

## Supplementary Material

**Table S1: Summary of cohort characteristics and the types of cutaneous swabs collected**

<b>Cohort Characteristics</b>			<b>n=9</b>
<b>Age</b>	Median,y		25
	Range, y		24-53
<b>Female:Male Ratio</b>			7:2
<b>Sampling Date</b>	Median		Aug 20, 2013
	Range		Aug 19, 2013 - Sept 4, 2013

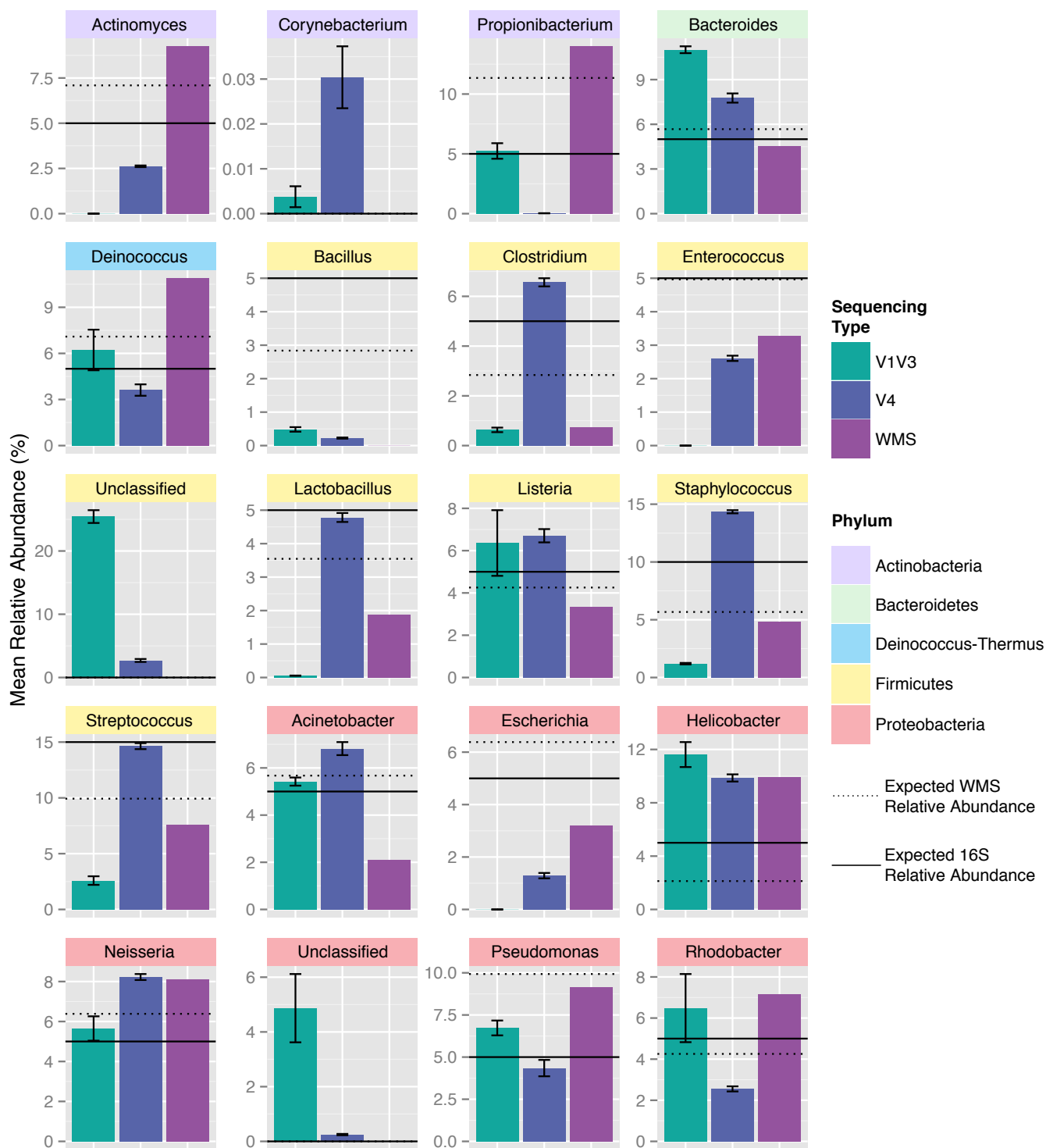
  

<b>Cutaneous Swabs</b>			<b>n=70</b>
<b>Site Microenvironment</b>	Sebaceous		26
	Intermittently Moist		18
	Moist		18
<b>Even Mock Community</b>	V1-V3		3
	V4		4
	WMS		1

**Table S2: Summary of samples with sequence counts for each sequencing technique**

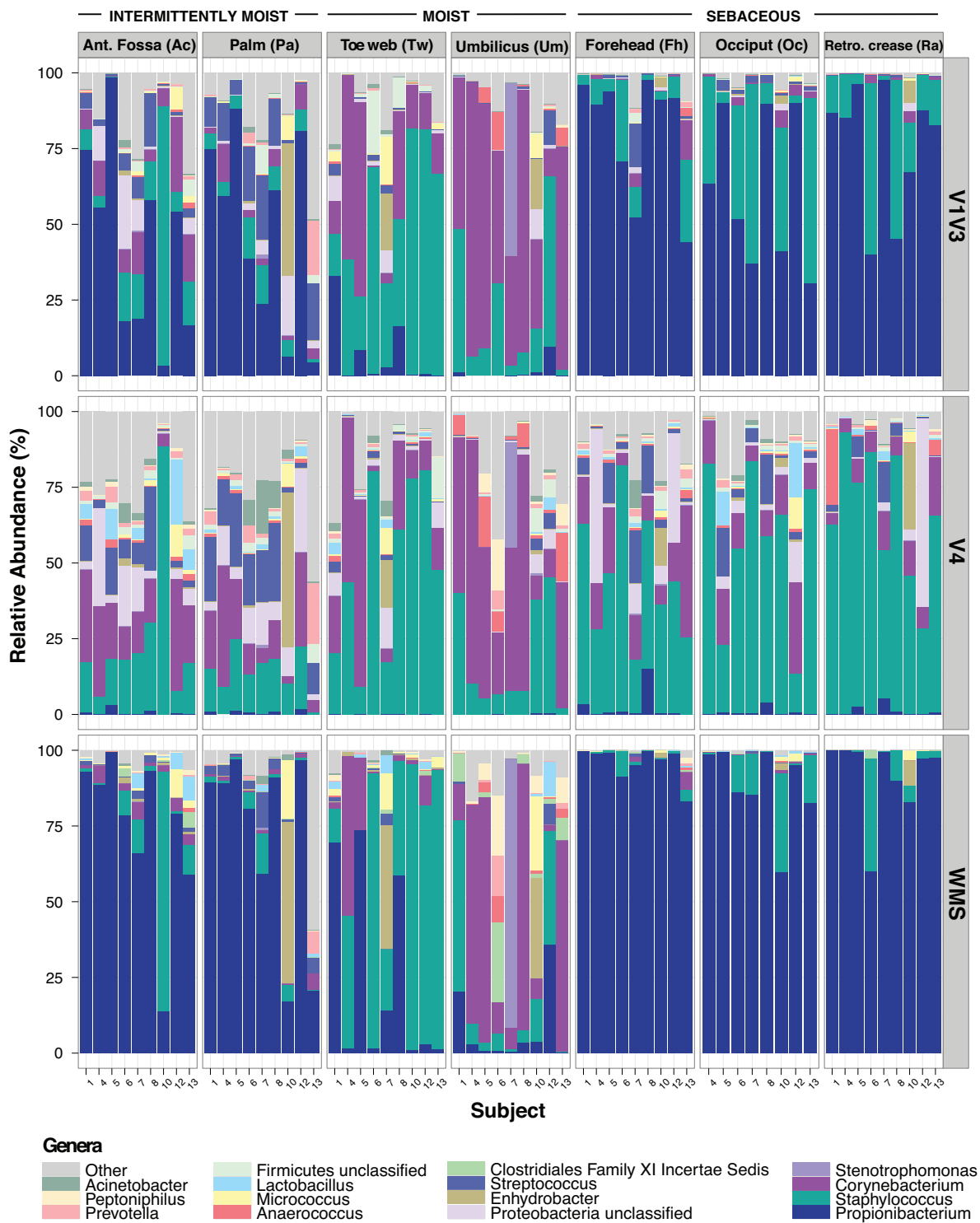
WMS SampleID	V4 SampleID	V1V3 SampleID	Subject ID	Age	Sex	Visit Date	Site Symbol	Site Category	V1V3 Sequence Counts	V4 Sequence Counts	WMS Sequence Counts
MG100171	102094	103071	1	24	F	8/19/13	Ra	Sebaceous	15699	54226	12313
MG100174	102130	103074	4	53	M	9/4/13	Ra	Sebaceous	40903	92500	24267
MG100175	102142	103075	5	30	F	8/20/13	Ra	Sebaceous	41334	47680	159595
MG100176	102150	103076	6	34	F	8/20/13	Ra	Sebaceous	20123	87232	34960
MG100177	102151	103077	7	24	F	9/3/13	Ra	Sebaceous	64298	48472	531901
MG100178	102091	103078	8	25	F	8/23/13	Ra	Sebaceous	90117	91823	73605
MG100180	102154	103080	10	25	F	8/19/13	Ra	Sebaceous	52433	85448	470524
MG100182	102156	103082	12	25	M	9/4/13	Ra	Sebaceous	48029	62534	474496
MG100183	102157	103083	13	26	F	8/19/13	Ra	Sebaceous	79184	99811	439125
MG100315	102093	102933	1	24	F	8/19/13	Fh	Sebaceous	92023	61387	1473353
MG100318	102101	102936	1	24	F	8/19/13	Um	Moist	32738	62695	1809124
MG100319	102102	102937	1	24	F	8/19/13	Tw	Moist	28204	71329	2668448
MG100321	102129	102939	4	53	M	9/4/13	Fh	Sebaceous	53314	83972	808779
MG100322	102131	102940	4	53	M	9/4/13	Oc	Sebaceous	70003	92151	1753679
MG100324	102137	102942	4	53	M	9/4/13	Um	Moist	37092	107392	1790933
MG100325	102138	102943	4	53	M	9/4/13	Tw	Moist	44715	64637	1764942
MG100327	102141	102945	5	30	F	8/20/13	Fh	Sebaceous	93971	76785	368729
MG100328	102143	102946	5	30	F	8/20/13	Oc	Sebaceous	32427	49238	423085
MG100330	102149	102948	5	30	F	8/20/13	Um	Moist	18482	117477	1442634
MG100331	102089	102949	5	30	F	8/20/13	Tw	Moist	18614	100382	428779
MG100333	102261	102951	6	34	F	8/20/13	Fh	Sebaceous	15232	70681	613509
MG100334	102263	102952	6	34	F	8/20/13	Oc	Sebaceous	10486	76023	204861
MG100336	102269	102954	6	34	F	8/20/13	Um	Moist	7047	149825	387543
MG100337	102270	102955	6	34	F	8/20/13	Tw	Moist	24292	125601	1296398
MG100339	102273	102957	7	24	F	9/3/13	Fh	Sebaceous	26173	64300	839201
MG100340	102275	102958	7	24	F	9/3/13	Oc	Sebaceous	39345	102790	1222102
MG100342	102281	102960	7	24	F	9/3/13	Um	Moist	15161	104701	1889565
MG100343	102282	102961	7	24	F	9/3/13	Tw	Moist	17503	85112	409067
MG100345	102285	102963	8	25	F	8/23/13	Fh	Sebaceous	43897	57904	786456
MG100346	102287	102964	8	25	F	8/23/13	Oc	Sebaceous	35935	54094	727641
MG100348	102293	102966	8	25	F	8/23/13	Um	Moist	15839	77618	973762
MG100349	102294	102967	8	25	F	8/23/13	Tw	Moist	31170	108378	1984989
MG100351	102573	102969	10	25	F	8/19/13	Fh	Sebaceous	35324	58284	947538
MG100352	102575	102970	10	25	F	8/19/13	Oc	Sebaceous	14955	51476	1434156
MG100354	102581	102972	10	25	F	8/19/13	Um	Moist	18183	83720	1163590
MG100355	102582	102973	10	25	F	8/19/13	Tw	Moist	29663	78358	2353225
MG100357	102597	102975	12	25	M	9/4/13	Fh	Sebaceous	20198	112034	2312001
MG100358	102599	102976	12	25	M	9/4/13	Oc	Sebaceous	16117	74434	1342272
MG100360	102605	102978	12	25	M	9/4/13	Um	Moist	21777	84801	1287361
MG100361	102606	102979	12	25	M	9/4/13	Tw	Moist	22947	105177	1605875
MG100363	102609	102981	13	26	F	8/19/13	Fh	Sebaceous	39053	79147	1238964
MG100364	102611	102982	13	26	F	8/19/13	Oc	Sebaceous	68733	79202	1233172
MG100366	102617	102984	13	26	F	8/19/13	Um	Moist	19384	85212	2815779
MG100367	102618	102985	13	26	F	8/19/13	Tw	Moist	34789	87142	7844414
MG100507	102098	104144	1	24	F	8/19/13	Ac	Intermittently_Moist	21971	60223	3515149
MG100508	102100	104146	1	24	F	8/19/13	Pa	Intermittently_Moist	30014	96226	2126984
MG100513	102134	104148	4	53	M	9/4/13	Ac	Intermittently_Moist	19777	29773	2011016
MG100514	102136	104153	4	53	M	9/4/13	Pa	Intermittently_Moist	19650	49066	2065821
MG100519	102146	104149	5	30	F	8/20/13	Ac	Intermittently_Moist	44336	61431	1425536
MG100520	102148	104154	5	30	F	8/20/13	Pa	Intermittently_Moist	40376	50105	1104271
MG100525	102266	104150	6	34	F	8/20/13	Ac	Intermittently_Moist	16750	72763	341628
MG100526	102268	104155	6	34	F	8/20/13	Pa	Intermittently_Moist	14607	115621	507266
MG100531	102278	104151	7	24	F	9/3/13	Ac	Intermittently_Moist	13623	79883	1176579
MG100532	102280	104156	7	24	F	9/3/13	Pa	Intermittently_Moist	19871	104244	1500649
MG100537	102290	104152	8	25	F	8/23/13	Ac	Intermittently_Moist	24504	76177	739144
MG100538	102292	104157	8	25	F	8/23/13	Pa	Intermittently_Moist	20929	72424	1249950
MG100543	102578	104158	10	25	F	8/19/13	Ac	Intermittently_Moist	8420	75235	1257309
MG100544	102580	104159	10	25	F	8/19/13	Pa	Intermittently_Moist	78984	68314	752084
MG100549	102602	104160	12	25	M	9/4/13	Ac	Intermittently_Moist	18584	70168	1027823
MG100550	102604	104165	12	25	M	9/4/13	Pa	Intermittently_Moist	36928	76984	1306366
MG100555	102614	104161	13	26	F	8/19/13	Ac	Intermittently_Moist	11098	78238	1608004
MG100556	102616	104166	13	26	F	8/19/13	Pa	Intermittently_Moist	8448	73998	1258881
NA	101803	NA	NA	NA	NA	NA	Mock	Control	NA	56555	NA
NA	102080	NA	NA	NA	NA	NA	Mock	Control	NA	146691	NA
NA	NA	103978	NA	NA	NA	NA	Mock	Control	24891	NA	NA
NA	102355	NA	NA	NA	NA	NA	Mock	Control	NA	114469	NA
NA	102570	NA	NA	NA	NA	NA	Mock	Control	NA	86442	NA
NA	NA	103122	NA	NA	NA	NA	Mock	Control	30426	NA	NA
NA	NA	103496	NA	NA	NA	NA	Mock	Control	23743	NA	NA
MG100410	NA	NA	NA	NA	NA	NA	Mock	Control	NA	NA	2711863

<b>Total # of Sequences</b>	2,124,836	5,328,215	81,553,035
<b>Median # Sequences Per Sample</b>	24,891	77,928	1,233,172



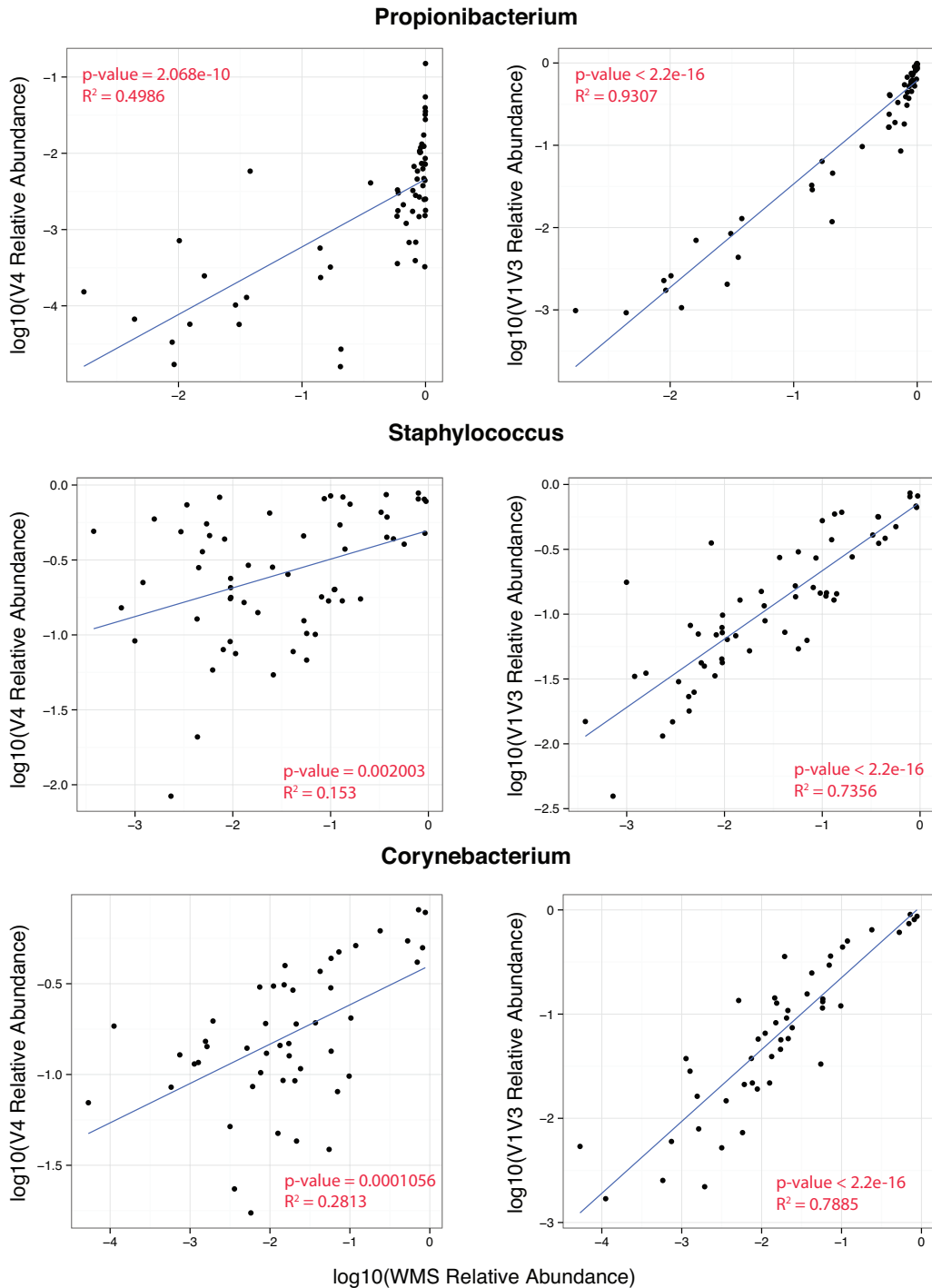
**Fig S1: Composition of the bacterial mock community sample**

Relative abundances, calculated using OTU-based methods, of the genera included in the even bacterial mock community (HM-782D from BEI). Each box represents a different genera (colored by phylum) and the x-axis denotes the sequencing technique employed. Expected relative abundances are denoted by a dotted line for WMS and a solid line for 16S tag sequencing.



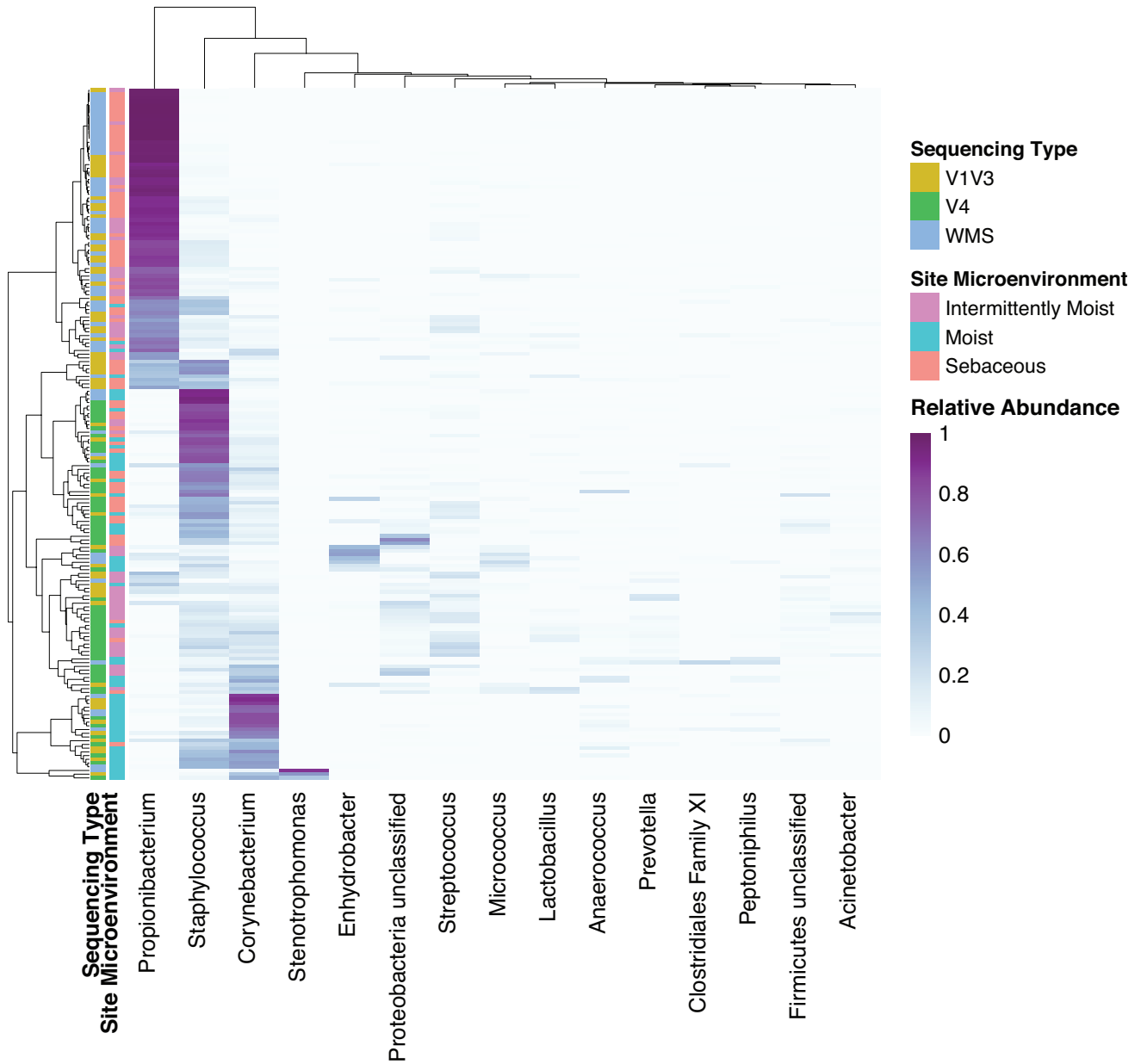
**Fig S2: Taxonomic variation between sequencing strategies**

Taxonomic relative abundance of the bacterial communities by site. Each bar represents a single sample from a subject, separated by sequencing method and anatomical location.



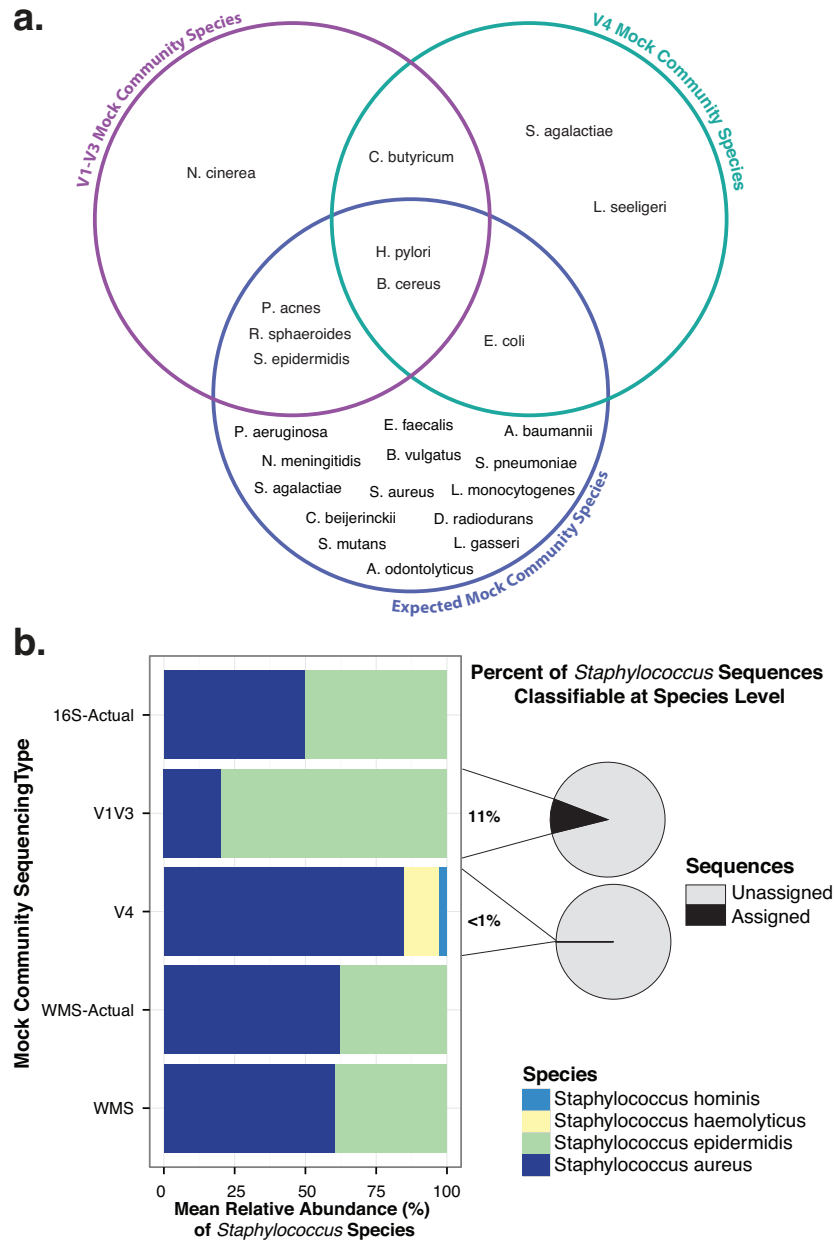
**Fig S3: Correlation of common skin commensals reveals differences in the 16S tag sequencing strategies**

Basic linear regression analysis was used to correlate the relative abundance of the three most prominent skin genera in 16S tag sequencing samples (y-axis) compared to their relative abundance in the whole metagenome shotgun sequenced samples (x-axis). Samples where the commensal was not identified by one or both of the sequencing techniques being correlated were removed from the analysis, since we are only comparing samples that contain information. The relative abundance values were log transformed to get a more even spread of the data.



**Fig S4: Hierarchical clustering reveals differences in the 16S tag sequencing strategies**

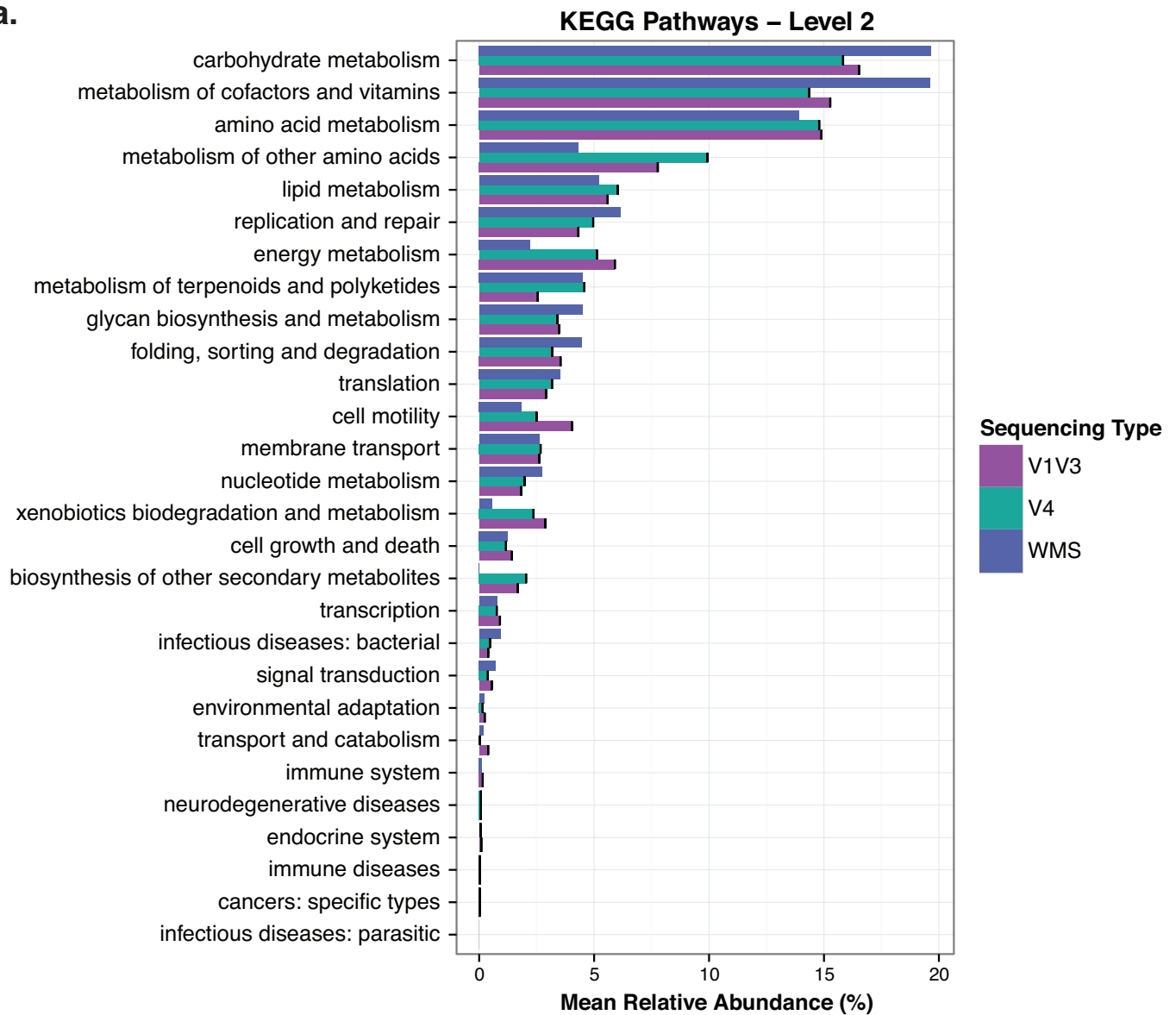
Heatmap depicting the relative abundance of the top 15 bacterial genera in the samples. Each row represents a single swab sample and every column a different bacterial genera. To the left of the heatmap, the first column denotes the different sequencing strategy applied to each sample and the second column highlights the site microenvironment sampled.



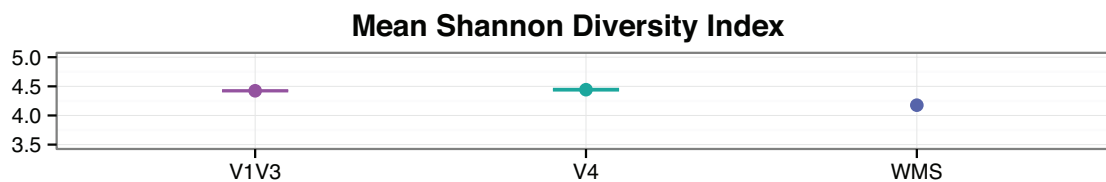
**Fig S5: Species level classification of mock community *Staphylococcus* sequences**

**(a)** Venn diagram highlighting the 20 species contained in the mock community and whether they were identified by V1-V3 and/or V4 tag sequencing OTU-based methods. Species in the center circle were identified in the mock community sample by both 16S tag sequencing methods, whereas species in the outer circles are “contaminants” (species identified by sequencing, but not expected members of the mock community). **(b)** Relative abundance of *Staphylococcus* sequences in the mock community that are able to be classified at the species level. *Staphylococcus* species in the WMS dataset were classified by MetaPhlAn, and in the 16S datasets using pplacer. Pie-charts depict the percentage of sequences that were classified as *Staphylococcus* at the genus level that were further classified at the species level. “Actual” refers to the expected abundance of *Staphylococcus* species based on mock community composition.

a.



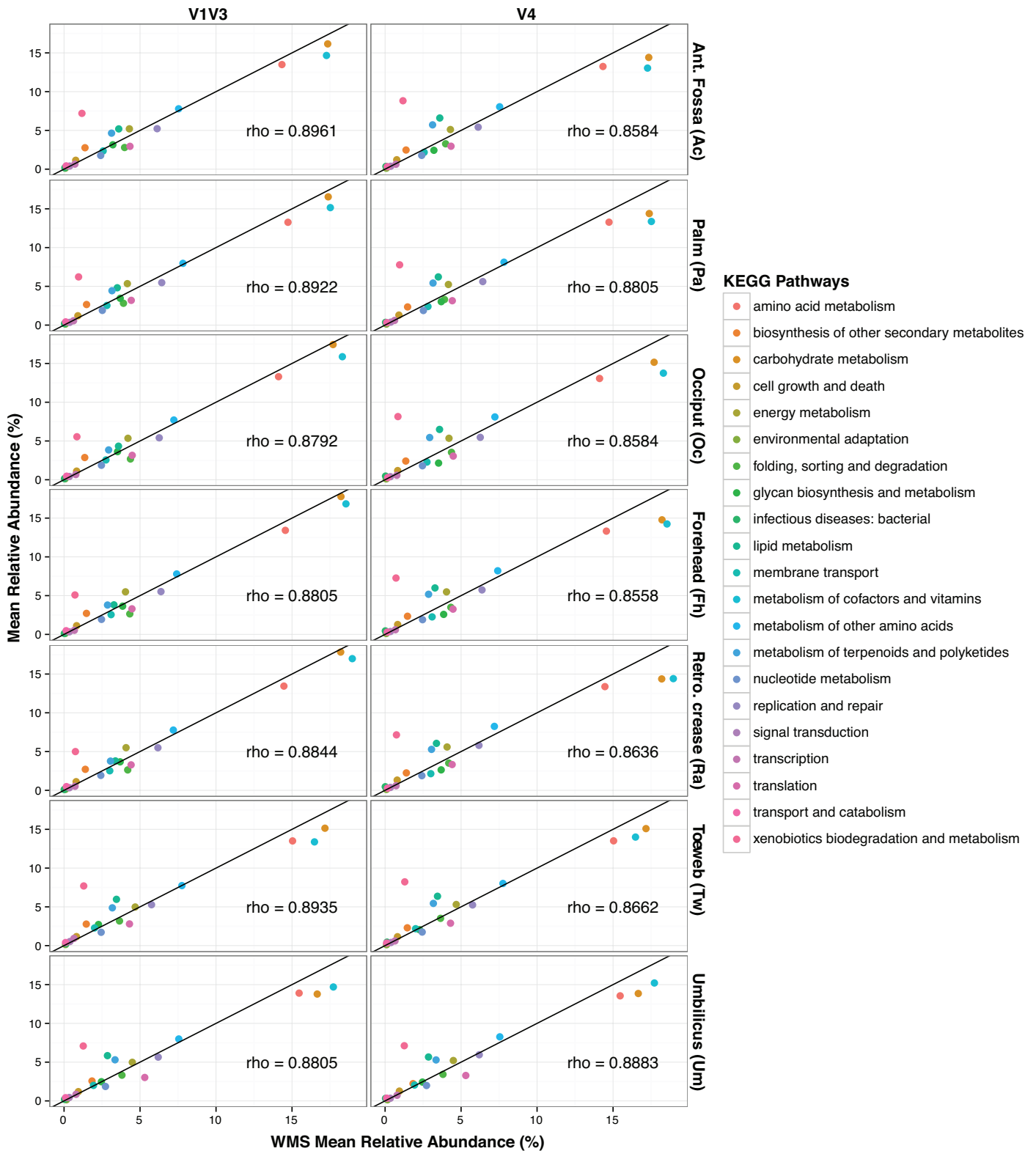
b.



**Fig S6: Comparison of KEGG pathways across sequencing methods**

(a) Comparison of the mean relative abundance of KEGG pathways in the mock community sample across sequencing techniques. (b) Shannon functional diversity of KEGG pathways in the mock community sample is presented by sequencing technique. No significant differences were detected using the Kruskal-Wallis and multiple comparison post hoc test.





**Fig S7: Comparison of KEGG pathways across sequencing methods**  
 Spearman correlations of the mean relative abundance of 16S predicted KEGG pathways to the mean relative abundance of KEGG pathways in the WMS samples.