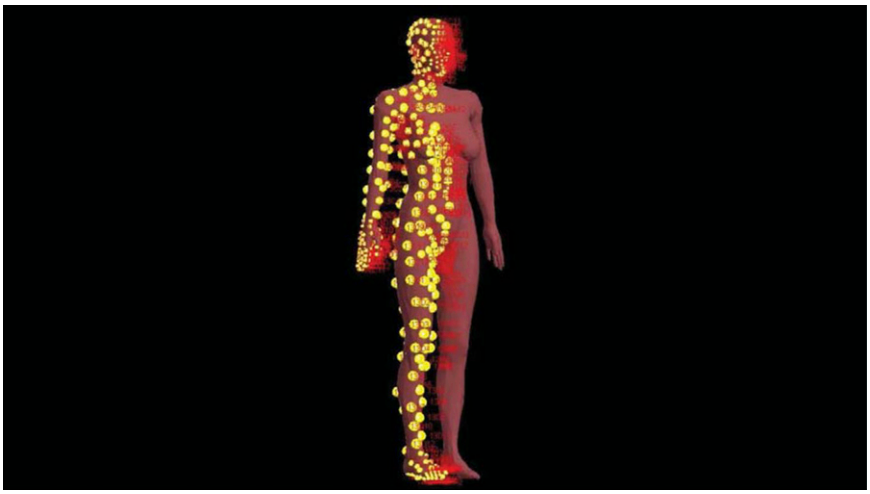


Fig. S9. Correlation between molecular and bacterial body distribution in person 2. (*A*) Groin molecular and bacterial communities in person 2. Colocalization analysis reveals more than 1,000 molecules (UPLC-QTOF data) and six bacteria genera (*Anaerococcus*, *Peptoniphilus*, *Negativicoccus*, *Peptostreptococcus*, *Prevotella*, and *Sutterella*) in the groin area of person 2. (*B*) Representative map constructed from MALDI-TOF data, showing person 2's body distribution of *m/z* 616 corresponding to the Heme. High abundance of Heme in the groin area of person 2. (*C*) Heme group identification in person 2 swabs using ESI MS/MS. MS/MS fragmentation of Heme *m/z* 616 $[M+H]^+$ (*Upper*) and MS/MS fragmentation of product ion *m/z* 557 $[M+H]^+$ (*Lower*). (*D*) Representative map generated from UPLC-QTOF data that shows the localization of Lyso-PC 16:1 (*m/z* 494.327, RT 327.6 s) in the groin area of person 2's body.



Movie S1. 3D_Person1_annotation.

Movie S1



Movie S2. 3D_Person2_annotation.

Movie S2

Dataset S1. 16S rRNA sequencing results

Dataset S1

Dataset S2. Person 1 sample list with body locations

Dataset S2

Dataset S3. Person 2 sample list with body locations

Dataset S3

Dataset S4. Person 1 and person 2 beauty product samples

Dataset S4

Dataset S5. List of chemical standards used for molecule identification

Dataset S5

Dataset S6. List of cultured human skin cells

Dataset S6

Dataset S7. List of cultured human bacteria and yeast strains

Dataset S7

Dataset S8. List of human skin tissues

Dataset S8

Dataset S9. MS/MS fragmentation parameters

Dataset S9

Dataset S10. LC-MS feature detection. (To open this file in TOPPAS/OpenMS, rename the file extension to .toppas.)

Dataset S10