

Changes in the Human Skin Microbiome Over One Year's Time

^{1,2}Richard W. Hyman, ¹Farbod Babrzadeh,
¹Curtis Palm, ¹Chunlin Wang, ¹Marilyn Fukushima and ^{1,2,3}Ronald W. Davis

¹Stanford Genome Technology Center, 855 S California Avenue, Palo Alto, CA 94304, USA

²Department of Biochemistry,

³Department of Genetics,

Stanford University College of Medicine, Stanford, CA 94305, USA

ABSTRACT

Human skin comprises a large number of distinguishable ecological niches. To describe fully the human skin microbiome, it will be necessary to identify the bacteria in each niche and to distinguish the commensal bacteria from the temporary residents. To contribute to the description of the human skin microbiome and employing a gene-based technology, we have identified the bacteria in two niches: the front and back of the base of the neck and over the course of one year. There were 50 volunteers and a total of 232 neck skin swabs. Roche 454 Tag pyrosequencing was employed to sequence a short hypervariable sequence region (V6) of the 16S ribosomal RNA gene. To identify the bacteria corresponding to the front and back of the neck for each volunteer, the "Classifier" software in the "Pyrosequencing" section of the Ribosomal Database Project was employed. The bacteria on virtually all 232 neck skin swabs were classified into bacterial Class. The skin microbiome of these two niches was composed principally of a mixture of five Classes of bacteria: *Actinobacteria*, *Alphaproteobacteria*, *Bacilli*, *Betaproteobacteria* and *Gammaproteobacteria*. The fraction of each Class could change over time. We could not distinguish the skin microbiome from the front of the base of the neck from the back of the base of the neck. At these two positions, we could not distinguish the male from the female skin microbiome. The principal variable was the time point. We concluded that the skin microbiome at the front and back of the base of the neck was composed principally of a mixture of five Classes of bacteria. The proportion of each Class could change over time.

Keywords: Bacteria, Skin, Human Microbiome, Time

1. INTRODUCTION

Human skin is a large organ with a multitude of distinguishable niches in contact with the external environment. In order to delineate the entire human skin microbiome, it will be necessary to identify the microbes in every niche, as a function of time and with people living in different geographical areas (Peterson *et al.*, 2009). An excellent start has been made (Grice and Segre, 2011; Kong, 2011; HMPC, 2012).

Palo Alto, California, U.S.A., has a mild climate. However, that climate does change with the seasons. The season is part of the interactive environment of the skin bacteria and might have a significant impact on the

quantitative composition of the skin microbiome. To investigate this possibility, we have identified the bacteria on two human skin niches (the front and back of the base of the neck) for fifty volunteers and over the course of one year.

2. MATERIALS AND METHODS

2.1. Human Subjects

Volunteers were recruited under a protocol approved by the Stanford University Institutional Review Board. All volunteers were healthy adults and gave written informed consent. Anyone who took antibiotics during the previous three months or had a significant underlying

Corresponding Author: Richard W. Hyman, Stanford Genome Technology Center, 855 S California Avenue, Palo Alto, CA 94304, USA

medical condition was ineligible to volunteer for this study. In total, there were 50 volunteers (**Table 1**). Forty volunteers were enrolled at the first time point (February, 2010; **Table 1A**). An additional 10 volunteers were enrolled at the second time point (July, 2010; **Table 1B**). Although this was their first swab, we gave their swabs the designation “2” (for the time point) to keep the focus on the time point. No new volunteers were accepted at the third time point (February, 2011). The volunteers were not asked to change any part of their routine: e.g., neither their standard hygiene procedures nor the clothing that they chose to wear. This point may be important because, for example, Staudinger *et al.* (2011) demonstrated that the use of facial powder significantly increased the diversity of the human forehead skin microbiome. We have previously published the results from the first time point, albeit in a very different form (Hyman *et al.*, 2012). We included those results here for the purpose of comparisons.

2.2. Amplifying and Purifying the V6 DNAs

Each volunteer took her/his own swabs from the base of the front of the neck (overlying the suprasternal notch) and the base of the back of the neck (overlying the posterior cervical vertebrae: the nape, below the hair line) (Hyman *et al.*, 2012). Total DNA was isolated from each front and back neck swab employing a Qiagen DNeasy Blood and Tissue Kit. The total DNA was dialyzed and concentrated by the use of Amicon Ultra Centrifugal Filters (Millipore Corp). For PCR amplification of the V6 region of the 16S ribosomal RNA gene (rDNA), the forward and reverse primers were from Dethlefsen *et al.* (2008). Five identical reactions were run in parallel for each template. The PCR conditions were from Hyman *et al.* (2005). Following amplification, the five identical reactions for each template were pooled. The V6 DNA was purified by gel electrophoresis. We have employed Qiagen DNeasy Blood and Tissue Kits, Amicon Ultra Centrifugal Filters, PCR and purification of the amplicons by gel electrophoresis in previously reported experiments without problems of environmental contamination (Hyman *et al.*, 2012; 2005).

2.3. Pyrosequencing (Roche 454 Life Sciences)

For the first time point only (February, 2010), the front and back amplicons for each volunteer shared the same bar code (Hyman *et al.*, 2012). Therefore, the combined front V6s were pyrosequenced separately from the back V6s. For the second time point, the front and back V6s had different barcodes (Wang *et al.*, 2012), so they were combined before pyrosequencing, as was also the case for the V6s for the third time point.

Table 1. Volunteers

Volunteer ID	Male/Female	Long hair in back
A. Volunteers enrolled in February, 2010		
01	M	
02	M	
03	M	
04	F	
05	M	
06	M	
07	F	Y
08	F	Y
09	F	Y
10	F	Y
11	F	Y
12	M	
13	M	
14	M	
15	M	
16	M	
17	M	
18	M	
19	M	Y/N
20	F	Y
21	M	
22	M	
23	F	Y
24	M	
25	M	
26	F	
27	M	
28	M	
29	F	
30	M	
31	M	
32	M	
33	F	Y
34	M	
35	M	
36	F	
37	M	
38	M	
39	M	
40	F	Y
B. Volunteers enrolled in July, 2010		
Volunteer ID	Male/Female	Long hair in back
51	M	
57	F	
58	F	Y
63	F	Y
66	F	
70	F	
71	M	
75	F	Y
76	F	Y
82	F	Y

M = Male. F = Female. Y = yes. Y/N = intermediate long hair in back

Our detailed pyrosequencing procedures were published previously (Hyman *et al.*, 2012).

The sequence reads were sorted by barcode and, thereby, assigned to the front or back of a specific volunteer. Then, the pyrosequencing reads were stripped of extraneous sequences. A data set was created that consisted of each unique sequence obtained for that sample and the number of times that sequence was represented in the sample. To identify the bacteria corresponding to the front and back of the neck for each volunteer, the “Classifier” software (Wang *et al.*, 2007) in the “Pyrosequencing” section of the Ribosomal Database Project (RDP) was employed (Cole *et al.*, 2009). Only reads that could be classified were considered further. There were reads too short to be classified. For example, for 3F, these short reads averaged 1.2% (n = 35) of the total. The DECIPHER software was employed to identify chimeras (Wright *et al.*, 2012). In virtually all cases, the software identified the bacteria by Class. A very few reads in some sets were identified not by Class but by genus: e.g., TM7. In these cases, the reads were subsumed into Class. In a minority of cases, the software also identified the bacteria by Order. *Actinobacteria* were divided into Subclasses rather than Order. These were subsumed into Order.

2.4. Analyses

Chao1 (Colwell and Coddington, 1994; Hughes *et al.*, 2001), Principal Component Analysis (PCA) and Shannon Diversity Index (SDI) analyses employed the QIIME, “Quantitative Insights into Microbial Ecology”, software (Caporaso *et al.*, 2010) with UniFrac distances (Lozupone *et al.*, 2011). The SDI (Shannon, 1948) was calculated for the microbiome of each neck swab (<http://math.hws.edu/javamath/ryan/DiversityTest.html>).

3. RESULTS AND DISCUSSION

From the 50 volunteers, we achieved data for a total of 232 neck swabs, divided between the front (115) and the back (117) of the base of the neck (**Table 1**). We have front and back swabs from three time points for 29 volunteers; two time points for 11 volunteers; and only one time point for 10 volunteers. The three time points were designated as 1, 2 and 3, respectively. Bacteria were identified by 454 Tag pyrosequencing of a short hypervariable sequence region (V6) of the 16S ribosomal RNA gene (Huse *et al.*, 2008).

3.1. Class

In sampling an ecological niche, cost and time considerations prevent exhaustive sequencing. Therefore,

statistical tests have been derived that indicate the closeness to saturation. Chao1 is one such test (Colwell and Coddington, 1994). The results of Chao1 analyses of the data are shown in **Fig. 1** and demonstrate that the data are not close to saturation (see Discussion).

The bacteria on all 232 neck skin swabs were classified into Class. The complete data for all 232 swabs are presented on the Stanford Genome Technology Center (SGTC) web site (Table S1; http://med.stanford.edu/sgtc/research/skin_microbiome.html). **Table 2** presents examples of the bacterial identifications. To make the data visually comprehensible, only those bacteria supported by, at least, 1% of the reads are reported in the tables. The overall average per swab (n = 232) of the total percent of the bacteria supported by more than 1% of the reads was 98.2+/-1.0% (**Table 2**). We concluded that this very high percentage demonstrated that little information was lost when the focus was on those bacteria supported by, at least, 1% of the reads.

Fifteen Classes of bacteria were identified (**Table 3**). The presence of five bacterial Classes was supported by more than 1% of the reads on virtually all of the swabs, albeit in very different amounts: *Actinobacteria*, *Alphaproteobacteria*, *Bacilli*, *Betaproteobacteria* and *Gammaproteobacteria* (**Table 3**). *Gammaproteobacteria* were supported by the most reads for 85% of the swabs, followed by *Actinobacteria* (10%) and *Bacilli* (5%). On the second and third time points, *Flavobacteria* and *Sphingobacteria* appear on virtually all swabs, also in very different amounts. On the other hand, the presence of some bacteria was supported by more than 1% of the reads only peripatetically. *Bacteroidia* were found on only 11 swabs (4.7%) and only in small amounts, as were *Deinococci* (one swab, 0.4%), *Deltaproteobacteria* (3 swabs; 1.3%), *Fusobacteria* (5 swabs; 2.2%) and *Verrucomicrobia* (one swab; 0.4%). *Cyanobacteria* were not found on swab 3, while Chloroplasts were found only on swab 3. The presence of *Proteobacteria* were supported by reads in the amounts of *Gammaproteobacteria* >>> *Betaproteobacteria* > *Alphaproteobacteria*. The blank spaces in the data tables (**Table 2** and **Table S1**; http://med.stanford.edu/sgtc/research/skin_microbiome.html) signify that that particular bacterium's presence was supported by <1% of the sequence reads, if present at all. As one example, there were three blank spaces for *Bacilli* in the complete data table (Table S1; http://med.stanford.edu/sgtc/research/skin_microbiome.html). However, there were sequence reads supporting the presence of *Bacilli* on all three of those swabs: 33-1F, 0.43% of the reads; 33-1B, 0.70%; and 34-1F, 0.75%. Analogously, there were reads supporting the presence of

Alphaproteobacteria and *Betaproteobacteria* on all swabs, albeit sometimes at a frequency of < 1%. As additional examples for what was not reported in **Table 2**, for 01-3F, the presence of *Deinococci* was supported by three reads (0.41%) and the presence of *Deltaproteobacteria* was supported by four reads (0.54%).

Analyses were undertaken for four variables: location (front versus back), gender (male versus female), hair length in back (long hair versus short hair) and time point. For an analysis of the variable of location, we

calculated the average percent per swab (front and back separately) for each of the five major bacteria at each of the three time points. The results of these calculations are presented in **Fig. 2**. For all five bacterial Classes, there was no statistically significant difference between the paired front and back swabs. As examples, for *Gammaproteobacteria*, 1F was not statistically significantly different from 1B ($p = 0.88$); 2F was not statistically significantly different from 2B ($p = 0.62$); and 3F was not statistically significantly different from 3B ($p = 0.38$).

Table 2. Examples of Class data. The bacterial Class composition as percent of four skin swabs as a function of time and location. Only those bacterial Classes supported by >1% of the sequence reads appear in the Table. The skin swabs were from volunteers 01, 02, 03 and 04. Time points 1, 2 and 3. F = front of the base of the neck. B = back of the base of the neck

Class (%)	volunteer ID	<i>Actinobacteria</i>	<i>Alphaproteobacteria</i>	<i>Betaproteobacteria</i>	<i>Chloroplast</i>	<i>Clostridia</i>	<i>Cyanobacteria</i>	<i>Flavobacteria</i>	<i>Gammaproteobacteria</i>	<i>Sphingobacteria</i>	Total (%)	Total reads
Winter	01-1F	5.8	1.3	27.1	1.3				63.4		98.9	8036
Summer	01-2F	8.6	3.4	10.9	9.2	1.2	1.3	4.7	58.2		97.5	595
Winter	01-3F	7.8	8.3	10.5	17.1	1.0		9.7	36.7	6.0	97.1	735
Winter	01-1B	9.9	1.4	3.9	1.7				81.5		98.4	7113
Summer	01-2B	7.6	4.2	3.0	11.6	1.0	1.0	6.3	62.7	2.3	99.7	432
Winter	01-3B	7.1	8.7	5.5	17.5			9.2	42.9	5.7	96.6	750
	02-1F	22.8	2.7	3.9	1.3				67.9		98.6	6052
	02-2F	17.2	2.9	15.2	7.4		2.5	3.1	49.2	2.0	99.5	551
	02-3F	26.2	2.9	12.6	7.7			4.7	42.2	2.4	98.7	803
	02-1B	18.3	2.6	2.4	1.3				74.7		99.3	3405
	02-2B	20.0	2.4	4.9	7.7		1.4	4.9	55.9	2.0	99.2	506
	02-3B	21.2	7.0	5.5	9.8	1.1		6.8	43.9	3.5	98.8	711
	03-1F	6.2	1.3	3.7	1.5				85.6		98.3	2752
	03-1B	12.8	1.2	4.6					79.2		97.8	1923
	04-1F	no data										
	04-2F	5.9	3.3	4.8	9.6	1.5		5.6	63.8	2.8	97.3	539
	04-3F	4.9	3.7	3.2	11.7			2.5	68.5	2.2	96.7	596
	04-1B	no data										
	04-2B	17.8	3.8	8.5	5.7	1.0		5.9	53.6	1.4	97.7	422
	04-3B	12.2	2.4	7.5	9.3	1.5		7.2	55.8	3.2	99.1	787

Table 3. Bacterial Class total percentages. The percent of the swabs with the bacterial Class supported by >1% of the reads as a function of time and location. Time points 1, 2 and 3. F = front of the base of the neck. B = back of the base of the neck

Class	1	2	3	1F	2F	3F	1B	2B	2B
	Winter Total	Summer Total	Winter Total	Winter Front	Summer Front	Winter Front	Winter Back	Summer Back	Winter Back
<i>Actinobacteria</i>	98.7	100.0	100.0	100.0	100.0	100.0	97.4	100.0	100.0
<i>Alphaproteobacteria</i>	87.2	96.5	100.0	84.6	97.6	100.0	89.7	95.3	100.0
<i>Bacilli</i>	93.6	100.0	100.0	92.3	100.0	100.0	94.9	100.0	100.0
<i>Bacteroidia</i>	0.0	8.2	7.1	0.0	11.9	11.8	0.0	4.7	2.8
<i>Betaproteobacteria</i>	91.0	100.0	100.0	94.9	100.0	100.0	87.2	100.0	100.0
<i>Chloroplast</i>	0.0	0.0	44.3	0.0	0.0	38.2	0.0	0.0	50.0
<i>Clostridia</i>	2.6	43.5	64.3	2.6	50.0	73.5	2.6	37.2	55.5
<i>Cyanobacteria</i>	6.4	38.8	0.0	5.1	35.7	0.0	7.7	41.9	0.0
<i>Deinococci</i>	0.0	0.0	1.4	0.0	0.0	2.9	0.0	0.0	0.0
<i>Deltaproteobacteria</i>	0.0	0.0	4.3	0.0	0.0	2.9	0.0	0.0	5.6
<i>Flavobacteria</i>	44.9	98.8	100.0	43.6	100.0	100.0	46.2	97.7	100.0
<i>Fusobacteria</i>	0.0	2.4	1.4	0.0	4.8	2.9	0.0	0.0	0.0
<i>Gammaproteobacteria</i>	98.7	100.0	100.0	100.0	100.0	100.0	97.4	100.0	100.0
<i>Sphingobacteria</i>	48.7	82.4	100.0	48.7	78.6	100.0	48.7	86.0	100.0
<i>Verrucomicrobia</i>	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	2.8
Total swabs	77.0	85.0	70.0	39.0	42.0	34.0	38.0	43.0	36.0

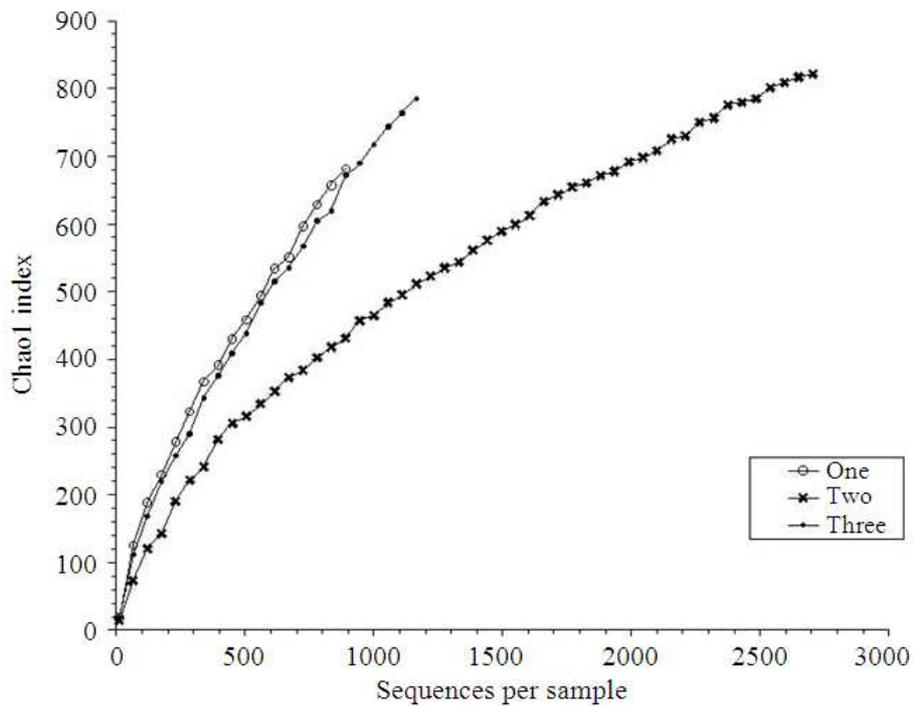


Fig. 1. Chao1 analysis of the data from the three time points. Time point 1, open circles; time point 2, crosses; and time point 3, closed circles

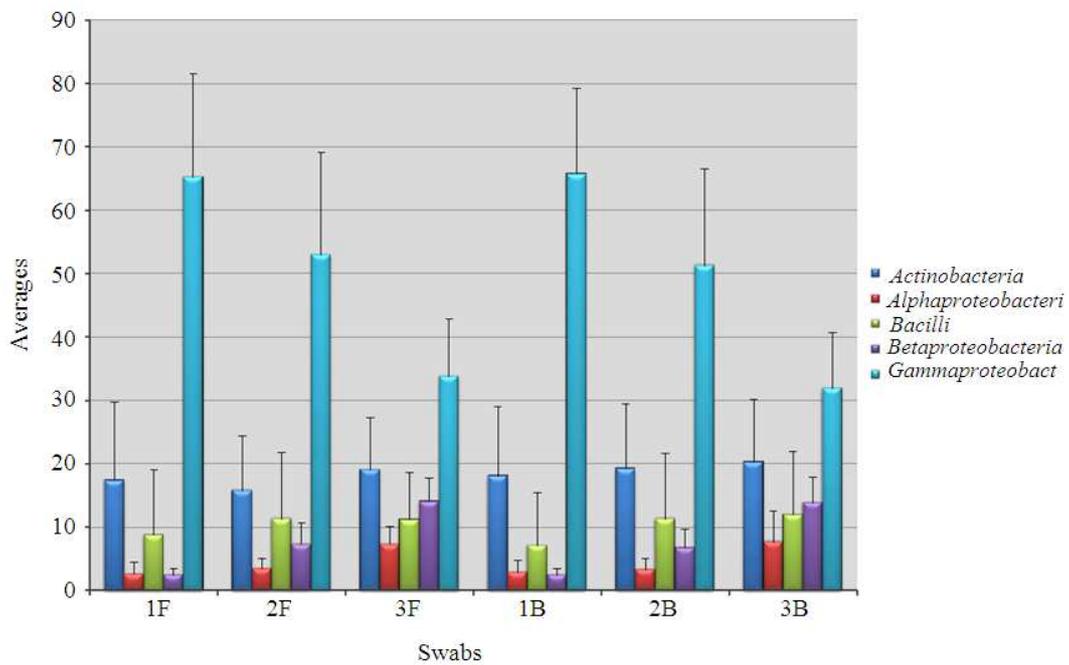


Fig. 2. Class time course. The average percentage and standard deviation of each of five bacterial Classes as a function of time and location. Time points 1, 2, and 3. Front (F) and back (B) of the base of the neck. Dark blue, *Actinobacteria*; red, *Alphaproteobacteria*; green, *Bacilli*; purple, *Betaproteobacteria*; light blue, *Gammaproteobacteria*

In addition, Principal Coordinate Analysis (PCA) did not distinguish Front from Back (Fig. S1; http://med.stanford.edu/sgtc/research/skin_microbiome.html).

For an analysis of the variable of gender, we compared the skin microbiome at the base of the neck of the 27 males to that of the 12 females. None of the gender comparisons yielded statistically significant differences (data not shown). As one example, we compared the percents of *Actinobacteria* and *Gammaproteobacteria* for men and women on swab 1F. There were no statistically significant differences ($p = 0.15$ and 0.12 , respectively). In addition, PCA did not distinguish females from males (Fig. S1; http://med.stanford.edu/sgtc/research/skin_microbiome.html).

Thirteen women had long hair in back and six women had short hair in back. To determine the effect of long hair in back on the skin microbiome at the base of the back of the neck, we compared the percents of *Actinobacteria* and *Gammaproteobacteria* for women with and without long hair in back for swab 2B. The differences were not statistically significant ($p = 0.70$ and 0.85 , respectively). In addition, PCA did not distinguish women with and without long hair in back (Fig. S1; http://med.stanford.edu/sgtc/research/skin_microbiome.html).

Lastly, we undertook an analysis of the time points, which did produce statistically significant results. For example, 1F was statistically significantly different from 2F ($p < 0.001$) and 2F was statistically significantly different from 3F ($p < 0.00001$) and the same for the back swabs. PCA was undertaken for the variable of time point. The results are shown in **Fig. 3**. The time points clustered and 34% of the variation in P1 versus P2 and P1 versus P3 could be ascribed to time point.

Bacterial diversity was examined by calculating the widely used Shannon Diversity Index (SDI) for each swab (Shannon, 1948). For example, when only one bacterium is present, the SDI = 0; there is no diversity. The SDI data are presented in **Table 4**, which also contains the average values and standard deviations for each group of swabs. There was no statistically significant difference between the average SDIs of 1F (1.37 ± 0.36) and 1B (1.38 ± 0.29 ; $p = 0.89$), 2F (1.95 ± 0.37) and 2B (1.94 ± 0.33 ; $p = 0.89$) and 3F (2.50 ± 0.22) and 3B (2.46 ± 0.16 ; $p = 0.39$). In addition, we were unable to distinguish the SDIs of males and females. As

examples, the SDI of the male ($n = 27$) front swab 1F was not statistically significant different from the female ($n = 12$) front swab 1F ($p = 0.23$). The SDI of the male ($n = 24$) back swab 2B was not statistically significant different from the female ($n = 18$) back swab 2B ($p = 0.46$). When we compared the SDIs of the swabs as a function of time point, an interesting pattern emerged: the average SDIs of the time points were statistically different from each other. As examples, the average SDI of the front swabs from the first time point was statistically significantly different from the average SDI of the front swabs of the second time point ($p < 0.00001$). The average SDI of the front swabs from the second time point was statistically significantly different from the average SDI of the front swabs from the third time point ($p < 0.00001$). The analogous statistically significant differences hold true for the SDIs of the back swabs.

To investigate the change in diversity as a function of time in more detail, we determined that we had data for 129 transitions: e.g., from 01-1F to 01-2F and from 01-2F to 01-3F (Table S2; http://med.stanford.edu/sgtc/research/skin_microbiome.html). We considered a change in the percentage of reads supporting the presence of a bacterium of less than an absolute 10% as “no change”, an arbitrary, but reasonable amount which we have employed previously (Hyman *et al.*, 2012). For example, for the 01-1 F to 01-2F transition, the percentage of reads supporting *Gammaproteobacteria* went from 63.4 to 58.2%. For these numbers to register as a change, 01-2F would have to be greater than 73.4% or less than 53.4%. On the other hand, within the same transition, the percentage of reads supporting *Bacilli* went from 27.1 to 10.9%. These numbers were different by an absolute 16.2%. Therefore, there was a “change”. With this definition, 29 of 129 transitions (23%) register as “no change”. An example is the 02-2F to 02-3F transition (Table S2; http://med.stanford.edu/sgtc/research/skin_microbiome.html), where the largest change is for *Actinobacteria*, which went from 17.2 to 26.2%, an absolute change of 9%. There was no series of 01-to-02-to-03 where there was no change. We considered a change of an absolute 20% or more as a large change: again, an arbitrary, but reasonable amount that we have employed previously (Hyman *et al.*, 2012). There were 47 out of 129 transitions (36%) that were more than an absolute 20%.

Table 4. Class SDI. The SDI was calculated from the bacterial Class data for each swab

ID	Winter 1F	Summer 2F	Winter 3F	Winter 1B	Summer 2B	Winter 3B
01	1.33	1.97	2.56	0.93	1.92	2.36
02	1.27	2.18	2.19	1.11	1.98	2.34
03	0.78			0.92		
04	No data	1.83	1.54	No data	2.05	2.10
05	0.76	1.84	2.75	1.32	1.80	2.32
06	0.94	1.74		1.06	2.20	
07	1.18	1.57		1.39	1.70	
08	0.89	1.74	2.42	0.88	1.80	2.25
09	1.39	1.45	2.60	1.38	1.63	2.52
10	0.78	1.73	2.40	1.27	1.87	2.53
11	0.78	1.50	2.56	0.90	1.52	2.43
12	1.45	1.61	2.63	1.02	1.75	2.38
13	1.54	1.49	2.35	1.50	1.67	2.52
14	1.28			1.27		
15	1.68	2.76	2.61	1.52	2.39	2.60
16	1.33	2.20	2.65	1.47	1.99	2.63
17	0.92	1.71		1.03	1.74	
18	1.41	1.65	2.37	1.59	1.91	2.25
19	1.04	2.11		1.43	1.93	
20	1.01	2.15	2.40	No data	2.14	2.42
21	1.42			1.64		
22	1.64	1.89	2.67	1.75	1.57	2.41
23	1.04			1.43		
24	1.82	1.99	2.31	1.92	1.96	2.51
25	1.65	2.12	2.54	1.65	2.22	2.49
26	1.77	1.98	2.67	1.72	1.90	2.44
27	1.63		2.56	1.51		2.52
28	1.38	2.20	2.56	1.38	2.17	2.36
29	1.65	1.60	No data	1.10	1.88	2.75
30	1.63	No data	2.63	1.22	1.84	2.81
31	1.86	1.98	2.44	1.89	1.37	2.48
32	1.88	1.81	2.45	1.60	1.97	2.50
33	1.01	2.00	2.63	0.96	1.68	2.23
34	1.18	2.19	2.50	1.35	2.00	2.46
35	1.73	1.93	2.34	1.71	1.85	2.40
36	2.07			1.46		
37	1.88	1.84	2.56	1.81	1.56	2.63
38	1.67	1.84	2.42	1.49	1.52	2.36
39	1.06			1.20		
40	1.54	2.29	No data	1.79	2.30	2.38
51		1.76			1.75	
57		1.71	2.48		2.12	2.61
58		1.89			2.13	
63		1.60			1.82	
66		1.61			1.56	
70		2.68	2.67		2.70	2.67
71		2.73	2.55		2.38	2.15
75		2.71	2.89		2.75	2.66
76		2.91	2.40		2.81	2.62
82		1.35	2.55		1.49	2.60
No. swabs	39	42	34	38	43.00	36.00
Average	1.37	1.95	2.50	1.38	1.94	2.46
SD	0.36	0.37	0.22	0.29	0.33	0.16

No = Number of; SD = Standard Deviation

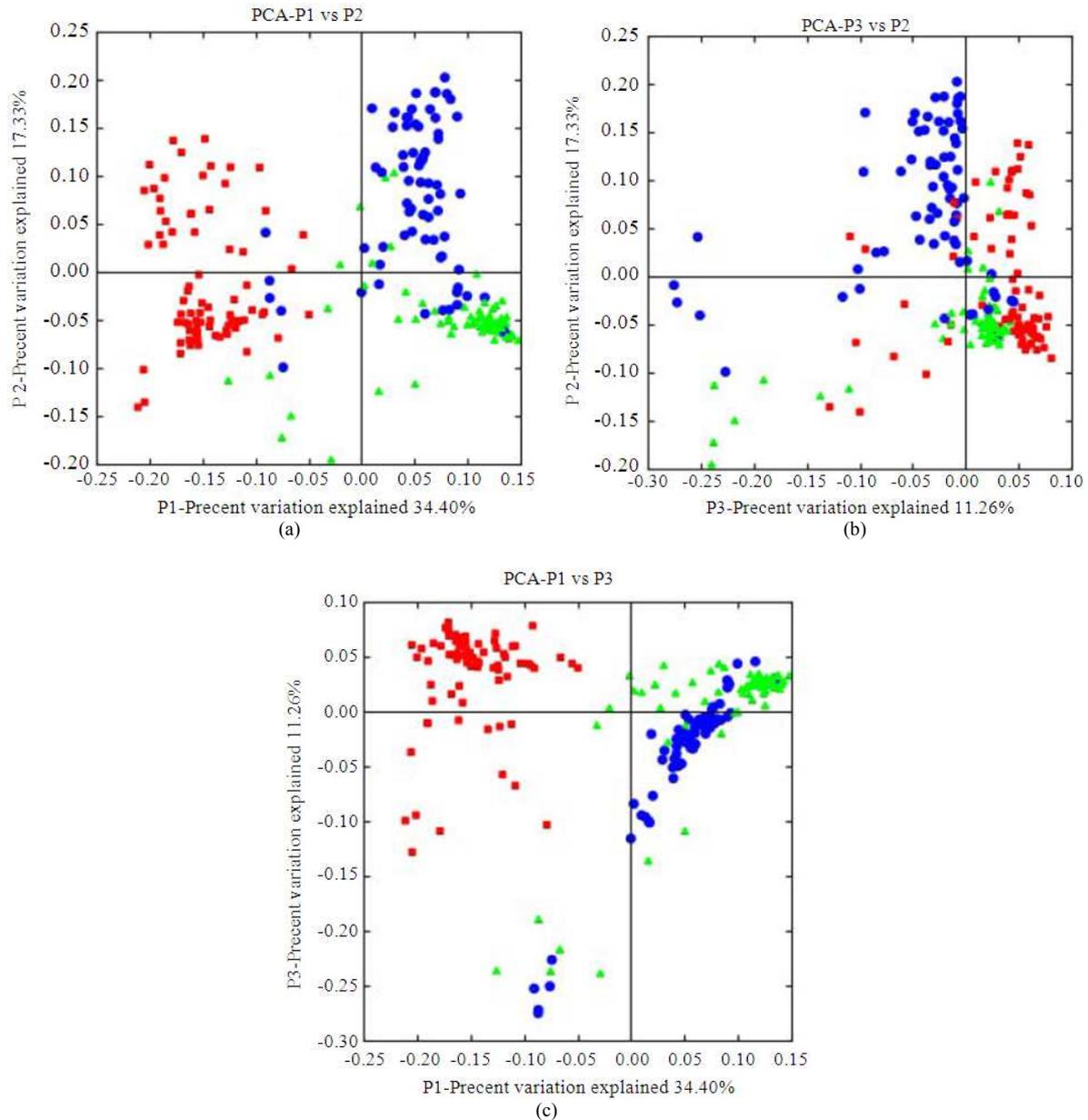


Fig. 3. PCA of time point. Time point 1, red squares; time point 2, green triangles; and time point 3, blue circles

There were three cases where there was more than an absolute 20% change in going from swab 1 to swab 2 to swab 3: e.g., 15-1F-to-15-2F-to-15-3F (Table S2; http://med.stanford.edu/sgtc/research/skin_microbiome.html). For this swab, the percentage of reads supporting the presence of *Gammaproteobacteria* went from 49.3-to-17.4-to-28.0%. There was no transition of greater than an absolute 30%.

The 129 transitions were divided into four groups: 1F-to-2F, 2F-to-3F, 1B-to-2B and 2B-to-3B. The changes in the composition of the three most abundant bacterial Classes were examined individually for the four types of transitions (Table 5). The percent of *Actinobacteria* and *Bacilli* seldom changed (up or down), while the percent of *Gammaproteobacteria* decreased by more than 50% in all four transitions (Table 5).

Table 5. Changes in Class percent for four time transitions: 1F to 2F, 2F to 3F, 1B to 2B, 2B to 3B. The average percent change, higher (Up), lower (Down), for four time transitions

Transition	No.	Up	Down	No change
1F to 2F	31			
<i>Actinobacteria</i>		6.5	12.9	80.6
<i>Bacilli</i>		19.4	6.5	74.2
<i>Gammaproteobacteria</i>		6.5	51.6	41.9
2F to 3F	32			
<i>Actinobacteria</i>		15.6	0.0	84.4
<i>Bacilli</i>		3.1	6.3	90.6
<i>Gammaproteobacteria</i>		6.3	68.8	25.0
1B to 2B	31			
<i>Actinobacteria</i>		16.1	9.7	74.2
<i>Bacilli</i>		19.4	3.2	77.4
<i>Gammaproteobacteria</i>		6.5	58.1	35.5
2B to 3B	35			
<i>Actinobacteria</i>		14.3	5.7	80.0
<i>Bacilli</i>		8.6	5.7	85.7
<i>Gammaproteobacteria</i>		0.0	74.3	25.7

Time points 1, 2 and 3. F = front of the base of the neck. B = back of the base of the neck

3.2. Order

The bacteria on 67 neck skin swabs (29% of the total swabs) were classified further into Order: 27 (40%) front swabs and 40 (60%) back swabs (Table S3; http://med.stanford.edu/sgtc/research/skin_microbiome.html). Presumably, our data and the RDP software were the determinants for which samples the bacteria could be identified by Order. As for the Class data, to make the Order data visually comprehensible, only those bacterial Orders supported by, at least, 1% of the reads are reported (Table S3; http://med.stanford.edu/sgtc/research/skin_microbiome.html). Thereby, 31 Orders of bacteria were identified, compared to 15 Classes of bacteria. (More Orders than Classes were expected because each Class is composed of more than one Order). Only one bacterial Order, *Actinobacteridae*, was supported by, at least, 1% of the reads on all of the 67 swabs (Table 6).

Table 6. Bacterial Order total percentages. The percent of the swabs with the bacterial Order supported by > 1% of the reads as a function of time and location

Order	Winter Total	Summer Total	Winter Total	Winter 1F	Summer 2F	Winter 3F	Winter 1B	Summer 2B	Winter 3B
<i>Acidimicrobidae</i>	0.0	2.6	(0) 0.0	0.0	(0)	0.04.8	(0)		
<i>Actinobacteridae</i>	100.0	100.0	(100)	100.0	100.0	(100)	100.0	100.0	(100)
<i>Aeromonadales</i>	0.0	57.9	(33.3)	0.0	70.6	(50)	0.0	47.6	(0)
<i>Alteromonadales</i>	76.9	60.5	(100)	62.5	64.7	(100)	83.3	61.9	(100)
<i>Bacillales</i>	96.2	100.0	(100)	100.0	100.0	(100)	94.4	100.0	(100)
<i>Bacteroidales</i>	0.0	5.3	(33.3)	0.0	5.9	(50)	0.0	4.8	(0)
<i>Bdellovibrionales</i>	0.0	2.6	(0) 0.0	5.9	(0)	0.00.0	(0)		
<i>Burkholderiales</i>	80.8	100.0	(100)	87.5	100.0	(100)	77.8	100.0	(100)
<i>Campylobacteriales</i>	0.0	2.6	(0) 0.0	5.9	(0)	0.00.0	(0)		
<i>Caulobacteriales</i>	3.8	50.0	(66.6)	12.5	58.8	(100)	0.0	42.9	(0)
Chloroplast	7.7	71.1	(66.6)	12.5	76.5	(100)	5.6	66.7	(0)
<i>Chromatiales</i>	73.1	78.9	(0)62.5	76.5	(0)	77.8	81.0	(0)	
<i>Clostridiales</i>	0.0	52.6	(66.6)	0.0	41.2	(100)	0.0	61.9	(0)
<i>Coriobacteridae</i>	0.0	2.6	(0) 0.0	5.9	(0)	0.00.0	(0)		
<i>Enterobacteriales</i>	100.0	97.4	(100)	100.0	94.1	(100)	100.0	100.0	(100)
<i>Flavobacteriales</i>	50.0	100.0	(100)	25.0	100.0	(100)	61.1	100.0	(100)
<i>Holophagales</i>	0.0	2.6	(0) 0.0	0.0	(0)	0.04.8	(0)		
<i>Hydrogenophilales</i>	0.0	2.6	(0) 0.0	5.9	(0)	0.00.0	(0)		
<i>Lactobacillales</i>	19.2	86.8	(100)	37.5	82.4	(100)	11.1	90.5	(100)
<i>Legionellales</i>	0.0	5.3	(0) 0.0	5.9	(0)	0.04.8	(0)		
<i>Neisseriales</i>	0.0	42.1	(33.3)	0.0	41.2	(50)	0.0	42.9	(0)
<i>Oceanospirillales</i>	0.0	7.9	(0) 0.0	5.9	(0)	0.09.5	(0)		
<i>Pseudomonadales</i>	69.2	100.0	(100)	62.5	100.0	(100)	72.2	100.0	(100)
<i>Rhizobiales</i>	65.4	73.7	(100)	50.0	76.5	(100)	72.2	71.4	(100)
<i>Rhodobacteriales</i>	3.8	7.9	(0) 0.0	5.9	(0)	11.1	9.5	(0)	
<i>Rhodocyclales</i>	0.0	31.6	(0) 0.0	35.3	(0)	0.028.6	(0)		
<i>Rhodospirillales</i>	3.8	2.6	(0) 0.0	0.0	(0)	11.1	4.8	(0)	
<i>Sphingobacteriales</i>	46.2	97.4	(100)	50.0	94.1	(100)	55.6	100.0	(100)
<i>Sphingomonadales</i>	3.8	39.5	(33.3)	0.0	35.3	(50)	11.1	42.9	(0)
<i>Verrucomicrobiae</i>	0.0	5.3	(0) 0.0	11.8	(0)	0.00.0	(0)		
<i>Xanthomonadales</i>	26.9	73.7	(100)	37.5	100.0	(100)	22.2	100.0	(100)
Total swabs	26.0	38.0	3	8.0	17.0	2	18.0	21.0	1

Time points 1, 2 and 3. F = front of the base of the neck. B = back of the base of the neck. The parentheses enclose percentages calculated from three or fewer swabs

Table 7. Bacterial Order total percentages. The percent of the swabs with the bacterial Order supported by >1% of the reads as a function of time and location. Time points 1 and 2. Time point 3 was not included because of a paucity of swabs

	1F	2F	1B	2B
Average (%)	95.0	96.0	95.4	95.8
SD	1.0	2.4	1.5	1.7
Number of swabs	8.0	17.0	18.0	22.0

F = front of the base of the neck. B = back of the base of the neck. SD = Standard Deviation

Table 8. Order SDI. The SDI was calculated from the bacterial Order data for each swab

Volunteer ID	Winter 1F	Summer 2F	Winter 3F	Winter 1B	Summer 2B	Winter 3B
01					3.71	
02		2.80			3.09	
03				1.63		
04		3.58				
05	1.75	2.64				2.53
06		3.20				
07	1.91			1.90	1.62	
08			3.40		3.45	
09					2.74	
10					2.74	
11		3.77				
12		3.38				
13		3.70				
15	2.20			1.66		
16				2.02	2.61	
17	1.50				3.30	
18		1.62				
19	1.89			1.58		
22					3.21	
23				2.31		
24		2.59			2.09	
25				1.84		
26		3.36			2.13	
27	1.50			1.92		
28				1.71		
29		3.34		2.04		
30	1.39			2.02		
31					2.79	
32		3.57			2.78	
34				1.92	3.62	
35	1.74	1.65		1.62	2.27	
36				1.90		
37				1.88	2.95	
38			3.17	1.88	3.34	
39				1.81		
a40				2.23		
51					3.34	
57		3.21			2.77	
58		3.44				
63		2.26				
66		3.43			3.52	
82					2.90	
No. swabs	8.00	17.00	2	18.00	22.00	1.00
Average	1.74	3.04		1.89	2.92	
SD	0.25	0.63		0.19	0.53	

No. = Number of; SD = Standard Deviation

Three additional bacteria Orders were supported by, at least, 1% of the reads on > 90% of all of the swabs: *Bacillales* (98.5%), *Enterobacteriales* (97.0%) and *Burkholderiales* (91.0%).

The average percentage of each bacterial Order supported by more than 1% of the reads on each neck swab is given in **Table 6**. *Actinobacteridae* were supported by the most reads for 58% of the swabs, followed by *Enterobacteriales* (29%), followed by *Bacillales* (10%), followed by *Burkholderiales* and *Flavobacteriales* (both at 1.4%). (Unfortunately, there were only three swabs with Order data for the third time point: two Front swabs and one Back swab. Therefore, for our further analyses, we ignored the Order data for the third time point.)

The average percentage of sequence reads for each swab supporting a bacterial Order by more than 1% of the reads is given in **Table 7**. None of the values was statistically significantly different from the others. For example, the value for 1F (95.0+/-1.0%) was not statistically significantly different from the value for 2F (96.0+/-2.4%; $p = 0.27$). The overall average percentage ($n = 67$) of each bacterial Order supported by more than 1% of the reads was 95.7+/-1.8%. This percentage was statistically significantly different ($p < 0.00001$) from the analogous average for the Class data. This difference was not surprising as the reads were spread over 31 Orders compared to 15 Classes.

The SDI was calculated for the Order data for each swab (**Table 8**). The average values and standard deviations for each group of swabs were also determined and reported in **Table 8**. The average SDI for 1F was not statistically different from the average SDI for 1B ($p = 0.10$). The average SDI for 2F was not statistically different from the average SDI for 2B ($p = 0.52$). In contrast, the average SDI of 1F was statistically significantly different from the average SDI of 2F ($p < 0.00001$) and the average SDI of 1B was statistically significantly different from the average SDI of 2B ($p < 0.00001$).

There were only eight transitions within the Order data: two 1F to 2F (volunteers 05 and 35) and four 1B to 2B (volunteers 07, 16, 34, 35, 37 and 38). For six of the eight transitions, the percentage of *Actinobacteridae* increased substantially, while for all eight transitions the percentage of *Enterobacteriales* decreased substantially.

4. CONCLUSION

The skin microbiome at the base of the front and back of the neck was composed principally of the same five

Classes of bacteria over the course of one year: *Actinobacteria*, *Alphaproteobacteria*, *Bacilli*, *Betaproteobacteria* and *Gammaproteobacteria*. We could not distinguish the front microbiome from the back microbiome, nor men from women. However, as demonstrated by SDI and PCA analyses, the skin microbiome at each of the three time points differed from the other two time points. That is, each time a neck swab was taken, the microbiome was unique, not as to the bacteria present but rather the relative percentages of the five principal Classes of bacteria. However, it must be noted that the mix of volunteers was different at each time point.

We chose to investigate the skin microbiome at the front and back of the base of the neck because no one else had published the skin microbiome at these two body sites, nor did we know of anyone else investigating those two sites. That makes comparisons of our results to published data difficult. Nevertheless, some comparisons may be made.

Our two studied skin sites may be classified as sebaceous sites (Kong, 2011). The closest previously studied sebaceous sites are the upper chest and the upper back. (The occiput, the back of the scalp, is not comparable to the nape, as our back swabs were taken below the hairline and the occiput is above the hairline). The microbiome of the upper chest is composed almost entirely of *Actinobacteria* (Kong, 2011). The microbiome of the upper back is composed principally of *Actinobacteria* with some *Proteobacteria* (Grice *et al.*, 2009). These comparisons again emphasize that small differences in skin location may give rise to substantial differences in the composition of the skin microbiome.

There were two important technical limitations in our data collection. The first limitation was the universality of the PCR primers. The PCR primers employed to amplify the V6 hypervariable region of the rDNA may have some mismatches to the rDNAs of relevant bacteria (Soergel *et al.*, 2012; Frank *et al.*, 2008). Presumably, such mismatch resulted in less amplification, especially in the critical first round. The second limitation was the depth of sequencing. A minimum of 1% of the sequence reads supporting the presence of a bacterium was required to report the presence of that bacterium. The Chao1 analysis (**Fig. 1**) suggested that there were bacteria present at concentrations below this limit. In addition, scientists working within the Human Microbiome Project have reported some very idiosyncratic results from employing rDNA sequence to identify bacteria (JCHMPDGWG, 2012; Haas *et al.*, 2011).

It is clear from this and previously published enumerations of the skin microbiome that many skin sites on many people in many geographic areas need to be studied and over a long period of time (Grice and Segre, 2011; Kong, 2011; Gao *et al.*, 2007; Fierer *et al.*, 2008; Costello *et al.*, 2009; Caporaso *et al.*, 2011). A good start has been made, but there is much more to be accomplished (Peterson *et al.*, 2009; HMPC, 2012).

5. ACKNOWLEDGEMENT

We thank the 50 volunteers who contributed skin swabs to this study. We thank Monica Trebo for posting the data on the Stanford Genome Technology Center website. This study was supported by Public Health Service grant PO1 HG000205 (R.W.D.) from the National Human Genome Research Institute. All authors declare that they have no financial conflicts of interest.

6. REFERENCES

- Caporaso, J.G., C.L. Lauber, E.K. Costello, D. Berg-Lyons and A. Gonzalez *et al.*, 2011. Moving pictures of the human microbiome. *Genome Biol.*, 12: R50-R50. DOI: 10.1186/gb-2011-12-5-r50
- Caporaso, J.G., J. Kuczynski, J. Stombaugh, K. Bittner and F.D. Bushman *et al.*, 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*, 7: 335-336. DOI: 10.1038/nmeth.f.303
- Cole, J.R., Q. Wang, E. Cardenas, J. Fish and B. Chai *et al.*, 2009. The ribosomal database project: Improved alignments and new tools for rRNA analysis. *Nucl. Acids Res.*, 37: D141-D145. DOI: 10.1093/nar/gkn879
- Colwell, R.K. and J.A. Coddington, 1994. Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. Royal Soc. B: Biol. Sci.*, 345: 101-118. DOI: 10.1098/rstb.1994.0091
- Costello, E.K., C.L. Lauber, M. Hamady, N. Fierer and J.I. Gordon *et al.*, 2009. Bacterial community variation in human body habitats across space and time. *Science*, 326: 1694-1697. DOI: 10.1126/science.1177486
- Dethlefsen, L., S. Huse, M.L. Sogin and D.A. Relman, 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.*, 6: e280-e280. DOI: 10.1371/journal.pbio.0060280
- Fierer, N., M. Hamady, C.L. Lauber and R. Knight, 2008. The influence of sex, handedness and washing on the diversity of hand surface bacteria. *Proc. Natl. Acad. Sci. USA.*, 105: 17994-17999. DOI: 10.1073/pnas.0807920105
- Frank, J.A., C.I. Reich, S. Sharma, J.S. Weisbaum and B.A. Wilson *et al.*, 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Applied Environ. Microbiol.*, 74: 2461-2470. DOI: 10.1128/AEM.02272-07
- Gao, Z., C.H. Tseng, Z. Pei and M.J. Blaser, 2007. Molecular analysis of human forearm superficial skin bacterial biota. *Proc. Natl. Acad. Sci. U.S.A.*, 104: 2927-2932. DOI: 10.1073/pnas.0607077104
- Grice, E.A. and J.A. Segre, 2011. The Skin Microbiome. *Nat. Rev. Microbiol.*, 9: 244-53. DOI: 10.1038/nrmicro2537
- Grice, E.A., H.H. Kong, S. Conlan, C.B. Deming and J. Davis *et al.*, 2009. Topographical and temporal diversity of the human skin microbiome. *Science*, 324: 1190-1192. DOI: 10.1126/science.1171700
- Haas, B.J., D. Gevers, A.M. Earl, M. Feldgarden, D.V. Ward and G. Giannoukos *et al.*, 2011. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.*, 21: 494-504. DOI: 10.1101/gr.112730.110
- HMPC, 2012. Structure, function and diversity of the healthy human microbiome. *Nature*, 486: 207-214. DOI: 10.1038/nature11234
- Hughes, J.B., J.J. Hellmann, T.H. Ricketts and B.J.M. Bohannan, 2001. Counting the uncountable: Statistical approaches to estimating microbial diversity. *Applied Environ. Micro.*, 67: 4399-4406. DOI: 10.1128/AEM.67.10.4399-4406.2001
- Huse, S.M., L. Dethlefsen, J.A. Huber, D.M. Welch and D.A. Relman *et al.*, 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.*, 4: e1000255. DOI: 10.1371/journal.pgen.1000255
- Hyman, R.W., F. Babrzadeh, C. Palm, C. Wang and M. Fukushima *et al.*, 2012. Bacteria on Human Skin. *Am. J. Micro.*, 2: 25-34. DOI: 10.3844/ajmsp.2011.25.34
- Hyman, R.W., M. Fukushima, L. Diamond, J. Kumm and L.C. Giudice *et al.*, 2005. Microbes on the human vaginal epithelium. *Proc. Natl. Acad. Sci. U.S.A.*, 102: 7952-7957. DOI: 10.1073/pnas.0503236102

- JCHMPDGWG, 2012. Evaluation of 16S rDNA-based community profiling for human microbiome research. *PLoS One*, 7: e39315-e39315. DOI: 10.1371/journal.pone.0039315
- Kong, H.H., 2011. Skin Microbiome: Genomics-Based Insights into the Diversity and Role of Skin Microbes. *Trends Mol. Med.*, 17: 320-8. DOI: 10.1016/j.molmed.2011.01.013
- Lozupone, C., M.E. Lladser, D. Knights, J. Stombaugh and R. Knight, 2011. UniFrac: An effective distance metric for microbial community comparison. *ISME J.*, 5: 169-172. DOI: 10.1038/ismej.2010.133
- Peterson, J., S. Garges, M. Giovanni, P. McInnes and L. Wang *et al.*, 2009. The NIH Human Microbiome Project. *Genome Res.*, 19: 2317-2323. DOI: 10.1101/gr.096651.109
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27: 379-423.
- Soergel, D.A., N. Dey, R. Knight and S.E. Brenner, 2012. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *ISME J.* DOI: 10.1038/ismej.2011.208
- Staudinger, T., A. Pipal and B. Redl, 2011. Molecular analysis of the prevalent microbiota of human male and female forehead skin compared to forearm skin and the influence of make-up. *J. Applied Microbiol.*, 110: 1381-1389. DOI: 10.1111/j.1365-2672.2011.04991.x
- Wang, C., S. Krishnakumar, J. Wilhelmy, F. Babrzadeh and L. Stepanyan *et al.*, 2012. High-throughput, high-fidelity HLA genotyping with deep sequencing. *Proc. Natl. Acad. Sci. U.S.A.*, 109: 8676-8681. DOI: 10.1073/pnas.1206614109
- Wang, Q., G.M. Garrity, M.J. Tiedje and J.R. Cole, 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied Environ. Microbiol.*, 73: 5261-5267. DOI: 10.1128/AEM.00062-07
- Wright, E.S., L. Safak Yilmaz and D.R. Noguera, 2012. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Applied Environ. Microbiol.*, 78: 3 717-725. DOI: 10.1128/AEM.06516-11