





QIIME 2 Example Moving Pictures tutorial















Moving Pictures

QIIME2 online tutorial

Human microbiome samples

2 individuals

Four body sites

Gut, left palm, right palm, and tongue

Five time points (first immediately before antibiotic usage)

Will look at the difference between the body sites











Questions

- 1. Is there a difference between the communities of the different body sites?
- 2. Which body sites are different and which are similar?
- 3. In what ways are the communities different?
- 4. Are there any biomarkers for the different body sites?











table.qzv

Table summary

Metric	Sample
Number of samples	34
Number of features	759
Total frequency	157,298











table.qzv

Number of reads per sample

Frequency per sample	
	Frequency
Minimum frequency	917.0
1st quartile	1,862.5
Median frequency	4,139.0
3rd quartile	7,057.25
Maximum frequency	10,095.0
Mean frequency	4,626.411764705882

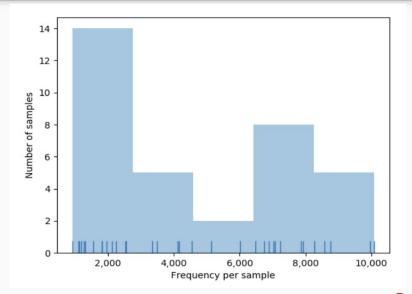








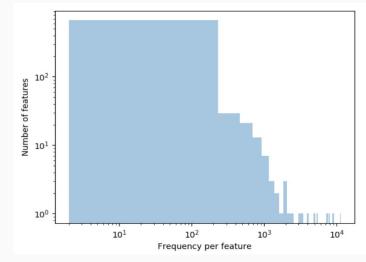




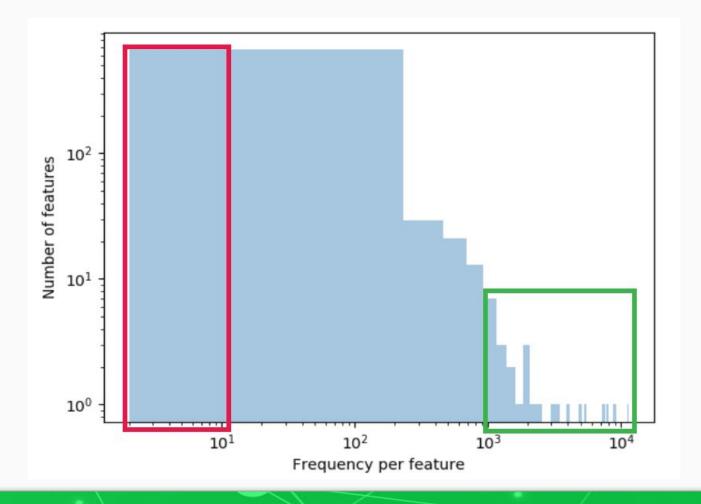
table.qzv

Feature = ASV = Amplicon Sequence Variant Shows how prevalent different features are

Frequency per feature	
	Frequency
Minimum frequency	2.0
1st quartile	9.0
Median frequency	25.0
3rd quartile	83.5
Maximum frequency	11,497.0
Mean frequency	207.2437417654809







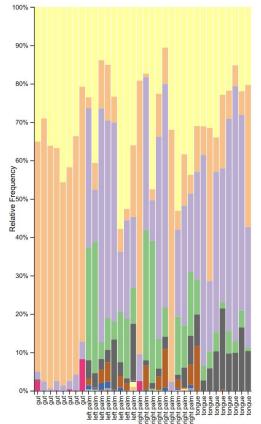
Taxonomy

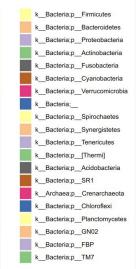












- QIIME2 plots are interactive
- Can choose different taxa levels to display
- Good for sanity checking
- Does this contain the taxa I expect?
- Any noticeable changes?











Alpha diversity

- Evenness
 - How even a community is
 - High evenness = all groups have v. similar abundances
 - Low evenness = groups have very dissimilar abundances
- Faith's Phylogenetic Diversity (PD)
 - A qualitative measure of community richness that incorporates phylogenetic relationships between the features
 - In essence a higher PD indicates more diverse ASVs
 - Calculated as the sum of all branch lengths of the ASVs in the group

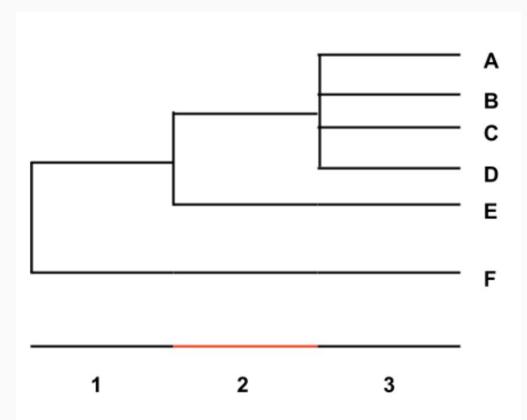












- Not actual calculation
- Sample 1
 - A & E
 - o **PD** =
- Sample 2
 - A,B,C & D
 - o **PD** =
- Sample 3
 - o A, E & F
 - o **PD** =

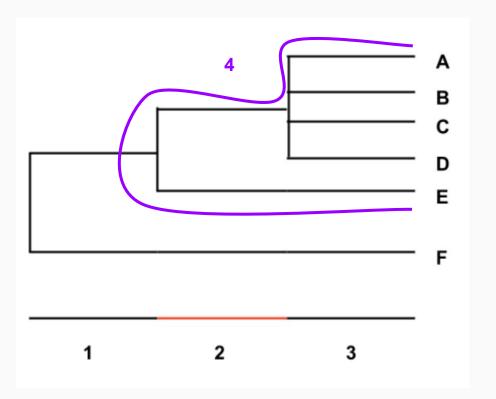












- Not actual calculation
- Sample 1
 - A & E
 - o PD = 4
- Sample 2
 - A,B,C & D
 - **PD** =
- Sample 3
 - A, E & F
 - **PD** =

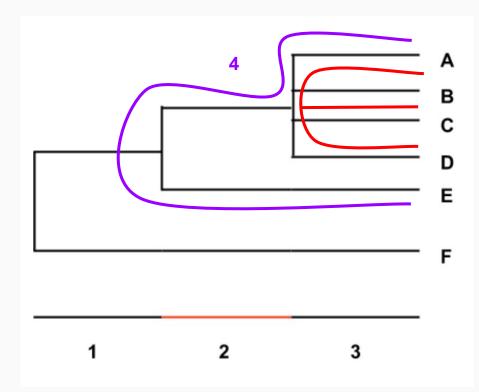












- Not actual calculation
- Sample 1
 - A & E
 - o **PD = 4**
- Sample 2
 - A,B,C & D
 - o PD = 4
 - Sample 3
 - o A, E & F
 - o **PD** =

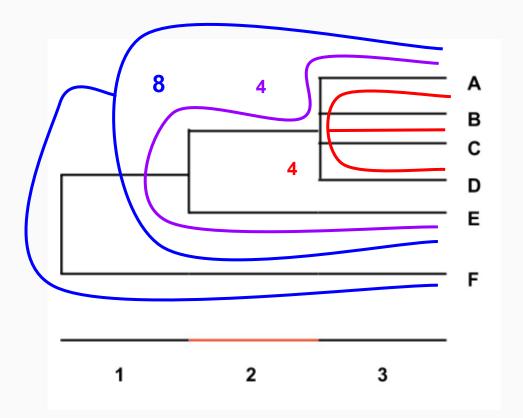












- Not actual calculation
- Sample 1
 - A & E
 - o PD = 4
- Sample 2
 - A,B,C & D
 - o PD = 4
- Sample 3
 - A, E & F
 - o PD = 8



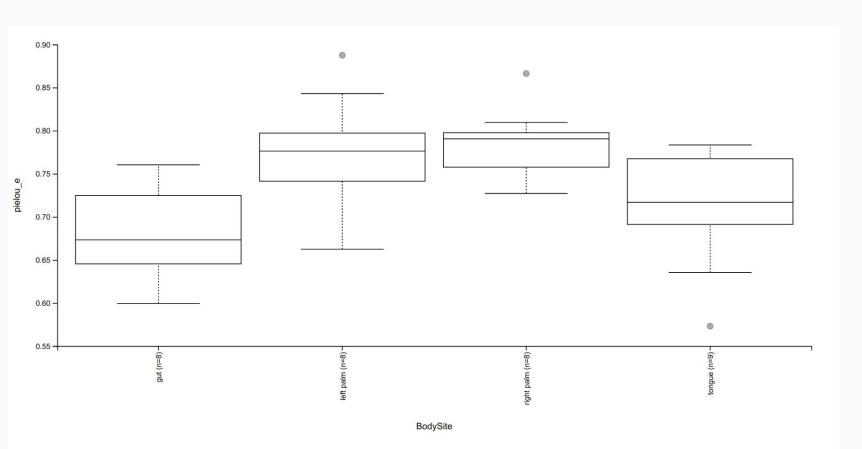
Body Site Evenness Box plot













Body Site Evenness P & Q values









Kruskal-Wallis (all groups)

(a. 3. cap c)	
	Result
н	13.238265002970877
p-value	0.00414867758580111

Kruskal-Wallis (pairwise)

Download CSV

2	_		
n (n=8)	6.352941	0.011719	0.024709
Im (n=8)	8.647059	0.003276	0.019655
(n=9)	1.333333	0.248213	0.297856
Im (n=8)	0.397059	0.528612	0.528612
(n=9)	3.000000	0.083265	0.124897
(n=9)	6.259259	0.012355	0.024709
n (i	m (n=8) m (n=8) n=9) m (n=8) n=9)	(n=8) 6.352941 m (n=8) 8.647059 n=9) 1.333333 m (n=8) 0.397059 n=9) 3.000000	(n=8) 6.352941 0.011719 m (n=8) 8.647059 0.003276 n=9) 1.333333 0.248213 m (n=8) 0.397059 0.528612 n=9) 3.000000 0.083265











Body site evenness conclusions

- Statistically the evenness of:
 - Gut and tongue are no different
 - Right palm and left palm are no different
 - Gut is different compared to right palm and left palm
 - Tongue is different compared right palm but not left palm
- Two main groupings by evenness
 - Left and right palm higher evenness
 - Gut and tongue lower evenness



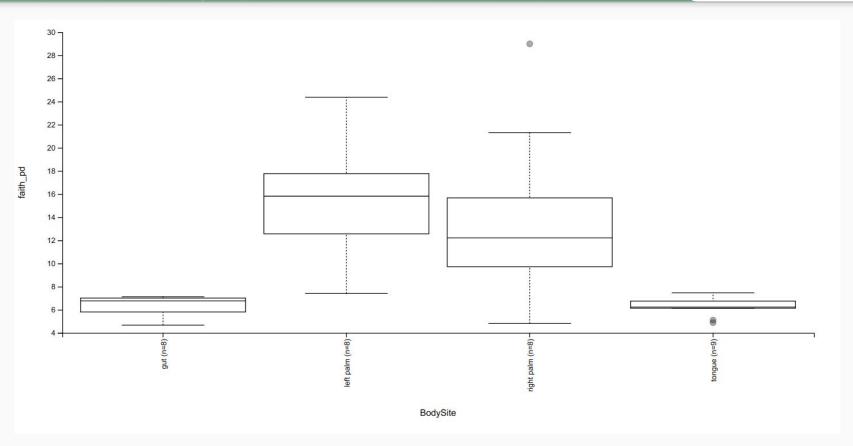
Body Site PD box plot













Body Site PD P & Q values









Kruskal-Wallis (all groups)

	Result
н	18.87240047534165
p-value	0.0002905198262917197

Kruskal-Wallis (pairwise)

Download CSV

		Н	p-value	q-value
Group 1	Group 2			
gut (n=8)	left palm (n=8)	11.294118	0.000778	0.002333
	right palm (n=8)	6.893382	0.008652	0.014062
	tongue (n=9)	0.083333	0.772830	0.772830
left palm (n=8)	right palm (n=8)	0.397059	0.528612	0.634335
	tongue (n=9)	11.342593	0.000757	0.002333
right palm (n=8)	tongue (n=9)	6.750000	0.009375	0.014062









Body site PD conclusions

- Statistically the PD of:
 - Gut and tongue are no different
 - Right palm and left palm are no different
 - PD of gut is lower than right palm and left palm
 - PD of tongue is lower than right palm and left palm
- Two groupings by PD
 - Left and right palm higher diversity
 - Gut and tongue lower diversity



Beta diversity measures

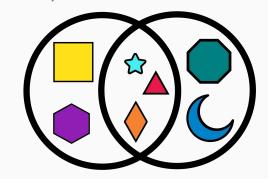








- Jaccard distance
 - measure of community richness dissimilarity (presence/absence)
- Bray-Curtis distance
 - measure of community diversity dissimilarity (abundance)
- unweighted UniFrac distance
 - measure of community richness dissimilarity that incorporates phylogenetic relationships between the features
- weighted UniFrac distance
 - measure of community diversity dissimilarity that incorporates phylogenetic relationships between the features





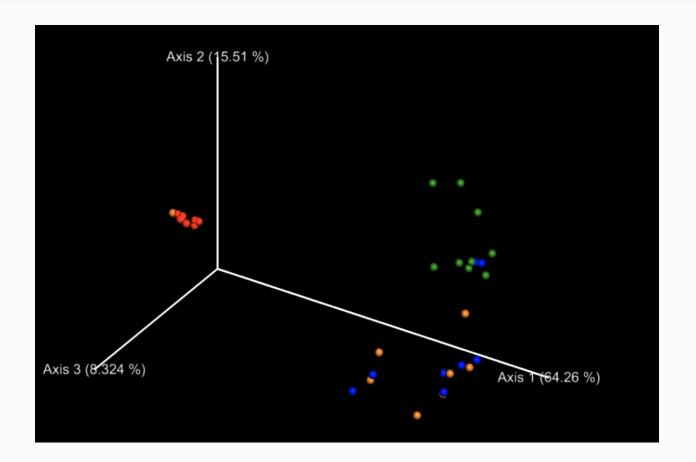
Weighted Unifrac plot























Weighted Unifrac PERMANOVA

Pairwise per	manova results					
		Sample size	Permutations	pseudo-F	p-value	q-value
Group 1	Group 2					
gut	left palm	16	999	78.500534	0.001	0.0015
	right palm	16	999	28.100258	0.002	0.0024
	tongue	17	999	104.356036	0.001	0.0015
left palm	right palm	16	999	0.633407	0.883	0.8830
	tongue	17	999	6.658468	0.001	0.0015
right palm	tongue	17	999	5.231644	0.001	0.0015











Weighted unifrac conclusions

- All the different body sites form statistically distinct clusters (bar the palms)
- This is based on Weighted unifrac which uses:
 - Shared abundance of the different ASVs
 - Phylogenetic information of ASVs
- Left palm and right palm are not statistically different
- 3 Groups: gut, tongue and palm
- Higher community similarity within these groups than between
- Gut communities over time and between patients are very similar compared to tongue and palm communities
- I.e. Gut communities are maintained more consistently











Differential abundance analysis

- Discover biomarkers
- a specific organism/taxa by which a particular type of community can be identified
- Disease example:
 - Helicobacter pylori is a biomarker of stomach communities with Ulcers compared to healthy stomachs
 - H. pylori causes the disease
- Environmental example:
 - Thermophilic bacteria are biomarkers of hot springs compared to cold springs
 - The temperature of the springs causes this difference
- Note: A biomarker does not mean the organisms/taxa is present in only one type of sample. Biomarkers are determined by difference in abundance.

Differential abundance analysis - ancom









Percentile abundances of features by group

Percentile	0.0	25.0	50.0	75.0	100.0		25.0	50.0	75.0	100.0	0.0	25.0	50.0	75.0	100.0	0.0			75.0	
Group	gut	gut	gut	gut	gut	left pal	m left palı	n left palı	n left paln	n left palm	right palm	right palm	right paln	n right palm	right palm	tongu	etongue	a tongue	e tongue	e tong
${\bf k_Bacteria; p_Actinobacteria; c_Actinobacteria; o_Actinomycetales; f_Corynebacteriaceae; g_Corynebacterium}$	1.0	1.00	1.0	1.00	8.0	36.0	144.50	229.0	549.75	1443.0	1.0	56.0	409.0	944.0	1111.0	1.0	1.0	1.0	1.0	5.0
k_Bacteria;p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micrococcaceae;g_Rothia	1.0	1.00	1.0	1.00	1.0	8.0	14.50	21.0	44.25	179.0	1.0	1.0	10.0	49.0	85.0	32.0	46.0	83.0	109.0	225.
k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides	3044.0	4206.7	4577.5	4637.25	4838.	0 1.0	9.25	14.5	33.50	67.0	1.0	7.0	14.0	65.0	783.0	1.0	6.0	9.0	10.0	11.0
k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Prevotellaceae;g_Prevotella	1.0	1.00	1.0	15.25	86.0	1.0	15.25	24.0	170.00	442.0	1.0	5.0	14.0	47.0	349.0	109.0	173.0	280.0	483.0	196
k_Bacteria;p_Firmicutes;c_Bacilli;o_Bacillales;f_Staphylococcaceae;g_Staphylococcus	1.0	1.00	1.0	1.00	10.0	1.0	74.75	106.5	253.00	352.0	1.0	30.0	76.0	492.0	1579.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Carnobacteriaceae;g_Granulicatella	1.0	1.00	1.0	1.00	1.0	1.0	1.00	10.5	14.50	54.0	1.0	1.0	7.0	54.0	86.0	24.0	29.0	36.0	42.0	133
k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus	1.0	1.00	2.0	7.75	31.0	134.0	210.50	477.0	662.75	934.0	4.0	40.0	81.0	568.0	2125.0	206.0	309.0	421.0	480.0	728
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;;_	4.0	6.25	12.0	20.50	23.0	1.0	1.00	1.0	1.00	4.0	1.0	1.0	1.0	1.0	3.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_;g_	1.0	14.50	34.5	37.50	39.0	1.0	1.00	1.0	3.50	9.0	1.0	1.0	1.0	1.0	3.0	1.0	1.0	3.0	4.0	13.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;_	118.0	130.00	136.5	489.50	749.0	1.0	3.75	7.5	23.50	43.0	1.0	1.0	1.0	26.0	37.0	1.0	1.0	6.0	50.0	81.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_	20.0	62.00	74.0	133.75	371.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	1.0	15.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Blautia	32.0	62.25	75.0	147.50	291.0	1.0	1.00	1.0	9.25	20.0	1.0	1.0	1.0	24.0	55.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Coprococcus	1.0	21.75	51.0	106.00	228.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	1.0	9.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Lachnospira	14.0	250.50	379.5	656.25	669.0	1.0	1.00	1.0	1.00	5.0	1.0	1.0	1.0	1.0	22.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Roseburia	128.0	178.50	233.5	319.25	408.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	1.0	32.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_[Ruminococcus]	12.0	12.75	17.0	21.75	64.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_	1.0	38.25	62.5	94.25	248.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	6.0	15.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Faecalibacterium	131.0	455.75	645.5	769.00	915.0	1.0	1.00	9.5	14.50	20.0	1.0	1.0	6.0	15.0	185.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Oscillospira	85.0	102.50	168.5	223.75	286.0	1.0	1.00	1.0	1.50	11.0	1.0	1.0	1.0	1.0	57.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Ruminococcus	1.0	21.75	58.5	183.75	267.0	1.0	1.00	1.0	1.50	6.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Veillonellaceae;g_Phascolarctobacterium	87.0	132.00	151.0	303.75	537.0	1.0	1.00	1.0	1.75	6.0	1.0	1.0	1.0	7.0	68.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Veillonellaceae;g_Veillonella	1.0	1.00	1.0	1.00	8.0	1.0	13.75	30.5	91.25	298.0	1.0	1.0	9.0	25.0	249.0	99.0	146.0	148.0	281.0	51
${\bf k_Bacteria; p_Firmicutes; c_Clostridia; o_Clostridiales; f_[Tissierellaceae]; g_Anaerococcus}$	1.0	1.00	1.0	1.00	1.0	8.0	40.00	52.0	122.75	173.0	1.0	11.0	48.0	135.0	425.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_[Tissierellaceae];g_Peptoniphilus	1.0	1.00	1.0	1.00	11.0	5.0	27.50	41.5	65.25	106.0	1.0	1.0	25.0	81.0	520.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Erysipelotrichaceae;g_	1.0	7.00	23.5	43.50	69.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
k_Bacteria;p_Fusobacteria;c_Fusobacteriia;o_Fusobacteriales;f_Fusobacteriaceae;g_Fusobacterium	1.0	1.00	1.0	1.75	32.0	1.0	20.25	28.0	188.50	633.0	1.0	1.0	14.0	38.0	391.0	45.0	170.0	204.0	626.0	92
k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Neisseriales;f_Neisseriaceae;g_Neisseria	1.0	1.00	1.0	1.00	6.0	25.0	50.25	80.0	149.75	1232.0	1.0	23.0	24.0	165.0	617.0	20.0	449.0	490.0	1368.0	35
$\underline{ \text{Bacteria;} p_Proteobacteria;} c\underline{ \text{Deltaproteobacteria;}} o\underline{ \text{Desulfovibrionales;}} f\underline{ \text{Desulfovibrionaleae;}} g\underline{ \text{Desulfovibrionaleae;}}$	22.0	27.25	33.0	47.75	80.0	1.0	1.00	1.0	1.00	1.0	1.0	1.0	1.0	1.0	19.0	1.0	1.0	1.0	1.0	1.0
${f Bacteria}; {f p_Proteobacteria}; {f c_Epsilon proteobacteria}; {f o_Campylobacterales}; {f f_Campylobacteraceae}; {f g_Campylobacterales}; {f o_Campylobacterales}; {f o_Campylobacterales}$	er 1.0	1.00	1.0	1.00	1.0	1.0	5.00	6.5	15.50	69.0	1.0	1.0	1.0	10.0	35.0	8.0	11.0	22.0	39.0	10
k_Bacteria;p_Proteobacteria;c_Gammaproteobacteria;o_Pasteurellales;f_Pasteurellaceae;g_Haemophilus	1.0	1.00	1.0	2.25	6.0	21.0	38.50	71.0	102.75	1210.0	1.0	6.0	21.0	124.0	677.0	256.0	390.0	550.0	728.0	129
Bacteria;p Proteobacteria;c Gammaproteobacteria;o Pseudomonadales;f Pseudomonadaceae;g Pseudomona	s 1.0	1.00	1.0	1.00	1.0	1.0	34.00	64.5	174.25	1612.0	1.0	17.0	26.0	237.0	727.0	1.0	1.0	1.0	1.0	1.0

Greengenes Level 6



Differential abundance analysis - ancom









Percentile abundances of features by group

Percentile	0.0	25.0	50.0	75.0	100.0	0.0	25.0	50.0	75.0	100.0
Group	gut	gut	gut	gut	gut	left palm	left palm	left palm	left paln	left palm
$\underline{Bacteria}; \underline{p}\underline{Actinobacteria}; \underline{c}\underline{Actinobacteria}; \underline{o}\underline{Actinomycetales}; \underline{f}\underline{Corynebacteriaceae}; \underline{g}\underline{Corynebacterium}$	1.0	1.00	1.0	1.00	8.0	36.0	144.50	229.0	549.75	1443.0
k_Bacteria;p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Micrococcaceae;g_Rothia	1.0	1.00	1.0	1.00	1.0	8.0	14.50	21.0	44.25	179.0
k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Bacteroidaceae;g_Bacteroides	3044.0	4206.75	4577.5	4637.25	4838.0	1.0	9.25	14.5	33.50	67.0
k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_Prevotellaceae;g_Prevotella	1.0	1.00	1.0	15.25	86.0	1.0	15.25	24.0	170.00	442.0
k_Bacteria;p_Firmicutes;c_Bacilli;o_Bacillales;f_Staphylococcaceae;g_Staphylococcus	1.0	1.00	1.0	1.00	10.0	1.0	74.75	106.5	253.00	352.0
$k_Bacteria; p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Carnobacteriaceae; g_Granulicatella$	1.0	1.00	1.0	1.00	1.0	1.0	1.00	10.5	14.50	54.0
k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus	1.0	1.00	2.0	7.75	31.0	134.0	210.50	477.0	662.75	934.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;_;_	4.0	6.25	12.0	20.50	23.0	1.0	1.00	1.0	1.00	4.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_;g_	1.0	14.50	34.5	37.50	39.0	1.0	1.00	1.0	3.50	9.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;_	118.0	130.00	136.5	489.50	749.0	1.0	3.75	7.5	23.50	43.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_	20.0	62.00	74.0	133.75	371.0	1.0	1.00	1.0	1.00	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Blautia	32.0	62.25	75.0	147.50	291.0	1.0	1.00	1.0	9.25	20.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Coprococcus	1.0	21.75	51.0	106.00	228.0	1.0	1.00	1.0	1.00	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Lachnospira	14.0	250.50	379.5	656.25	669.0	1.0	1.00	1.0	1.00	5.0
${f k_Bacteria}; {f p_Firmicutes}; {f c_Clostridia}; {f o_Clostridiales}; {f f_Lachnospiraceae}; {f g_Roseburia}$	128.0	178.50	233.5	319.25	408.0	1.0	1.00	1.0	1.00	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_[Ruminococcus]	12.0	12.75	17.0	21.75	64.0	1.0	1.00	1.0	1.00	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_	1.0	38.25	62.5	94.25	248.0	1.0	1.00	1.0	1.00	1.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Faecalibacterium	131.0	455.75	645.5	769.00	915.0	1.0	1.00	9.5	14.50	20.0
k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Oscillospira	85.0	102.50	168.5	223.75	286.0	1.0	1.00	1.0	1.50	11.0

Pseudocounts are added as ANCOM can't handle zeroes, Therefore minus 1 from all values to get real value











Conclusions

- Alpha diversity
 - The PD of Palms are significantly higher than Gut and Tongue
- Beta diversity
 - The communities of Palms, Gut and Tongue are significantly different
 - There appears to be no significant difference between right and left palms
 - Gut communities are more similar between individuals and time than palms or tongues
- Biomarker detection
 - Most biomarkers between the 4 body sites are genus that are in either high or low abundance in gut and low in the others
 - This matches the clustering seen in the Weighted Unifrac plot