Methods for the analysis of skin microbiomes: a comparison of sampling processes and 16S rRNA hypervariable regions.

Authors: Alyami R. Y.[1], Cleary D. W.[1,2], Forster J.[3], Feelisch M.[1,2], Ardern-Jones M. R.[1]*

Author Affiliations:

[1] School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK, SO16 6YD

[2] Southampton NIHR Biomedical Research Centre, University Hospital Southampton NHS Trust, Southampton, UK, SO16 6YD

[3] Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, UK, SO16 6YD

***Corresponding author**

Ardern-Jones M. R.

m.aj@soton.ac.uk

Running head:

Optimisation of skin microbiome analysis

Funding

RA was funded by a grant from the Royal Embassy of Saudi Arabia. MF acknowledges funding from the Medical Research Council UK, the Wellcome Trust and AOBiome. None of these funding sources had any input in the direction of study, selection of results and/or the decision to publish.

Declaration of Interests

None of the interests described below are considered to be conflicts with regard the work presented here. This work was carried out in partial fulfilment of RYA's PhD thesis supported by the Royal Embassy of Saudi Arabia. DWC acknowledges support from Pfizer and GSK. MF was a consultant and member of the Scientific Advisory Board of AOBiome. MRAJ has been an adviser/consultant/speaker for AbbVie, Pfizer, Sanofi Genzyme, Lilly, Ducentis, Hosei Septares, Unilever.

What's already known about this topic?

- Previous microbiome studies have utilised many different methodologies, but a detailed comparison of their utility in skin research has not been undertaken
- Methodological choices in skin microbiome analysis affect both sensitivity to detect diversity and sensitivity to identify at a species level
- Optimisation to detect diversity and species identification can require different approaches

What does this study add?

- Our study compared a matrix the four most widely utilised skin sampling methodologies and the two principle means for hypervariable gene analysis.
- We show that alpha diversity is optimal with tape sampling
- Additionally, in contrast to gut microbiome analyses, in the skin hypervariable region V1-3 analysis is superior to V3-4 when the species of interest are staphylococci.

Abstract:

Background

The skin microbiome is increasingly recognised as a critical component of the innate skin immune response and is central to the pathogenesis of many inflammatory skin disorders including atopic dermatitis. Previous studies have not looked in detail at the impact of changing both sampling methodology and hypervariable region sequencing for skin microbiome analysis.

Objectives

We set out to undertake a detailed analysis comparing microbe population diversity and resolution at the single species level by swab, tape, scrape and scrape then swab sampling assayed by hypervariable region sequencing.

Methods

In triplicate samples from the antecubital fossa were taken from healthy volunteers for 16s RNA analysis by primer amplification for hypervariable regions 1-3, and 3-4.

Results

Alpha (phylogenetic) diversity was the greatest with tape sampling and V1-3 analysis, whereas for V3-4, tape and swab were equivalent. Scrape sampling showed lower alpha diversity at both V1-3 and V3-4. All measures of beta diversity showed the scrape methodology yielded a lesser diversity than the others. Phyla composition was similar across all sampling methodologies. Minor differences in composition were noted between V1-3 and V3-4 sequencing, but V1-3 was optimal for identification of firmicutes (including staphylococci).

Conclusions

In the methodological planning of skin microbiome analysis skin scientists need to consider the microbes of interest to choose the optimal hypervariable region to sequence, and harmonisation of methodological approaches would be beneficial to the field. For detection of staphylococci on flexural skin, we would recommend tape sampling with analysis of V1-3.

85 **Introduction**

86

87 The skin is among the largest organs of the human body and is populated by a diverse array
88 of microorganisms (1). There is increasing evidence that the skin microbiome plays a crucial
89 role in the defence against pathogens, the education of the immune system, and skin
90 homoeostasis (1, 2). Unsurprisingly, shifts in skin microbiota composition are associated with
91 skin disorders like acne, atopic dermatitis, eczema, and psoriasis (3). For example, the
92 presence and/or abundance of pathobionts such as *Staphylococcus aureus* have been shown
93 to correlate with disease severity of atopic dermatitis (3).

94

95 To better understand the microbiome-host crosstalk in the context of specific cutaneous
96 diseases, and as an avenue to assess potential intervention strategies, it is important to
97 develop reliable tools for characterisation of the composition of the skin microbiome beyond
98 the limitations of culture. Whilst several molecular profiling approaches based on high-
99 throughput sequencing of microbial 16S ribosomal RNA have proved successful in addressing
100 this (4, 5), specific technical challenges remain such as low microbial biomass and the impact
101 of extensive background human DNA material that have implications for choosing the most
102 adequate method to reconstruct the microbial community composition.

103

104 The first decision investigators have to make is to select the most appropriate sampling
105 methodology. Although, arguably, skin biopsy would be ideal and allow further spatial
106 selection of compartments within skin (e.g. using laser-capture microscopy), it is an invasive
107 procedure, complicating sampling from multiple body regions of the same individual (6, 7).
108 Whilst previous reports of sampling using pre-moistened swabbing (4, 8, 9) suggest that the
109 relative proportions of bacterial species this technique recovers are comparable to a skin
110 biopsy (6), alternative methods for skin microbiome sampling may be preferable to wet
111 swabbing. Moreover, it is challenging to minimise methodological variation with wet
112 swabbing for quantitative analyses. For example, it is difficult to control for swab pressure,
113 frequency and direction of swabbing (10). In contrast, tape adhesion sampling, is a
114 straightforward technique for investigators, which yields high sample biomass and uniformly
115 samples a specific surface area of the skin (11, 12). Skin scrape methods further increase the
116 captured biomass and are therefore likely to be useful for lower abundance microorganisms
117 (6, 13). Cup scrubbing and cylinder suspension are alternative sampling techniques that have

been reported to yield higher levels of viable bacteria from a well-defined surface area compared to traditional swabbing methods (14, 15).

The second decision to make before engaging in a skin microbiome study relates to the molecular methodology employed for sequencing. One of the main challenges for taxonomic analysis of 16S rRNA sequencing data is the relatively short read lengths (up to 300bp for single-end and 600bp for paired-end) obtained from the most commonly employed next-generation sequencing approaches. Thus, sequence resolution down to species level is enhanced by choosing the most appropriate hypervariable region of the 16S rRNA gene relevant to the expected microbiota in the sampled region and is therefore to some degree dependent upon the tissue of interest. In the skin, differentiation within a genus is critically important because of the recognised differences between microbial species. For example, within staphylococci, distinguishing between homeostasis regulating *S. epidermidis* and inflammation inducing *S. aureus (SA)*is important for many studies (16). Of the nine 16S rRNA gene hypervariable regions (19), V1 to V9, region four (V4) has been used most widely in studying human gastrointestinal tract microbiome. However, recent work has shown that 16S rRNA V1-3 region demonstrate a good taxonomic classification result including an increased resolution for *Staphylococcus'* species (17). Furthermore, V1-3 has the advantage of showing similar skin microbial communities and genetic functional profiles to whole metagenome shotgun (WMS) sequencing result as compared to V4 region (5, 18) and is more sensitive at detecting other skin microbiota (e.g. Propionibacterium)(19).

Previous analysis using V4 sequencing of skin has shown good concordance between swabbing, tape stripping and cup scrubs (12), but because different biomass yields are returned by the various sampling methodologies, it is likely that the choice of sampling will interact with the choice of 16S rRNA sequencing target. As a result, detailed comparisons between sampling methodology and 16S rRNA target are important.

The aim of the present study was to compare microbiome composition using two 16S rRNA hypervariable molecular targets across sampling techniques, to select the optimal methodology for identification of skin microbial diversity.

146 **Methods**

147 **Ethics**

148 Ethical approval was obtained from the Northern Ireland regional ethics committee (NRES
149 15/NI/0180). Local ethical approval was granted by the University of Southampton (ERGO ID
150 17473) and the University Hospital Southampton NHS Foundation Trust (R&D CRI 0320).

151 **Skin Microbiota Sampling**

152 Four healthy volunteers (male, aged 26-40) were recruited for this study. Four methodologies,
153 each performed in triplicate, were compared for sampling from the same antecubital fossa.
154 Relevant control samples were analysed for each sampling method. Prior to sampling, the
155 investigator's gloves were treated with 70% v/v EtOH and air dried prior to contact with the
156 subject.

157 Method 1 - Swab: Tubed sterile dry tipped swabs™ (MWE medical wire®, UK), coated in 0.15
158 M NaCl + 0.1% Tween20 (SigmaAldrich, UK) were turned back and forth over a defined 4 cm^2
159 area of skin for 30 secs.

160 Method 2 - Scrape: A 4 cm^2 area of skin adjacent to the swabbed region was outlined and then
161 scraped gently with a surgical scalpel blade (No.15, Williams Medical, UK) with the blade curve
162 horizontal to the skin surface to obtain microbial DNA(9). Materials were collected from the
163 blade using sterile dry swab tips (MWE medical wire®, UK), coated in 0.15 M NaCl + 0.1%
164 Tween20.

165 Method 3 - Swab after scrape: Following the scrape (Method 2), the same area was swabbed
166 exactly as described in Method 1.

167 Method 4 – Tape adhesion: Sampling tape (D-Squame, CuDerm Corporation) was sterilised by
168 submersion in 70% v/v EtOH. The adhesive portion of the 4cm^2 tape disc was placed on the
169 surface of the skin and even pressure applied down with the gloved thumb. The disc was then
170 peeled off the skin and the same single tape disc was reapplied to the same area of skin 50
171 times over a period of 2 min.

172 **DNA Extraction**

173 Sterile forceps were used to place swab tip or sealed tape into a PowerBead® DNA extraction
174 tube (Qiagen, UK). Total DNA extraction was accomplished using PowerSoil® DNA Isolation Kit
175 (Qiagen, UK) according to the manufacturer's instructions.

176 **16S rRNA PCR**

177 The 16S rRNA PCR was done using primers 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 534R
178 (5'-GTGCCAGCAGCCGCGGTAA-3') for variable regions 1 to 3 (V1-3) and 515F (5'-
179 GTGCCAGCMGCCGCGGTAA-3') and 806R (5'- GGACTACCGGGGTATCT-3') for V3-4 (20). PCR

180 was performed using a GeneAmp® PCR System 9700 (Applied Biosystems, UK) with the
181 following cycling conditions: 3 min of denaturation at 98°C followed by 35 cycles of 98°C for
182 30 secs, 55°C of annealing for 30 secs, and extension at 72°C for 30 secs, followed by a final
183 extension at 72°C for 5 min. Each sample was amplified in triplicate 25 µl reactions consisting
184 of primers at final concentrations of 0.2 µM and 1X PCR reagent (Q5® Hot Start High-Fidelity
185 2X Master Mix, New England BioLabs, UK). Triplicate PCR reactions for each sample were then
186 pooled and purified using Agencourt® AMPure® XP beads (#A63881, Beckman Coulter, UK).

187 **Library Preparation**

188 Adapter ligation and barcoding were performed using the Nextera XT Index Kit (#FC-131-1002,
189 Illumina®, UK). PCR conditions involved 3 min of denaturation at 98°C, then 12 cycles at 98°C
190 for 30 secs, 55°C of annealing for 30 secs, 30 secs of extension at 72°C, followed by a final
191 extension at 72°C for 5 minutes. The number of PCR cycles was kept low to reduce
192 amplification bias, and reduce the risk of amplifying contaminants present in kit reagents. This
193 PCR amplification was followed by another purification using the Agencourt® AMPure® XP
194 beads (#A63881, Beckman Coulter, UK). A ratio of 6.5 of Solid Phase Reversible Immobilisation
195 beads (SPRI) to DNA was used for size selection. The library was quantified using a Qubit™
196 double stranded DNA (High Sensitivity) Assay Kit (#Q32866, Life Technologies Ltd, UK) with
197 the Qubit® 2.0 Fluorometer (Life Technologies Ltd, UK).

198 **Sequencing**

199 The DNA library sample pool (2nM) was denatured using 0.2N NaOH, diluted to 10pM, and
200 combined with 25% denatured 4nM PhiX. Samples were sequenced locally (within the Wessex
201 Investigational Sciences Hub (WISH) facility, University Hospital Southampton NHS Trust)
202 using a MiSeq (Illumina®, UK) to generate 2 x 300 bp paired-end reads using V3 chemistry,
203 following standard sequencing protocols.

204 **Analysis**

205 Initial processing and analysis was performed using the Quantitative Insights in Microbial
206 Ecology (QIIME2) package (21). Individual forward and reverse read files were imported into
207 QIIME2 as a single merged file using the provided import tool. Reads were demultiplexed and
208 the trimmed using a quality score threshold of <Q35. Only forward reads were processed due
209 due to low quality of reverse reads. After trimming, the reads were denoised using DADA2
210 (22), exact amplicon sequence variants (ASVs) inferred, aligned and finally given a taxonomic
211 label using the Greengenes database (release 2013-08; gg_13_8_otus). The output QIIME
212 artefacts, namely the feature table, taxonomy, phylogenetic tree and metadata, were then
213 used to create a phyloseq object using qza_to_phyloseq from the package qiime2R (23)(v

214 0.99.12). Subsequent analysis was done using phyloseq (24) (v 1.28.0). Contaminant ASVs
215 were identified in the controls and excluded using the R package 'decontam' (25). All analysis
216 was done in R (v 3.6.1) using RStudio (v 1.2.5019).

217 **Statistical Analysis**

218 Alpha diversity were tested using (one-way ANOVA, $p > 0.05$) and beta dissimilarity score
219 tested using PERMANOVA test, $p > 0.05$) in R (v 3.6.1) using RStudio (v 1.2.5019). Figures
220 were produced using the package gplot2 (v.3.3.3).

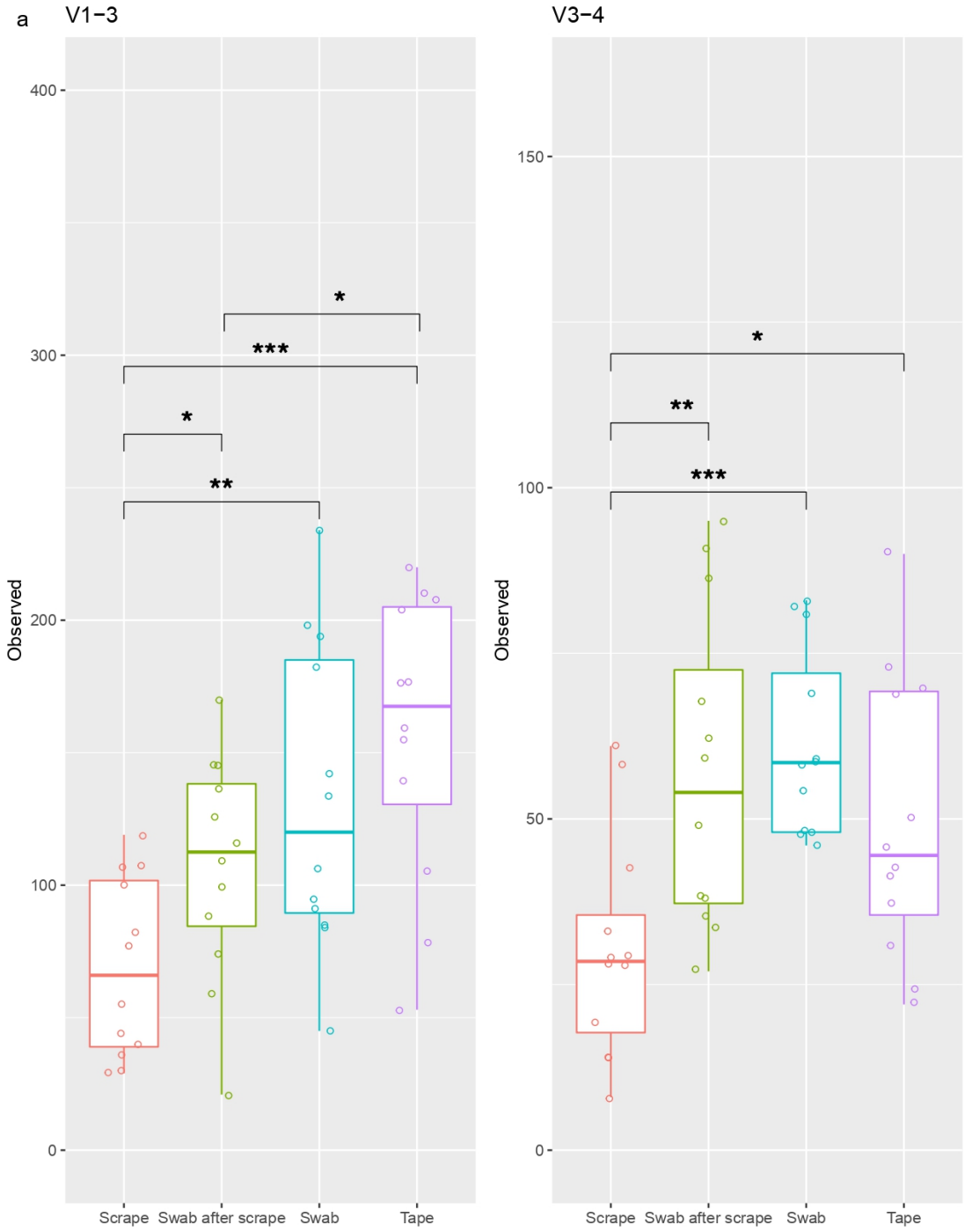
221 Further analysis was conducted using GraphPAD Prism® software (v9, GraphPad Software,
222 Inc., La Jolla, CA, USA). Statistical difference was evaluated using paired two-tailed t-test and
223 one-way analysis of variance (one-way ANOVA) test to determine any statistically significant
224 differences between groups and results were considered significant when $p<0.05$.

225

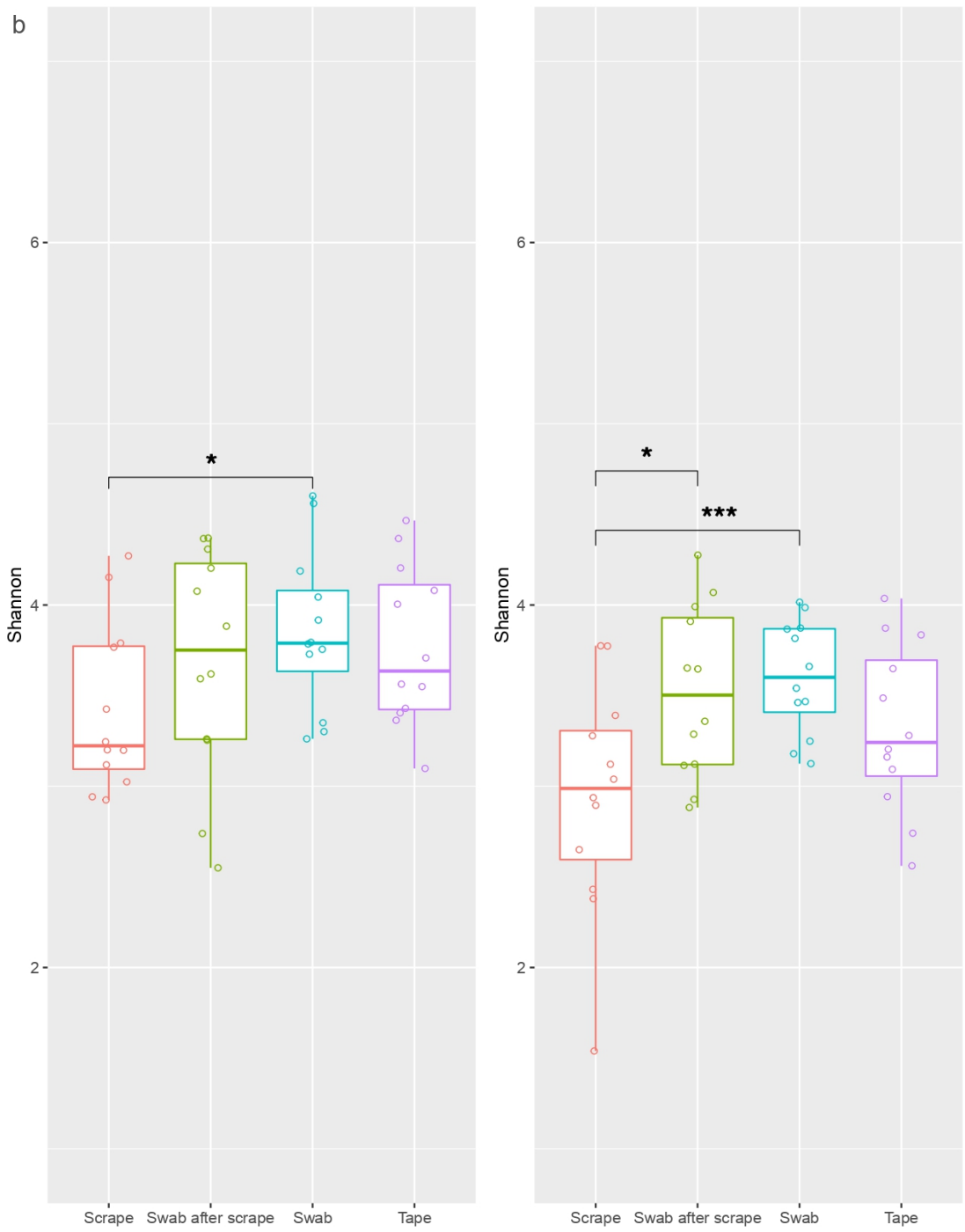
226 **Results**

227 After filtering, the total number of sequences identified across all sampling methodologies
228 was considerably higher for V1-3 (1,848,741) as compared to V3-4 (627,088). The mean per
229 sample sequence counts excluding negative controls, for V1-3 was 38,515 (range 13,215 to
230 121,731) vs 13,064 (range 3,990 to 31,786) for V3-4. In accordance with the higher number of
231 sequence counts, higher numbers of Amplicon Sequence Variants (ASVs) were assigned for
232 V1-3 (2724 taxa) than V3-4 (1385 taxa) across all sampling methodologies. A two-tailed paired
233 t- test performed on V1-3 and V3-4 sequence count showed a statistically significant
234 difference ($p < 0.0001$). Despite the low biomass of skin microbiomes and the inherent risks of
235 contamination (26) we observed a negligible amount of contamination, as confirmed by the
236 sequence counts from the negative control samples for V1-3 and V3-4 (3, 1 respectively).

237 Alpha diversity (the within-sample ASV richness and evenness) was measured using three
238 separate estimators. The total number of ASVs per sample was significantly higher in samples
239 collected by swabbing, swab after scrape, or tape collection than scrape collection alone (Fig
240 1a). Similarly, skin microbiome evenness (the relative proportions of identified species,
241 measured using the Shannon index) showed that Scrape sampling identified a less even
242 species distribution as compared to swabbing ($p=0.024$ V1-3, $p = 0.0018$ V3-4, Fig 1b).
243 Community diversity (a quantitative measure accounting for number and abundance of the
244 identified species, measured using Simpsons 1-D index) observed following scrape sampling
245 was significantly lower than both swabbing and swabbing after scraping for V3-4 ($p = 0.0036$,
246 $p= 0.045$ respectively, Fig 1c), but this was not seen with V1-3. Comparison of rarefaction for
247 both regions (Faith-pd) also showed scrape sampling reported the least diverse community,
248 whereas tape sampling was consistently the most diverse ($p<0.0001$, Supplementary Figure
249 1). Overall, the V1-3 region showed an approximately 2-fold greater diversity of observed
250 ASVs as compared to V3-4 with all sampling methodologies (Fig 1a-c).


a
V1−3
V3−4
Observed
400
300
200
100
0
150
100
50
0
*
***
**
Scrape
Swab after scrape
Swab
Tape

251


b
Shannon
6
4
2
*
Scrape
Swab after scrape
Swab
Tape
Shannon
6
4
2
*
***
Scrape
Swab after scrape
Swab
Tape

252

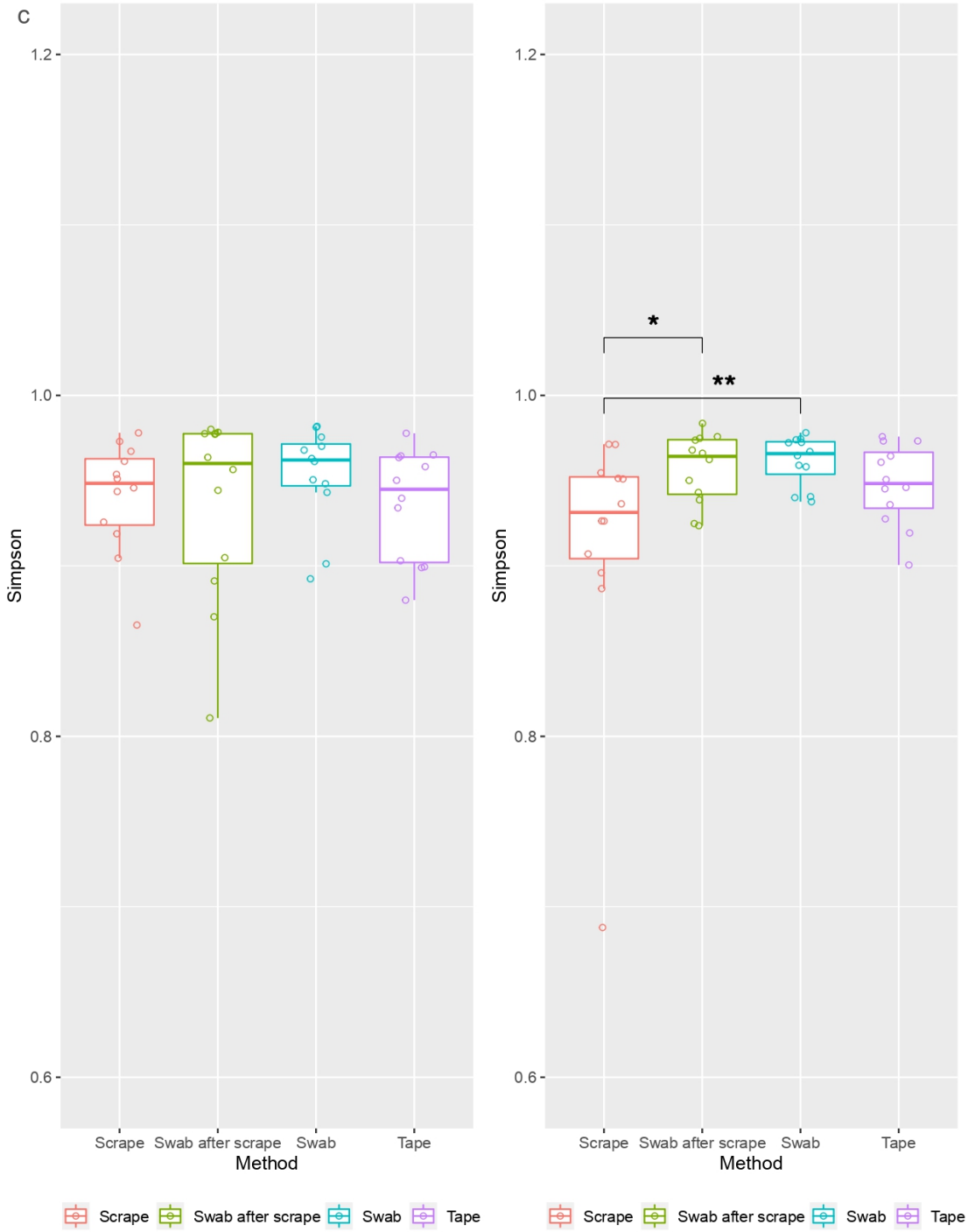
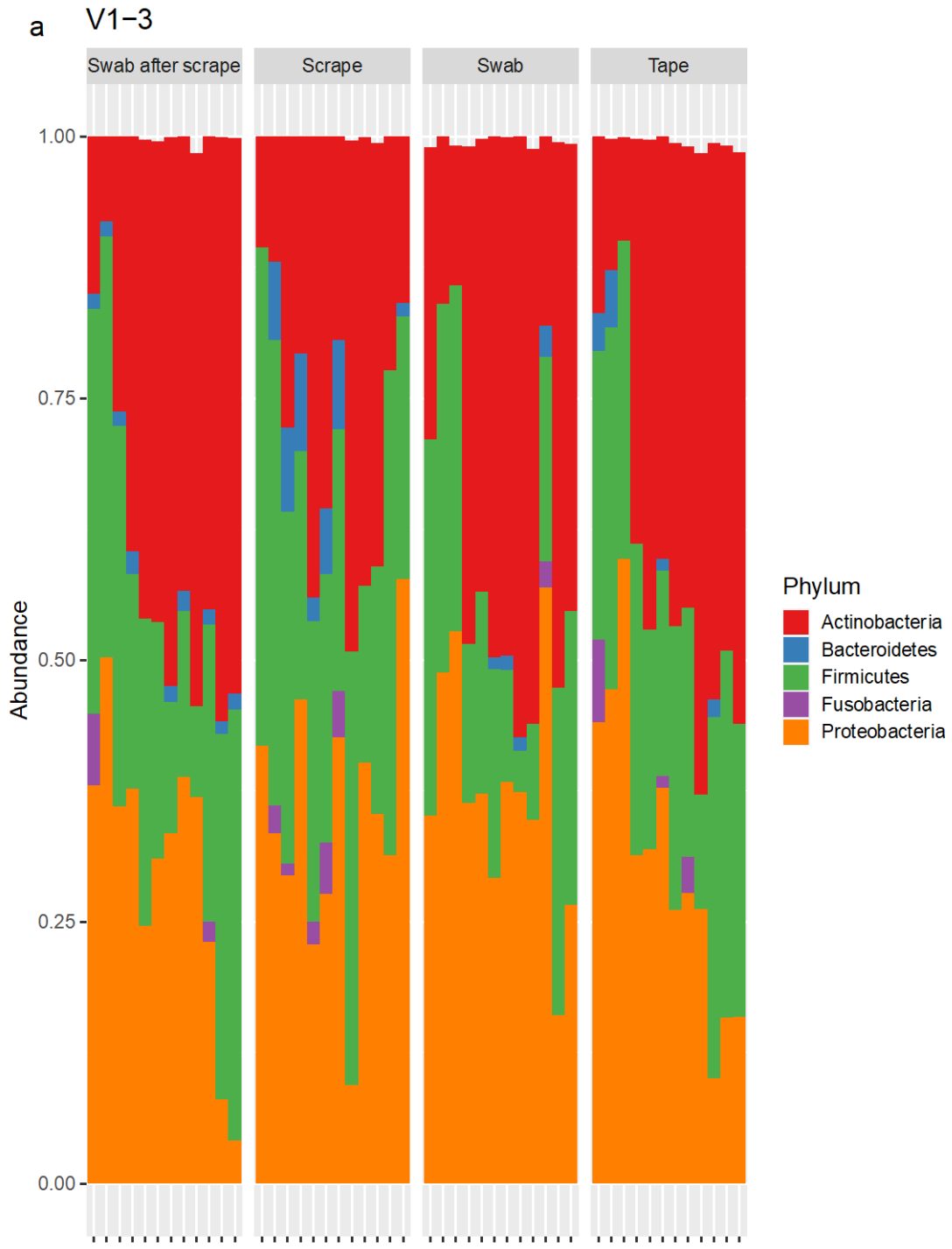


Figure 1. Box and Whisker plot showing alpha-diversity of the microbiomes acquired using four different skin sampling methods from four healthy volunteers. Diversity was estimated using Observed ASVs (a), Shannon (b) and Simpson 1-D (c). Kruskal-Wallis ANOVA for V1-3 (left) and V3-4 (right) are shown.

To compare compositional differences between samples acquired using the different sampling strategies, principal coordinate analysis (PCoA) of the Bray-Curtis distance

261 dissimilarity score and tested with PERMANOVA (P >0.05) analysis using pairwise adonis tool
262 with 999 permutations (27) (Supplementary Figure 2a, b). There was no statistical difference
263 when swabbing (with or without prior skin scrape) was compared to tape or scrape sampling.
264 However, community composition was significantly different when tape sampling was
265 compared to scrape for both V1-3 and V3-4 (respectively p = 0.009 and p = 0.048,
266 Supplementary Figure 2a, b).
267 Regardless of which 16S rRNA variable region or sampling technique the presence of bacterial
268 phyla comprising *Actinobacteria, Firmicutes, Bacteroidetes*, *Proteobacteria* and *Fusobacteria*,
269 was generally as expected (Figure 2a, b).


a
V1−3
Swab after scrape
Scrape
Swab
Tape
1.00
0.75
0.50
0.25
0.00
Abundance
Phylum
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria

270

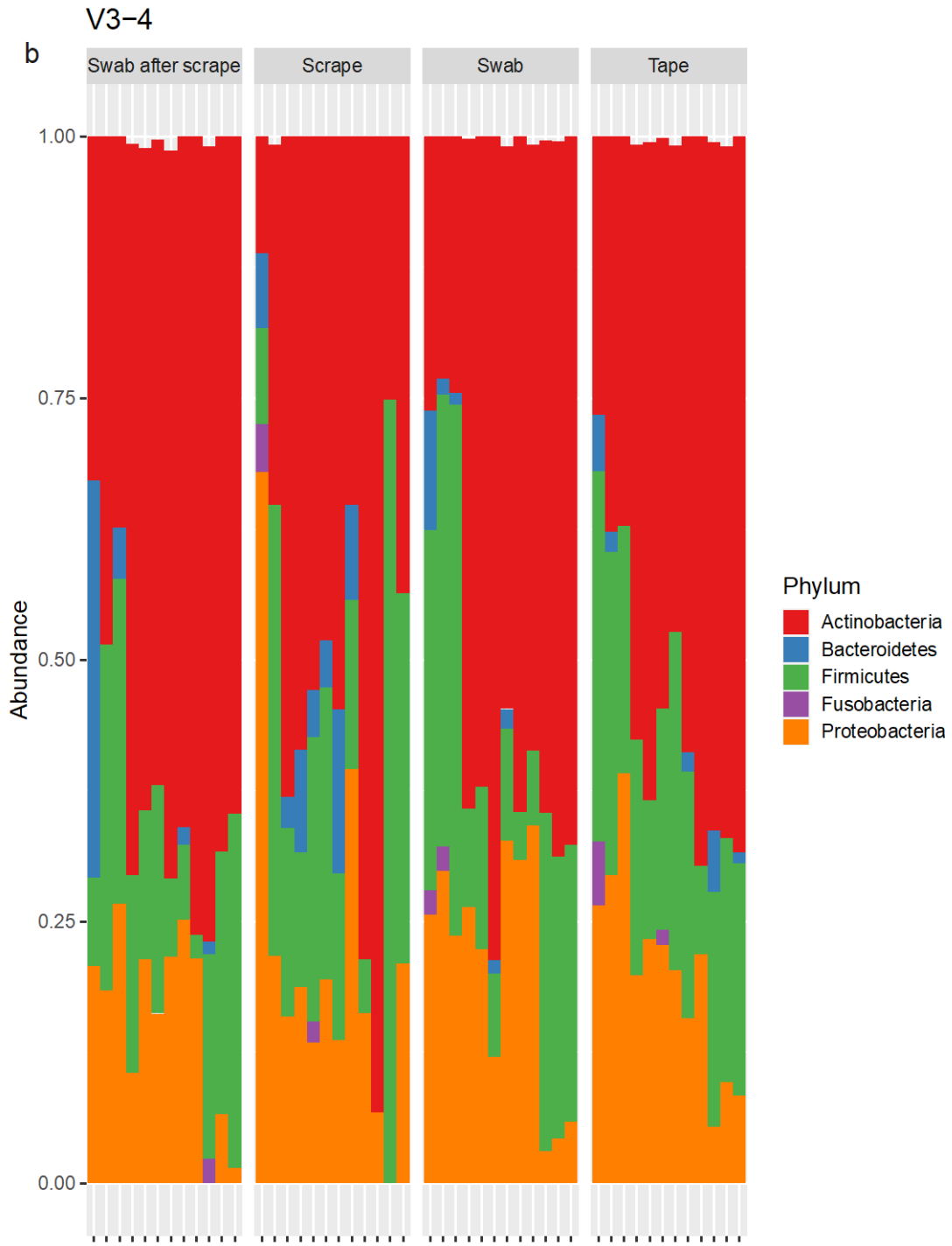


Figure 2. Proportional Bar plots showing the relative abundance of the dominant bacterial phyla using V1-3 across sampling methods. *Actinobacteria* (red), *Bacteriodetes* (blue), *Firmicutes* (green), *Fusobacteria* (purple) and *Proteobacteria* (orange) are shown. (a) V1-3. (b) V3-4.

Identification of *Bacteroidetes* and *Proteobacteria* was significantly enhanced by scrape sampling as compared to other methodologies on V1-3 analysis (Figure 3). However, this apparent methodological difference was not detected by analysis of V3-4 (Figure 4).

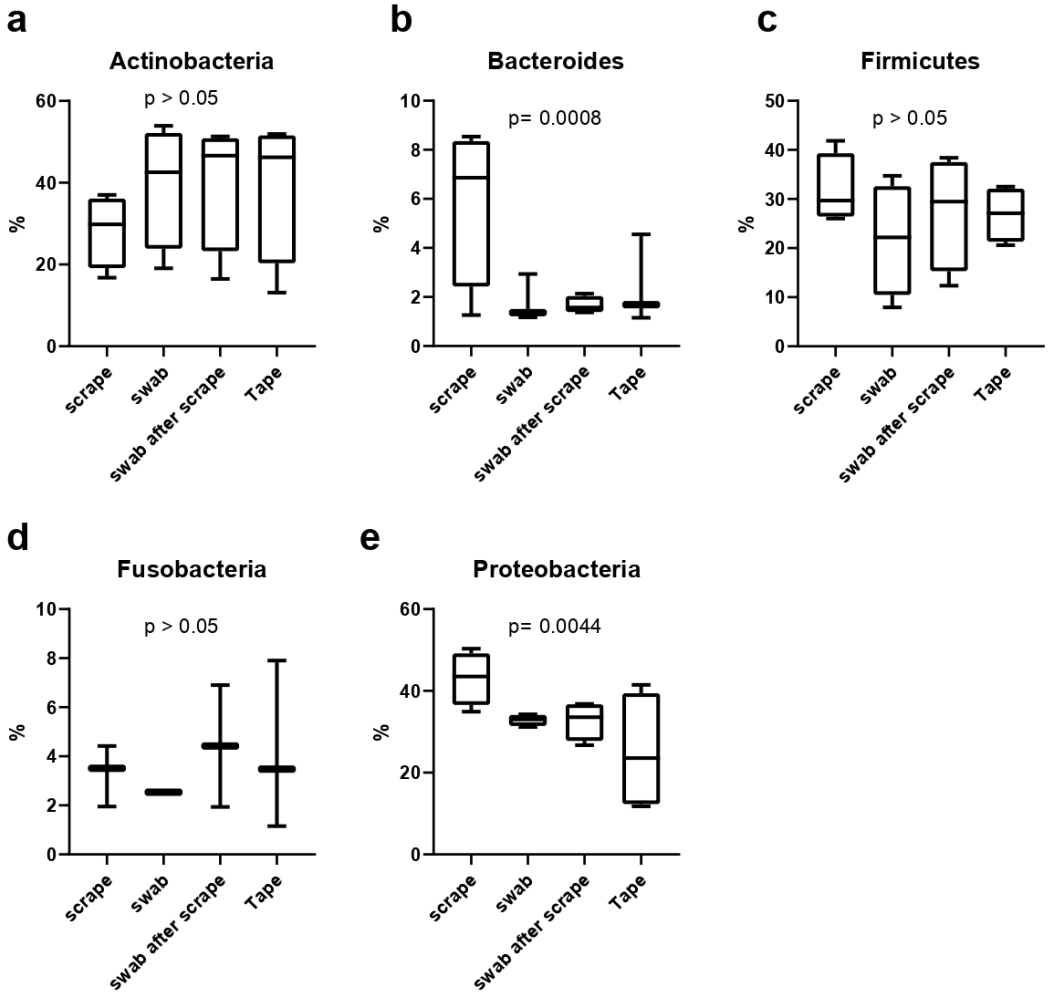


Figure 3. Box plots showing V1-3 relative abundance of identified phyla (a, Actinobacteria; b, Bacteroides; c, Firmicutes; d, Fusobacteria; e, Proteobacteria) by the four used sampling methods: Scrape, Swab, Swab after scrape and Tape.

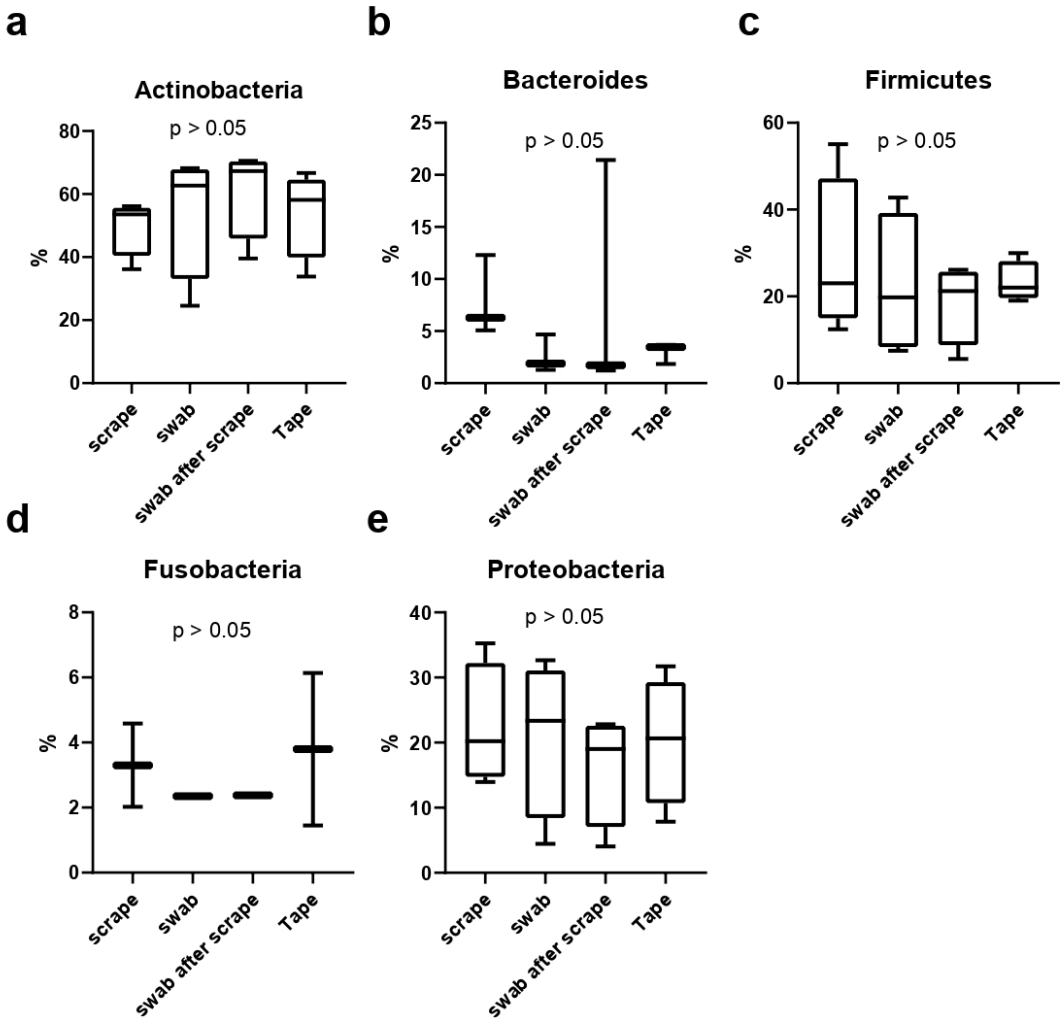


Figure 4. Box plots showing V3-4 relative abundance of identified phyla (a, Actinobacteria; b, Bacteroides; c, Firmicutes; d, Fusobacteria; e, Proteobacteria) by the four used sampling methods: Scrape, Swab, Swab after scrape and Tape.

Comparing V1-3 and V3-4 sampling analysis per phyla, V3-4 was more sensitive in detection of *Actinobacteria* ($P<0.05$ Scrape, Swab after scrape, and Tape) but resulted in reduced identification of *Firmicutes* and *Proteobacteria* ($P<0.05$ Swab, Swab after scrape, and Tape) (Figure. 5).

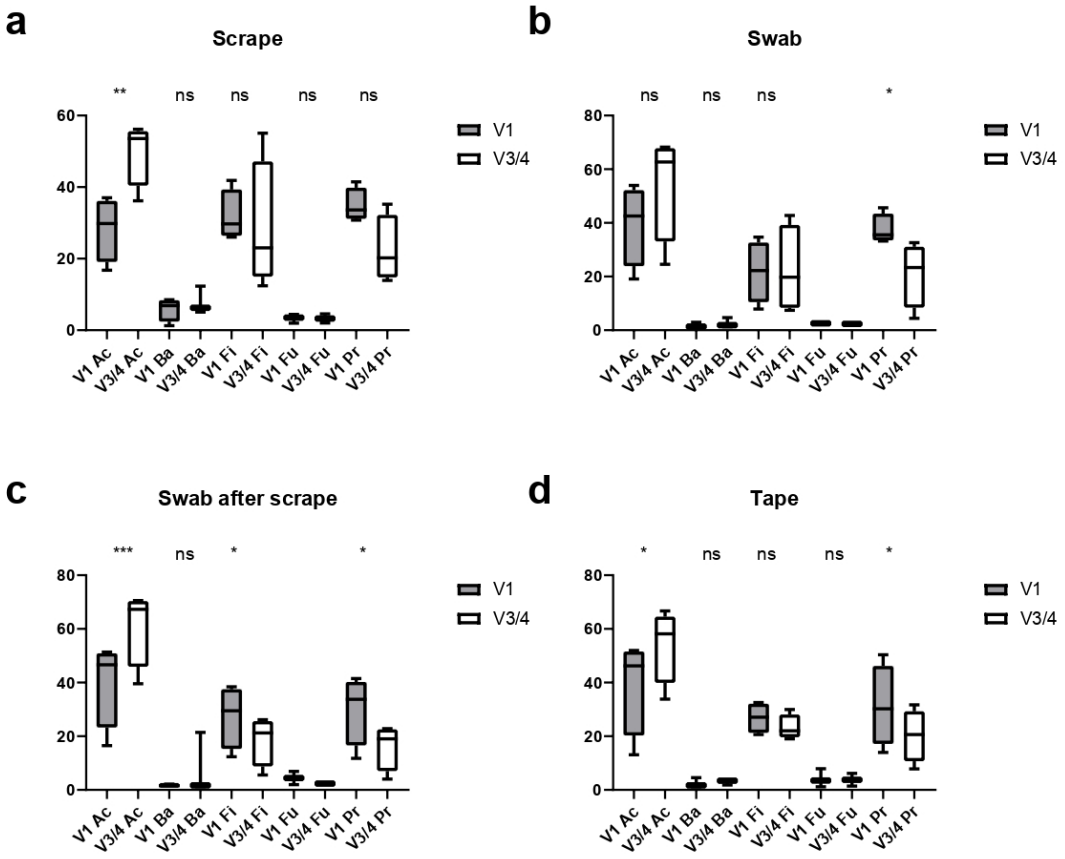


Figure 5. Bar plots showing the relative abundance of identified phyla by V1-3 (grey) or V3-4(white) analysis. Ac, Actinobacteria; Ba, Bacteroides; Fi, Firmicutes; Fu, Fusobacteria; Pr, Proteobacteria.

Given their importance in skin disease, we next compared V1-3 and V3-4 for species level identification of *Staphylococcus* sp. (Table 1). Analysis revealed a high proportion of ASVs aligned as belonging to staphylococci. Because *Staphylococcus epidermidis* is ubiquitous in human skin, detection of *S. epidermidis* is expected in all samples. As expected, with V1-3, the dominant staphylococcal species was aligned to *S. epidermidis* with a mean abundance of 18.7-21.5%, and no significant difference between the sampling methodologies was detected ($p>0.05$, Kruskal-Wallis ANOVA). However, in contrast, no *S. epidermidis* were aligned using V3-4 analysis. Instead, other staphylococci were classified: *S. equorum* and *S. haemolyticus.* With V1-3 sequencing, S. aureus was identified in two of the four volunteers and where identified the relative abundance of S. aureus was highest in tape samples (3.7%). Whereas targeting V3-4 detected S. aureus in only one of the four volunteers.

307 **Table 1:** Relative abundance of Staphylococcal ASVs identified to species level in skin
308 microbiomes using four different skin sampling methods, with V1-3 or V3-4 16S rRNA analysis.
309 Values represent mean relative abundance with ranges in parentheses.

V1-3	***S. aureus***	***S. epidermidis***	***S. haemolyticus***
Tape	3.7%	20.8% (8.8%- 39.7%)	NA
Scrape	2.0%	19.7% (3.8%- 41.7%)	NA
Swab	NA	18.7% (4.4%- 36.1%)	NA
Scrape after swab	NA	21.5% (1.5%- 39.9%)	NA
V3-4	***S. aureus***	***S. equorum***	***S. haemolyticus***
Tape	5.5%	NA	24.4% (7.9% -52.9%)
Scrape	NA	8.0%	10.6%
Swab	NA	5.5%	21.2% (13.0%-37.3%)
Scrape after swab	NA	4.2%	18.9% (5.1%-38.2%)

310 **Discussion**

311

312 Analysis of community compositions in microbiome research is reliant on the use of sampling
313 and analysis methods that give the most accurate representation of the resident microbiota.
314 It is recognised there is no 'one size fits all' approach in this context and in this respect the
315 optimal methods the characterisation of human skin microbiota have not been fully
316 determined. Nevertheless, most recent studies used the pre-moistened swabbing technique
317 for sample collection (4, 9, 28). There are various reasons to question whether this is the
318 optimal approach considering, for example, that tape stripping typically yields greater sample
319 biomass and viability (11, 12, 29). Other methods that have been recommended include skin
320 scraping and dry swabbing; while both provide lower biomass, each of them may yield greater
321 specificity (30). Furthermore, whilst recent publications have tended to focus on amplification
322 of the V3-4 region, direct comparisons of this approach with V1-3 in skin are limited (5, 31,
323 32).

324

325 To determine the optimal sampling method for 16S rRNA hypervariable region analysis we
326 matrixed variable region selection (V1-3 or V3-4) with replicate sampling methodology. Using
327 Bray-Curtis distance measures to assess beta-diversity, only the scrape samples were
328 significantly different from the other sampling techniques. The data also showed that the
329 scrape method revealed a significantly less diverse and more uneven microbiota composition,
330 suggesting that this approach is inferior for characterising the skin microbiome as compared
331 to the other three methods. Of the other three, swabbing after scrape yielded the most
332 heterogeneous results, but overall, good uniformity between the approaches was detected,
333 with similar measures of skin microbiome evenness and diversity.

334

335 Composition of the microbiome in our analyses was broadly as expected for both V1-3 and
336 V3-4 based on previous work, showing that *Actinobacteria* dominated (51.8%), followed by
337 *Firmicutes* (24.4%), *Proteobacteria* (16.5%) and *Bacteroidetes* (6.3%) (33). However,
338 *Actinobacteria* were identified to a higher abundance with V3-4 analysis. This might suggest
339 that there is preferential amplification of a member(s) of this Phylum using the V3-4 primers,
340 but there is little data to support this in the context of skin microbiota (34-36) and we found
341 no clear differences at lower taxonomic levels to support this hypothesis. V3-4 also showed
342 reduced relative abundance of *Proteobacteria* and importantly *Firmicutes* which includes
343 staphylococcal species.

A greater number of ASVs classified as *Staphylococcus* sp. were identified in V1-3 analysis. Indeed, it is important for skin researchers to note that, in our study of healthy volunteers where S. epidermidis colonisation can be assumed, S. epidermidis was also only identified with V1-3 amplification. It is hard to determine whether this is a feature of the known improvements in taxonomic classification performance of this region for staphylococci (37) or an issue with potential biases in amplification as noted above for V3-4. Nonetheless, no significant differences between methods were detected for V1-3. The lack of *S. epidermidis* when V3-4 was used suggests a significant limitation of this approach (38). We did not pre-screen for *S. aureus* carrier status, but the staphylococcal carriage in our population has been previously found to be 28% (95 % CI 26.1–30.2 %) in healthy individuals (39). Here we identified *S. aureus* in 50% of individuals studied with V1-3 analysis, whereas V3-4 identified *S. aureus* in only 25%. Taken together, this strongly suggests V3-4 is not suitable if the goal is to detect and distinguish species of staphylococci, a finding that mirrors previous research (5).

Some limitations exist with the study presented here. These include the use of only a small number of healthy volunteers (n = 4), only one body site for sampling and the lack of skin biopsy which was deemed too invasive to justify its inclusion in such a study. Inevitably, this makes extrapolation to different skin types difficult. Additionally, a lack of data on bacterial biomass recovered, which could have been achieved through use of a 16S rRNA qPCR for example, would have added quantitative data and made inter-sampling method comparisons more informative. Pre-screening recruits for *S. aureus* and *S. epidermidis* colonisation would also have been useful to enhance the comparative analysis between sampling approaches and their sensitivity for the detection of these species.

In conclusion, our data provide useful information for investigators in studying the skin microbiome. We show important differences between four commonly used sampling methods, evaluated using 16S rRNA sequencing of the V1-3 and V3-4 hypervariable regions. To maximise the potential for skin scientists to be able to compare different studies on the microbiome, it is important that technical aspects of the methodology are harmonised as far as possible. We suggest that the optimal approach to characterising diversity, with potential advantages for the recovery of *S. aureus*, incorporates tape or swab sampling followed by sequencing analysis of the V1-3 16S rRNA hypervariable region.

Acknowledgments

We thank all volunteers for their time and involvement in this study. We also thank the research nursing staff at the Wellcome Clinical Research Facility for their support.

379 **Figure Legends**

380

381 Figure 1. Box and Whisker plot showing alpha-diversity of the microbiomes acquired using
382 four different skin sampling methods from four healthy volunteers. Diversity was estimated
383 using Observed ASVs (a), Shannon (b) and Simpson 1-D (c). Kruskal-Wallis ANOVA for V1-3
384 (left) and V3-4 (right) are shown.

385

386 Figure 2. Proportional Bar plots showing the relative abundance of the dominant bacterial
387 phyla using V1-3 across sampling methods. *Actinobacteria* (red), *Bacteriodetes* (blue),
388 *Firmicutes* (green), *Fusobacteria* (purple) and *Proteobacteria* (orange) are shown. (a) V1-3. (b)
389 V3-4.

390

391 Figure 3. Box plots showing V1-3 relative abundance of identified phyla (a, Actinobacteria; b,
392 Bacteroides; c, Firmicutes; d, Fusobacteria; e, Proteobacteria) by the four used sampling
393 methods: Scrape, Swab, Swab after scrape and Tape.

394

395 Figure 4. Box plots showing V3-4 relative abundance of identified phyla (a, Actinobacteria; b,
396 Bacteroides; c, Firmicutes; d, Fusobacteria; e, Proteobacteria) by the four used sampling
397 methods: Scrape, Swab, Swab after scrape and Tape.

398

399 Figure 5. Bar plots showing the relative abundance of identified phyla by V1-3 (grey) or V3-
400 4(white) analysis. Ac, Actinobacteria; Ba, Bacteroides; Fi, Firmicutes; Fu, Fusobacteria; Pr,
401 Proteobacteria.

402

403 Table 1. Relative abundance of *Staphylococcal* ASVs identified to species level in skin
404 microbiomes using four different skin sampling methods, with V1-3 or V3-4 16S rRNA analysis.
405 Values represent mean relative abundance with ranges in parentheses.

406

407 Supplementary Figure 1. Alpha Diversity. Faith-pd diversity was used to compare community
408 rarefaction curves (an area under the curve analysis) for both regions. Scrape sampling is
409 shown to be generally lower in terms of diversity with tape sampling conversely the highest
410 ($p<0.0001$).

411

412 Supplementary Figure 2. Nonmetric Multidimensional scaling (NMDS) of Bray-Curtis
413 dissimilarity distance of skin microbiomes. The effect of sampling method (coloured dots) is
414 shown for V1-3 (left panel) and V3-4 (right panel). Ellipses represent 95% CI. Tape sampling
415 showed community composition to be statistically significantly different between tape and
416 scrape for both V1-3 and V3-4, p=0.026 and p=0.002 respectively.
417
418

References

1. Ferretti P, Farina S, Cristofolini M, Girolomoni G, Tett A, Segata N. Experimental metagenomics and ribosomal profiling of the human skin microbiome2016.
2. Chen YE, Tsao H. The skin microbiome: current perspectives and future challenges. J Am Acad Dermatol. 2013;69(1):143-55.
3. Kong HH, Oh J Fau - Deming C, Deming C Fau - Conlan S, Conlan S Fau - Grice EA, Grice Ea Fau - Beatson MA, Beatson Ma Fau - Nomicos E, et al. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. 2012(1549-5469 (Electronic)).
4. HMP C. A framework for human microbiome research. Nature. 2012;486(7402):215-21.
5. Meisel JS, Hannigan GD, Tyldsley AS, SanMiguel AJ, Hodkinson BP, Zheng Q, et al. Skin Microbiome Surveys Are Strongly Influenced by Experimental Design. J Invest Dermatol. 2016;136(5):947-56.
6. Grice EA, Kong Hh Fau - Renaud G, Renaud G Fau - Young AC, Young AC, Bouffard Gg Fau - Blakesley RW, Blakesley Rw Fau - Wolfsberg TG, et al. A diversity profile of the human skin microbiota. 2008(1088-9051 (Print)).
7. Nakatsuji T, Chiang Hi Fau - Jiang SB, Jiang Sb Fau - Nagarajan H, Nagarajan H Fau - Zengler K, Zengler K Fau - Gallo RL, Gallo RL. The microbiome extends to subepidermal compartments of normal skin. 2013(2041-1723 (Electronic)).
8. Huttenhower C, Gevers D, Knight R, Abubucker S, Badger J, Chinwalla A. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-14.
9. Aagaard K, Petrosino J, Keitel W, Watson M, Katancik J, Garcia N, et al. The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. The FASEB Journal. 2013;27(3):1012-22.
10. Van Horn KG, Audette CD, Sebeck D, Tucker KA. Comparison of the Copan ESwab system with two Amies agar swab transport systems for maintenance of microorganism viability. J Clin Microbiol. 2008;46(5):1655-8.
11. Zeeuwen PL, Boekhorst J, van den Bogaard EH, de Koning HD, van de Kerkhof PM, Saulnier DM, et al. Microbiome dynamics of human epidermis following skin barrier disruption. Genome Biol. 2012;13(11):R101.
12. Chng KR, Tay AS, Li C, Ng AH, Wang J, Suri BK, et al. Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare. Nat Microbiol. 2016;1(9):16106.
13. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, et al. Topographic diversity of fungal and bacterial communities in human skin. Nature. 2013;498(7454):367-70.
14. Whyte W, Carson W, Hambraeus A. Methods for calculating the efficiency of bacterial surface sampling techniques. J Hosp Infect. 1989;13(1):33-41.
15. Hambraeus A, Hoborn J, Whyte W. Skin sampling-validation of a pad method and comparison with commonly used methods. J Hosp Infect. 1990;16(1):19-27.

16. Byrd AL, Belkaid Y, Segre JA. The human skin microbiome. Nature Reviews Microbiology. 2018;16:143.
17. Conlan S, Kong HH, Segre JA. Species-level analysis of DNA sequence data from the NIH Human Microbiome Project. PLoS One. 2012;7(10):e47075.
18. Kong HH. Details Matter: Designing Skin Microbiome Studies. J Invest Dermatol. 2016;136(5):900-2.
19. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, et al. Experimental and analytical tools for studying the human microbiome. Nat Rev Genet. 2011;13(1):47-58.
20. Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers. Journal of microbiological methods. 2003;55(3):541-55.
21. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology. 2019;37(8):852-7.
22. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581-3.
23. Bisanz JE. qiime2R: Importing QIIME2 artifacts and associated data into R sessions. 2018.
24. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013;8(4):e61217.
25. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome. 2018;6(1):226.
26. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014;12:87.
27. Oksanen Jea. Vegan: Community Ecology Package. R package version 2. 5–7. 2020 [Available from: https://CRAN.R-project.org/package=vegan.
28. Paulino LC, Tseng CH, Strober BE, Blaser MJ. Molecular analysis of fungal microbiota in samples from healthy human skin and psoriatic lesions. J Clin Microbiol. 2006;44(8):2933-41.
29. Ogai K, Nagase S, Mukai K, Iuchi T, Mori Y, Matsue M, et al. Corrigendum: A Comparison of Techniques for Collecting Skin Microbiome Samples: Swabbing Versus Tape-Stripping. Front Microbiol. 2018;9:2812.
30. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe. 2010;7(6):509-15.
31. Castelino M, Eyre S, Moat J, Fox G, Martin P, Ho P, et al. Optimisation of methods for bacterial skin microbiome investigation: primer selection and comparison of the 454 versus MiSeq platform. BMC Microbiol. 2017;17.
32. Huse SM, Ye Y, Zhou Y, Fodor AA. A core human microbiome as viewed through 16S rRNA sequence clusters. PloS one. 2012;7(6):e34242.
33. Grice EA, Kong Hh Fau - Conlan S, Conlan S Fau - Deming CB, Deming Cb Fau - Davis J, Davis J Fau - Young AC, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. 2009(1095-9203 (Electronic)).

510 34. Farris MH, Olson JB. Detection of Actinobacteria cultivated from
511 environmental samples reveals bias in universal primers. Lett Appl Microbiol.
512 2007;45(4):376-81.
513 35. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. Critical
514 evaluation of two primers commonly used for amplification of bacterial 16S rRNA
515 genes. Applied and Environmental Microbiology. 2008;74(8):2461-70.
516 36. Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ, Scott KP. 16S rRNA gene-
517 based profiling of the human infant gut microbiota is strongly influenced by sample
518 processing and PCR primer choice. Microbiome. 2015;3:26.
519 37. Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of
520 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J
521 Microbiol Methods. 2007;69(2):330-9.
522 38. Otto M. Staphylococcus epidermidis--the 'accidental' pathogen. Nat Rev
523 Microbiol. 2009;7(8):555-67.
524 39. Gamblin J, Jefferies JM, Harris S, Ahmad N, Marsh P, Faust SN, Fraser S,
525 Moore M, Roderick P, Blair I, Clarke SC. Nasal self-swabbing for estimating the
526 prevalence of Staphylococcus aureus in the community. J Med Microbiol. 2013;62(Pt
527 3):437-40.
528

529

Plain Language Summary

Title: A detailed analysis of the best methods for analysis of skin bacteria

Healthy skin is covered with bacteria at a high density (approximately one million bacteria per square centimetre), which are known as the skin microbiome. The skin is used to having the bacteria on its surface and to some degree the bacteria in healthy skin are important in keeping the skin healthy. It is increasingly recognised that diseases affecting the skin often show an imbalance in the types and quantities of bacteria on the skin.

Historically, measurement of skin bacterial populations has relied on growth of the bacteria in a laboratory to quantities that become visible to the eye but is dependent upon the laboratory growing the bacteria efficiently. This approach has been superseded by a technique which does not rely on the bacteria growing and instead looks for the genetic signature of bacteria, and is therefore much better able to fully characterise the repertoire of bacteria from a sample. Because this new technology was largely developed in research on the gut microbiome, methodologies applied to skin, may not be appropriate.

Therefore, we set out to examine the skin microbiome using four sampling method (swabbing, scraping, swabbing after scraping and tape adhesion), as well as two different bacterial gene analysis approaches. Whilst many skin researchers have opted to follow the same protocol as gut analysis, we show that this is inferior for skin researchers. Additionally, we show that the optimal sampling method is tape adhesion. As a skin research community, we should harmonise these basic approaches to the study of skin microbiome, so that technical differences in methodology to not make it hard to compare different studies.