Skin Microbiome Surveys Are Strongly Influenced by Experimental Design

Jacquelyn S. Meisel[1], Geoffrey D. Hannigan[1], Amanda S. Tyldsley[1], Adam J. SanMiguel[1], Brendan P. Hodkinson[1], Qi Zheng[1] and Elizabeth A. Grice[1]

Culture-independent studies to characterize skin microbiota are increasingly common, due in part to affordable and accessible sequencing and analysis platforms. Compared to culture-based techniques, DNA sequencing of the bacterial 16S ribosomal RNA (rRNA) gene or whole metagenome shotgun (WMS) sequencing provides more precise microbial community characterizations. Most widely used protocols were developed to characterize microbiota of other habitats (i.e., gastrointestinal) and have not been systematically compared for their utility in skin microbiome surveys. Here we establish a resource for the cutaneous research community to guide experimental design in characterizing skin microbiota. We compare two widely sequenced regions of the 16S rRNA gene to WMS sequencing for recapitulating skin microbiome community composition, diversity, and genetic functional enrichment. We show that WMS sequencing most accurately recapitulates microbial communities, but sequencing of hypervariable regions 1–3 of the 16S rRNA gene provides highly similar results. Sequencing of hypervariable region 4 poorly captures skin commensal microbiota, especially *Propionibacterium*. WMS sequencing, which is resource and cost intensive, provides evidence of a community's functional potential; however, metagenome predictions based on 16S rRNA sequence tags closely approximate WMS genetic functional profiles. This study highlights the importance of experimental design for downstream results in skin microbiome surveys.

INTRODUCTION

Research devoted to the skin microbiome has surged in the past decade, due in large part to accessible, affordable high-throughput DNA sequencing technology and the realization that the microbiome may modulate the pathogenesis of many cutaneous disorders. The majority of protocols for characterizing microbial communities were initially developed and optimized to survey the gastrointestinal tract or the environment, niches that harbor distinct sets of microbiota compared to the skin. A standardized methodology for skin microbiome studies is lacking, although these protocols are often pivotal to their outcome (Albertsen et al., 2015; Guo et al., 2013; Nelson et al., 2014).

A common approach to characterize cutaneous microbial communities relies on amplification, sequencing, and analysis of the prokaryotic 16S ribosomal RNA (rRNA) gene. This approach has been used in multiple studies of skin bacterial communities and their association with health and disease (Hannigan and Grice, 2013). Initial studies used full-length 16S rRNA gene sequences ($\sim$1500 kb) generated by Sanger sequencing methods. Next-generation sequencing platforms that allow for vastly increased sequencing depth at a fraction of the cost generate shorter read lengths, making it impractical to sequence the full-length gene. Therefore, one or more hypervariable regions, or 16S tags, are selected for sequencing as a proxy for the full-length gene. No single hypervariable region is able to distinguish among all bacteria, and primer biases may differentially affect amplification efficiency of different types of bacteria. However, specific regions may be optimal for capturing the diversity and composition of different ecosystems.

More recently, whole metagenomic shotgun (WMS) sequencing has been used for both taxonomic and functional annotation of skin microbial communities (Hannigan et al., 2015; Human Microbiome Project Consortium, 2012b; Oh et al., 2014). This approach reduces amplification bias, captures multikingdom communities, and allows for strain-level analysis. WMS datasets, although information rich, are more expensive to generate and require greater computational knowledge and resources to store, process, and analyze. Although gene content can be extracted from WMS data to provide insight into functional processes of the microbial community, bioinformatic tools now exist to predict functional content from 16S tag sequences (Langille et al., 2013) and in some cases may be superior to WMS sequencing for microbial community classification (Xu et al., 2014).

Here we present a comparison of experimental strategies to identify optimal parameters for capturing the composition, diversity, and genetic content of the cutaneous microbiome.

[1]Department of Dermatology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

Correspondence: Elizabeth A. Grice, 421 Curie Boulevard, 1007 BRB II/III, Philadelphia, Pennsylvania 19104, USA. E-mail: egrice@upenn.edu

We applied 16S rRNA tag sequencing to cutaneous swabs and a publicly available mock community control of 20 bacterial species in known concentration. We sequenced two 16S tag regions commonly utilized in microbiome studies, including hypervariable regions 1−3 (V1−V3) and region 4 (V4) (Caporaso et al., 2011) to compare their utility in accurately characterizing skin microbiota diversity and composition. Additionally, we performed WMS sequencing on the same swab samples and controls to identify any additional utility of WMS sequencing over 16S tag sequencing for characterizing skin microbiota and identifying genetic functional enrichment.

RESULTS

Sampling, sequencing, and quality control

Sixty-two cutaneous skin swabs were collected from nine healthy volunteers (for cohort characteristics, see Supplementary Table S1 online). Sampled skin sites consisted of diverse microenvironments with respect to moisture (sweat) and sebum: sebaceous (retroauricular crease,

occiput, forehead), moist (toe web, umbilicus), and intermittently moist (antecubital fossa, palm) (Figure 1a).

Whole genomic DNA was extracted from the swab samples and subjected to microbiome profiling using three different approaches: (i) V1−V3 tag sequencing; (ii) V4 tag sequencing; and (iii) WMS sequencing (Figure 1b and c). V1−V3 tag sequencing was used by the Human Microbiome Project (Aagaard et al., 2013; Human Microbiome Project Consortium 2012a, 2012b). The V4 region was used in the Earth Microbiome Project (Gilbert et al., 2014) and is widely utilized to characterize microbiota of other body habitats. We did not include regions further 3′ in the 16S rRNA gene, as these have been documented to generally perform less well for a variety of analyses (Conlan et al., 2012; Jumpstart Consortium Human Microbiome Project Data Generation Working 2012; Wu et al., 2010) and/or are not widely used for characterization of microbial communities. All sequencing was performed on either the Illumina MiSeq or HiSeq 2500 platforms. A publicly available mock community control (MCC) was sequenced in parallel with the skin samples.

Figure 1. Study design for analyzing cutaneous bacterial communities. (a) Seven skin sites were sampled from a healthy human cohort of nine volunteers. (**b**) DNA was isolated from cutaneous swabs and sequenced for downstream bioinformatics analyses. (**c**) Schematic illustrating the primers used for the two different targeted hypervariable regions on the 16S ribosomal RNA (rRNA) gene. Ac, antecubital fossa; Fh, forehead; Oc, occiput; Pa, palm; Ra, retroauricular crease; Tw, toe web; Um, umbilicus; WMS, whole metagenome shotgun.
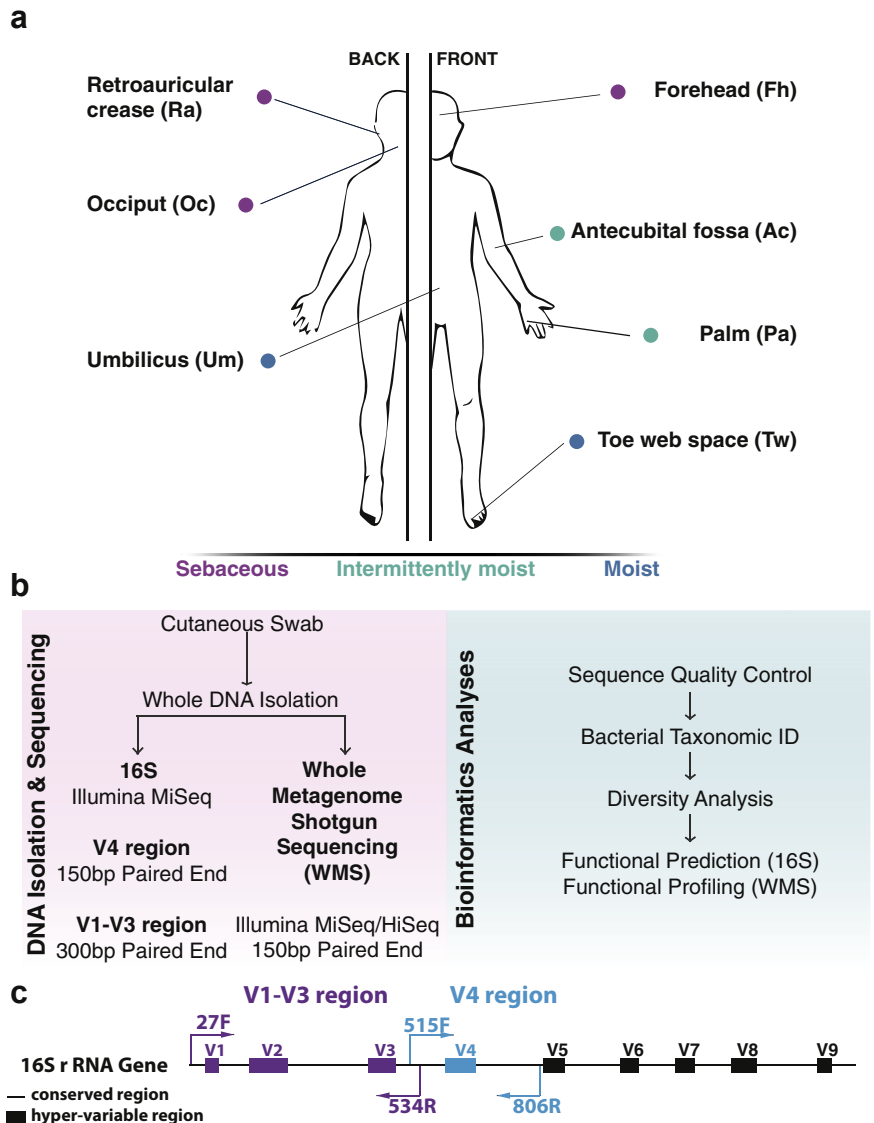
The V1–V3 dataset contained 2,124,836 total high-quality sequence reads, with a median of 24,891 sequence reads per sample. The V4 dataset contained 5,328,215 total high-quality sequence reads, with a median of 77,928 sequence reads per sample. The WMS dataset contained 81,553,035 total high-quality sequence reads, with a median of 1,233,172 sequence reads per sample (for per sample sequence counts, see Supplementary Table S2 online).

## Skin bacterial community composition varies by sequencing technique

We compared each sequencing method to determine how well they recapitulated the taxonomic relative abundance of the MCC, which contained 100,000 rRNA operon copies per organism per microliter. Therefore, each of the 20 species contained in the MCC should account for 5% of the community by 16S tag sequencing. For WMS sequencing, expected MCC abundances must take into account the

concentration of the input DNA. We first mapped our sequences to the expected MCC species to identify community composition (Figure 2a). Hierarchical clustering of taxonomic profiles indicated that WMS provided a close approximation of the MCC (Figure 2a). V1–V3 tag sequencing provided the best proxy for 16S-based profiling, whereas V4 tag sequencing severely underrepresented *Staphylococcus epidermidis* and *Propionibacterium acnes* and overrepresented *Staphylococcus aureus*. When using operational taxonomic unit (OTU)-based methods to characterize the MCC, similar trends were observed, however, V1–V3 was unable to classify all taxa to the genus level (see Supplementary Figure S1 online).

Propionibacterium (including *P. acnes*), *Staphylococcus* (including *S. epidermidis* and *S. hominis*), and *Corynebacterium* were the dominant bacterial genera on healthy human skin (Figure 2b and Supplementary Figure S2 online). The most notable observation was that *Propionibacterium* was

a

Figure 2. Taxonomic profiles of cutaneous bacterial communities vary. (**a**) Heat map depicting relative abundances of the 20 bacterial species in the mock community control. Rows display bacterial species, and columns denote sequencing technique used for analysis. "Actual" refers to the expected abundance based on the community composition. Dendrogram (x-axis) clusters each sequencing type by similar taxonomic profiles. (**b**) Pie charts depicting the mean relative abundances of the top 15 taxa in the cutaneous samples. The innermost circle represents the V4 samples, the middle circle the V1–V3 samples, and the outermost whole metagenome shotgun (WMS) sequenced samples.

b
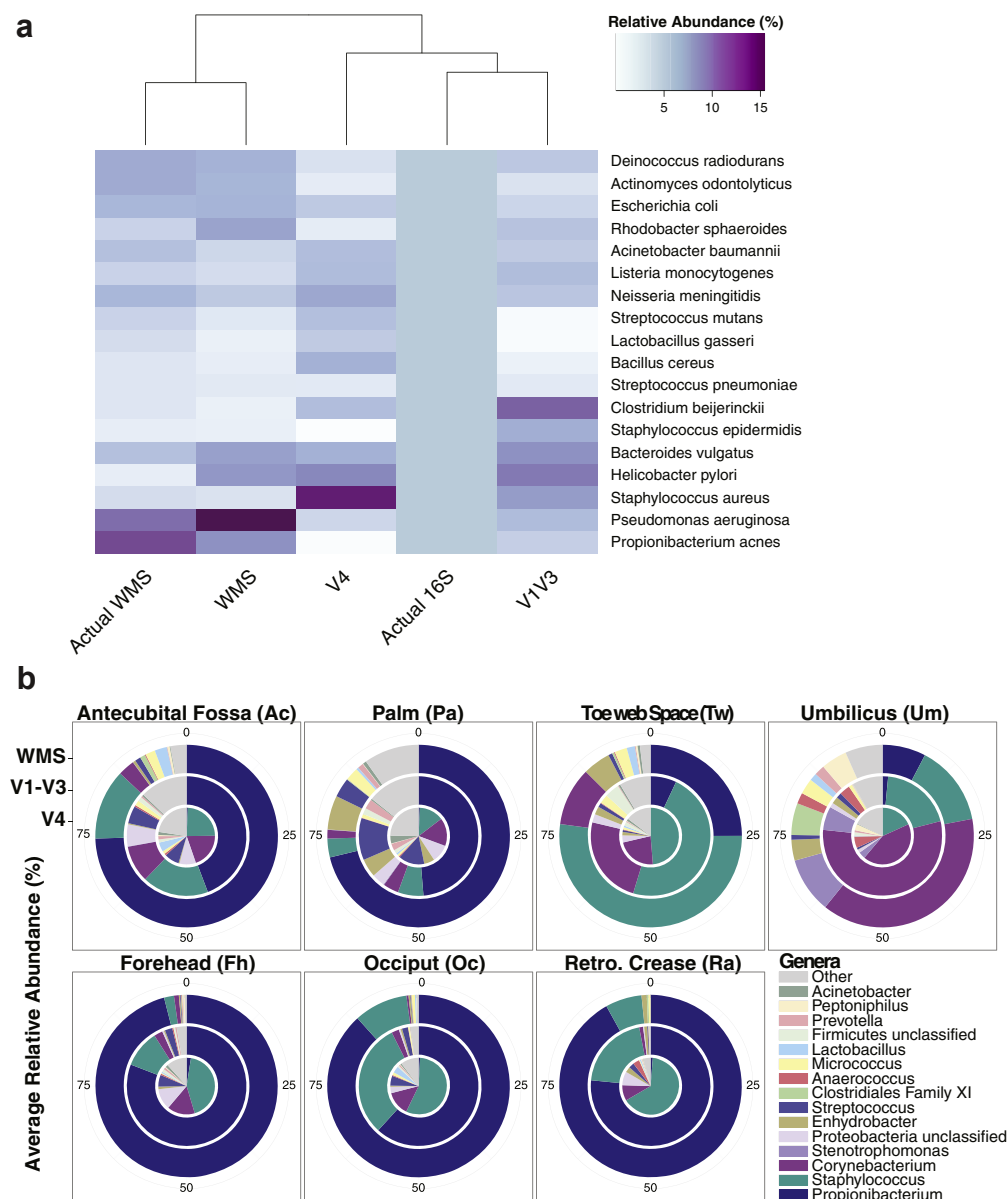
vastly underrepresented in the V4 dataset. We used basic linear regression analysis to correlate the relative abundance of three prominent skin bacteria in V1−V3 and V4 datasets compared to their relative abundance in the WMS dataset, which fairly accurately recapitulated the composition of the MCC (see Supplementary Figure S3 online). The relative abundances represented by the V1−V3 dataset had much higher positive correlations to WMS relative abundances than were observed with the V4 dataset for *Propionibacterium* ($R^2 = 0.931$ vs. 0.499), *Staphylococcus* ($R^2 = 0.736$ vs. 0.153), and *Corynebacterium* ($R^2 = 0.789$ vs. 0.281). These data indicate that V4 representations of skin microbiome composition are severely biased against bacteria that are present in great prevalence and abundance on the human skin.

Hierarchical clustering revealed that this bias was not equal across all microenvironments. Intermittently moist and sebaceous samples from the V1−V3 and WMS datasets cluster together, but V4 samples were less similar (see Supplementary Figure S4 online). This clustering appears to be driven largely by underrepresentation of *Propionibacterium* in V4 tags. Moist

sites were most taxonomically similar across all sequencing methods and clustered together regardless of method.

Staphylococcus species level classification in 16S datasets is enabled by phylogenetic placement algorithms

A tradeoff when using cost-effective next-generation sequencing is the short read lengths that these platforms generate, presenting a challenge for accurate genus-, species- and strain-level classification. Using OTU-based methods, both V1−V3 and V4 tag sequencing failed to accurately identify >30% of the species in the MCC (see Supplementary Figure S5a online). Moreover, only 13.7% of the V1−V3 and 7.6% of the V4 OTUs were classified to the species level in the cutaneous swab samples (Figure 3a).

Species-level resolution of skin microbiota is especially important when trying to differentiate between commensals (i.e., *S. epidermidis*) and pathogens (i.e., *S. aureus*). Using the RDP classifier in QIIME (Quantitative Insights into Microbial Ecology), we were unable to speciate *Staphylococcus* in the V4 samples and only identified *S. epidermidis* in the V1−V3

Figure 3. Species level classification of *Staphylococcus* sequences. (**a**) Bar plot of the total number of operational taxonomic units (OTUs) identified in cutaneous samples by V1−V3 and V4 sequencing, highlighting the number of OTUs named at the species level. (**b**) Mean relative abundance of *Staphylococcus* species identified by 16S tag sequencing of skin samples at >1%. (**c−e**) Relative abundance of *Staphylococcus* sequences able to be classified at the species level. (**c**) *Staphylococcus* species in the whole metagenome shotgun (WMS) sequencing dataset were classified using MetaPhlAn. (**d, e**) V1−V3 and V4 species level classifications were determined by pplacer. Pie charts depict the percentage of sequences classified as *Staphylococcus* at the genus level that were further classified at the species level.
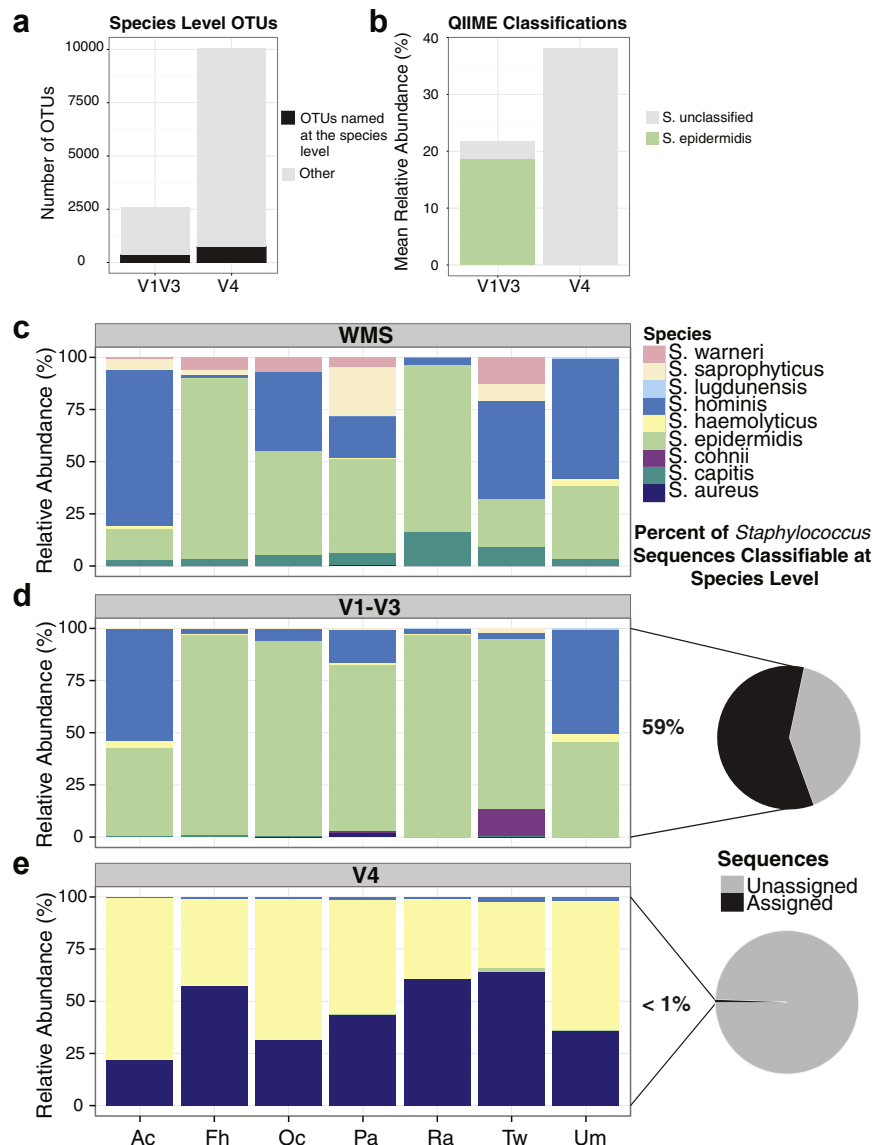
samples (Figure 3b), despite evidence that additional *Staphylococcus* species live on the skin (Figure 3c). An approach to improve taxonomic resolution of 16S rRNA tag sequence data is to use phylogenetic information. We attempted to classify *Staphylococcus* species in the 16S datasets by using pplacer (Matsen et al., 2010), an algorithm that uses maximum likelihood criteria to place sequences on a fixed phylogenetic reference tree.

WMS accurately identified the two *Staphylococcus* species, *S. epidermidis* and *S. aureus*, in the MCC (see Supplementary Figure S5b). Only 11% and <1% of *Staphylococcus* sequences using V1−V3 and V4 tags, respectively, were classifiable at the species-level using pplacer (see Supplementary Figure S5b). The pplacer classification of V1−V3 tags identified the correct species but overrepresented the relative abundance of *S. epidermidis*. V4 tag species-level classification identified *S. aureus* but also falsely identified *S. hominis* and *S. haemolyticus*.

The pplacer analysis of the skin swabs revealed agreement between the V1−V3 and WMS datasets, but not the V4 dataset. WMS identified the predominant *Staphylococcus* species to be *S. epidermidis*, *S. hominis*, and *S. capitis* (Figure 3a). Of the sequences identified as *Staphylococcus* at the genus level in the V1−V3 dataset, 59% were classified at the species level. *S. epidermidis* and *S. hominis* were identified, but *S. capitis* was absent (Figure 3b). Less than 1% of the V4 *Staphylococcus* sequences were classified by pplacer, and they were predominantly characterized as *S. aureus* and *S. haemolyticus* (Figure 3c).

Computationally predicted versus observed functional profiles

A perceived advantage of WMS approaches for skin microbiome studies is the functional insight gained through analysis of genetic enrichment. However, functional genetic profiles can be predicted from 16S rRNA sequences with the program PICRUSt (Phylotypic Investigation of Communities by Reconstruction of Unobserved States) (Langille et al., 2013), which uses reference genomes to infer a composite metagenome and predict abundance of gene families. Therefore, we compared functional genetic profiles obtained by WMS to PICRUSt-predicted functional genetic profiles of V4 and V1−V3 tag sequence datasets.

Functional enrichment analysis of the MCC identified variation in KEGG (Kyoto Encyclopedia of Genes and Genomes) Pathway enrichment by sequencing technique (see Supplementary Figure S6a online) but did not reveal significant differences in Shannon Diversity (see Supplementary Figure S6b). Notably, several metabolic pathways, including "metabolism of cofactors and vitamins" and "carbohydrate metabolism," were more abundant, and "energy metabolism" and "biosynthesis of other secondary metabolites" were less abundant in the WMS dataset than in metagenomes predicted from 16S tag sequence data. Functional profiles of each skin swab generated from the WMS dataset also differed compositionally from their matched V1−V3 and V4 predicted metagenomes. We focused on the 102 pathways identified across all datasets, in at least four samples, and at >0.5% abundance. We grouped these pathways into 28 higher-level KEGG categories, 21 of which were shared in all

datasets and significantly differentially enriched between either of the 16S and the WMS datasets (FDR corrected paired Wilcoxon test, $P < 0.05$; Figure 4a). The KEGG category "xenobiotics biodegradation and metabolism" was enriched in both 16S-predicted functional profiles (Figure 4b), with the greatest differences observed in pathway ko00930 ("Caprolactam degradation"). Conversely, the "translation" category was more prominent in the WMS dataset, with significant differences in ribosomal (ko03008, ko03010) and tRNA (ko00970) pathways (Figure 4b). Several KEGG categories also differed between V1−V3 and V4 sequencing techniques, including "glycan biosynthesis and metabolism," which is significantly different between the V4 and WMS datasets, but not the V1−V3 and WMS datasets (Figure 4b). Despite these observed differences, Spearman correlations revealed strong trends between the mean relative abundances of higher-level KEGG pathways in the predicted functional profiles compared to the WMS dataset across all body sites sampled (see Supplementary Figure S7 online).

Diversity trends are dependent on methodology

We estimated and compared taxonomic alpha diversity of skin bacterial communities using the Shannon diversity index, which takes into account both the total number of species in the community (richness) and the evenness of the species present. All three sequencing approaches identified sebaceous sites as the least diverse and significantly less diverse than intermittently moist sites ($P < 0.05$ Kruskal-Wallis and multiple comparison post hoc test; Figure 5a). Whereas V1−V3 and WMS sequencing identified significant diversity trends between moist and sebaceous sites, V4 tag sequencing did not. Alternatively, V4 found intermittently moist sites to be significantly more diverse than moist sites, a trend that was not confirmed by the other methods.

Cutaneous functional diversity, calculated based on predicted gene functions, has previously been shown to vary by biogeography (Oh et al., 2014). However, we identified conflicting trends in the skin microbiome based on the microenvironment of the site sampled. Both V4 and V1−V3 tags identified significant differences dependent on microenvironment that were not found in the WMS sequencing dataset ($P < 0.05$ Kruskal-Wallis and multiple comparison post hoc test; Figure 5b).

We also compared beta diversity, or bacterial community structure, as recapitulated by V1−V3 and V4 tag sequencing. We applied Procrustes analysis to Bray-Curtis dissimilarity matrices in order to determine whether the use of different 16S rRNA sequence tags would derive similar beta diversity conclusions. Although significant, Procrustes analysis showed very weak congruence between the datasets (m_{12} squared = 0.6338, $P = 0.0001$; Figure 5c).

DISCUSSION

As microbiome sequencing surveys become increasingly common, effective study design is crucial for the development of meaningful datasets. We make the following recommendations for studying skin microbiota from swab samples. (i) Regarding choosing a 16S rRNA gene amplicon, the V1−V3 region provides more accurate assessments of human skin microbiota compared to the V4 region. (ii) WMS

Figure 4. 16S predictions differ from whole metagenome functional enrichment. (**a**) Heat map depicting the log2 fold change of statistically significantly different KEGG categories (FDR corrected paired Wilcoxon test, $P < 0.05$) between 16S and whole metagenome shotgun (WMS) sequencing functional profiles. Purple depicts enrichment in WMS samples. Green depicts enrichment in the 16S samples. Each column represents a different sample and each row a KEGG pathway. The colors above the columns indicate sequencing of the V1−V3 (gray) or V4 (black) region of the 16S rRNA gene. (**b**) Box plots depicting mean relative abundances of significantly different KEGG pathways (FDR corrected paired Wilcoxon test, $P < 0.05$) between 16S and WMS functional profiles.
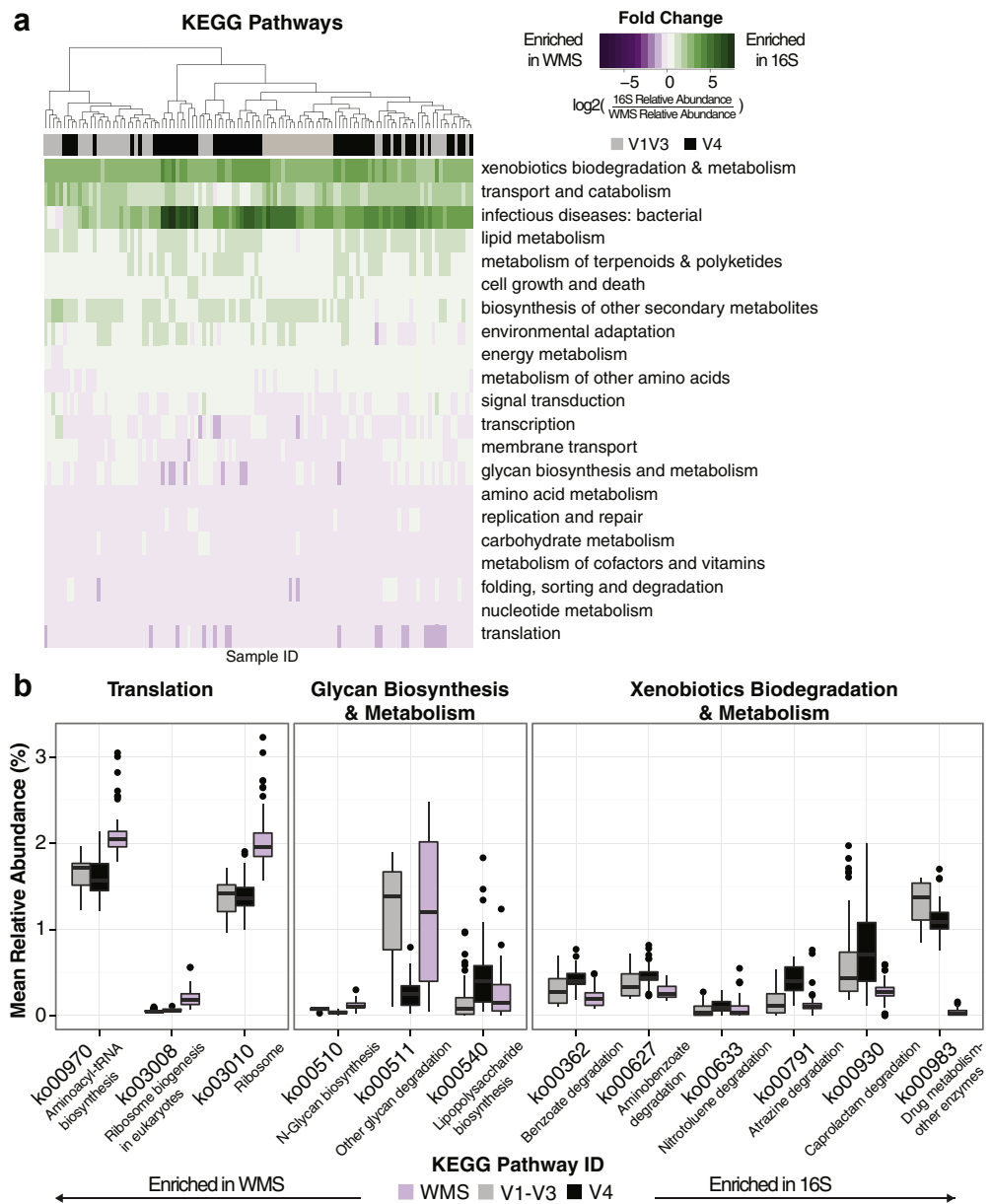
sequencing is superior for species-level taxonomic classification, and previous reports have demonstrated the utility in strain-level analysis and capturing nonprokaryotic elements of the skin microbiome (Oh et al., 2014). However, V1−V3 tags provide reasonable proxies for taxonomic composition and diversity at a much lower cost and effort. The goal of the experiment should be carefully considered in addition to available resources for generating and analyzing the resulting datasets. (iii) Functional genetic predictions based on 16S rRNA tags are remarkably similar to those provided by WMS sequencing and in some cases may provide a reasonable estimate of functional enrichment when the expertise and/or resources are not available to perform WMS. However, because of strain variability and widespread horizontal gene transfer between bacteria, results of predictive analyses should be interpreted with caution. Predictive analyses are also limited in their ability to identify antibiotic resistance

and virulence genes that may be of interest and could be inferred from WMS sequencing.

Primers amplifying the V4 variable region, as used here, were not able to recover *Propionibacterium* or reliably speciate *Staphylococcus*. This is not unexpected because the V4 hypervariable region is much shorter than the V1−V3 region and has a higher degree of sequence conservation (Chakravorty et al., 2007). A separate study also remarked on the absence of *Propionibacterium* in V4 libraries, suggesting that a single nucleotide difference between the 27F forward primer and annealing site in the *P. acnes* 16S rRNA gene may impair detection (Nelson et al., 2014). Our findings underscore the importance of thoroughly vetting primers for their ability to capture microbiota of importance to the skin habitat.

A recent study noted that biogeography of the skin, as well as individuality, shaped strain-level cutaneous diversity (Oh et al., 2014). *P. acnes* was also found to be differentially
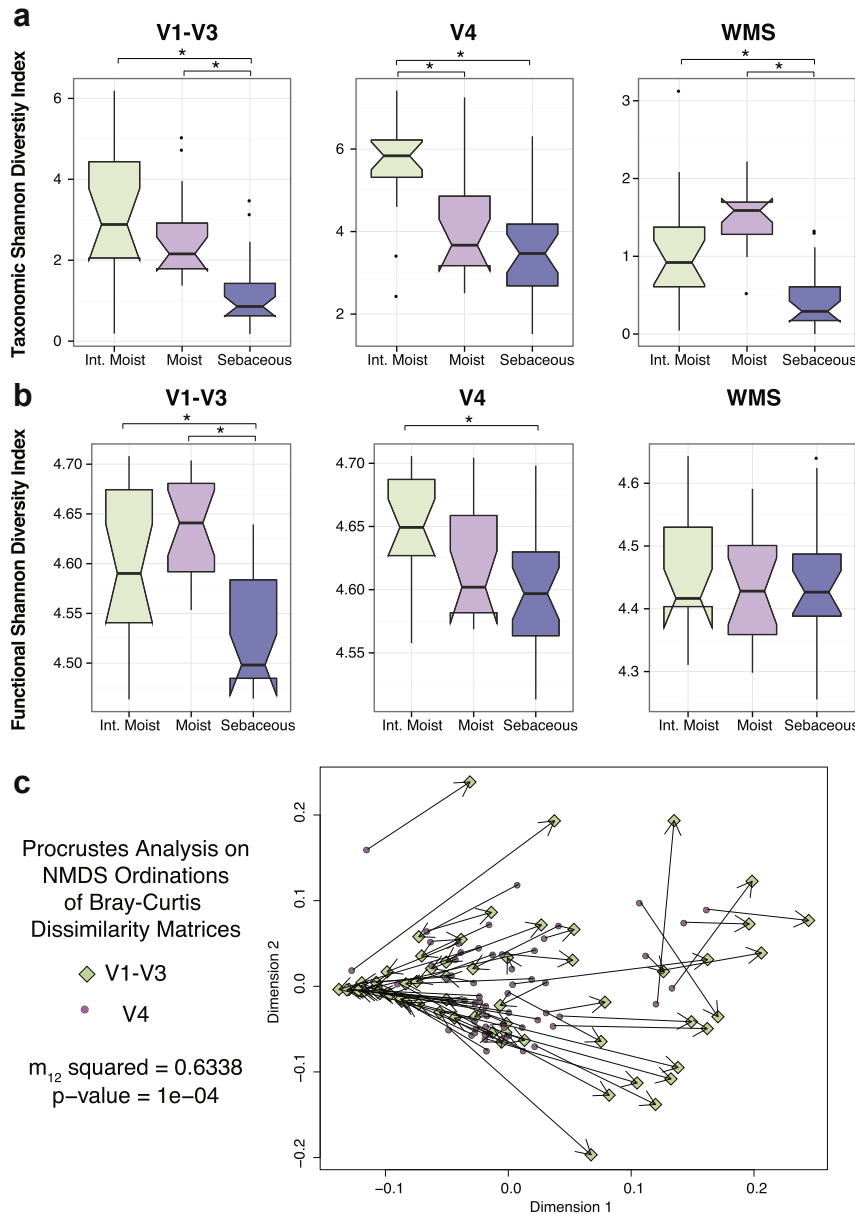
Figure 5. Cutaneous taxonomic and functional diversity trends depend on sequencing method. Shannon diversity of (**a**) taxonomic and (**b**) functional profiles for each sequencing technique presented by site microenvironment. Asterisks indicate significance of $P < 0.05$ using Kruskal-Wallis and multiple comparison post hoc tests. Box plots were calculated using the ggplot2 R package. (**c**) Procrustes analysis revealing congruence between NMDS ordinations of the V1−V3 (target) and V4 (rotated) Bray-Curtis dissimilarity matrices. Circles indicate V4 samples and diamonds indicate V1−V3 samples, with matched samples connected by a line. Shorter lines reflect greater clustering similarity.

associated with acne (Fitz-Gibbon et al., 2013). Speciation and strain-level identification of microbiota may be important if the ultimate goal is to identify a putative causal microbe/microbiota for downstream studies to examine mechanism. We noted that the V1−V3 region was able to speciate the majority of *Staphylococcus* sequences based on phylogenetic placement against a curated reference database of *Staphylococcus* species. However, WMS sequencing would be a superior approach if one wished to identify overall strain-level variability and/or did not have access to a reliable curated reference database for the genus or species of choice.

Based on the striking differences in taxonomic composition of the datasets, the strong correlations of KEGG pathway abundance across sequencing methods is surprising but may indicate shared functionality among different microorganisms in cutaneous communities. A question that remains is whether functional units provide additional insight and are more effective at characterizing microbiome datasets than the taxonomic units currently in use. Xu et al. (2014) found that taxonomic profiles are better at classifying samples into biologically meaningful categories. However, this may change with technological advances, including improvement in predictive tools and database annotations.

Although our study focused on the effect of primer selection and sequencing approach, there are many other factors to consider when designing a skin microbiome survey. First, the sample collection technique should be consistent throughout the entire study. Here, we used a skin swab method that is minimally invasive. Other studies have reported the utility of deeper sampling of the skin layers (Nakatsuji et al., 2013). The utility of these sampling methods for WMS sequencing, however, probably is limited because the amount of human DNA present in these samples would greatly overwhelm the microbial DNA present.

Second, studies investigating low-biomass sites, such as the skin, must also take sequencing depth into account and

use appropriate controls (Salter et al., 2014). Controls accounting for reagent contamination are critical for interpreting results. We recommend eliminating potential contaminants whenever possible during sample preparation by purchasing high-quality, ultrapure, DNA-free reagents, treating equipment and reagents with UV, and performing all experiments in a hood.

Third, computational analysis and selection of variables, such as OTU picking method and alpha diversity metric, can greatly impact the interpretation of results. We used default and commonly used variables, when possible, to make the analysis widely applicable.

Finally, we did not obtain cultures in parallel to collecting skin swabs for microbiome analysis to compare our results. Although it would be a point of interest to compare cultures to 16S tag sequencing for deciphering community composition, we expect that, as reported previously in several examples (Findley et al., 2013; Gardner et al., 2013), cultures would greatly underestimate the diversity and composition of the skin microbiota.

Overall, our comparison of three different DNA sequencing methods indicates that 16S tag sequencing of the V1—V3 region is a reasonable, cost-effective approach to simply profiling the composition of a skin microbial community or identifying biomarkers associated with skin disease.

MATERIALS AND METHODS
Sample collection
The University of Pennsylvania Internal Review Board approved all human subject recruitment and sample collection. Healthy adult human volunteers residing in Philadelphia, Pennsylvania, and surrounding areas were recruited to provide cutaneous swabs. Sample collection was performed after written, informed consent was obtained from the subjects. Exclusion criteria included self-reported antibiotic treatment (oral or systemic) 6 months before enrollment, observable dermatologic diseases, and significant comorbidities, including HIV and other immunocompromised states. Subjects were instructed to avoid hand sanitizers and antimicrobial soaps and skincare products for 1 week before the sample collection appointment. Subjects were also instructed not to shower for 24 hours before the sample collection appointment. Cutaneous swabs (Epicentre, Madison, WI) were collected as described previously (Grice et al., 2009) and stored in 300 μl of yeast cell lysis solution (from Epicentre MasterPure Yeast DNA Purification kit) at −20 °C immediately after collection. Swabs were incubated for 1 hour at 37 °C with shaking and 10,000 units of ReadyLyse Lysozyme solution (Epicentre). Samples were subjected to bead beating for 10 minutes at maximum speed on a vortex mixer with 0.5-mm glass beads (MoBio, Carlsbad, CA), followed by a 30-minute incubation at 65 °C with shaking. As previously described (Gardner et al., 2013), protein precipitation reagent (Epicentre) was added, and samples were spun at maximum speed. The supernatant was removed, mixed with isopropanol, and applied to a column from the PureLink Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA). Instructions for the Invitrogen PureLink kit were followed exactly, and DNA was eluted in 50 ml of elution buffer (Invitrogen). At each sampling event, swab control samples that never came into contact with the skin were collected, prepared, and sequenced exactly as the experimental samples. No significant background contamination from either reagents and/or collection procedures was recovered.

16S rRNA sequencing, sequence processing, and analysis
Sequencing libraries were prepared using the Invitrogen Accuprime for PCR, the AMPure kit (Beckman Coulter, Brea, CA) for PCR product cleanup and normalization, and the Qiagen (Valencia, CA) MinElute column for pooled PCR product purification. Sequencing was performed at the Penn Next Generation Sequencing Core on the Illumina MiSeq. The mock community control (MCC; obtained from BEI Resources, NIAID, NIH, as part of the Human Microbiome Project: Genomic DNA from Microbial Mock Community B [Even, Low Concentration], version 5.1L, for 16S rRNA Gene Sequencing, HM-782D) was sequenced in parallel with experimental samples. Sequencing of the V4 region was performed using 150-bp paired-end chemistry, and reads between 248 and 255 nucleotides long were retained for analysis (99.58% of total sequences). Sequencing of the V1—V3 region was performed using 300-bp paired-end chemistry, and reads between 465 and 535 nucleotides long were retained (96.74% of total sequences). Samples were processed in QIIME 1.8.0 (Caporaso et al., 2010), and statistical analysis and visualization was performed in the R statistical computing environment (R Core Team, 2015) as follows. Sequences were clustered into OTUs with a 97% similarity threshold by reference-based Uclust clustering (Edgar, 2010), using the Greengenes database 13_8 (DeSantis et al., 2006). Taxonomic classification was assigned using the RDP classifier (Wang et al., 2007). Chimeric sequences were identified using ChimeraSlayer (Haas et al., 2011) and removed along with those identified as Unclassified or Cyanobacteria. OTUs were removed if they only represented one sequence or were present in only one sample. Samples were rarefied to an even depth of 2,500 sequences per sample, after which alpha and beta diversities were calculated. In addition to the OTU-based methods, the MCC datasets were blasted against a custom database (blastn, max_target_seqs 1, e$<10^{-10}$; alignment length >300 for V1—V3 and >150 for V4 samples) to calculate community composition. Sequences classified as *Staphylococcus* at the genus level were analyzed using the pplacer algorithm with "—keep-at-most 100 —max-pitches 100" (Matsen et al., 2010) and a curated phylogenetic reference package (Conlan et al., 2012). Taxonomic classifications were generated using the guppy program, and species-level classifications with a maximum likelihood >0.75 were retained. "Closed-reference" OTU picking against the Greengenes database, with OTUs assigned at 97% identity, was used to generate biom-formatted OTU tables for functional prediction with PICRUSt (Langille et al., 2013) that were subsequently annotated with HUMAnN (HMP unified metabolic analysis network) version 0.99 (Abubucker et al., 2012). Kruskal-Wallis and multiple comparison post hoc tests were calculated in R with the pgirmess package (Giraudoux, 2015). Procrustes analysis was performed in R using beta diversity Bray-Curtis dissimilarity matrices generated in QIIME and the metaMDS and protest functions in the vegan package (Oksanen et al., 2015).

Whole metagenome sequencing and analysis
Libraries were prepared using the NexteraXT (Illumina, San Diego, CA) library preparation kit according to the manufacturer's instructions, except that PCR cycles were increased to 15. Additionally, instead of using the manufacturer's NexteraXT bead-based normalization protocol, we manually normalized and pooled based on DNA concentration and average fragment lengths. Sequencing was performed at the Penn Next Generation Sequencing Core on the Illumina MiSeq and/or HiSeq2500 rapid chemistry to obtain 150-bp paired-end reads.

Sequence data were obtained in fastq format. Adapters were removed using cutadapt (version 1.4.1) with an error rate of 0.1 and overlap of 10. Low-quality sequences (quality score <33) were removed using the standalone FASTX toolkit (version 0.0.14) with default parameters. Sequences mapping to the human genome were removed from the quality-trimmed dataset using the standalone DeconSeq toolkit (version 0.4.3) with default parameters and the human reference GRCh37 (Schmieder and Edwards, 2011). Because a 1% spike-in of PhiX Control was added to the sequencing runs for quality control purposes, any sequences mapping to the PhiX174 genome (NCBI Accession: NC_001422) were also removed using DeconSeq. Sequences <80 nucleotides long were removed from the quality trimmed, DeconSeq filtered fastq files, and one of the paired reads (SE1) was input into MetaPhlAn version 1.7.7 (Segata et al., 2012, 2013) for taxonomic classification. One of the paired ends (SE1) from the MCC sample was blasted against a custom database of genomes from the 20 expected bacterial species (blastn, max_target_seqs 1, e<10^{-10}; alignment length >50) to calculate community composition. Alpha diversity was calculated in vegan (Oksanen et al., 2015) using the biom table generated from MetaPhlAn output. For functional annotation and comparison, one set of the paired-end reads (SE1) for each sample was sub-sampled to 200,000 sequences, queried against a reduced KEGG reference database version 56 (blastx; max_target_seqs 1, e < 10^{-10}), and input into HUMAnN version 0.99 (Abubucker et al., 2012).

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at www.jidonline.org, and at http://dx.doi.org/10.1016/j.jid.2016.01.016.

REFERENCES
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. FASEB J 2013;27:1012—22.

Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol 2012;8:e1002358.

Albertsen M, Karst SM, Ziegler AS, Kirkegaard RH, Nielsen PH. Back to basics: the influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. PloS One 2015;10. e0132783—15.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335—6.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America. 2011;108:4516—22.

Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Biocrobiol Methods 2007;69:330—9.

Conlan S, Kong HH, Segre JA. Species-level analysis of DNA sequence data from the NIH Human Microbiome Project. PLoS One 2012;7:e47075—7.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 2006;72:5069—72.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26:2460—1.

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;498:367—70.

Fitz-Gibbon S, Tomida S, Chiu B-H, Nguyen L, Du C, Liu M, et al. Propionibacterium acnes strain populations in the human skin microbiome associated with acne. J Invest Dermatol 2013;133:2152—60.

Gardner SE, Hillis SL, Heilmann K, Segre JA, Grice EA. The neuropathic diabetic foot ulcer microbiome is associated with clinical factors. Diabetes 2013;62:923—30.

Gilbert JA, Jansson JK, Knight R. The Earth Microbiome project: successes and aspirations. BMC Biol 2014;12:69.

Giraudoux P. Pgirmess: data analysis in ecology. R package version 1.6.4, http://CRAN.R-project.org/package=pgirmess; 2015.

Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. Science 2009;324:1190—2.

Guo F, Ju F, Cai L, Zhang T. Taxonomic precision of different hypervariable regions of 16S rRNA gene and annotation methods for functional bacterial groups in biological wastewater treatment. PloS One 2013;8:e76185.

Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res 2011;21:494—504.

Hannigan GD, Grice EA. Microbial ecology of the skin in the era of metagenomics and molecular microbiology. Cold Spring Harbor Perspect Med 2013;3:a015362—2.

Hannigan GD, Meisel JS, Tyldsley AS, Zheng Q, Hodkinson BP, SanMiguel AJ, et al. The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. MBio 2015;6:e01578—15.

Human Microbiome Project Consortium. A framework for human microbiome research. Nature 2012a;486:215—21.

Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012b;486:207—14.

Jumpstart Consortium Human Microbiome Project Data Generation Working Group. Evaluation of 16S rDNA-based community profiling for human microbiome research. PLoS One 2012;7:e39315.

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 2013;31:814—21.

Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformat 2010;11:538.

Nakatsuji T, Chiang H-I, Jiang SB, Nagarajan H, Zengler K, Gallo RL. The microbiome extends to subepidermal compartments of normal skin. Nat Commun 2013;4:1431.

Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J. Analysis, optimization and verification of illumina-generated 16S rRNA gene amplicon surveys. PloS One 2014;9:e94249—14.

Oh J, Byrd AL, Deming C, Conlan S, NISC Comparative Sequencing Program, Kong HH, et al. Biogeography and individuality shape function in the human skin metagenome. Nature. 2014;514: 59—64.

Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara RB, et al. Vegan: community ecology package. R package version 2.3-3, 2015. http://CRAN.R-project.org/package=vegan.

R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: Austria; 2015.

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.

Schmieder R, Edwards R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PloS One 2011;6:e17288.

Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C. Computational meta'omics for microbial community studies. Mol Syst Biol 2013;9:666.

Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods 2012;9:811—4.

Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 2007;73:5261—7.

Wu GD, Lewis JD, Hoffmann C, Chen Y-Y, Knight R, Bittinger K, et al. Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. BMC Microbiol 2010;10:206—14.

Xu Z, Malmer D, Langille MGI, Way SF, Knight R. Which is more important for classifying microbial communities: who's there or what they can do? ISME J 2014;8:2357—9.