Large-scale skin metagenomics reveals extensive prevalence, coordination, and functional adaptation of skin microbiome dermotypes across body sites

Chengchen Li[1,2*], Aarthi Ravikrishnan[1*], Indrik Wijaya[1*], Ahmad Nazri Mohamed Naim[1], Jean-Sebastien Gounot[1], Stephen Wearne[3], Yue Yuan On[1], Eliza Xin Pei Ho[1], Qi Yan Ang[1], Minghao Chia[1], Ng Hui Qi Amanda[1], Srinivas Ramaswami[3], Yik Weng Yew[4,5], Marie Loh[1,4,5,6], Ruo Yan Ong[3], Faith Ng[3], Sze Han Lee[3], James Chan[3,7], The HELIOS Study team, Viduthalai Rasheedkhan Regina[3,8], Thomas Dawson[3,9#], John Chambers[5#], John Common[3#], Niranjan Nagarajan[1,2#]

*[1]Genome Institute of Singapore (GIS), Agency for Science, Technology and Research (A*STAR), Singapore, Republic of Singapore*

[2]Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore

*[3]A*STAR Skin Research Labs (A*SRL), Agency for Science, Technology and Research (A*STAR) & Skin Research Institute of Singapore (SRIS), Singapore, Republic of Singapore*

[4]National Skin Centre, National Healthcare Group, Singapore, Republic of Singapore

[5]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Republic of Singapore

[6]Department of Epidemiology and Biostatistics, Imperial College London, South Kensington, London, UK

*[7]A*STAR Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), Singapore, Republic of Singapore*

[8]Singapore Centre for Environmental Life Sciences and Engineering, Nanyang Technical University, Singapore, Republic of Singapore

[9]Department of Drug Discovery, College of Pharmacy, Medical University of South Carolina, Charleston, SC, USA

**Joint First Authors*

[#]Corresponding Authors

Lead Contact: nagarajann@gis.a-star.edu.sg

Abstract

The human skin microbiome is increasingly recognized to influence skin health, immune function and disease susceptibility. However, large-scale, multi-site metagenomic studies in the general population remain scarce, limiting our understanding of the full diversity and variability of skin microbial communities across different individuals and body sites. Here, we present the largest-to-date skin microbiome dataset (>3,550 shotgun metagenomes across 18 sampling sites), generated using standardized cohort-wide protocols and rigorous contaminant removal. We define and characterize 17 robust, site-specific dermotypes—distinct microbial configurations—across seven out of nine body sites, whose coordination across disparate skin regions suggests systemic microbiome regulation. We also uncover striking bilateral concordance (median Pearson correlation>0.95 except for axilla) in microbial composition, further suggesting that host-intrinsic factors may contribute more than behavioral influences to shape skin microbiota. We trained near-perfect classifiers for dermotype labels for dermotype identification with as few as five signature species (AUC-ROC>0.98). Dermotype-dependent microbial interactions such as hypoxia-dependent inhibition of *S. hominis* by *S. epidermidis* and *M. luteus*, and metabolic resource utilization (e.g., galactose and histidine metabolism), reveal adaptive specialization within dermotypes, with implications for microbial assembly and host health. Notably, we associate a *S. hominis*-dominated elbow crease dermotype with skin irritation and itch, suggesting a potential link between microbial shifts and variations in barrier functions. By establishing dermotypes as a key organizing principle of skin microbial ecology, our findings provide a framework for microbiome-based prognostics and personalized therapeutic strategies.

Introduction

The human skin harbors a diverse community of bacterial, viral, and eukaryotic microbial species whose specific composition is known to vary substantially across body sites[1–4]. Within an individual and body site, community composition is however surprisingly stable over time[5]. Several factors could influence skin microbiome composition, including skin physiology[1,2], microbial ecology[6–9], lifestyle choices[10–13], and environmental influences[14–18], but the extent to which they shape microbiome heterogeneity across individuals within a population remains unknown.

The presence of distinct ecological community types is a key organizing principle and risk factor associated with host health, particularly in the gut and the vaginal tract[37–45]. For the gut microbiome, initial studies revealed three distinct so-called enterotypes—stable microbiome patterns within the gut—each characterized by the predominance of either *Bacteroides, Prevotella,* or *Ruminococcus* species*,* and further associated with long-term dietary patterns[37–39]. A recent study identified five entero-signatures (*Bacteroides, Firmicutes, Prevotella, Bifidobacterium,* and *Escherichia*), with combinations of these signatures accurately describing gut microbial composition in individuals across diverse human populations[22]. Similarly, vaginal microbiome studies have consistently identified five distinct community types dominated by various *Lactobacillus* species and other anaerobes (e.g., *Gardnerella*, *Prevotella,* and *Bifidobacterium*) and linked them with personal hygiene, lifestyle, and reproductive health factors[23–27].

Skin microbiome studies, in comparison, have primarily focused on disease-associated or single-site studies[17,18,28–30]. In previous work, we showed that atopic dermatitis patients have one of two distinct community compositions—so-called dermotypes—in the elbow crease, one of which is associated with more severe itch, higher IgE levels, and increased disease severity[28]. Similarly, a study on psoriasis identified a distinctive enrichment of *Firmicutes* and *Actinobacteria*-rich communities on the skin of affected individuals, independent of body site[29]. Finally, two distinct microbial communities—co-called cutotypes—on the face have been associated with air pollution exposure, and further linked to differences in skin physiological measurements and abundance of antimicrobial resistance genes[17,18]. Despite these initial studies, the heterogeneity of skin microbiome configurations across body sites in the general population has not been explored.

Here we address these knowledge gaps by comprehensively studying the skin microbiomes in 18 different sampling sites of a multi-ethnic cohort of 200 adults from the general population that is not focused on a specific disease. We captured a broad range of parameters by combining shotgun metagenomic sequencing data (n=3,600 samples), demographic information (e.g., age, ethnicity, etc.), skin physiological measurements (e.g., skin pH, hydration, etc.) and extensive skin health questionnaire data (e.g., skin care and skin discomfort/disease related questions). Analysis of body site-specific microbiome variability revealed strong bilateral concordance within most individuals, but substantial inter-host variability. We uncover the existence of distinct ecological configurations—dermotypes—in seven out of nine body sites, including five different dermotypes in the axilla. Each dermotype was found to contain a core set of taxa that were not necessarily core to its respective skin site, including species with median relative abundances <0.5%, together with ancillary species that have dermotype-specific blooms in some individuals (e.g., *Pseudomonas stutzeri*, *Acinetobacter pittii/calcoaceticus/nosocomialis*). Microbiome-intrinsic factors associated with dermotypes include differences in species co-abundance patterns, interspecies interactions, and microbial metabolic pathways within a site. Host factors such as demographics were found to be informative for dermotype prediction in only a few body sites (e.g., elbow crease, mean AUC-ROC=0.79), while intrinsic factors such as metabolic functions and dermotypes from other body sites typically had greater predictive power. We trained near-perfect classifiers for dermotype labels (AUC-ROC>0.98), with as few as five species in all cases, highlighting the potential for dermotype identification with just a few signature species. Overall, this work establishes the importance of accounting for population heterogeneity in skin microbiome investigations, and highlights the diverse intrinsic and extrinsic factors that can shape skin microbiome variability and its impact on host phenotypes. We further provide a clinically actionable framework for assigning skin microbiome states to individuals, potentially facilitating the development of stratification and personalized treatments.

Results

Analysis of inter-individual heterogeneity in the skin microbiome identifies robust dermotype states across multiple body sites

To capture skin microbiome heterogeneity across individuals and different body sites in the general population, we obtained a large collection of skin samples (n=3,600) from 18 different sampling

110 sites in 200 adults recruited from the Health For Life In Singapore (HELIOS) precision medicine
111 cohort[31] (**Figure 1A**, **Table 1**; **Methods**). The samples included the left and right sides of nine
112 skin sites: Scalp – Sc, Parietal scalp – Ps, Forehead – Fo, Cheek – Ch, Upper back – Ub, Axilla –
113 Ax, Antecubital fossa (elbow crease) – Ac, Volar forearm – Vf, and Lower leg – Ll (**Figure 1A**).
114 99.6% of the samples (3,586/3,600) were successfully and consistently processed for the
115 construction of metagenomic libraries and sequenced on an Illumina HiSeq platform to
116 generate >70 billion reads and >10Tbp of data in total (2×150bp reads, a median of 21.6 million
117 reads and 3.3Gbp of data per library). We processed the resulting metagenomic datasets to filter
118 reads for quality, identify the non-human fraction (median of 7.4 million reads and 1.1Gbp per
119 library), and conduct microbial taxonomic profiling. Data from negative controls (n=46;
120 **Supplementary File 1**) were integrated to identify and exclude contaminant taxa (**Supplementary**
121 **File 2**) and remove poor quality profiles (**Supplementary Figure 1**; **Methods**). Separately, we
122 collected extensive participant data to investigate associations between inter-individual skin
123 microbiome variations and host factors, including demographic information (age, gender, ethnicity,
124 etc.), skin physiological measurements (sebum level, pH, etc.), skin-care practice (hygiene,
125 personal-care product usage), and skin health data (skin discomfort and diseases; **Table 1**). This
126 collection provides the largest consistently processed skin metagenome dataset for a population-
127 wide cohort to date (n=3,572; **Supplementary File 3**), enabling in-depth comparisons of the skin
128 microbiome across individuals and analysis of its implication for skin health.

129 Leveraging data availability from the left and right sides of different body sites of many
130 individuals, we first investigated whether the skin microbiomes exhibited similarity between sides.
131 While several skin physiological properties, including sebum level[32] and trans-epidermal water
132 loss (TEWL)[33], are known to exhibit bilateral symmetry, human behaviors such as dominant hand
133 use or skin product application are less so. Limited previous work suggested that in a few
134 individuals[35] and some sites[34,28] bilateral consistency of microbial community composition was
135 observed. However, our data suggests this is a strong and general phenomenon across body sites
136 and for most individuals (median left-right Pearson correlation>0.95, except for axilla; **Figure 1B**).
137 At all nine body sites studied, skin microbiomes from the left and right sides of the body were
138 significantly more similar than across individuals (Wilcoxon test p-value<0.001; **Figure 1B**). This
139 pattern is robust and seen with the Yue-Clayton similarity index as well, which factors in shared
140 and non-shared species[36] (Wilcoxon test p-value<0.001, **Supplementary Figure 2A**). As we

141 found that intra-individual similarity is inversely correlated with community (within-sample)
142 diversity (Shannon index, **Figure 1B**), which in turn is influenced by *Cutibacterium acnes*
143 dominance[3], we reanalyzed the data after removing *C. acnes* from the profiles and confirmed that
144 bilateral concordance is a uniformly consistent property across body sites (Wilcoxon test p-
145 value<0.001, **Supplementary Figure 2B**; **Methods**). Despite the general concordance observed,
146 14% of sites (255/1779) exhibited discordance in 121 individuals across different body sites
147 (**Supplementary Figure 3A**; **Methods**). Intriguingly, bilateral discordance was enriched in a few
148 individuals (n=5, binomial test BH-adjusted p-value<0.05; **Supplementary Figure 3B**), and
149 multiple body sites exhibited correlated patterns of bilateral discordance (e.g., head sites and upper
150 back; **Supplementary Figure 3C**). Specific associations were detected between bilateral
151 microbiome discordance and various host attributes, including gender (cheek discordance in
152 women; **Supplementary Figure 3D**), personal habits (scalp discordance and daily hair covering;
153 **Figure 1C**), and skin conditions (forehead discordance and higher itch score; **Figure 1C**). Hence,
154 while the adult skin microbiome is typically strongly bilaterally concordant, discordance at specific
155 sites may be triggered by host-specific differences in the skin microenvironment arising out of
156 behavioral and/or skin condition changes.

157 Aggregating skin microbiome taxonomic profiles across body sites confirmed that skin
158 microbiome composition is influenced by skin microenvironment[1,2], with profiles from the same
159 site largely clustering together, and axilla (Ax) sites with their hypoxic and moist environments
160 being the most distinct (**Supplementary Figure 4A**). While sebaceous head and scalp sites (Sc,
161 Ps, Fo, and Ch) exhibited the highest similarity, even the two most similar scalp sites (Ps, Sc) had
162 statistically significant differences in beta (between-sample) diversity (PERMANOVA adonis
163 R^2=0.003, p-value<0.05; **Supplementary Figure 4B**). Site-specific differentially abundant
164 species included known enrichment in lipophilic *C. acnes* in sebaceous head sites[1], the fungus
165 *Malassezia globosa* in the elbow crease (Ac) and upper back[37] (Ub), and *Malassezia restricta* in
166 scalp sites, as well as new observations including *Cutibacterium avidum*, *Cutibacterium*
167 *granulosum*, and various *Staphylococcus* and *Corynebacterium* species in the axilla (consistent
168 with their preference for moist environments), and *Micrococcus luteus, Moraxella osloensis*, and
169 common oral taxa (*Streptococcus pneumoniae*, *Corynebacterium matruchotii*, and *Rothia*
170 *mucilaginosa*) in the forearm (Vf) and elbow crease (Ac; **Supplementary File 4**).

171 Despite these consistent differences across sites, we found that the individual skin
172 microbiomes from a given body site exhibit multimodal distributions across principal axes of
173 variation (e.g., Ac and Ax sites; **Supplementary Figure 4A**), suggesting substantial inter-
174 individual heterogeneity in the population. Exploiting the availability of a large number of
175 consistently processed taxonomic profiles for each site (n>395), we developed a systematic
176 clustering strategy to identify robust stratification in the data and assess stability of cluster
177 membership via bootstrapping (n=100; **Supplementary Figure 5**; **Methods**). Overall, seven out
178 of nine sites (i.e. Ax, Ac, Vf, Ub, Fo, Ps, Sc) showed clear evidence for multiple robust clusters
179 we term dermotypes, with all clusters having a median stability >90%. Of note, axilla sites
180 displayed the greatest diversity with five robust dermotypes, each represented by a high-density
181 region of samples (**Figure 1D**). Additionally, we identified two dermotypes at each of the
182 remaining sites except cheek and lower leg sites, where only a single robust cluster was detected
183 (**Supplementary Figure 6A-F**). In all seven sites with dermotypes identified, a *C. acnes*-enriched
184 dermotype with high prevalence (Ax-2, Ac-1, Vf-1, Ub-1, Fo-1, Ps-1, and Sc-1; **Figure 1E**) and
185 lower community diversity was observed (except Ac-1, Wilcoxon test p-value<0.05;
186 **Supplementary Figure 7A**). In the axilla, the remaining four dermotypes were dominated either
187 by combinations of *Corynebacterium* and *Dermabacter* species with higher α-diversity (Ax-1, Ax-
188 4) or two different *Staphylococcal* species (*Staphylococcus hominis* at Ax-3, *Staphylococcus*
189 *epidermidis* at Ax-5) with intermediate levels of community diversity. Intriguingly, the elbow
190 crease exhibited a *Staphylococcus*-dominant dermotype (Ac-2) seen in >15% of the general
191 population, similar to our previous observation in atopic dermatitis patients[28], while scalp sites
192 (Sc, Ps) exhibited an *M. restricta* enriched dermotype in ~30% of the population, with its known
193 association with dandruff[38] (**Supplementary Figure 7B**). In other body sites, while *M. luteus* and
194 *M. osloensis* were the most abundant species after *C. acnes* in the less-prevalent dermotypes, a
195 broader set of *Corynebacterium*, *Dermabacter, Streptococcus*, *Finegoldia*, *Brevibacterium,* and
196 *Rothia* species were found to be differentially abundant, highlighting their potential contributions
197 to dermotype structure and function (**Supplementary Figure 7B**; **Supplementary File 5**).
198 Species-level relative abundances enabled the training of near-perfect classifiers for dermotype
199 labels, with as few as five species in all cases, highlighting the potential for dermotype
200 identification with just a few signature species (AUC-ROC>0.98; **Supplementary Figure 7C**;
201 **Supplementary File 6**).

202 Building on the observation that skin microbiomes are bilaterally concordant, we
203 investigated whether dermotypes across different body sites exhibited consistent associations
204 within individuals (**Figure 1F**). Indeed, remarkably strong associations were detected between
205 dermotypes across various body sites, particularly those between physically proximal sites (e.g.,
206 scalp and parietal scalp), those with similar physiological characteristics (e.g., scalp and forehead),
207 but also more distant sites with different physiology (e.g., dry volar forearm and sebaceous
208 forehead sites; Fisher's exact test adjusted p-value<0.05; **Figure 1F**). These associations do not
209 seem to be fully explained just by preferences for specific species; for example, *Staphylococcus*-
210 dominant dermotypes in the elbow crease and axilla were negatively associated with each other
211 (**Supplementary Figure 8A**), even though *Staphylococcus* strains are known to be transmissible
212 via skin contact[39–41]. Furthermore, no strong correlations were observed for the abundance of
213 *Staphylococcus* species between these sites (**Supplementary Figure 8B**), indicating that other
214 systemic factors play a role in the observed dermotype associations. In summary, distinct
215 dermotype configurations exist both across individuals and body sites. Within individuals, these
216 exhibit high levels of bilateral and cross-site concordance, likely as a complex reflection of shared
217 host, environmental, and microbial factors that we explore next.

218 **Dermotype-specific core and ancillary species exhibit coordinated patterns across body sites**

219 To gain deeper insights into dermotype composition, we categorized species into dermotype-
220 specific core (with ≥50% prevalence) and ancillary (<10% prevalence, ≥3 occurrences) species
221 (**Figure 2A**; **Methods**). Across all dermotypes, this identified 15 core bacterial species from the
222 genera *Cutibacterium*, *Staphylococcus*, *Corynebacterium*, *Micrococcus,* and *Moraxella*, and two
223 core fungal species from the genus *Malassezia* (**Figure 2B**). No archaea or viruses were core
224 among the 560 species analyzed, emphasizing their sporadic distribution in the human skin
225 microbiome[5]. Each dermotype contained 2-12 core species, collectively accounting for >70% of
226 community abundance, with *C. acnes* being a universal core species across all dermotypes and
227 sites. Most taxa were, however, core in only some sites or dermotypes, e.g., *Dermabacter sp.*
228 *HFH0086,* which was categorized as a core species in the most common dermotype in the axilla
229 (Ax-1, prevalence>77%) but rarely seen in the other dermotype (Ax-2, prevalence<2%; **Figure**
230 **2B**). Even abundant species such as *S. epidermidis* and *S. hominis* were not universally core, while
231 some taxa with low relative abundances (median abundance<0.5%) were core in some dermotypes
232 (e.g., *Corynebacterium tuberculostearicum*, *Corynebacterium pseudogenitalium*, *Staphylococcus*

233 *caprae/capitis,* and *M. globosa*). This co-existence of high- and low-abundance core species within
234 communities suggests that their collective maintenance may arise from niche partitioning or
235 complementary functional roles, independent of an abundance effect[42] (**Figure 2B**).
236 Building on our observation of intra-individual dermotype correlation between different
237 body sites (**Figure 1F**), we analyzed intra-individual correlations in core species abundances
238 across the nine skin sites. Several taxa, such as *M. restricta*, *S. epidermidis,* and *C. acnes*, exhibited
239 modest positive correlations when aggregated across all sites spanning diverse microenvironments
240 (mean ρ=0.26), and higher correlations when excluding a few sites (e.g. *M. restricta* excluding Ax,
241 mean ρ=0.37; **Supplementary Figure 9A-D**). This suggests that core species shared across skin
242 sites are systemically coordinated within individuals, potentially mediated by global sebum
243 secretion favoring lipophilic taxa, by horizontal gene transfer disseminating antibiotic resistance
244 genes and facilitating strain dispersal[43], or by immune-driven niche filtering[44,45]. In contrast, *S.*
245 *hominis* showed positive correlations limited to the non-scalp region (**Supplementary Figure 9E**),
246 while *M. globosa* displayed divergent correlation patterns, with positive associations among facial
247 sites (ρ=0.28) but negative correlations between the cheek and upper back (ρ=-0.25;
248 **Supplementary Figure 9F**). These heterogeneous patterns imply that systemic host drivers might
249 interact with localized ecological factors, such as interspecies competition and
250 microenvironmental variation (e.g., pH, humidity gradients), further modulating cross-site
251 microbial assembly patterns.
252 The 25 ancillary species we identified were disproportionally abundant at various sites and
253 dermotypes (in some cases >50% relative abundance) despite being, by definition, not overall very
254 prevalent (<10%; **Figure 2C**; **Methods**). This group includes bacteria primarily from the
255 *Pseudomonas*, *Acinetobacter*, and *Gordonia* genera (often environmentally resilient and biofilm
256 forming), along with bacteriophages of common skin bacteria (e.g., *Cutibacterium*,
257 *Staphylococcus*), common skin viruses (e.g., polyomavirus and human papillomavirus), and a
258 fungus (*Malassezia slooffiae*; **Figure 2C**). Nearly half of ancillary taxa (12/25=48%) showed
259 specificity to a particular dermotype; for example, *Acinetobacter pittii/calcoaceticus/nosocomialis*
260 was disproportionally abundant (median abundance>50%) only in the less prevalent forehead
261 dermotype Fo-2, indicating that ancillary taxa may be adapted to specific niches available only in
262 some individuals and microbiome configurations. This was further seen in the form of significant
263 enrichment of ancillary species across multiple body sites of many individuals (≥2 sites in 87

264 individuals, Poisson binomial test BH-adjusted p-value<0.05; **Figure 2C**, **Supplementary Figure**
265 **10A**). A striking example of this is *polyomavirus HPyV6,* which was abundantly present (>10%)
266 in many skin sites of some individuals (n=3) despite being lowly abundant in the general
267 population (<0.1%; **Figure 2D**). Interestingly, we also detected co-occurrence patterns in these
268 "blooms" of ancillary species across various body sites and species (**Supplementary 10B**),
269 including statistically significant coordination between head and limb sites for *polyomavirus*
270 *HPyV6* (cheek and lower leg, scalp and forearm, Fisher's exact test p-value<0.05; **Figure 2E**). In
271 some cases, more frequent carriage of an ancillary species across body sites was also associated
272 with host attributes including ethnicity (Chinese and *Malassezia slooffiae*) and skin phenotypes
273 (itch severity and *Betapapillomavirus 2;* adjusted p-value<0.05; **Supplementary Figure 10C**).
274 These results highlight the dermotype-specific nature of core and ancillary species, their
275 coordinated distribution across body sites within individuals, and their unique associations with
276 host attributes.

277 **Species co-abundance patterns reveal dermotype-dependent microbial interactions**

278 To understand how dermotypes emerge at various skin sites, we analyzed microbial correlated
279 abundance (co-abundance) patterns with a view to gaining insights into resource competition and
280 metabolic cross-feeding interactions (**Methods**). The scale of our dataset enabled robust
281 identification of >100 species-level co-abundance relationships across the nine body sites and 16
282 distinct dermotypes (**Supplementary File 7**). These include known antagonism between *C. acnes*
283 and *S. epidermidis*[46–48], and cross-feeding interactions between *Corynebacterium* and
284 *Staphylococcal* species[49] (|r|≥0.3, CCREPE q-value<0.05; **Figure 3A**; **Supplementary Figure 11**).
285 Systematic comparison of co-abundance networks and hub species—highly connected taxa—
286 across dermotypes from the same site showed that while a core set of relationships are conserved,
287 8-48% of co-abundance relationships differed significantly between dermotypes, with highest
288 differences in the axilla and lowest in parietal scalp and forearm (Fisher's Z-test p-value<0.05;
289 **Supplementary Figure 12A**). For example, *C. acnes* and *M. luteus* abundances displayed a strong
290 negative correlation in the Ac-1 dermotype of the elbow crease, but positive correlation values
291 were obtained in the Ac-2 dermotype (**Figure 3A**).

292 This unexpected pattern extended to core species (significant changes in 10-68% of co-
293 abundance relationships; **Supplementary Figure 12B**), with *C. acnes* in particular frequently
294 exhibiting switched relationships across dermotypes (**Supplementary Figure 12C**). Given that *C.*

295 *acnes* is an aerotolerant anaerobe, its variable abundance and metabolism may reflect local oxygen
296 gradients. Therefore, we hypothesized that oxygen availability strongly (negatively) impacts *C.*
297 *acnes* growth and through its antagonistic interactions with other skin microbes plays a role in
298 shaping dermotype structure across many body sites. To explore this, we systematically conducted
299 pairwise antagonism assays among a subset of core skin bacteria (*C. acnes*, *S. hominis*, *S.*
300 *epidermidis* and *M. luteus*) under hypoxic, aerobic and mixed conditions (n=186 experiments, with
301 triplicates; **Supplementary File 8**; **Methods**). As expected, *C. acnes*, exhibited poor growth in
302 aerobic conditions but thrives under hypoxic conditions, while *M. luteus* (a strict aerobe) exhibited
303 the opposite pattern. Consistent with prior reports[50], we observed partial inhibition of *M. luteus* by
304 *S. epidermidis* (**Supplementary File 8**). Interestingly, we observed frequent oxygen-dependent
305 interspecies inhibitions, with *S. hominis*, a dominant species in the Ac-2 dermotype, being strongly
306 inhibited by *S. epidermidis* and *M. luteus* under hypoxic but not aerobic conditions (**Figure 3B**,
307 **Supplementary Figure 12D-E**). Further testing showed that the agar became alkaline (pH>7) in
308 aerobic conditions but turned acidic under hypoxic conditions (**Supplementary File 8**), indicating
309 that hypoxia-dependent inhibition could be mediated by acidification or fermentation-derived
310 short-chain fatty acids[51] that are unfavorable for *S. hominis* growth. Our findings suggest that the
311 common elbow crease dermotype (Ac-1) may represent a hypoxic niche where *C. acnes* flourishes,
312 but *S. hominis* remained suppressed due to inhibition by other skin species. In contrast, the less
313 common elbow crease dermotype (Ac-2) could provide a more aerobic environment enabling *S.*
314 *hominis* to evade inhibition and dominate, especially with reduced growth of *C. acnes*. Consistent
315 with this hypothesis, we noted that while *S. epidermidis* is a prevalent hub species in the Ac-1
316 dermotype, it is no longer so in the Ac-2 dermotype and is instead replaced by *S. hominis* as a hub
317 (**Figure 3C**), supporting the notion that *S. hominis* is more well-adapted to the environment of the
318 Ac-2 dermotype.

319 Further analysis of hub nodes identified key species such as *C. pseudogenitalium* and *C.*
320 *tuberculostearicum* in 4 out of 12 networks, highlighting that even low-abundance community
321 members can play a major role in determining community structure (**Figure 3C**). Lastly, we noted
322 that several bacterial species typically found as part of commensal oral flora were network hubs
323 for specific dermotypes in the forearm, upper back, and forehead (**Figure 3C**). Further
324 investigation of their species correlations revealed a guild of co-abundant oral bacteria such as
325 *Veillonella parvula*, *Rothia dentocariosa*, and *Corynebacterium matruchotii*, that were seen in

326 several dermotypes and were strongly co-abundant in some (e.g. Ac-1 & Vf-2; **Figure 3A**,
327 **Supplementary Figure 11**). These species are known to form a "hedgehog" structure in the
328 supragingival biofilm, which regulates pH in response to changes in nutrient availability associated
329 with food intake[52,53]. While the co-abundance patterns in these dermotypes are reminiscent of oral
330 interactions, it remains unknown whether they perform similar pH-stabilizing functions, act as
331 transient co-colonizers, or potentially exacerbate pathogenicity on skin—analogous to the role of
332 oral microbes in gastrointestinal diseases when translocating to the gut[54,55]. Overall, the analysis
333 of species co-abundance networks, supported by the scale of our datasets and experimental data
334 from pairwise antagonism assays, revealed dermotype-specific inter-species interactions and
335 switches in hub species that may be a function of the aerobicity of the environment and likely play
336 an essential role in dermotype assembly.

337 **Metabolic variations between dermotypes reflect distinct host microenvironments and**
338 **influence on skin phenotypes**

339 Based on the observation that microbes residing on skin adapt to limited nutrient availability in
340 the form of carbohydrates, lipids, and proteins present in sweat, sebum, and the stratum
341 corneum[3,56], we hypothesized that differences in metabolic function between dermotypes of the
342 same body site could provide insights into how they assemble and interact with the host.
343 Systematic comparisons of metabolic pathways present in skin metagenomes (**Supplementary**
344 **File 9**) highlighted significant differences in pathway diversity across dermotypes
345 (**Supplementary Figure 13A; Supplementary File 10**), with a large fraction of pathways (>35%)
346 being differentially abundant in at least one dermotype of a body site, particularly relating to lipid,
347 carbohydrate, amino acid, nucleotide, and energy metabolism, as well as vitamin and cofactor
348 synthesis (**Supplementary Figure 13B**). We further observed a suite of metabolic pathways that,
349 while consistently present across distinct dermotypes of a given site, vary significantly in
350 abundance. For example, enrichments in arginine, tryptophan, and branched-chain amino acid
351 biosynthesis (e.g., L-isoleucine, L-valine; **Supplementary Figure 14**), that suggest potential
352 microbial compensation for amino acid auxotrophy under nutrient-limited conditions[57–59].

353 Focusing on key differences in microbial functions for the degradation of specific sugars
354 and amino acids, and oxidation of fatty acids, highlighted the strong enrichment of lactose and
355 galactose degradation pathways in the elbow crease dermotype Ac-2 (**Figure 3D**). Galactose is a
356 common component of human sweat[60] and complex stratum corneum glycoconjugates where it is

357 released by corneodesmosome degradation[61], and it is absorbed by skin commensals via diverse
358 phosphotransferase systems[62]. We therefore hypothesized that inter-individual differences may
359 play an important role in dermotype structure. Indeed, our data indicates that specific
360 *Staphylococcus* species, particularly *S. hominis*, contribute strongly to the enrichment of this
361 pathway (LACTOSECAT-PWY, **Figure 3E**), potentially due to a competitive advantage under
362 these conditions. To test this hypothesis, we compared growth kinetics of *S. hominis* and *S.*
363 *epidermidis* in mono- and co-culture, under baseline (BHI) and galactose-supplemented conditions
364 (**Supplementary Figure 15A**). While galactose enhanced overall growth in both species, *S.*
365 *hominis* exhibited disproportionally greater fitness gains in co-culture (>100% relative to *S.*
366 *epidermidis*; **Supplementary Figure 15B**), suggesting that preferential galactose metabolism by
367 *S. hominis* can provide it a competitive edge and that host-derived metabolites could shape
368 dermotype structure.

369 Similarly, histidine, a product from epidermal filaggrin breakdown and a common amino
370 acid detected on human skin[63], can be degraded by skin microbes[64]. We observed significant
371 variations in the L-histidine degradation pathway across dermotypes (**Supplementary Figure 14**).
372 For example, the *Cutibacterium* dominant dermotype Ac-1 in the elbow crease exhibited
373 enrichment in L-histidine degradation, with *M. luteus* being a key contributor (HISDEG-PWY;
374 **Figure 3F**; **Supplementary Figure 16A**). Further analysis of metagenomic and metabolomic data
375 collected at the elbow crease from a separate cohort (n=100; **Methods**), showed significant
376 differences in histidine derivatives, including enrichment for cis-urocanic acid (cis-UCA) and
377 depletion for trans-UCA in Ac-1 subjects (**Supplementary Figure 16B**), consistent with more
378 effective degradation of trans-UCA by *M. luteus* (urocanate hydrolase; **Supplementary Figure**
379 **16A**). *M. luteus* abundance on skin exhibited marginal correlation with the levels of histidine and
380 its derivates (ρ=0.21; **Supplementary Figure 16C**), as well as a notable growth advantage with
381 L-histidine supplementation *in vitro* (OD=1.8 vs 1.5; **Supplementary File 12**). Intriguingly, we
382 observed a negative correlation between lactose metabolism and histidine catabolism pathways
383 specific to Ac-2 subjects, potentially reflecting niche-specific usage of these energy sources
384 (**Supplementary Figure 16D**). Overall, enrichment of the histidine degradation pathway was
385 negatively associated with skin sensitivity (p-value<0.05; **Supplementary Figure 16E**;
386 **Supplementary File 11**), consistent with the known immunosuppressive effects of the histidine
387 derivative cis-urocanic acid[65]. These findings therefore highlight that dermotypes, beyond

388 representing mere taxonomic differences, are niches with distinct metabolic properties that can be
389 shaped by differences in host-derived nutrients, and in turn may impact host skin conditions.

390 **Dermotype composition is linked to demographic factors, along with physiological,**
391 **behavioral, and skin condition attributes**

392 Skin microbial composition can be shaped by a combination of endogenous and exogenous host
393 factors (e.g., genetics, skin physiology, lifestyle, and skin conditions) and in turn, influences some
394 of them. We thus leveraged the scale of our cohort to evaluate associations between dermotypes
395 and host demographics, skin physiological measurements, >15 personal care behaviors, and >25
396 skin conditions (**Methods**). Focusing on attributes relevant to each body site, we identified >35
397 significant associations (**Supplementary File 13)**.

398 Among demographic factors, ethnicity exhibited strong associations with axilla and elbow
399 crease dermotypes. In particular, the axilla dermotypes dominated by *C. acnes* (Ax-2), *S. hominis*
400 (Ax-3), and *C. kroppenstedii* (Ax-4) were significantly enriched in ethnic Malay, Chinese, and
401 Indian subjects, respectively (adjusted p-value<0.05; **Figure 4A**). In contrast, the *S. hominis*-
402 enriched elbow crease dermotype (Ac-2) was more frequent in Malay and less common in Chinese
403 subjects (adjusted p-value<0.05; **Figure 4A**). Beyond ethnicity, the *S. hominis*-enriched elbow
404 crease dermotype (Ac-2) and the *Cutibacterium*-depleted upper back dermotype (Ub-2) were
405 significantly associated with female and older subjects (adjusted p-value<0.05). Strikingly, no
406 significant demographic associations were observed for head sites (forehead, scalp, parietal scalp;
407 **Supplementary File 13**), indicating that other factors (e.g. localized skincare practices or
408 environmental exposures) may drive microbiome variation in the head region.

409 We next assessed if skin physiological measurements revealed distinct trends across
410 dermotypes. Of note, individuals with an *S. hominis*-enriched elbow crease dermotype (Ac-2)
411 exhibited lower forearm surface hydration (p-value<0.05; **Figure 4B**), consistent with our prior
412 report of dry skin in atopic dermatitis (AD) patients harboring a *Staphylococcus*-enriched
413 dermotype[28], and indicating that Ac-2 may arise from compromised skin barrier function.
414 Additionally, elevated cheek sebum levels, a proxy for overall sebum production in the body, were
415 observed in individuals with *C. acnes*-enriched dermotypes at the elbow crease and the volar
416 forearm (i.e., Ac-1, Vf-1), consistent with the robust growth of *C. acnes* in sebum-rich
417 environments. In contrast, higher cheek sebum levels were associated with the *C. kroppenstedii*-
418 enriched dermotype in the axilla (i.e., Ax-4; p-value<0.05; **Figure 4B**). While both species are

419 lipophilic[66], this divergence may reflect differences in local lipid composition and species-specific
420 substrate preferences. More broadly, to explore whether site-specific skin physiological
421 measurements reflect body-wide intrinsic states, we conducted site-agnostic analyses. Intriguingly,
422 we observed strong associations between lower trans-epidermal water loss (TEWL) levels on the
423 cheek and *C. acnes*-depleted dermotypes in the upper back (Ub-2) and forehead (Fo-2; p-
424 value<0.05; **Supplementary Figure 17A**). Similarly, lower TEWL on the forearm associated with
425 the *C. acnes*-depleted parietal scalp dermotype (Ps-2; p-value<0.05; **Supplementary Figure**
426 **17A**), highlighting the systemic connections between skin barrier functionality and skin microbial
427 community structure.

428 Leveraging skincare routine information based on questionnaire data, several noteworthy
429 associations with dermotypes were observed. For example, dermotypes in the axilla were found to
430 frequently exhibit statistically significant associations with underarm product use, even after
431 accounting for demographic factors. For example, routine deodorant use was positively associated
432 with Ax-2 (~60% prevalence), while antiperspirant use was linked to Ax-4 (~30% prevalence);
433 conversely, both product uses were negatively correlated with the most common axilla dermotype
434 Ax-1, where deodorant and antiperspirant use were notably lower (11% and 4.9%, respectively; p-
435 value<0.05; **Figure 4C**). Consistent with the antimicrobial effects of these products, Ax-2 and Ax-
436 4 exhibited reduced alpha diversity compared to Ax-1 (**Supplementary Figure 7A**). Similarly,
437 routine moisturizer use and hair treatments correlated with the *C. acnes*-depleted forehead
438 dermotype (Fo-2), while daily hair covering was associated with the *C. acnes*-depleted parietal
439 scalp dermotype (Ps-2), indicating that skincare practices can significantly influence facial,
440 parietal scalp, and axilla microbiomes.

441 Several skin condition-related attributes, including sensitivity, itch, and eczema history,
442 were associated with the *Staphylococcus*-enriched elbow crease dermotype while adjusting for
443 demographic factors (Ac-2; p-value<0.05; **Figure 4C**), consistent with our previous atopic
444 dermatitis-focused study[28]. While skin sensitivity and itch were more common in those with
445 eczema history, Ac-2 subjects showed more severe irritation and itch, even in the absence of
446 eczema (**Supplementary Figure 17B-C**). Treatment of human keratinocytes with supernatants
447 from *S. hominis* strains revealed that several of them (5/8; as well as *S. epidermidis* strains: 9/12)
448 induced significantly elevated pro-IL-33 levels (cytokine signal in itch[67,68]) compared to controls
449 (**Supplementary Figure 17D**), suggesting that enrichment of *Staphylococcus* species on skin (as

450 is seen in the Ac-2 dermotype) can play a role in modulating host immune signaling, heightened
451 skin irritation and itch.

452 Finally, we applied machine learning models to test whether the information in the
453 demographic data, skin physiological measurements, and questionnaire responses could predict
454 dermotype labels across body sites (**Supplementary Figure 18A**; **Methods**). Machine learning
455 models that were significantly better than random were obtained in all sites except for scalp sites,
456 with the highest classification performance observed in the elbow crease (Ac random forest AUC-
457 ROC=0.8; **Figure 4D**). Among various host attributes, demographic factors emerged as the most
458 predictive in four out of seven sites analyzed (axilla, scalp, upper back, elbow crease), with
459 ethnicity frequently being among the most important features (e.g. for elbow crease; **Figure 4D**,
460 **Supplementary Figure 18B**). In summary, our findings underscore the intricate interplay between
461 host demographics (particularly ethnicity), skin physiology, skincare practices, and skin microbial
462 composition, suggesting that dermotypes could serve as biomarkers for skin health, disease risk
463 and therapeutic eligibility in precision medicine.

464 **Metabolic functions and dermotypes across body sites are more strongly predictive of**
465 **dermotypes**

466 Having shown that certain host attributes can predict dermotype occurrence, we wondered if
467 metabolic functions could serve this purpose. To ensure predictions were not driven by species-
468 specific pathway signatures, we excluded pathways highly correlated with taxa abundances
469 (Spearman $\rho \geq 0.9$, p-value<0.05) and trained multiple classifiers to explore key metabolic
470 functions that could jointly predict dermotype structure; among them, ridge classifiers achieved
471 the highest performance for many sites (AUC-ROC>0.95, except for Ps and Sc; **Figure 4E**). This
472 analysis was further restricted to conserved pathways (prevalence≥95% across all dermotypes;
473 **Supplementary Figure 19A**), which still retained strong predictive ability (**Figure 4E**). We
474 identified a small number of strongly predictive pathways (**Supplementary Figure 19B**), and their
475 aggregation into metabolic superclasses highlighted the role of carboxylate degradation, amino-
476 acid degradation, carbohydrate biosynthesis, fermentation, and secondary metabolite biosynthesis
477 in being predictive of dermotype labels (**Supplementary Figure 19C**). Hence, a small number of
478 microbial metabolic functions are more predictive of dermotype identity than most of the host
479 attributes analyzed here.

As dermotypes were strongly correlated across body sites (**Figure 1F**), we investigated if dermotype labels could be predicted based on taxonomic profiles from other sites. Remarkably, high predictive performance was obtained in many sites (AUC-ROC>0.8, except for Ax, Ac; **Supplementary Figure 20A**), even when only data from a single site (**Supplementary Figure 20B-C)**, or a small number of taxa were used (best AUC-ROC=0.89; **Supplementary Figure 20D**). Intriguingly, data from the lower leg (Ll) consistently ranked highest in feature importance among top-performing classifiers, despite no robust dermotypes being discovered at this site. Several species including *C. acnes*, *S. epidermidis, M. restricta*, *M. globosa,* and *M. luteus,* consistently ranked high in terms of feature importance for dermotype classification across sites (**Figure 4F**), reflecting the importance of their systemic coordination within the host for defining dermotype structure. Restricting to just dermotype labels, we again noted that cross-site labels were strongly predictive and in many cases more so than host attributes (AUC-ROC>0.84, except for Ax, Ac; **Figure 4G**).

Ridge classifiers that integrated cross-site dermotype labels, metabolic pathway profiles, and demographic factors were used to evaluate their relative utility for dermotype prediction (**Figure 4H**), with uniformly high performance across all sites (AUC-ROC>0.9). Notably, cross-site dermotype labels ranked highest by feature importance at the parietal scalp (Ps), upper back (Ub), and volar forearm (Vf; **Figure 4H**), highlighting the importance of systemic microbial assembly patterns underlying dermotype structure across body sites. Although metabolic pathways were generally secondary to cross-site dermotype labels, they provided complementary information and even ranked highest in terms of feature importance for the scalp (Sc), suggesting that microbial metabolic capabilities may have greater influence on dermotype assembly for some sites. Similarly, while demographic factors had lesser feature importance in general, they had significant predictive value in select sites (e.g., elbow crease [Ac] and forehead [Fo]; **Figure 4H**), showcasing the influence of host attributes in defining microbial community structure.

Overall, our findings emphasize the unexpected strength and predictive power of dermotype coordination across body sites and the critical role of key microbial metabolic functions and host demographic factors in shaping dermotypes on skin.

Discussion

Despite growing efforts to understand the role of the skin microbiome in skin health, disease-agnostic studies in the general population to help define baseline variations and features have been

511 limited to a few early studies with small cohort sizes[1,5] (n<20). Here we present a metagenomic
512 characterization of skin microbiomes from multiple body sites (n=18) in an adult population cohort
513 that is an order of magnitude larger (n=200). This provides a well-powered, rich dataset that has
514 been consistently generated and processed (n=3,572 metagenomes), and complemented by
515 detailed phenotypic information about the subjects, creating a unique resource for investigating
516 skin microbial ecology and function in relation to host attributes and skin health.

517 We systematically identified the presence of two or more robust dermotypes in seven out
518 of nine body sites, revealing significant yet consistent heterogeneity in the general population,
519 even within a well-defined skin site. The dermotypes identified here seem to represent distinct
520 microbiome states in regions with higher density of support (**Figure 1D**, **Supplementary Figure**
521 **6**), unlike the continuum seen in gut enterotypes[69,70], and are defined by the abundances of multiple
522 species rather than just one signature species (**Supplementary File 6**). This allowed the training
523 of near-perfect classifiers for dermotype labels, with as few as five species in all cases. Our work
524 thus goes beyond prior studies which have identified skin microbial community types (i) on the
525 face in the context of air pollution[18] and elbow crease for eczema patients[28], or (ii) as site-agnostic
526 patterns that overlook key site-specific differences[17], by showing that there is clear intra-site
527 heterogeneity in multiple body sites (**Figure 1D-E**). For example, the axilla demonstrates striking
528 heterogeneity that has never been described before, with evidence for at least five robust
529 dermotypes dominated by multiple *Staphylococcus* and *Corynebacterium* species, providing a
530 basis for further studying their ecological interactions and associations with body odor[71] and
531 underarm product use[72]. Similarly, we identified a *Staphylococcus*-dominant elbow crease
532 dermotype (Ac-2) as a feature of the general population and not just atopic dermatitis patients[28],
533 defining an at-risk population for further study. Additional observations of *M. luteus*-enriched
534 dermotypes in dry sites, and *M. restricta*-enriched dermotypes in sebaceous regions, such as the
535 scalp, provide the foundation for further investigations into the skin microbiome's role in skin
536 conditions such as dryness and dandruff, potentially guiding new treatment approaches. Our
537 dermotype classifier models provide an accessible starting point for launching such studies that
538 could be based on cost-effective and clinically-applicable strategies (e.g. using qPCR panels) to
539 assign dermotypes to individuals.

540 Leveraging the multi-site design of our study, we revealed remarkable cross-site
541 correlations in dermotype patterns, even between skin sites that are physically and physiologically

542 distinct (e.g. dry forearm and sebaceous forehead sites; **Figure 1F**). This goes beyond the site
543 specificity of skin microbiome composition that has been previously demonstrated[1,2], as well as
544 the high bilateral concordance in skin microbiomes across all 9 sites and most individuals as we
545 establish here (**Figure 1B**, **Supplementary Figure 2**), and indicates a deeper systemic
546 coordination across the body within individual hosts. This coordination is also evident from the
547 ability to classify dermotypes at one site using information for as little as 10 species or a few
548 dermotype labels from another site (**Figure 4F**). Such patterns may partially result from correlated
549 abundances of a few core species shared across multiple sites (e.g., *C. acnes*, *S. epidermidis*, and
550 *M. restricta*; **Figure 2B**, **Supplementary Figure 9**), perhaps due to strain-specific ecological
551 functions[73–75]. However, the preference for specific species/strains alone does not fully explain the
552 patterns observed, as some taxa show no significant correlations between specific sites (e.g.,
553 *Staphylococcus* species in axilla and elbow crease; **Supplementary Figure 9**). This suggests that
554 other shared factors (e.g. host genetics, global sebum production levels) or behaviors (e.g.
555 showering habits) might dominantly shape the skin microbiome in a non-site-specific, bilaterally-
556 symmetric manner. The observation that loss of bilateral concordance can also be correlated across
557 body sites further suggests that such dysbiosis might also be shaped by coordinated influences
558 across the body (**Supplementary Figure 3**). Interestingly, we also observed body-wide
559 enrichment of ancillary species, particularly skin eukaryotic viruses (e.g., polyomaviruses or
560 papillomaviruses; **Figure 2C-D**), within some individuals, indicating that immune dysregulation
561 could also play a role[1,2] (**Figure 2E**). Overall, these observations provide the key hypothesis that
562 the colonization and adaptation of microbial species, and the corresponding emergence of
563 dermotypes, is a coordinated phenomenon across various skin sites that is strongly shaped by host
564 factors, and not purely a stochastic assembly process.

565 Systematically investigating this further based on dermotype-dependent co-abundance
566 patterns and pairwise antagonism assays between core species, we identified varying oxygen
567 availability (shaped by skin physiology[76,77]) as a key factor impacting interspecies interactions,
568 such as hypoxia-dependent inhibition of *S. hominis* by *S. epidermidis* and *M. luteus* (**Figure 3B**).
569 Such oxygen-dependent antagonism may arise from differential toxin production (e.g., *S.*
570 *epidermidis* phenol-soluble modulins produced under hypoxia[78]) or redox-sensitive metabolic
571 cross-feeding[79]. Similar modulation was also seen in a previous study of tri-species interaction
572 dynamics in healthy and dandruff scalp microbiomes[80], highlighting the role of oxygen availability

573 in defining microbial interactions on skin. While oxygen gradients explain part of this dynamic,
574 nutrient availability (lipids, carbohydrates and amino acids) may synergistically shape interactions.
575 For example, in sebaceous environments *C. acnes* may grow well by hydrolyzing sebum
576 triglycerides into free fatty acids and thus acidifying the skin surface. This acidification, combined
577 with lipid reduction and potential antimicrobial activities, likely contributes to unfavorable
578 conditions for *M. luteus* and other skin bacteria[81]. Conversely under lipid-depleted conditions, the
579 competitive dominance of *C. acnes* may diminish (**Supplementary File 8**), while species such as
580 *S. hominis* and *M. luteus* may competitively leverage availability of other nutrients such as
581 galactose and histidine to grow well (**Figure 3E-F**). Future work in human cohorts would benefit
582 from integration of metatranscriptomic[87], metabolomic[86], and lipidomic[84,82] data in a longitudinal
583 study design, while *in vitro* studies could use multi-species consortia co-cultured with skin models
584 to further investigate the relative contributions of oxygen gradients, nutrient availability, and pH
585 in shaping dermotype structures across individuals.

586 Several of our observations in the *S. hominis*-dominated elbow crease dermotype (Ac-2)
587 including associations with older age, reduced surface hydration, skin sensitivity, and itch, might
588 indeed reflect a bidirectional interplay between microbial ecology, metabolic activities, and barrier
589 dysfunction (**Supplementary Figure 21**). Compromised skin barrier, whether due to dryness,
590 inflammation, or mechanical damage from scratching, may enhance oxygen penetration, creating
591 aerobic microenvironments where *S. hominis* escapes inhibition by *S. epidermidis* and *M. luteus*.
592 Notably, *S. hominis* may further circumvent *S. epidermidis* antagonism via galactose metabolism,
593 which disproportionately enhances its growth relative to *S. epidermidis* (**Supplementary Figure**
594 **15**). Thriving under these niche conditions in the skin, *S. hominis* can exacerbate itch by inducing
595 IL-33 expression in keratinocytes, establishing an itch-scratch cycle that further damages the
596 barrier and reinforces the aerobic conditions that favor its dominance (**Supplementary Figure**
597 **21**). In contrast, enhanced *M. luteus*-mediated histidine degradation in Ac-1 dermotypes could
598 promote production of immunosuppressive cis-urocanic acid and thus contribute to reduced skin
599 sensitivity in Ac-1 individuals. Together, such metabolic cross-talk underscores the role of host-
600 derived nutrient availability and microbial substrate preferences in shaping distinct ecological
601 niches, and consequently, skin phenotypes (**Supplementary Figure 21**). Further work is thus
602 warranted to disentangle this complex interplay in contributing to skin itch and inflammatory skin
603 conditions such as atopic dermatitis. As a result, dermotype status could serve to stratify subjects

604 in terms of their disease risk, severity and treatment response, while targeting microbial pathways
605 (e.g., IL-33-inducing proteases), modulating oxygen gradients (e.g., via topical emollients or
606 occlusion therapy), or leveraging microbial metabolic pathways (e.g., through probiotic or
607 metabolite supplementation) may offer more effective personalized therapeutic avenues to
608 interrupt this pathogenic cycle.

609 Overall, our study also helps delineate the potential contributions of host demographics,
610 skin physiology, and lifestyle factors in shaping skin microbiome variability across the population.
611 Leveraging our multi-ethnic cohort spanning a broad age range, we demonstrate that ethnicity,
612 age, and gender have site-specific associations with dermotypes, particularly in moist sites such as
613 the axilla and elbow crease (**Figure 4A**). While the strong ethnicity-specific patterning of axillary
614 dermotypes may arise from genetic predispositions to lipid and sweat content on skin[85,86],
615 differences in skincare practices, or cultural norms for product usage could also play a role. The
616 absence of demographic associations at head sites on the other hand underscore the influence of
617 localized practices—such as washing, chemical treatment, or hair covering—in overriding broader
618 host factors in these regions (**Figure 4C**), aligning with growing evidence that personal care
619 products can shape microbial resilience on skin[87,88,11]. However, by integrating all these factors
620 into a machine learning framework we found that cross-site microbial coordination and core
621 microbial metabolic functions are stronger predictors of dermotype identity than host attributes in
622 most regions (**Figure 4H**). The prominence of shared pathways such as amino acid, carbohydrate
623 and nucleotide biosynthesis suggests that microbial metabolic function (shaped by nutrient
624 availability from the host), rather than taxonomic identity, drives niche specialization
625 (**Supplementary Figure 19**). What is even more striking is the predictive power of dermotype
626 information across sites, with AUC-ROC values as high as 0.9 for some sites (**Figure 4F-G**),
627 pointing to systemic coordination in the host that transcends local differences in ecology.

628 In conclusion, we demonstrate the use of a large-scale, multi-site skin microbiome resource
629 of >3,550 metagenomes across 18 sampling sites to establish key features including bilateral
630 symmetry, dermotype prevalence, and coordinated occurrence of dermotypes across body sites.
631 We show that beyond mere differences in species abundance, dermotypes represent distinct
632 ecological niches where microbial populations adapt and specialize in exploiting their unique
633 resource availability and environmental conditions, reflected in dermotype-dependent microbial
634 interactions (e.g., *M. luteus* and *S. hominis*) and metabolic activities (e.g., galactose and histidine

degradation pathways). Integrating the availability of extensive host phenotype data, we provide new insights into how host attributes may shape dermotypes, which in turn can influence host skin conditions. Our results, dermotype prediction models and metagenomic datasets lay the groundwork for future research into therapeutic strategies for skin conditions based on microbiome-informed stratification and personalized skin interventions.

640 # Methods

641 ## Subject recruitment and microbiome sampling

642 Subjects were recruited as part of the Asian Skin Microbiome Programme (ASMP) sub-study for
643 the Health For Life In Singapore (HELIOS) study (https://www.healthforlife.sg/), a large
644 population cohort that is central to the National Precision Medicine program (NPM) in Singapore.
645 The HELIOS cohort study specifically recruits subjects to ensure that their demographics
646 accurately match that of Singapore's population, considering factors such as ethnicity, age, gender
647 and socio-economic status. Informed consent was obtained from all subjects, and associated
648 protocols for the study were approved by the NTU Institutional Review Board (IRB reference
649 number: IRB-2016-11-030). Subjects were excluded from the study if they were currently
650 pregnant, breastfeeding, had any ongoing acute illness, had recent major surgery, were under the
651 age of 30, or had any mental incapacities that would prevent them from understanding and giving
652 informed consent. Upon arrival to the HELIOS clinic, subjects were given time to acclimatize to
653 the temperature of the clinic. Following this, subjects were asked to complete a skin health
654 questionnaire. This extensive questionnaire has questions related to several topics including
655 overall skin health, skin-related medical treatments, personal care routines, product usage, lifestyle
656 habits, and finally a series of questions relating to menopause for all female subjects
657 (**Supplementary File 14**). Subjects were then moved to a climate-controlled room for sampling
658 by a trained research coordinator/clinician.

659 A diverse set of body sites were chosen for skin surface sampling, based on their
660 importance for consumer care studies, and to capture the breadth of the human skin microbiome.
661 These consisted of nine different body sites, sampled on both the left and right sides (**Figure 1A**),
662 including: scalp (Sc), forehead (Fo), cheek (Ch), upper back, (Ub), antecubital fossa (elbow crease,
663 Ac), volar forearm (Vf), and front of lower leg (Ll), for which an adhesive tape disc was used for
664 sampling (D-Squame Standard Sampling Discs, D100, CuDerm). The tape was repeatedly placed
665 on the same location with pressure and peeled 25 times to saturate discs with skin surface material.
666 For the remaining two sites, sterile swabs (FLOQSwabs Flocked Swabs, 552C, Copan) were used
667 to sample the skin microbiome on the parietal scalp (5cm above the ear, Ps) and axilla (armpit,
668 Ax), by soaking in 0.9% saline and rubbing back and forth on the skin for 30 seconds. One tape
669 disc or swab was collected from each side (left and right) of the skin sites, giving 14 tape discs and

670 4 swabs per subject. All collected samples were stored at -80°C before sample processing.
671 Negative control samples were collected in the form of blank tape discs (n=34) and swabs (n=12)
672 that were processed in the same way as skin samples.

673 **Skin physiological measurements**

674 Trans-epidermal water loss (TEWL) was measured using a VapoMeter (Delfin) on the forearm
675 (left or right) and on the cheek (just below the left eye). pH was measured using a pH meter (Delfin)
676 on the forearm (left or right). Stratum corneum surface hydration was measured using a
677 MoistureMeter (Delfin) on the forearm (left or right). Skin sebum secretion was measured using a
678 SebumScale (Delfin) on the cheek (just below the left eye). Three consistent consecutive readings
679 were taken for all measurements and averaged. Measurements were performed in a controlled
680 temperature (20-22°C) and humidity (relative humidity 40-60%) environment after at least 20
681 minutes of acclimatization.

682 **DNA extraction**

683 Genomic DNA was extracted from skin tapes and swabs using a previously established
684 protocol[35,28], with modifications to improve yield. Briefly, samples were transferred into Lysing
685 Matrix E tubes containing 500 µL of Buffer ATL (Qiagen) and subjected to bead-beating at 4 m/s
686 for 30 seconds, repeated twice, using a FastPrep-24 Automated Homogenizer (MP Biomedicals).
687 Following homogenization, tubes were centrifuged at maximum speed for 5 minutes, and 200 µL
688 of supernatant was transferred to 2 mL EZ1 sample tubes (Qiagen). Proteinase K (12 µL) was
689 added, and samples were vortexed and incubated at 56 °C for 15 minutes. DNA was then purified
690 using the EZ1 Advanced XL Instrument (Qiagen) with the EZ1 DNA Tissue Kit and eluted in
691 50 µL of EB buffer.

692 **Construction and sequencing of metagenomic libraries**

693 Purified genomic DNA underwent NGS library construction steps using NEBNext Ultra II FS
694 DNA Library Prep Kit according to manufacturer's instructions. 26 µl of DNA added to
695 fragmentation reagents were mixed well and placed in a thermocycler for 10 mins at 37°C,
696 followed by 30 mins at 65°C to complete enzymatic fragmentation. Post fragmentation, adaptor-
697 ligation was performed using Illumina-compatible adaptors, diluted 10-fold as per kit's
698 recommendations prior to use. Post-ligation, size selection or cleanup of adaptor-ligated DNA was

699 performed using Ampure XP beads in a 7:10 beads-to-sample volume ratio. Unique barcode
700 indexes were added to each sample and amplified for 12 cycles under recommended kit conditions
701 to achieve multiplexing within a batch of samples. Finally, each library sample was assessed for
702 quality based on fragment size and concentration using the Agilent D1000 ScreenTape system.
703 Samples that passed the quality-checks were adjusted to identical concentrations employing
704 dilution and volume-adjusted pooling. The multiplexed sample pool was paired-end (2×151bp)
705 sequenced on an Illumina HiSeq X Ten system. All libraries were sequenced at the Novogene
706 sequencing facility in the Genome Institute of Singapore following standard Illumina sequencing
707 protocols.

708 **Taxonomic and functional profiling**

709 Raw Illumina sequencing reads were processed using an in-house Nextflow pipeline
710 (https://github.com/CSB5/shotgunmetagenomics-nf). Quality control and adapter trimming were
711 performed with FastQC v0.20.0 (default parameters). Human reads were removed by mapping to
712 the hg19 human reference genome using BWA-MEM (v0.7.17-r1188, default parameters), and
713 filtering with SAMtools (v1.7). Taxonomic profiling of the remaining reads was conducted with
714 MetaPhlAn2 (v2.7.7, default parameters). During sequencing data generation (2020 January),
715 MetaPhlAn3 (v3.0, 2020 April release) was tested but frequently omitted eukaryotic taxa,
716 including fungal species commonly observed and cultured from human skin; additionally,
717 MetaPhlAn4 (v4.0, 2022 April release) outputs viral profiles separately, without integration into
718 the non-viral taxonomic profile. Due to these limitations, we opted to use MetaPhlAn2 consistently
719 in this long-running project. To address the limited representation of *Malassezia* species in
720 MetaPhlAn2 (a commonly found fungal member of the human skin microbiome[89,90]), we
721 integrated results from PathoScope[91] (v2.0) for *Malassezia* species. Briefly, after human read
722 removal, we aligned the remaining reads to 14 *Malassezia* genomes (*M. furfur, M. obtusa, M.*
723 *yamatoensis, M. japonica, M. sympodialis, M. pachydermatis, M. restricta, M. globosa, M.*
724 *slooffiae, M. caprae, M. cuniculi, M. dermatis, M. equina,* and *M. nana*) downloaded from NCBI
725 using bpipe (v0.9.8.6_rc), and normalized mapped read counts by the total number of reads to
726 estimate the relative abundance of *Malassezia* species. Lastly, MetaPhlAn2 taxonomic profiles
727 (excluding *Malassezia*) were rescaled by multiplying by the proportion of reads that were not
728 mapped to *Malassezia* (calculated from PathoScope results). The rescaled MetaPhlAn2 profile and
729 PathoScope-derived *Malassezia* abundances were merged and renormalized using Total Sum

730 Scaling (TSS) to generate final taxonomic profiles (n=3,586) for downstream analyses. In addition,
731 we reannotated *Enyhydrobacter aerosaccus* as *Moraxella osloensis*, consistent with a previous
732 study showing that 99.4% of sequences labeled as *E. aerosaccus* in the MetaPhlAn2 database were
733 identified as *M. osloensis* using BLASTN against the NCBI nt database[17].

734 Sequencing data from negative control samples (n=46; **Supplementary File 1**) and
735 information from existing databases (PubMed, DISBIOME[92], MicroPhenoDB[93]) were used to
736 identify potential artifacts in the taxonomic profiles, particularly *kitome* species derived from DNA
737 extraction kits and sequencing reagents (**Supplementary Figure 1**). In the first tier, species
738 commonly found in negative controls (median abundance≥0.1%) were evaluated for negative
739 correlation with library concentration, a signature of a contaminant species[94] (negative Spearman
740 correlation, p-value<0.05). These kitome candidates were then used to identify additional taxa
741 highly correlated with them, potentially also arising from contamination (Spearman ρ≥0.7, p-
742 value<0.05). The second tier refines this initial selection by restricting to genera not previously
743 reported on human skin or associated with skin disease in three different databases (PubMed,
744 DISBIOME[92], MicroPhenoDB[93]). Poor-quality profiles were removed (n=14, 0.4%; kitome total
745 abundance>50%). Kitome taxa (**Supplementary File 2**) and extremely low-abundance species
746 (relative abundance<0.1%) were excluded from taxonomic profiles (set to 0), and profiles were
747 TSS-renormalized (n=3,572; **Supplementary File 3**).

748 Functional profiles and detailed species-specific reconstructions of microbial metabolic
749 pathways (MetaCyc) were generated with HUMAnN2 (v2.8.1, default parameters). Pathways that
750 were unannotated, or highly correlated with kitome species (Spearman ρ≥0.7) were filtered out.
751 TSS-renormalized pathway abundance profiles were used for subsequent analyses (n=3,550;
752 **Supplementary File 9**). Downstream data analysis was performed in R v4.4.0 unless stated
753 otherwise.

754 **Identification of dermotypes**

755 Dermotypes were identified via a three-step protocol **(Supplementary Figure 6)**. Firstly,
756 Partitioning Around Medoid (PAM) clustering was performed using Bray-Curtis dissimilarity for
757 each site. The optimal number of clusters was determined using key indices (prediction strength,
758 silhouette index, and Calinski-Harabasz index). In cases of ambiguity (i.e., disagreement across
759 various indices), we employed Mean-Shift clustering (bandwidths determined by cross-validation)

760 to search the local maxima of kernel density estimated in PCoA space (PCoA 1-3). Next, we
761 assessed sample stability and cluster robustness via bootstrapping (n=100). Clusters were deemed
762 robust if over half of the samples were consistently grouped into the same cluster in over 90% of
763 iterations (median stability≥0.9). Only sites with at least 2 robust clusters detected were defined as
764 having distinct dermotypes for subsequent analyses. Finally, we relied on PAM to infer cluster
765 memberships for robust dermotypes, with samples with <90% stability excluded from further
766 analyses.

767 **Diversity and differential abundance analysis**

768 Metrics for α-diversity (Shannon index) and β-diversity (Bray-Curtis dissimilarity) were computed
769 using the vegan package (v2.6.4). Permutational multivariate analysis of variance
770 (PERMANOVA) was performed with Bray-Curtis dissimilarity values to quantify differences in
771 microbial compositions using the adonis2 function from vegan[2] (v2.6.4) with 1,000 permutations
772 and restrictive permutations within blocks of SubjectID. Multivariable analysis estimating which
773 microbiome attributes (species, pathways) differ significantly across comparison groups was
774 carried out using MaAsLin2 (linear model, "LOG" transformation, prevalence≥10%, BH-adjusted
775 p-value<0.05), including demographic covariates (age, gender, ethnicity) as fixed effects and
776 subject ID as a random effect when applicable.

777 **Construction of species co-abundance networks**

778 Species co-abundance networks were constructed with taxonomic profiles with high cluster
779 membership stability (≥90%) in each dermotype using the "netConstruct" function in NetCoMi[95]
780 (v1.1.0), and visualized using igraph[96]. Prevalence filtering was applied to focus on prevalent taxa
781 (≥30% for Vf, Ac, Ub, and Fo dermotypes, and ≥10% for Sc, Ps, and Ax dermotypes); this filtering
782 ensures that each network construction included at least 15 taxa, minimizing spurious findings and
783 allowing for biologically meaningful insights. Compositionally corrected Spearman ρ for each
784 species pair and corresponding p-values were obtained from CCREPE[97] (v1.34.0), and
785 downstream analysis were restricted to species pairs with strong co-abundant or co-exclusion
786 patterns in each dermotype (|ρ|>0.3, BH-adjusted p-value<0.05; **Supplementary Figure 12**).
787 Hubs were defined as species with correlation strength-weighted degree centrality above the 95th
788 percentile in each network. Differential edges between dermotypes for each skin site were
789 identified using Fisher's z-test[98] for correlation values.

790 **Antagonism co-culture assays**

791 The bacterial strains studied (*Cutibacterium acnes* 6919, *Micrococcus luteus* 10240,
792 *Staphylococcus epidermidis* 12228, and *Staphylococcus hominis* 27844) were obtained from
793 ATCC. *C. acnes* was cultured in half-strength BHI broth in an anaerobic chamber at 37°C
794 overnight with agitation, and other species were cultured under the same condition aerobically.
795 After measuring the optical density at 600 nm (OD600), each culture was diluted to OD600=0.1
796 with fresh broth (5mL total) and incubated at 37°C with agitation (220 rpm) until growing to an
797 OD that corresponds to the mid-log phase of the bacteria (i.e., *C. acnes* OD=0.6; *M. luteus* OD=1;
798 *S. epidermidis* OD=2; *S. hominis* OD=2).

799 Assay plates were prepared by dispensing 2mL of molten half-strength BHI agar into each
800 well of a 12-well plate and allowing it to solidify and dry thoroughly. Each bacterial pair was
801 assayed in triplicates, with one well per replicate. For hypoxic conditions, plates were placed in
802 compact plastic pouches with anaerobic sachets (Thermo Fisher Scientific Oxoid AnaeroGen, Cat.
803 No. AN0020D). A total of 40µL of test bacterial culture (bacteria being tested for inhibitory
804 activity; 10µL per droplet, sequentially spotted four times) was applied to each well, allowing each
805 droplet to dry before applying the next. The following day, cultures of target bacteria (bacteria to
806 be assessed for inhibition) were diluted 100× in 1× PBS (10× for *C. acnes*). A 10µL aliquot of the
807 diluted target culture was then spotted adjacent to the test bacteria in each well. The plates were
808 incubated at 37°C overnight, or two nights if the target bacteria was *C. acnes*, and clearing zones
809 were assessed the following day.

810 **Effect of galactose on *Staphylococcus* species growth**

811 The effect of galactose on the growth of *Staphylococcus* species was assessed by performing
812 monoculture and co-culture experiments with *Staphylococcus epidermidis* 12228 (SE) and
813 *Staphylococcus hominis* 27844 (SH) in Brain heart infusion (BHI) broth with and without 5%
814 galactose supplementation. Briefly, six different experimental conditions were designed: 1) BHI
815 + SE 2) BHI + SH 3) BHI + Galactose 5% + SE 4) BHI + Galactose 5% + SH 5) BHI + SE + SH
816 6) BHI + Galactose 5% + SE + SH, with two separate negative controls including, BHI and BHI
817 + Galactose 5%. Growth curve experiments were performed at 37ºC for 48h with high amplitude,
818 fast speed shaking in a 100-well Honeycomb plate (Growth Curves USA) with a working volume
819 of 200µl in each well. All the experiments were performed in quadruplicates. Optical Density (OD)

820 at 600nm measured using an automated Bioscreen C MBR (Growth Curves USA) reader at 30-
821 minute intervals was plotted to obtain growth curves.
822 Cell pellet was used for DNA extraction for qPCR analysis. Each sample was treated with
823 500µl Buffer ATL (Qiagen) in Lysing Matrix E tubes and homogenized twice at 4 m/s for 30
824 seconds using the FastPrep-24 (MP Biomedicals). After centrifugation at maximum speed for 5
825 minutes, 200µl of supernatant was transferred into 2ml EZ1 tubes (Qiagen), mixed with 12µl
826 Proteinase K, vortexed, and incubated at 56°C for 15 min. DNA was extracted in EZ1 Advanced
827 XL Instrument (Qiagen) using EZ1 DNA Tissue Kit (Qiagen) and eluted in 50µl EB buffer. DNA
828 concentration was determined using a Qubit 4 Fluorometer with the Qubit dsDNA BR Assay Kit
829 (Invitrogen). qPCR reaction mixture was prepared following the instructions in the PowerSYBR
830 Green MasterMix protocol (Applied Biosystems, Cat. No. 4368577) and 1µl of the extracted DNA
831 was used as the DNA template. Reactions were run in triplicates on a 96-well plate in ViiA 7 Real-
832 Time PCR System (Applied Biosystems) with the following cycling conditions: 95°C for 2 min;
833 40 cycles of 95°C for 15 s, 60°C for 1 min; followed by a dissociation step. Standard curves were
834 generated using gBlocks (IDT) diluted to 10^8–10^3 copies/µl. Primer and gBlock sequences used
835 for the qPCR are listed in **Supplementary File 17**.

836 **Effect of histidine on *Micrococcus luteus* growth**

837 The effect of histidine on *Micrococcus luteus* was assessed by culturing *M. luteus* in Brain heart
838 infusion (BHI) broth with and without L-histidine supplementation (50µg/ml concentration), with
839 BHI used as a negative control. Growth curve experiments were performed at 37ºC for 6 days at
840 high amplitude, normal speed shaking in a 100-well Honeycomb plate (Growth Curves USA) with
841 a working volume of 200µl in each well. All the experiments were performed in triplicates. Optical
842 Density (OD) at 600nm measured using an automated Bioscreen C MBR (Growth Curves USA)
843 reader at 30-minute intervals was plotted to obtain growth curves.

844 **Metabolomics sample collection and processing**

845 An additional cohort of 100 subjects was recruited and skin samples collected in a separate sub-
846 study of the HELIOS study for skin metagenomic and metabolomic analysis (Ac, elbow crease).
847 For metabolomics, D-Squame (D100, Clinical and Derm) skin tapes were collected in the same
848 manner as for metagenomics, except that they were pushed deep into their collection tube and
849 processed by adding 1 ml of extraction buffer per tube, consisting of 0.1% formic acid in water,

850 and 0.1μg/ml heavy isotope labelled standards. The tubes were then water bath sonicated (Elma)
851 for 15 minutes at 4°C. Following sonication, 120 μl from each sample was loaded onto 96 well
852 chimney plates (Greiner Bio-One) prior to acquisition by liquid chromatography mass
853 spectrometry (LCMS). An additional 150μl was used to determine the protein concentration with
854 the Micro BCA kit (Thermo Scientific). 96 well plates holding the samples were loaded into the
855 Vanquish LC system (Thermo Scientific), utilizing the Atlantis Premier BEH C18 AX 1.7 μm,
856 2.1×100 mm column (Waters) before being acquired on the TSQ Quantis triple quadrupole MS
857 machine (Thermo Scientific). All data files were analysed via the software Skyline. No heavy
858 isotope labelled standard were available for trans-UCA, so the cis-UCA heavy isotope labelled
859 standard was used as a proxy. Output values from each sample were then normalized against their
860 respective Micro BCA derived protein concentrations.

861 **Dermotype associations with host attributes**

862 Samples with unstable dermotype assignments (stability<90%) were excluded for each skin site;
863 additionally, subjects with highly stable yet discordant dermotypes at the same site were excluded,
864 to enhance consistency. Associations between dermotypes and demographic or physiological data
865 was assessed in two stages. In the first stage, categorical variables were analyzed using Fisher's
866 exact test, while continuous variables were evaluated using Wilcoxon tests. Only attributes that
867 demonstrated overall statistical significance were carried forward to the second stage. In the
868 second stage, categorical variables were analyzed using Fisher's exact test in a one-versus-the-rest
869 framework, with odds ratios recorded as effect sizes; for continuous variables, pairwise Wilcoxon
870 tests were performed, and rank-biserial correlations were reported as effect sizes.

871 Associations between dermotypes and questionnaire responses were assessed using logistic
872 regression models with a one-versus-rest approach to accommodate multi-label cases (e.g., with
873 five axilla dermotypes), with Beta coefficients recorded as effect sizes. When dermotypes
874 exhibited significant associations with demographic factors in prior analyses (**Figure 4A**), these
875 factors were included as covariates in the regression model. We prioritized responses relevant to
876 the body site (e.g., underarm product usage for axilla dermotypes), and included response groups
877 with at least five observations to maintain statistical power. Questions related to specific product
878 name were excluded.

879 **Machine learning-based dermotype classification**

880 To ensure comparability across questionnaire responses, all continuous variables (e.g., itch
881 severity) were z-standardized, and categorical variables were one-hot encoded. Binary-class
882 machine learning classifiers were trained to predict the dermotype of the sample based on its own
883 metagenomic profile. A one-vs-rest strategy was employed to accommodate multi-label
884 classification (i.e., axilla dermotype classification). To address class imbalance, we employed the
885 Synthetic Minority Over-sampling Technique (SMOTE) to up-sample minority class instances[99].
886 Samples with unstable clustering membership (stability<90%) were excluded from each skin site.
887 We selected six machine learning methods based on their distinct properties, such as linear
888 and non-linear capabilities. These included decision trees, logistic regression, random forests,
889 support vector classifiers[100] (linear and gaussian kernel) and XGBoost[101]. Model training and
890 performance reporting utilized a nested cross-validation approach to provide an unbiased
891 performance estimate by accounting for hyperparameter selection—this involved 4-fold training
892 with 3-fold inner cross-validation for hyperparameter tuning. Tuned models were evaluated on a
893 withheld 1-fold test, and the process was repeated five times with different data splits. Mean
894 accuracy and AUC-ROC scores were reported (one-vs-rest, macro average). Using the same
895 framework as described above, we trained additional dermotype classifiers with alternative input
896 features, including: 1) the sample's metabolic pathway profile, 2) metagenomic profiles from other
897 skin sites within the same host. Further analyses were based on the best-performing methods:
898 random forests for models using metagenomic profiles, and logistic regression for models using
899 metabolic pathway profiles.
900 Random forest classifiers were further trained to assess associations between host factors
901 and dermotypes using host attributes (i.e., demographics, skin physiological parameters, and
902 questionnaire data), as they provided the best performance. Only subjects with concordant
903 dermotype labels between bilateral sites were included. Lastly, to determine the relative
904 informativeness of different feature sets for dermotype classification at a given skin site, we trained
905 an inclusive ridge classifier using three broad categories of features: demographic factors,
906 conversed metabolic pathway profiles, and dermotype labels from other skin sites. Categorical
907 values were one-hot encoded and continuous values z-standardized.

908 **Commensal *Staphylococcus* strain isolation**

909 Commensal *Staphylococcus* strains were isolated from skin swabs for 4 healthy donors across 4
910 separate body sites: antecubital fossa, axilla, cheek and scalp. Swabs were immersed in 2.5mL

911 BHI broth (Oxoid), incubated at 37°C at 210rpm for 2 hours before being spread onto TSA-SB
912 (Thermo Fisher Scientific) and Baird Parker (Oxoid) plates and incubated at 37°C for 24 hours.
913 Colonies were then picked and grown in 4ml BHI broth for 16 hours overnight at 230rpm and
914 37°C. 200µL of the cultures were spun down at 5000rpm for 5 minutes. To extract gDNA the pellet
915 was resuspended in 50µL QuickExtract (Lucigen). Bacterial suspensions were then heated to 65°C
916 for 6 minutes, briefly vortexed, then heated to 98°C for 4 minutes and briefly vortexed again. PCR
917 reactions were setup with 10µL KAPA SYBR FAST (Merck), 6.4µL nuclease free water
918 (Promega), 3µL extracted gDNA and 0.6µL primer mix. Primer mixes were specific for 4 specific
919 *Staphylococcal* species (**Supplementary File 15**). PCR was run on an AriaMx Real Time PCR
920 System (Agilent Technologies). PCR products were sent for Sanger sequencing to confirm product
921 sequences. The identified *Staphylococcal* colonies were inoculated into 4ml BHI broth and grown
922 overnight for 16 hours at 230rpm and 37°C. Optical density (OD) was measured with a
923 SpectraMax M5 Microplate Reader (Molecular Devices). Overnight cultures were then grown
924 from OD 0.1 in fresh BHI for 26 hours at 230rpm and 37°C. Final OD measurements for all strains
925 after 26 hours was normalized to OD 2.75. Bacterial cultures were spun down at 5000rpm for 5
926 minutes, pellets were discarded and supernatants stored at 4°C.

927 **Human keratinocyte cell culture and Nano-Glo HiBiT**

928 A genetically modified N/TERT keratinocyte cell-line containing a 33 nucleotide HiBiT tag
929 directly adjacent to the IL33 gene, was utilized for Nano-Glo HiBiT experiments. Keratinocytes
930 were cultured in keratinocyte serum free medium (KSFM; Gibco) supplemented with bovine
931 pituitary extract (BPE) at a final concentration of 20µg/mL, epidermal growth factor (EGF) at a
932 final concentration of 0.2ng/mL, calcium chloride at a final concentration of 300µM and 1:1000
933 penicillin-streptomycin (pen-strep; Gibco). Keratinocytes were seeded onto 96-well cell culture
934 plates (Greiner) at a density of 20,000 cells per well and grown for 24 hours in 100µL KSFM.
935 After 24 hours, KSFM was removed and the keratinocytes were cultured in 100µL KSFM with
936 either 0.2% TritonX (Merck), 5% BHI, or 5% bacterial supernatants. Keratinocytes were cultured
937 overnight for 16 hours in treatment conditions. After 16 hours of treatment on the HiBiT N/TERT
938 keratinocytes, 50µL of the KSFM was transferred to a white-bottom 96-well plate (Sigma-Aldrich)
939 and mixed with 50µL Nano-Glo HiBiT extracellular reagent. The remaining 50µL of KSFM was
940 retained on the cell culture plate with the cells and mixed with 50µL CellTitre-Glo reagent. Both

941 plates were mixed on an orbital shaker for 1 minute, and incubated for 10 minutes at room
942 temperature before luminescence was measured with a SpectraMax M5 Microplate Reader.

943 **Data availability**

944 Shotgun metagenomic sequencing data is available from the European Nucleotide Archive (ENA,
945 https://www.ebi.ac.uk/ena/browser/home) under project accession number PRJEB87291.
946 Demographic data and skin physiological measurements can be found in **Supplementary File 18**.
947 Large datasets and metabolomics data are available on Figshare at
948 https://figshare.com/projects/ASMP_Dermotype/243932. The authors of this study do not own the
949 rights to the HELIOS dataset, and this dataset is under controlled access to ensure good data
950 governance, responsible data use, and that the dataset is only used for the intended research
951 purposes in compliance with HELIOS study cohort IRB and ethics approval. Access to individual
952 level data for study participants should be submitted to the HELIOS Data Access Committee
953 (email: helios_science@ntu.edu.sg).

954 **Code availability**

955 Source code for scripts used to analyze the data are available in a GitHub project at
956 https://github.com/CSB5/ASMP_Dermotype.

957 **Acknowledgments**

958 The work is supported by funding for the Asian Skin Microbiome Programme (Industry Alignment
959 Fund Pre-Positioning – IAF-PP; H18/01/a0/16) and Asian Skin Microbiome Programme 2.0 (IAF-
960 PP H22/J1/a0/040) from the National Research Foundation, Singapore. Additional support is
961 provided by HELIOS study (IRB No. 2016-11-030), which is supported by Singapore Ministry of
962 Health's (MOH) National Medical Research Council (NMRC) under its OF-LCG funding scheme
963 (MOH-000271-00), Singapore Translational Research (STaR) funding scheme
964 (NMRC/StaR/0028/2017), the National Research Foundation (Singapore) through the Singapore
965 MOH NMRC and the Precision Health Research (PRECISE) under the National Precision
966 Medicine programme (NMRC/PRECISE/2020), National Cohorts Office through the Singapore
967 MOH NMRC (P2022-01-03 and P2022-02-03) and intramural funding from Nanyang
968 Technological University, Lee Kong Chian School of Medicine and the National Healthcare
969 Group. The HELIOS study is also supported by an outstanding team of administrative and

operational staff (**Supplementary File 16**). The computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore (https://www.nscc.sg) and the A*STAR Computational Resource Center. We thank all investigators, staff members and study participants who made HELIOS and ASMP studies possible.

Competing interests

The authors declare no competing interests.

Table 1. Characteristics of study participants (n=200)

Demographics	
Age (years, median [IQR])	49.00 [40.00, 57.00]
Gender Male (%)	96 (48.0)
Ethnicity (%)	
Chinese	120 (60.0)
Indian	40 (20.0)
Malay	40 (20.0)
Monthly household income (SGD) (%)	
< 6K	79 (39.5)
≥ 6K	102 (51.0)
Unknown	19 (9.5)
Current smoker (%)	0 (0)
Skin physiological measurements	
Forearm	
Surface hydration (median [IQR])	30.02 [23.18, 39.00]
pH (median [IQR])	4.67 [4.38, 5.14]
TEWL[1] (median [IQR])	4.60 [3.67, 5.34]
Cheek	
TEWL[1] (median [IQR])	13.18 [11.64, 15.27]
Sebum (median [IQR])	7.00 [2.00, 19.25]
Skin health and conditions	
Itchy skin (%)	59 (29.5)
Sensitive skin (%)	69 (34.5)
Body odor (%)	44 (22.0)
Antibiotic usage in the last month (%)	12 (6.0)
Topical antimicrobial usage in the last month (%)	24 (12.0)
History of eczema (%)	44 (22.0)
Acne (%)	17 (8.5)
Psoriasis (%)	5 (2.5)
Vitiligo (%)	2 (1.0)
Skin infection[2] (%)	18 (9.0)
Skin cancer (%)	0 (0)
Alopecia areata (%)	3 (1.5)

[1]TEWL = Transepidermal water loss

[2]Skin infection includes bacterial infection, fungal infection (at feet, body, and groin), viral warts, scabies, chronic hives and ulcer wounds

979 **Figure Legends**

980 **Figure 1 | Heterogeneity of site-specific skin microbial community compositions within and**
981 **across individuals. A.** Overview of skin sampling strategy, illustrating bilateral sampling at nine
982 sites with swabs for Ps and Ax and tapes for others. **B.** Boxplots quantifying the similarity of the
983 skin microbiome within (left versus right) and between individuals across sites. Sites in the inset
984 figure for alpha diversity are presented in the same order. All p-values were calculated using the
985 one-sided Wilcoxon rank-sum test (* denotes BH-adjusted p-value<0.001). Boxplot whiskers
986 represent 1.5× interquartile range or the maximum/minimum data point within the range. **C.** Left
987 barplot comparing the percentage of subjects with their hair covered daily between those with
988 concordant and discordant parietal scalp microbiomes between left and right body sites; right
989 boxplot showing that individuals with discordant forehead microbiomes have significantly higher
990 itch severity. **D**. Perspective kernel density plot of Ax microbiome PCoA results with the five red
991 dots representing local maxima determined via Mean Shift analysis. **E.** Averaged relative
992 abundance plots of major species ranked by median abundance for each dermotype or site, with
993 the number of stable samples in each site and dermotype shown at the bottom. **F.** Heatmap showing
994 significant correlations for dermotype labels across sites. The color gradient denotes log10-
995 transformed BH-adjusted Fisher's exact test p-values, and odds ratios are annotated in the boxes
996 where appropriate.

997 **Figure 2 | Dermotypes are characterized by distinct core and ancillary species. A.** Overview
998 of the general trend relating species prevalence and abundances (lower leg microbiomes as an
999 example), and the schema to categorize species into *core species* (≥50% prevalence, in blue) and
1000 *ancillary species* (≥3 occurrences, <10% prevalence, ≥95th percentile non-zero median abundance,
1001 in orange). **B.** Core species distribution by site and dermotype, with sizes corresponding to
1002 prevalences and color intensities corresponding to non-zero median abundance. Circles with red
1003 boundaries mark species with <0.5% median abundance. **C.** Left: Ancillary species distribution
1004 and non-zero median abundance at each site and dermotype. Right: Boxplots showing the intra-
1005 individual occurrence and median abundance of each ancillary species, with the number of
1006 individuals indicated on the right (each overlaid point represents one subject). The x-axis shows
1007 the number of sites (out of nine) where a given ancillary species is observed, and the color indicates
1008 its within-subject median abundance. **D.** Stacked barcharts depicting relative abundance profiles

1009 for selected subjects with blooms of *HPyV6* across multiple sites. **E.** Significant inter-site co-
1010 abundance patterns of *HPyV6*, with the color gradient denoting log10-transformed BH-adjusted
1011 Fisher's exact test p-values.

1012 **Figure 3 | Dermotypes exhibit varied species co-abundance patterns and interactions, and**
1013 **present different metabolic niches. A.** Species co-abundance networks for dermotypes in the
1014 axilla and volar forearm, respectively. Nodes represent microbial species, sized by relative
1015 abundance and colored by microbial genera (names are bolded for hub species). Squares denote
1016 species typically present in oral cavities. Edges denote significant Spearman correlations
1017 (CCREPE $|r| \geq 0.3$, p-value<0.05): purple for positive and orange for negative correlations. Thin
1018 lines indicate shared edges, while thick solid and thick dashed lines denote differential edges with
1019 $|r| \geq 0.5$ and $0.3 \leq |r| < 0.5$, respectively. **B.** Representative agar-based antagonism co-culture assay
1020 results showing *S. epidermidis* and *M. luteus* interactions with *S. hominis* under hypoxic and
1021 aerobic conditions. **C.** Presence/absence patterns for hub species identified from species co-
1022 abundance networks across dermotypes. **D.** Selected degradation-related metabolic pathways that
1023 are differentially abundant across dermotypes. **E-F**. Boxplots depicting the variation in
1024 community-level relative abundance of, (**E**) Lactose and galactose degradation I pathway, and (**F**)
1025 L-histidine degradation I pathway between two elbow crease dermotypes (Ac-1 & Ac-2), with
1026 corresponding HUMAnN2-derived species contributions summarized in stacked barcharts. P-
1027 values for comparisons across dermotypes were calculated using a one-sided Wilcoxon rank-sum
1028 test (∗∗∗∗ denotes p−value<0.0001), and boxplot whiskers represent 1.5× interquartile range or the
1029 maximum/minimum data point.

1030 **Figure 4 | Comprehensive investigation of host attributes, same-site and cross-site intrinsic-**
1031 **microbiome factors for their ability to predict dermotype labels. A.** Heatmap depicting
1032 associations between dermotypes and host demographic factors. Effect sizes are represented as
1033 odds ratios (gender, ethnicity; Fisher's exact test) or rank-biserial correlations (age;
1034 Wilcoxon test). **B.** Boxplots displaying variations in skin physiological measurements across
1035 dermotypes. **C.** Heatmap of beta coefficients from one-versus-rest logistic regression models, with
1036 corresponding p-value labeled on each tile, showing associations between dermotypes and selected
1037 personal care behaviors and skin conditions. **D.** AUC-ROC values for dermotype classifiers using
1038 random forest models trained on multi-domain host attributes. The left panel compares feature

importance of various demographic factors for the Ac dermotype classifier. **E.** Boxplots depicting AUC-ROC values for ridge classifiers trained on different categories of metabolic pathways. **F.** Top boxplots comparing AUC-ROC values for cross-site random forest classifiers trained on the full taxonomic profile versus top 10 abundant species from a single site, with the bottom heatmap showing feature importance for the top 10 species. **G.** Top barplot showing AUC-ROC values for the best-performing cross-site dermotype classifiers trained on dermotype labels from other skin sites, with the heatmap displaying combinations of dermotype labels used for training. **H.** Feature importance comparison in an inclusive ridge classifier trained on host demographic features, cross-site dermotype labels, and conserved metabolic pathway profiles. For inter-group comparisons of continuous variables (e.g., skin physiology, AUC-ROC, feature importance), p-values were calculated using a Wilcoxon rank-sum test, and boxplot whiskers represent 1.5× interquartile range or the maximum/minimum data point. For all subfigures $*$, $**$, $***$ and $****$ denote $p-\text{value}<0.05$, 0.01, 0.001 and 0.0001, respectively.

1052 **References**

1053 1. Oh, J. *et al.* Biogeography and individuality shape function in the human skin metagenome.
1054 *Nature* **514**, 59–64 (2014).
1055 2. Perez, G. I. P. *et al.* Body Site Is a More Determinant Factor than Human Population
1056 Diversity in the Healthy Skin Microbiome. *PLOS ONE* **11**, e0151990 (2016).
1057 3. Byrd, A. L., Belkaid, Y. & Segre, J. A. The human skin microbiome. *Nat Rev Microbiol* **16**,
1058 143–155 (2018).
1059 4. Roth, R. R. & James, W. D. Microbial Ecology of the Skin. *Annu. Rev. Microbiol.* **42**, 441–
1060 464 (1988).
1061 5. Oh, J., Byrd, A. L., Park, M., Kong, H. H. & Segre, J. A. Temporal Stability of the Human
1062 Skin Microbiome. *Cell* **165**, 854–866 (2016).
1063 6. Claesen, J. *et al.* A Cutibacterium acnes antibiotic modulates human skin microbiota
1064 composition in hair follicles. *Sci Transl Med* **12**, eaay5445 (2020).
1065 7. Nakatsuji, T. *et al.* Antimicrobials from human skin commensal bacteria protect against
1066 Staphylococcus aureus and are deficient in atopic dermatitis. *Sci Transl Med* **9**, eaah4680
1067 (2017).
1068 8. Bomar, L., Brugger, S. D., Yost, B. H., Davies, S. S. & Lemon, K. P. Corynebacterium
1069 accolens Releases Antipneumococcal Free Fatty Acids from Human Nostril and Skin
1070 Surface Triacylglycerols. *mBio* **7**, e01725-01715 (2016).
1071 9. Iwase, T. *et al.* Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm
1072 formation and nasal colonization. *Nature* **465**, 346–349 (2010).
1073 10. Lee, H. J. *et al.* Effects of cosmetics on the skin microbiome of facial cheeks with different
1074 hydration levels. *Microbiologyopen* **7**, e00557 (2017).

1075 11. Bouslimani, A. *et al.* The impact of skin care products on skin chemistry and microbiome
1076 dynamics. *BMC Biol* **17**, 47 (2019).
1077 12. Murphy, B. *et al.* Alteration of barrier properties, stratum corneum ceramides and
1078 microbiome composition in response to lotion application on cosmetic dry skin. *Sci Rep* **12**,
1079 5223 (2022).
1080 13. Fyolek, J. *et al.* Modulation of the Skin Microbiome Induced by High-Fat Diet in Children
1081 with Food Allergies and Eczema. *Journal of Allergy and Clinical Immunology* **153**, AB66
1082 (2024).
1083 14. Wang, L. *et al.* Facial Skin Microbiota-Mediated Host Response to Pollution Stress Revealed
1084 by Microbiome Networks of Individual. *mSystems* **6**, e00319-21.
1085 15. van Mierlo, M. M. F. *et al.* The influence of treatment in alpine and moderate maritime
1086 climate on the composition of the skin microbiome in patients with difficult to treat atopic
1087 dermatitis. *Clinical & Experimental Allergy* **49**, 1437–1445 (2019).
1088 16. Leung, M. H. Y. *et al.* Changes of the human skin microbiota upon chronic exposure to
1089 polycyclic aromatic hydrocarbon pollutants. *Microbiome* **8**, 100 (2020).
1090 17. Li, Z. *et al.* Characterization of the human skin resistome and identification of two
1091 microbiota cutotypes. *Microbiome* **9**, 47 (2021).
1092 18. Leung, M. H. Y. *et al.* Skin microbiome differentiates into distinct cutotypes with unique
1093 metabolic functions upon exposure to polycyclic aromatic hydrocarbons. *Microbiome* **11**,
1094 124 (2023).
1095 19. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180
1096 (2011).

1097 20. Wu, G. D. *et al.* Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes.
1098 *Science* **334**, 105–108 (2011).
1099 21. Costea, P. I. *et al.* Enterotypes in the landscape of gut microbial community composition.
1100 *Nat Microbiol* **3**, 8–16 (2018).
1101 22. Frioux, C. *et al.* Enterosignatures define common bacterial guilds in the human gut
1102 microbiome. *Cell Host Microbe* **31**, 1111-1125.e6 (2023).
1103 23. Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A*
1104 **108**, 4680–4687 (2011).
1105 24. Gajer, P. *et al.* Temporal Dynamics of the Human Vaginal Microbiota. *Sci Transl Med* **4**,
1106 132ra52 (2012).
1107 25. De Seta, F., Campisciano, G., Zanotta, N., Ricci, G. & Comar, M. The Vaginal Community
1108 State Types Microbiome-Immune Network as Key Factor for Bacterial Vaginosis and
1109 Aerobic Vaginitis. *Front Microbiol* **10**, 2451 (2019).
1110 26. France, M. T. *et al.* VALENCIA: a nearest centroid classification method for vaginal
1111 microbial communities based on composition. *Microbiome* **8**, 166 (2020).
1112 27. Lebeer, S. *et al.* A citizen-science-enabled catalogue of the vaginal microbiome and
1113 associated factors. *Nat Microbiol* **8**, 2183–2195 (2023).
1114 28. Tay, A. S. L. *et al.* Atopic dermatitis microbiomes stratify into ecologic dermotypes enabling
1115 microbial virulence and disease severity. *Journal of Allergy and Clinical Immunology* **147**,
1116 1329–1340 (2021).
1117 29. Alekseyenko, A. V. *et al.* Community differentiation of the cutaneous microbiota in
1118 psoriasis. *Microbiome* **1**, 31 (2013).

1119 30. Sun, C. *et al.* Integrated analysis of facial microbiome and skin physio-optical properties
1120 unveils cutotype-dependent aging effects. *Microbiome* **12**, 163 (2024).
1121 31. Wang, X. *et al.* The Health for Life in Singapore (HELIOS) Study: delivering Precision
1122 Medicine research for Asian populations. 2024.05.14.24307259 Preprint at
1123 https://doi.org/10.1101/2024.05.14.24307259 (2024).
1124 32. Liu, Y. *et al.* An optimal method for quantifying the facial sebum level and characterizing
1125 facial sebum features. *Skin Res Technol* **29**, e13454 (2023).
1126 33. Akdeniz, M., Gabriel, S., Lichterfeld-Kottner, A., Blume-Peytavi, U. & Kottner, J.
1127 Transepidermal water loss in healthy adults: a systematic review and meta-analysis update.
1128 *Br J Dermatol* **179**, 1049–1055 (2018).
1129 34. Grice, E. A. *et al.* Topographical and Temporal Diversity of the Human Skin Microbiome.
1130 *Science* **324**, 1190–1192 (2009).
1131 35. Chng, K. R. *et al.* Whole metagenome profiling reveals skin microbiome-dependent
1132 susceptibility to atopic dermatitis flare. *Nat Microbiol* **1**, 16106 (2016).
1133 36. Yue, J. C. & Clayton, M. K. A Similarity Measure Based on Species Proportions.
1134 *Communications in Statistics - Theory and Methods* **34**, 2123–2131 (2005).
1135 37. Theelen, B. *et al.* Malassezia ecology, pathophysiology, and treatment. *Medical Mycology*
1136 **56**, S10–S25 (2018).
1137 38. Gupta, A. K., Batra, R., Bluhm, R., Boekhout, T. & Dawson, T. L. Skin diseases associated
1138 with *Malassezia* species. *Journal of the American Academy of Dermatology* **51**, 785–798
1139 (2004).

1140 39. Fritz, S. A. *et al.* Staphylococcus aureus Colonization in Children With Community-
1141 Associated Staphylococcus aureus Skin Infections and Their Household Contacts. *Archives*
1142 *of Pediatrics & Adolescent Medicine* **166**, 551–557 (2012).
1143 40. Zhou, W. *et al.* Host-specific evolutionary and transmission dynamics shape the functional
1144 diversification of Staphylococcus epidermidis in human skin. *Cell* **180**, 454-470.e18 (2020).
1145 41. Chia, M. *et al.* Shared signatures and divergence in skin microbiomes of children with atopic
1146 dermatitis and their caregivers. *Journal of Allergy and Clinical Immunology* **150**, 894–908
1147 (2022).
1148 42. Han, G., Luong, H. & Vaishnava, S. Low abundance members of the gut microbiome exhibit
1149 high immunogenicity. *Gut Microbes* **14**, 2104086.
1150 43. Zhou, W. *et al.* Host-Specific Evolutionary and Transmission Dynamics Shape the
1151 Functional Diversification of *Staphylococcus epidermidis* in Human Skin. *Cell* **180**, 454-
1152 470.e18 (2020).
1153 44. Oh, J. *et al.* The altered landscape of the human skin microbiome in patients with primary
1154 immunodeficiencies. *Genome Res* **23**, 2103–2114 (2013).
1155 45. Zhou, X. *et al.* Longitudinal profiling of the microbiome at four body sites reveals core
1156 stability and individualized dynamics during health and disease. *Cell Host & Microbe* **32**,
1157 506-526.e9 (2024).
1158 46. Christensen, G. J. M. *et al.* Antagonism between Staphylococcus epidermidis and
1159 Propionibacterium acnes and its genomic basis. *BMC Genomics* **17**, 152 (2016).
1160 47. Nakamura, K. *et al.* Short chain fatty acids produced by Cutibacterium acnes inhibit biofilm
1161 formation by Staphylococcus epidermidis. *Sci Rep* **10**, 21237 (2020).

1162 48. Ahle, C. M. *et al.* Interference and co-existence of staphylococci and Cutibacterium acnes
1163 within the healthy human skin microbiome. *Commun Biol* **5**, 1–14 (2022).
1164 49. Kwaszewska, A., Sobiś-Glinkowska, M. & Szewczyk, E. M. Cohabitation—relationships of
1165 corynebacteria and staphylococci on human skin. *Folia Microbiol* **59**, 495–502 (2014).
1166 50. Nakazono, K. *et al.* Complete sequences of epidermin and nukacin encoding plasmids from
1167 oral-derived Staphylococcus epidermidis and their antibacterial activity. *PLoS One* **17**,
1168 e0258283 (2022).
1169 51. Huang, T. Y., Jiang, Y. E. & Scott, D. A. Culturable bacteria in the entire acne lesion and
1170 short-chain fatty acid metabolites of *Cutibacterium acnes* and *Staphylococcus epidermidis*
1171 isolates. *Biochemical and Biophysical Research Communications* **622**, 45–49 (2022).
1172 52. Mark Welch, J. L., Rossetti, B. J., Rieken, C. W., Dewhirst, F. E. & Borisy, G. G.
1173 Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National*
1174 *Academy of Sciences* **113**, E791–E800 (2016).
1175 53. Radaic, A. & Kapila, Y. L. The oralome and its dysbiosis: New insights into oral
1176 microbiome-host interactions. *Computational and Structural Biotechnology Journal* **19**,
1177 1335–1360 (2021).
1178 54. Gevers, D. *et al.* The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell*
1179 *Host & Microbe* **15**, 382–392 (2014).
1180 55. Yachida, S. *et al.* Metagenomic and metabolomic analyses reveal distinct stage-specific
1181 phenotypes of the gut microbiota in colorectal cancer. *Nat Med* **25**, 968–976 (2019).
1182 56. Chen, Y. E., Fischbach, M. A. & Belkaid, Y. Skin microbiota–host interactions. *Nature* **553**,
1183 427–436 (2018).

1184 57. Joglekar, P. *et al.* Integrated genomic and functional analyses of human skin–associated
1185 *Staphylococcus* reveal extensive inter- and intra-species diversity. *Proc. Natl. Acad. Sci.*
1186 *U.S.A.* **120**, e2310585120 (2023).
1187 58. Kaiser, J. C. & Heinrichs, D. E. Branching Out: Alterations in Bacterial Physiology and
1188 Virulence Due to Branched-Chain Amino Acid Deprivation. *mBio* **9**, e01188-18 (2018).
1189 59. Kaiser, J. C. *et al.* Repression of branched-chain amino acid synthesis in Staphylococcus
1190 aureus is mediated by isoleucine via CodY, and by a leucine-rich attenuator peptide. *PLOS*
1191 *Genetics* **14**, e1007159 (2018).
1192 60. Roux, P.-F., Oddos, T. & Stamatas, G. Deciphering the Role of Skin Surface Microbiome in
1193 Skin Health: An Integrative Multiomics Approach Reveals Three Distinct Metabolite–
1194 Microbe Clusters. *Journal of Investigative Dermatology* **142**, 469-479.e5 (2022).
1195 61. Danzberger, J. *et al.* Glycan distribution and density in native skin's stratum corneum. *Skin*
1196 *Res Technol* **24**, 450–458 (2018).
1197 62. Li, Z. *et al.* Integrated Human Skin Bacteria Genome Catalog Reveals Extensive Unexplored
1198 Habitat-Specific Microbiome Diversity and Function. *Advanced Science* **10**, 2300050 (2023).
1199 63. Kezic, S. *et al.* Levels of filaggrin degradation products are influenced by both filaggrin
1200 genotype and atopic dermatitis severity. *Allergy* **66**, 934–940 (2011).
1201 64. Hug, D. H., Dunkerson, D. D. & Hunter, J. K. The degradation of l-histidine and *trans* and
1202 *cis*-urocanic acid by bacteria from skin and the role of bacterial *cis*-urocanic acid isomerase.
1203 *Journal of Photochemistry and Photobiology B: Biology* **50**, 66–73 (1999).
1204 65. Walterscheid, J. P. *et al.* Cis-urocanic acid, a sunlight-induced immunosuppressive factor,
1205 activates immune suppression via the 5-HT2A receptor. *Proceedings of the National*
1206 *Academy of Sciences* **103**, 17420–17425 (2006).

1207 66. Tauch, A., Fernández-Natal, I. & Soriano, F. A microbiological and clinical review on
1208 *Corynebacterium kroppenstedtii*. *International Journal of Infectious Diseases* **48**, 33–39
1209 (2016).
1210 67. Liu, B. *et al.* IL-33/ST2 signaling excites sensory neurons and mediates itch response in a
1211 mouse model of poison ivy contact allergy. *Proceedings of the National Academy of*
1212 *Sciences* **113**, E7572–E7579 (2016).
1213 68. Trier, A. M. *et al.* IL-33 signaling in sensory neurons promotes dry skin itch. *Journal of*
1214 *Allergy and Clinical Immunology* **149**, 1473-1480.e6 (2022).
1215 69. Jeffery, I. B., Claesson, M. J., O'Toole, P. W. & Shanahan, F. Categorization of the gut
1216 microbiota: enterotypes or gradients? *Nat Rev Microbiol* **10**, 591–592 (2012).
1217 70. Knights, D. *et al.* Rethinking "Enterotypes". *Cell Host Microbe* **16**, 433–437 (2014).
1218 71. Lam, T. H. *et al.* Understanding the microbial basis of body odor in pre-pubescent children
1219 and teenagers. *Microbiome* **6**, 213 (2018).
1220 72. Urban, J. *et al.* The effect of habitual and experimental antiperspirant and deodorant product
1221 use on the armpit microbiome. *PeerJ* **4**, e1605 (2016).
1222 73. Spittaels, K.-J., Ongena, R., Zouboulis, C. C., Crabbé, A. & Coenye, T. Cutibacterium acnes
1223 Phylotype I and II Strains Interact Differently With Human Skin Cells. *Front Cell Infect*
1224 *Microbiol* **10**, 575164 (2020).
1225 74. Conwill, A. *et al.* Anatomy promotes neutral coexistence of strains in the human skin
1226 microbiome. *Cell Host & Microbe* **30**, 171-182.e7 (2022).
1227 75. Severn, M. M. & Horswill, A. R. Staphylococcus epidermidis and its dual lifestyle in skin
1228 health and infection. *Nat Rev Microbiol* **21**, 97–111 (2023).

1229 76. Evans, S. M., Schrlau, A. E., Chalian, A. A., Zhang, P. & Koch, C. J. Oxygen Levels in
1230 Normal and Previously Irradiated Human Skin as Assessed by EF5 Binding. *Journal of*
1231 *Investigative Dermatology* **126**, 2596–2606 (2006).
1232 77. Wong, W. J., Richardson, T., Seykora, J. T., Cotsarelis, G. & Simon, M. C. Hypoxia-
1233 Inducible Factors Regulate Filaggrin Expression and Epidermal Barrier Function. *Journal of*
1234 *Investigative Dermatology* **135**, 454–461 (2015).
1235 78. Cheung, G. Y. C., Joo, H.-S., Chatterjee, S. S. & Otto, M. Phenol-soluble modulins – critical
1236 determinants of staphylococcal virulence. *FEMS microbiology reviews* **38**, 698 (2014).
1237 79. Ch'ng, J.-H. *et al.* Heme cross-feeding can augment *Staphylococcus aureus* and
1238 *Enterococcus faecalis* dual species biofilms. *The ISME Journal* **16**, 2015–2026 (2022).
1239 80. Regina, V. R. *et al.* Decoding scalp health and microbiome dysbiosis in dandruff.
1240 2024.05.02.592279 Preprint at https://doi.org/10.1101/2024.05.02.592279 (2024).
1241 81. Seo, Y., Sung, M., Hwang, J. & Yoon, Y. Minimum Inhibitory Concentration (MIC) of
1242 Propionic Acid, Sorbic Acid, and Benzoic Acid against Food Spoilage Microorganisms in
1243 Animal Products to Use MIC as Threshold for Natural Preservative Production. *Food Sci*
1244 *Anim Resour* **43**, 319–330 (2023).
1245 82. Pagac, M. P. *et al.* Life stage impact on the human skin ecosystem: lipids and the microbial
1246 community. *NPJ Biofilms Microbiomes* **11**, 13 (2025).
1247 83. Chia, M. *et al.* Skin metatranscriptomics reveals landscape of variation in microbial activity
1248 and gene expression across the human body. 2024.12.02.626500 Preprint at
1249 https://doi.org/10.1101/2024.12.02.626500 (2024).
1250 84. Li, M. *et al.* Microbiome and lipidomic analysis reveal the interplay between skin bacteria
1251 and lipids in a cohort study. *Front Microbiol* **15**, 1383656 (2024).

1252 85. Martin, A. *et al.* A Functional ABCC11 Allele Is Essential in the Biochemical Formation of
1253 Human Axillary Odor. *Journal of Investigative Dermatology* **130**, 529–540 (2010).
1254 86. Pappas, A., Fantasia, J. & Chen, T. Age and ethnic variations in sebaceous lipids.
1255 *Dermatoendocrinol* **5**, 319–324 (2013).
1256 87. Pierard, G. E., Arrese, J. E., Pierard-Franchimont, C. & DE Doncker, P. Prolonged effects of
1257 antidandruff shampoos - time to recurrence of Malassezia ovalis colonization of skin. *Int J*
1258 *Cosmet Sci* **19**, 111–117 (1997).
1259 88. Urban, J. *et al.* The effect of habitual and experimental antiperspirant and deodorant product
1260 use on the armpit microbiome. *PeerJ* **4**, e1605 (2016).
1261 89. Grice, E. A. & Dawson, T. L. Host–microbe interactions: *Malassezia* and human skin.
1262 *Current Opinion in Microbiology* **40**, 81–87 (2017).
1263 90. Ianiri, G., LeibundGut-Landmann, S. & Dawson, T. L. Malassezia: A Commensal, Pathogen,
1264 and Mutualist of Human and Animal Skin. *Annual Review of Microbiology* **76**, 757–782
1265 (2022).
1266 91. Hong, C. *et al.* PathoScope 2.0: a complete computational framework for strain identification
1267 in environmental or clinical sequencing samples. *Microbiome* **2**, 33 (2014).
1268 92. Janssens, Y. *et al.* Disbiome database: linking the microbiome to disease. *BMC Microbiology*
1269 **18**, 50 (2018).
1270 93. Yao, G. *et al.* MicroPhenoDB Associates Metagenomic Data with Pathogenic Microbes,
1271 Microbial Core Genes, and Human Disease Phenotypes. *Genomics Proteomics*
1272 *Bioinformatics* **18**, 760–772 (2020).
1273 94. de Goffau, M. C. *et al.* Recognizing the reagent microbiome. *Nat Microbiol* **3**, 851–853
1274 (2018).

95. Peschel, S., Müller, C. L., Von Mutius, E., Boulesteix, A.-L. & Depner, M. NetCoMi: network construction and comparison for microbiome data in R. *Briefings in Bioinformatics* **22**, bbaa290 (2021).

96. Csárdi, G. *et al.* igraph for R: R interface of the igraph library for graph theory and network analysis. Zenodo https://doi.org/10.5281/ZENODO.7682609 (2024).

97. Faust, K. *et al.* Microbial Co-occurrence Relationships in the Human Microbiome. *PLOS Computational Biology* **8**, e1002606 (2012).

98. Gill, R., Datta, S. & Datta, S. A statistical framework for differential network analysis from microarray data. *BMC Bioinformatics* **11**, 95 (2010).

99. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Int. Res.* **16**, 321–357 (2002).

100. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

101. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (ACM, San Francisco California USA, 2016). doi:10.1145/2939672.2939785.

Figure 1

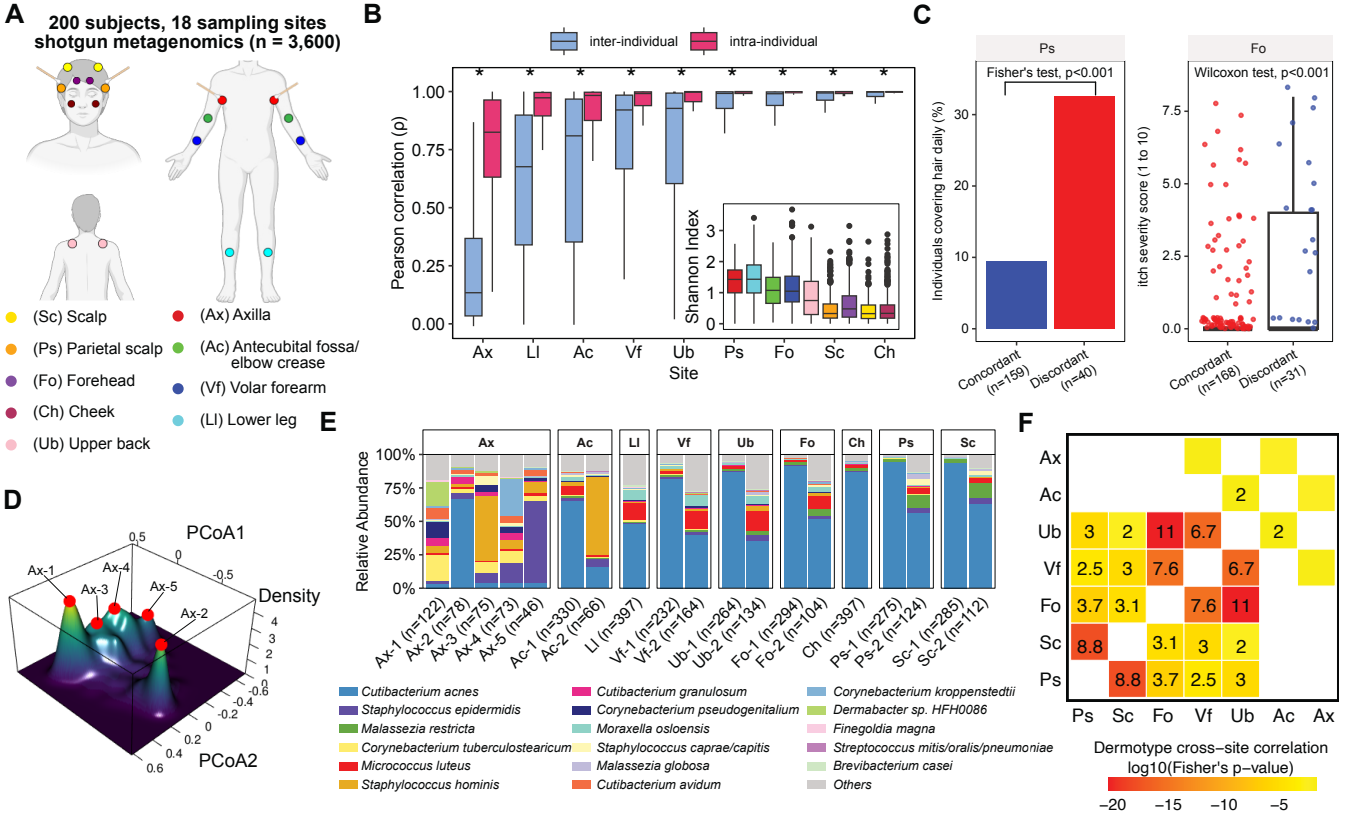

A
200 subjects, 18 sampling sites
shotgun metagenomics (n = 3,600)
(Sc) Scalp
(Ps) Parietal scalp
(Fo) Forehead
(Ch) Cheek
(Ub) Upper back
(Ax) Axilla
(Ac) Antecubital fossa/ elbow crease
(Vf) Volar forearm
(Ll) Lower leg
B
inter-individual
intra-individual
Pearson correlation (ρ)
Shannon Index
Ax
Ll
Ac
Vf
Ub
Ps
Fo
Sc
Ch
Site
C
Ps
Fo
Fisher's test, p<0.001
Wilcoxon test, p<0.001
Individuals covering hair daily (%)
itch severity score (1 to 10)
Concordant (n=159)
Discordant (n=40)
Concordant (n=168)
Discordant (n=31)
D
PCoA1
PCoA2
Density
Ax-1
Ax-2
Ax-3
Ax-4
Ax-5
E
Relative Abundance
Ax-1 (n=122)
Ax-2 (n=78)
Ax-3 (n=75)
Ax-4 (n=73)
Ax-5 (n=46)
Ac-1 (n=330)
Ac-2 (n=66)
Ll (n=397)
Vf-1 (n=232)
Vf-2 (n=164)
Ub-1 (n=264)
Ub-2 (n=134)
Fo-1 (n=294)
Fo-2 (n=104)
Ch (n=397)
Ps-1 (n=275)
Ps-2 (n=124)
Sc-1 (n=285)
Sc-2 (n=112)
Cutibacterium acnes
Cutibacterium granulosum
Corynebacterium kroppenstedtii
Staphylococcus epidermidis
Corynebacterium pseudogenitalium
Dermabacter sp. HFH0086
Malassezia restricta
Moraxella osloensis
Finegoldia magna
Corynebacterium tuberculostearicum
Staphylococcus caprae/capitis
Streptococcus mitis/oralis/pneumoniae
Micrococcus luteus
Malassezia globosa
Brevibacterium casei
Staphylococcus hominis
Cutibacterium avidum
Others
F
Dermotype cross–site correlation
log10(Fisher's p–value)

Figure 2

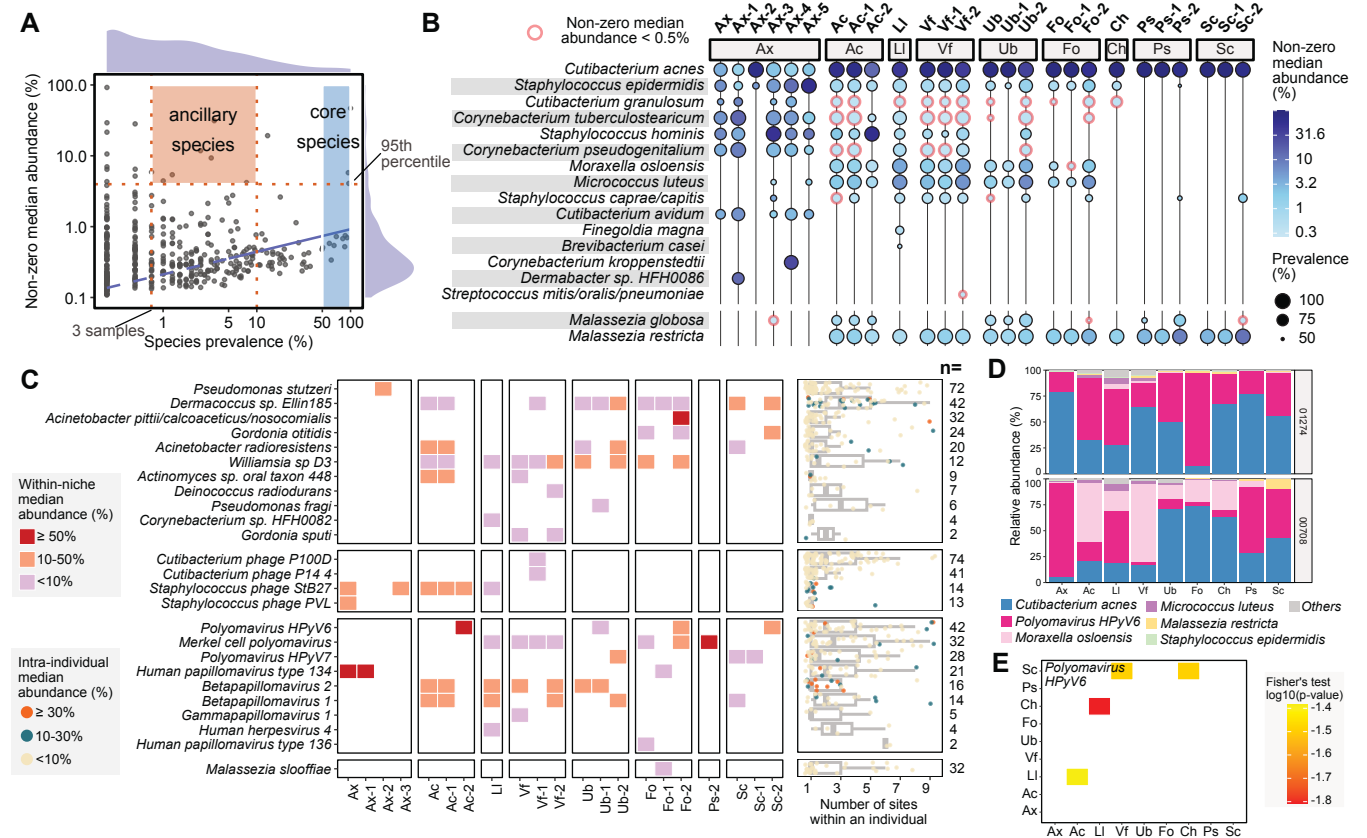

A
Non-zero median abundance (%)
100.0
10.0
1.0
0.1
ancillary species
core species
95th percentile
3 samples
1
5
10
50
100
Species prevalence (%)
B
Non-zero median abundance < 0.5%
Ax
Ac
LI
Vf
Ub
Fo
Ch
Ps
Sc
Cutibacterium acnes
Staphylococcus epidermidis
Cutibacterium granulosum
Corynebacterium tuberculostearicum
Staphylococcus hominis
Corynebacterium pseudogenitalium
Moraxella osloensis
Micrococcus luteus
Staphylococcus caprae/capitis
Cutibacterium avidum
Finegoldia magna
Brevibacterium casei
Corynebacterium kroppenstedtii
Dermabacter sp. HFH0086
Streptococcus mitis/oralis/pneumoniae
Malassezia globosa
Malassezia restricta
Non-zero median abundance (%)
31.6
10
3.2
1
0.3
Prevalence (%)
100
75
50
C
Within-niche median abundance (%)
≥ 50%
10-50%
<10%
Intra-individual median abundance (%)
≥ 30%
10-30%
<10%
Pseudomonas stutzeri
Dermacoccus sp. Ellin185
Acinetobacter pittii/calcoaceticus/nosocomialis
Gordonia otitidis
Acinetobacter radioresistens
Williamsia sp D3
Actinomyces sp. oral taxon 448
Deinococcus radiodurans
Pseudomonas fragi
Corynebacterium sp. HFH0082
Gordonia sputi
Cutibacterium phage P100D
Cutibacterium phage P14 4
Staphylococcus phage StB27
Staphylococcus phage PVL
Polyomavirus HPyV6
Merkel cell polyomavirus
Polyomavirus HPyV7
Human papillomavirus type 134
Betapapillomavirus 2
Betapapillomavirus 1
Gammapapillomavirus 1
Human herpesvirus 4
Human papillomavirus type 136
Malassezia slooffiae
n=
72
42
32
24
20
12
9
7
6
4
2
74
41
14
13
42
32
28
21
16
14
5
4
2
32
Number of sites within an individual
D
Relative abundance (%)
01274
00708
Cutibacterium acnes
Polyomavirus HPyV6
Moraxella osloensis
Micrococcus luteus
Malassezia restricta
Staphylococcus epidermidis
Others
E
Polyomavirus HPyV6
Fisher's test log10(p-value)
-1.4
-1.5
-1.6
-1.7
-1.8

Figure 3

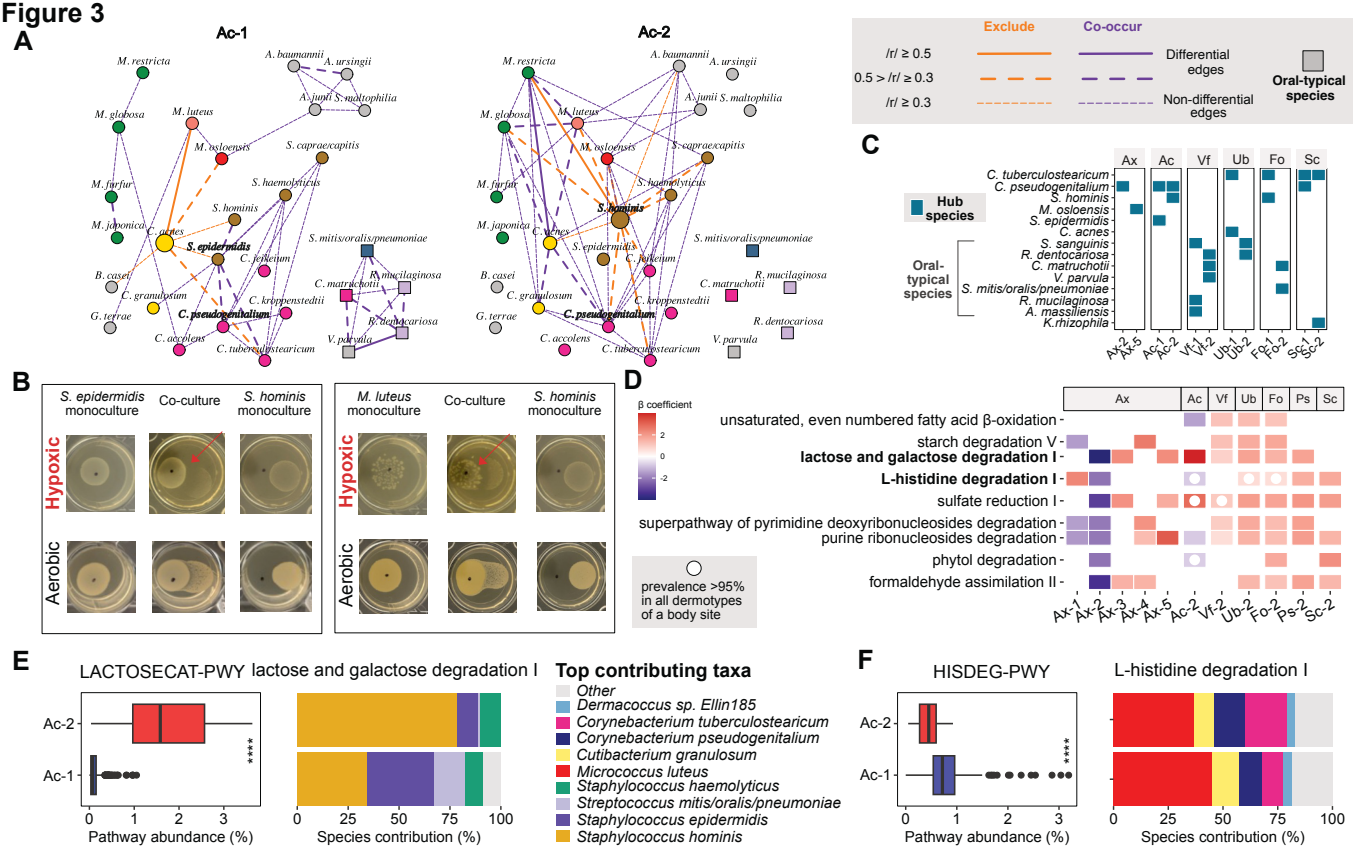
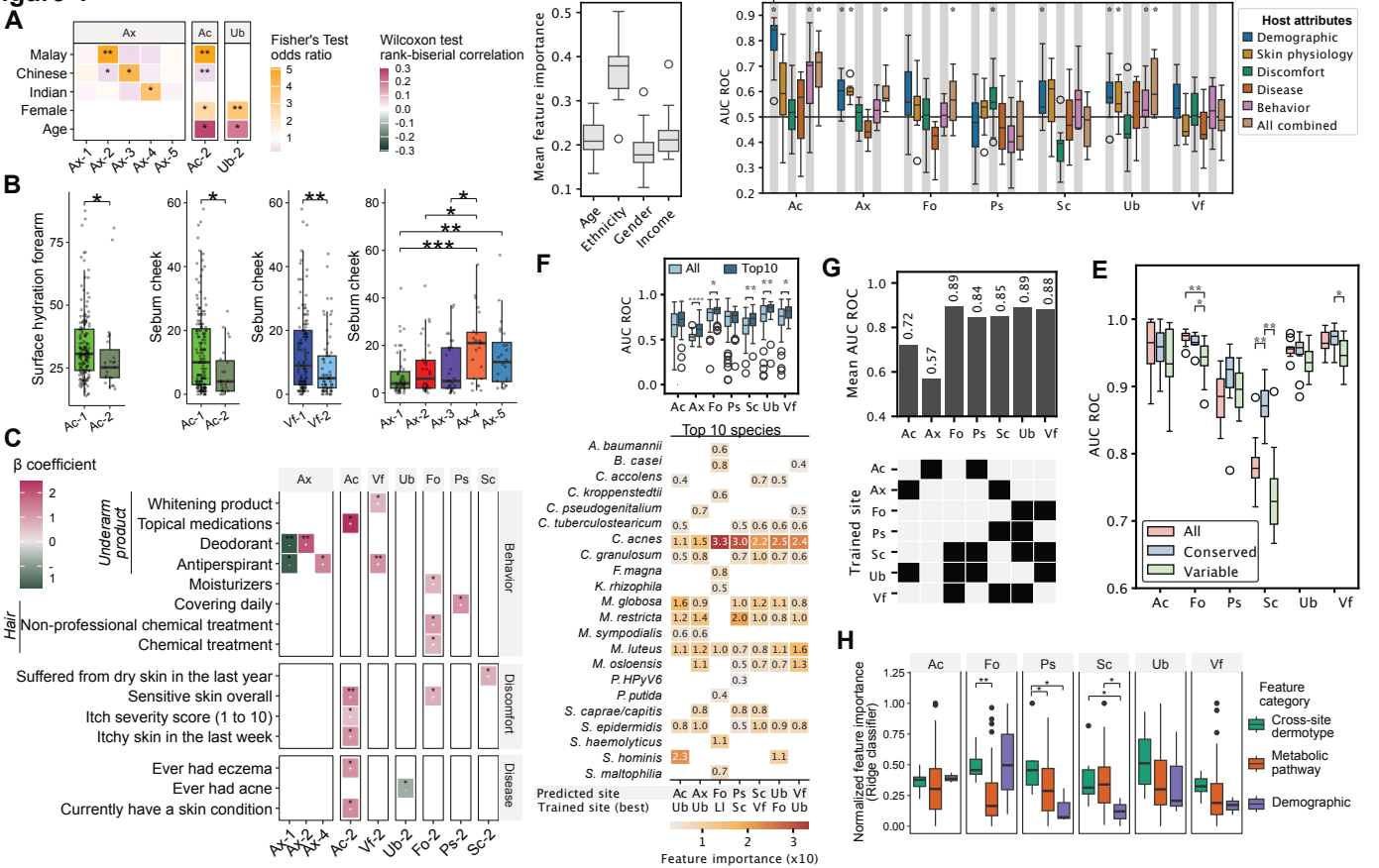

A
Ac-1
Ac-2
M. restricta
A. baumannii
A. ursingii
A. junii
S. maltophilia
M. globosa
M. luteus
M. osloensis
S. caprae/capitis
M. furfur
S. haemolyticus
S. hominis
M. japonica
C. acnes
S. epidermidis
S. mitis/oralis/pneumoniae
C. jeikeium
B. casei
R. mucilaginosa
C. matruchotii
C. granulosum
C. kroppenstedtii
G. terrae
C. pseudogenitalium
R. dentocariosa
C. accolens
C. tuberculostearicum
V. parvula
Exclude
Co-occur
/r/ ≥ 0.5
0.5 > /r/ ≥ 0.3
/r/ ≥ 0.3
Differential edges
Non-differential edges
Oral-typical species
B
S. epidermidis monoculture
Co-culture
S. hominis monoculture
M. luteus monoculture
Hypoxic
Aerobic
C
Hub species
Oral-typical species
Ax
Ac
Vf
Ub
Fo
Sc
C. tuberculostearicum
C. pseudogenitalium
S. hominis
M. osloensis
S. epidermidis
C. acnes
S. sanguinis
R. dentocariosa
C. matruchotii
V. parvula
S. mitis/oralis/pneumoniae
R. mucilaginosa
A. massiliensis
K.rhizophila
Ax-2
Ax-5
Ac-1
Ac-2
Vf-1
Vf-2
Ub-1
Ub-2
Fo-1
Fo-2
Sc-1
Sc-2
D
β coefficient
2
0
-2
prevalence >95% in all dermotypes of a body site
Ax
Ac
Vf
Ub
Fo
Ps
Sc
unsaturated, even numbered fatty acid β-oxidation
starch degradation V
lactose and galactose degradation I
L-histidine degradation I
sulfate reduction I
superpathway of pyrimidine deoxyribonucleosides degradation
purine ribonucleosides degradation
phytol degradation
formaldehyde assimilation II
Ax-1
Ax-2
Ax-3
Ax-4
Ax-5
Ac-2
Vf-2
Ub-2
Fo-2
Ps-2
Sc-2
E
LACTOSECAT-PWY
lactose and galactose degradation I
Ac-2
Ac-1
****
0
1
2
3
Pathway abundance (%)
0
25
50
75
100
Species contribution (%)
Top contributing taxa
Other
Dermacoccus sp. Ellin185
Corynebacterium tuberculostearicum
Corynebacterium pseudogenitalium
Cutibacterium granulosum
Micrococcus luteus
Staphylococcus haemolyticus
Streptococcus mitis/oralis/pneumoniae
Staphylococcus epidermidis
Staphylococcus hominis
F
HISDEG-PWY
L-histidine degradation I
Ac-2
Ac-1
****
0
1
2
3
Pathway abundance (%)
0
25
50
75
100
Species contribution (%)

Figure 4


A
Ax
Ac
Ub
Malay
Chinese
Indian
Female
Age
Ax-1
Ax-2
Ax-3
Ax-4
Ax-5
Ac-2
Ub-2
Fisher's Test odds ratio
Wilcoxon test rank-biserial correlation
B
Surface hydration forearm
Sebum cheek
Ac-1
Ac-2
Vf-1
Vf-2
Ax-1
Ax-2
Ax-3
Ax-4
Ax-5
C
β coefficient
Underarm product
Hair
Whitening product
Topical medications
Deodorant
Antiperspirant
Moisturizers
Covering daily
Non-professional chemical treatment
Chemical treatment
Suffered from dry skin in the last year
Sensitive skin overall
Itch severity score (1 to 10)
Itchy skin in the last week
Ever had eczema
Ever had acne
Currently have a skin condition
Behavior
Discomfort
Disease
Ax
Ac
Vf
Ub
Fo
Ps
Sc
D
Mean feature importance
Age
Ethnicity
Gender
Income
AUC ROC
Ac
Ax
Fo
Ps
Sc
Ub
Vf
Host attributes
Demographic
Skin physiology
Discomfort
Disease
Behavior
All combined
F
All
Top10
AUC ROC
Top 10 species
A. baumannii
B. casei
C. accolens
C. kroppenstedtii
C. pseudogenitalium
C. tuberculostearicum
C. acnes
C. granulosum
F. magna
K. rhizophila
M. globosa
M. restricta
M. sympodialis
M. luteus
M. osloensis
P. HPyV6
P. putida
S. caprae/capitis
S. epidermidis
S. haemolyticus
S. hominis
S. maltophilia
Predicted site
Trained site (best)
Ac Ax Fo Ps Sc Ub Vf
Ub Ub Ll Sc Vf Fo Ub
Feature importance (x10)
G
Mean AUC ROC
0.72
0.57
0.89
0.84
0.85
0.89
0.88
Trained site
E
AUC ROC
All
Conserved
Variable
H
Normalized feature importance (Ridge classifier)
Feature category
Cross-site dermotype
Metabolic pathway
Demographic