Supplemental Table 2 Synopsis of microbiome results and metrics that were reviewed to serve as a resource to readers who wish to explore detailed results from the reviewed manuscripts. General study information is detailed in Table 2a, alpha and beta diversity metrics are presented in Table 2b, individual bacterial taxa differences are presented in Table 2c, and the results from studies using shotgun sequencing are presented on Table 2d.

Supplemental Table 2a. Results synthesis: Methods

Author Year	Study Aim	Groups Studied	Specimen Collection	HMP Yes/No	Participant age	Shotgun vs 16S	Region	Extraction	Database
Costello 2009	The goal of the research is to address questions regarding biogeography of the human microbiota in healthy adults including how bacterial diversity is partitioned across body habitats, people and time; how diversity at a variety of skin locales compares to that found in other body habitats, and if skin communities assemble differently at different sites.	Healthy controls	9 healthy adults (6M, 3F) recruited to donate samples on 2 consecutive days x 1 (2 oral sites: **dorsal tongue and oral cavity rinse** & stool) and then repeated 3 months later x 2 consecutive days (dorsal tongue and stool only)	No	Mostly 30-35, 1 subject 60 years old	16S- Roche 454 FLX Titanium	V2	MoBio PowerSoil	Greengenes
Ding 2014	Partitioned HMP data into community types for each body site sampled using Dirichlet multinomial mixture models in order to map enterotypes by body sites.	Healthy controls	2 time points for 300 healthy subjects (50% female) and third time point for 100 subjects. Interval between repeat sampling varied from 30-451 days (median = 224 days). **7 samples from mouth (buccal mucosa, keratinized gingiva, hard palate, saliva, tongue, 2 surfaces along tooth), 2 oropharyngeal sites (throat and palatine tonsils)**, colon (stool)	Yes	Median 25 (range 18-40)	16S- Roche 454 FLX Titanium	V3-V5 (V35)	MoBio PowerSoil	SILVA

Huse 2012	Identified a set of core OTUs common across individuals and body sites- focus on the more abundant organisms that are common across individuals - focused on the more abundant organisms that are common across individuals.	Healthy controls	112 female and 127 male subjects. Samples were collected from 18 body sites including **7 samples from mouth (buccal mucosa, keratinized gingiva, hard palate, saliva, tongue, 2 surfaces along tooth), 2 oropharyngeal sites (throat and palatine tonsils**), colon (stool)	Yes	Median 25 (range 18-40)	16S- Roche 454 FLX Titanium	V1-V3 & V3-V5	MoBio PowerSoil	SILVA
Iwauchi 2019	Evaluate if the transition of oral bacteria to the GI tract is more prevalent in the elderly vs. adults. Hypothesis was that inflammation in the oral cavity can cause systemic inflammation by changing the gut microbiome.	Healthy controls	1 time point for both oral and stool. 2 oral samples (**1 subgingival plaque, 1 tongue coating**). 29 elderly subjects and 30 adults. Fecal sampling conducted within a week after oral sampling. Females > males in elderly group	No	Elderly subjects (age 80.2 +/- 9.1 years) and adults (age 35.9 +/- 5.0 years).	16S- Illumina MiSeq	V3-V4	Custom bead beating protocol using extraction buffer (100 mM Tris-HCl and 40 mM EDTA at pH 9.0) with 50 µL of 10% SDS	Greengenes
Li 2012	Investigate microbiome diversity across body habitats and individuals from 16S profiles generated by the HMP.	Healthy Controls	Healthy controls who passed a screening systemic health exam had samples obtained from 15 male or 18 female body sites including **7 samples from mouth (buccal mucosa, keratinized gingiva, hard palate, saliva, tongue, 2 surfaces along tooth), 2 oropharyngeal sites (throat and palatine tonsils),** colon (stool). 2 recruitment centers (Baylor College of Medicine, Houston, TX and Washington University, St. Louis, MO).	Yes	Median 25 (range 18-40)	16S- Roche 454 FLX Titanium	V3-V5 (V35)	MoBio PowerSoil	SILVA
Llyod-Price 2017	To further knowledge of baseline human microbial diversity across body sites.	Healthy Controls	265 Individuals (female n=128, male n-137), repeated measures and primarily targeted **buccal mucosa, subgingival plaque, tongue dorsum** and stool	Yes	Median 25 (range 18-40)	Both Analyzed: 16S- Roche 454 FLX Titanium WGS- Illumina Gaiix platform	V3-V5 (V35)	MoBio PowerSoil	SILVA

Schmidt 2019	Aim to test the hypothesis that microbial transmission along the GI tract (oral to gut) is more common that previously appreciated.	Healthy controls Colorectal CA Type 1 DM Rheumatoid Arthritis	Data was merged from healthy controls collected for this study (DE-CTR, FR-CRC, LU-T1D) and healthy controls from other publicly available datasets(FJ-CTR, CN-RA). **Control samples total= 340 stool (114 M, 225 F, 1 unknown) and 259 oral saliva (93M, 166F).** Subsets= CN-RA: 97 gut (67 F, 29 M, 1 unknown), 47 oral saliva (33F, 14 M); DE-CTR: 10 gut (6M, 4F), 10 oral saliva(6M, 4F); FJ-CTR: 166 gut (56M, 100F), 140 oral saliva (53M, 87F); FR-CRC: 16 gut (8M, 8F), 16 oral saliva (8M, 8F), LU-T1D: 51 gut (15M, 36F), 46 oral (12 M, 34F). **1 timepoint for**: CN-RA, FJ-CTR, FR-CRC, **2 timepoints for** DE-CTR, and **2-3 timepoint**s for LU-T1D.	No	Not all studies had age listed. Subsets= CN-RA: 42.69 +/- 9.50; DE-CTR: age unknown; FJ-CTR: age unknown; FR-CRC: 63.88 +/- 3.82; LU-T1D: 39.60 +/- 21.07 (large age range).	Shotgun. CN-RA: Illumina (model not spec.). FJ-CTR, DE-CTR & FR-CRC: Illumina HiSeq 2000. LU-T1D: Illumina HiSeq4000 & Illumina NextSeq 500	N/A	GNOME DNA Isolation Kit (MP Biomedicals)= DE-CTR & FR-CRC. Qiagen Allprep Kit (Qiagen)= LU-T1D. MoBio PowerSoil= FJ-CTR (HMP protocol).	Taxonomic profiling: NGless (https://ngless.readthedocs.io/en/latest/). Phylogenetic tree: ETE3 toolkit (http://etetoolkit.org/)
Segata 2012	Measure and compare the composition, relative abundance, phylogenetic and metabolic potential of the bacterial populations inhabiting multiple sites along the digestive tract in the defined adult reference HMP subject population.	Healthy controls	209 subjects (147 had samples from all 10 sites). **7 samples from mouth (buccal mucosa, keratinized gingiva, hard palate, saliva, tongue, 2 surfaces along tooth), 2 oropharyngeal sites (throat and palatine tonsils),** colon (stool). Shotgun sequencing in 98 subjects	Yes	Median 25 (range 18-40)	Both analyzed 16S- Roche 454 FLX Titanium WGS- Illumina Gaiix platform (98 subjects)	V3-V5 (V35)	MoBio PowerSoil	SILVA
Stearns 2011	Aim to address the changing microbial communities along the GI tract and identify relevant bacterial communities.	Healthy controls	4 individuals studied (2F, 2M). Samples: **mouth plaque (L&R supra-gingival and sub-gingival, tongue),** stomach (antrum and body), duodenum, colon (transverse and descending), rectum and stool	No		16S- Illumina	V3	MoBio PowerSoil - 40 second bead beating step and heating to 70 degrees C for 10 min added.	PyNAST aligner in QIIME v 1.2.0 aligned to Greengenes core set

Vasapolli 2019	Aim was to characterize the transcriptionally active bacteria (i.e. alive and capable of reproducing) in saliva from the oral cavity; corpus and antrum from the stomach; and duodenum, terminal ileum, ascending and descending colon and feces of the GI tract in healthy subjects.	Healthy controls	**Saliva,** mucosal and fecal samples were collected from healthy adults (10 men and 11 women) who underwent upper and lower GI tract endoscopy. Biopsies were taken from the upper (corpus, antrum and duodenum) and lower (terminal ileum, ascending and descending colon) GI tract	No	Mean 59 +/- 12.3 years	16S- Illumina MiSeq	V1-V2	RNA from samples were extracted using Rneasy kit and cDNA was synthesized to selectively identify the transcriptionally active bacteria	aligned with MOTHUR using the SILVA ref. database

Supplemental Table 2b. Results synthesis: Alpha, beta, differential abundance

Author Year	Alpha Method	Oral Results	Gut Results	Comparison	Beta Method	Oral Results	Gut Results	Beta diversity Comparison
Costello 2009					Weighted and unweighted UniFrac	Unweighted UniFrac distance: variation within people (day-to-day)=0.5, variation between people (on any given day)=0.58	Unweighted UniFrac distance: variation within people (day-to-day)= 0.6, variation between people (on any given day)= 0.73	UniFrac showed strong primary clustering by body habitat versus host sex, individual or day. weighted UniFrac PcoA plots Fig. S5: oral cavity and gut closer on plot. Percent of variation explained in PC1:27% and PC2: 12%. Figure S22, when just oral vs. gut compared with unweighted & weighted UniFrac (using 2 specimen collection methods) appears farther in space but percent of variation explained much lower **(agrees with observations of UniFrac distances and PC plots seen in Iwauchi paper)**
Ding 2014					Associations studied using Fisher's exact test			Community-type associations are strongest within a body region, but **also exist between stool and the oral cavity-- also agrees with the fact that oral and stool are not completely distinct communities** (throat P<10^-1, saliva P<10^-2, tongue dorsum P<10^-1, hard palate P<10^-1, buccal mucosa P<10^-2, supragingival plaque P<10^-1, keratinized gingiva P<10^-2). Strongest association was between stool and saliva.
Huse 2012	O.T.U. richness (total number of species)	Oral richness estimates: V1-V3: 3,793 (lowest, hard palate) to 14,410 (highest, subgingival plaque). V3-V5: 3,135 (lowest, hard palate) to 11,501 (highest, subgingival plaque).	Stool Richness estimates: V1-V3: 23,665 and V3-V5: 33,627	Stool had much higher richness estimates vs. oral samples for both the V1-V3 and V3-V5 data. (this is the opposite of the other alpha diversity measures when evenness is taken into account)				

Iwauchi 2019					UniFrac distance. PCoA based on unweighted (presence/absence) and unweighted (abundance of observed taxa) UniFrac distances performed using QIIME 1.8.0			Values approximate based on Figure 3: Between-group distances by Unweighted UniFrac (mean): **Fecal vs. subgingival plaque**= 0.79, **Fecal vs. tongue**= 0.77 (note, unsure if error bars SEM or SD). **Unweighted UniFrac PCoA plots with much clearer separation between oral vs. gut as compared to weighted PCoA plots** (Figure 2). Weighted plots better explain the percentage of variation of the data (PC1: 31.15% and PC2: 21.72%) vs. unweighted (PC1: 24.16% and PC2: 8.78%). **Agrees with Costello and Stearns papers.** Also note that the analysis based on unweighted UniFrac distance showed a higher similarity between the fecal and subgingival plaque microbiota in the elderly vs. adult groups. This suggests the transition of subgingival plaque bacteria to the gut is more prevalent in the elderly than in adults.
Li 2012	Shannon Entropy	**Genera-based taxonomic units** (values are median with 95% CI)= Buccal mucosa: 1.664 (95% CI: 0.783, 2.541), Hard palate: 2.098 (95% CI: 1.068, 2.695), keratinized gingiva: 1.588 (95% CI: 0.463, 2.495), palatine tonsils: 2.412 (95% CI: 1.476, 2.810), saliva: 2.655 (95% CI: 1.957, 3.008), subgingival plaque: 2.634 (95% CI: 1.948, 3.049), supragingival plaque: 2.589 (95% CI: 1.803, 3.002), throat: 2.420 (95% CI: 1.261, 2.888), tongue dorsum: 2.304 (95% CI: 1.552, 2.755). **OTU-based taxonomic units** (values are median with 95% CI)= Buccal mucosa: 1.903 (95% CI: 0.709, 3.039), Hard	Genera-based taxonomic units (values are median with 95% CI)= Stool: 1.663 (95% CI: 0.412, 2.615). OTU-based taxonomic units (values are median with 95% CI)= Stool: 2.583 (95% CI: 1.072, 3.849).					

		palate: 2.432 (95% CI: 1.390, 3.160), keratinized gingiva: 1.721 (95% CI: 0.810, 2.811), palatine tonsils: 2.877 (95% CI: 1.697, 3.448), saliva: 3.143 (95% CI: 2.498, 3.682), subgingival plaque: 3.175 (95% CI: 2.231, 3.676), supragingival plaque: 3.005 (95% CI: 1.935, 3.615), throat: 2.830 (95% CI: 1.349, 3.370), tongue dorsum: 2.609 (95% CI: 1.786, 3.216).						
Lloyd-Price, 2017								
Schmidt 2019					Bray-Curtis dissimilarity			Overall oral and fecal microbiome communities appear independent of each other based on Bray at species level

Segata 2012	inverse Simpson	oral group 1 average 5.3, group 2 ranged from 7.3 +/-3 (tonsils) to 10.6 +/-3.1 (saliva), plaque in group 3 had average diversity of 9.6 +/- 3.1 and 9.8 +/- 3.0	Group 4 (stool) had average diversity of 4.6 +/- 2.9	Oral alpha diversity much higher than gut/stool	Bray- Curtis and MDS plot	**Within-group Bray-Curtis index** (mean) and SD: **oral group 1**= 0.58 +/- 0.14 , **oral group 2**= 0.51 +/- 0.14, **oral group 3** = 0.49 +/- 0.14	**Within-group Bray-Curtis index** (mean) and SD: **fecal group 4**= 0.53 +/- 0.17	For beta diversity summary statistics, within group distance was significantly lower than between-group distance for all 4 groups, i.e. community structure similarity is higher for samples in same group vs. samples in different groups. **Between group Bray values (mean +/- SD): Oral G1 vs Fecal G4=** 0.02 +/- 0.03, **Oral G2 vs Fecal G4=** 0.05 +/- 0.05, **Oral G3 vs Fecal G4=** 0.03 +/- 0.04 [note, higher Bray indices= more alike). Gut community appears very distinct from oral communities on MDS plot which aligns with data.
Stearns 2011	Chao1, phylogenic diversity and Shannon diversity	Shannon diversity metrics were highest in samples from the mouth. Shannon diversity not significantly different between mouth and colon/stool. Phylogenetic diversity was significantly higher and much more consistent across individuals than in any of the other locations tested.	Colon and stool samples were highly variable.	Chao1 curve did not level off	unweighted and weighted UniFrac analysis. PCoA plots of weighted and unweighted UniFrac were calculated and plotted in QIIME.	PcoA plots shown but no numbers or statistics provided	PcoA plots shown but no numbers or statistics provided	Figure S3, using unweighted UniFrac to cluster sample sites- colon and mouth appear discrete/distinct communities. Percent of variance explained PC1: 17%, PC2: 7.9%. With weighted uniFrac, samples are much closer together in 2D space, percent of variance explained= PC1: 38%, PC2 14%. **(agrees with observations of UniFrac distances and PC plots seen in Iwauchi & Costello papers)**

Vasapolli 2019	Simpson, Shannon, taxonomy index and rarity index	Shannon diversity index was significantly higher in saliva vs. descending colon. Richness significantly higher in saliva vs. descending colon and feces.		Phylogenetic richness: In saliva mean numbers of phylotypes were 634 +/- 131 vs. 382 +/- 73 in the feces.	Bray-Curtis similarity matrix distances and PCoA plots	Significant differences between the consecutive sampling regions: saliva and corpus (PERMANOVA: t=3.02, P=0.001; ANOSIM: R= 0.329, P=0.001), duodenum and terminal ileum (PERMANOVA: t=2.74, P=0.001; ANOSIM: R= 0.888, P=0.001). Could this be because RNA/transcriptionally active (putative) bacteria analyzed? No differences between communities in corpus, antrum and duodenum	Communities from the terminal ileum, ascending colon and descending colon of each individual were similar, but distinct from those of the corresponding feces (did pts undergo colon prep?). Significant differences between the consecutive sampling regions: descending colon and feces (PERMANOVA: t=1.71, P=0.001; ANOSIM: R= 0.313, P=0.001). Could this be because RNA/ transcriptionally active (putative) bacteria analyzed? No differences between communities in terminal ileum, ascending colon and descending colon.	Discrete clustering between upper and lower GI. Feces, terminal ileum, colon ascendant and colon descent clustered together while saliva, duodenum, corpus and antrum clustered together. Saliva to feces comparison of regional groups by ANOSIM significantly different (P = 0.0001). There were significant overall differences in community structure by region (pseudo-F = 7.24; P= 0.001) and dependent on H pylori infection (pseudo-F = 5.83, P= 0.001). There was also an interaction between region x H pylori infection (pseudo-F = 2.04 P=0.001). The communities from the corpus and antrum were shown to be truly dependent on H pylori infection. H pylori constituted approximately 8-99% and 42%-97% of the total gastric bacterial community in the antrum and corpus, respectively. Lower GI tract and fecal microbiota communities did not different based on H pylori infection.

Supplemental Table 2c. Results synthesis: Taxa

Author Year	Phylum Abundant Taxa-Gut	Phylum Abundant Taxa-Oral	Class/Family Abundant Taxa-Gut	Class/Family Abundant Taxa-Oral	Genera Abundant Taxa-Gut	Genera Abundant Taxa-Oral	Species Abundant Taxa-Gut	Species Abundant Taxa-Oral	Variability of community
Costello 2009	Reported in text: Bacteroidetes: 59.8%, Firmicutes 35.2%; Estimates from figure S2: Higher Bacteroidetes (60%), similar Firmicutes 35%, less Proteobacteria (5%), little to none Actinobacteria, Fusobacteria. Figure S3: Less common bacterial phyla (percent of sequences)- Verrucomicrobia 0.25%, Lentisphaerae 0.025%. In figure S4, most abundant bacterial sequences (listed as highest taxonomic to which they could be assigned): Bacteroidetes (phylum), Proteobacteria (phylum)	**Oral sites: dorsal tongue and oral cavity rinse.** Estimates from figure S2: higher representation of Proteobacteria (20%), Fusobacteria (5%) and Actinobacteria (10%), Firmicutes 40%, less Bacteroidetes (20%). Figure S3: Less common bacterial phyla (percent of sequences)- TM7 0.45%, SE1 0.2%, Acidobacteria 0.05%. In figure S4, most abundant bacterial sequences (listed as highest taxonomic to which they could be assigned): TM7 (phylum)	In figure S4, most abundant bacterial sequences (listed as highest taxonomic to which they could be assigned): Lachnospiraceae (family), Lachnospiraceae (family), Ruminococcaceae (family), Clostridiales (order), Prevotellaceae (family), Bacteriodales (order), Porphyromonadaceae (family), Veillonellaceae (family), Burkholderiales (order),	In figure S4, most abundant bacterial sequences (listed as highest taxonomic to which they could be assigned): pasteurellaceae (family), micrococcineae (family), actinomycineae (family), carnobacteriaceae (family), bacteroidales (order), coriobacterineae (family)	In figure S4, most abundant bacterial sequences (listed as highest taxonomic to which they could be assigned): Bacteroides (genus), Prevotella (genus), Faecalibacterium (genus), Parabacteroides (genus), Alistipes (genus), Roseburia (genus), Sutterella (genus), Anaerotruncus (genus)	*Streptococcus(*Firmicutes) 18.7%; *Pasteruella* (gamma Proteobacteria) 15.9%, *Veillonella* (Firmicutes) 13.9%, *Prevotella* (Bacteroidetes) 13.6%, *Neisseria* (beta Proteobacteria) 9.9%. In figure S4, most abundant bacterial sequences (listed as highest taxonomic to which they could be assigned): streptococcus (genus), veillonella (genus), prevotella (genus), neissera (genus), fusobacterium (genus), pasteruella (genus), prophyromonas (genus), camplobacter (genus), oribacterium (genus), gemella (genus), leptotrichia (genus), megasphaera (genus)			Oral community were significantly less variable in terms of membership, both within and between people vs. other habitats. Gut community structure was highly variable among people but exhibited minimal variability within people over time. **Core (shared) phylotypes among individuals likely to be larger in the oral cavity vs. the gut**

Ding 2014				Saliva dominated by Pasteruellaceae family (gram negative)	*Prevotella* also abundance in stool communities (other genus seen in saliva did not have similar abundance)	Saliva dominated by *Prevotella, Streptococcus, Veillonella, Fusobacterium*			Community types from sites within the oral cavity were least stable (supragingival plaque), community types from gut/stool samples most stable. Taxonomic compositions of the oral and gut microbiomes were different but **community types at these sites predictive of each other.** For ex: subjects with stool type D (highest level of *Prevotella*) were 2.1 times more likely to harbor saliva communities type A and C which were also high in *Prevotella* (relative to saliva communities type B and D). **Suggest oral populations may seed the gut**
Huse 2012			Stool samples contained 7 core OTUs representing several members of the Lachnospiraceae family	V3-V5 core OTUs across all oral site: Pasteruellaceae (family)			The stool had less core OTUs vs. the oral cavity (5 V3-V5, 7 V1-V3 OTUs). Stool samples contained 7 core OTUs representing several members of the Lachnospiraceae family as well as *Faecalibacterium, Oscillibacter* and two *Bacteroides*. *Bacteroides* were the most	The oral cavity had a consistently richer core microbiome vs. other sites. Core OTUs; mouth samples=ranging from 7 core V3-V5 OTUs in the keratinized gingiva and subgingival plaque to 22 in the saliva. The buccal mucosa, hard palate, palatine tonsils, supragingival plaque, throat and tongue dorsum all had similar core richness of 11 to 16 V3-V5 OTUs.	Despite their prevalence across subjects, the relative abundance of core OTUs vary dramatically between subjects.

								abundant OTUs comprising on average 21% of the sequences.	The V1-V3 OTUs showed similar patterns with the oral sites having a core V1-V3 richness of 10-15 OTUs except the keratinized gingiva with only 3 OTUs. V3-V5 core OTUs across all oral site: Fusobacterium (genus), Streptococcus (genus), Pasteruellaceae (family) and Veillonella (genus, 95%). Granulicatella (genus) and Gemella (genus) were present in all samples from all oral sites but in some cases with a prevalence of only 92-94%. Only 2 OTUs were present in all V1-V3 oral sites and both were Streptococcus (genus).	

Iwauchi 2019					**These relative abundance values are approximate based off Figure 1. Adult feces**: ***Bacteroides*15%, *Bifidobacterium* 17%**, *Streptococcus* 1%, *Prevotella* 1%, ***Unclassified_Lachnospiraceae* 18%**, *Unclassified_Clostridiales* 2%, ***Blautia* 11%**, *Faecalibacterium* 9%, *Unclassified_Ruminococcaceae* 6%, *Lactobacillus* 0.3%, others 25%	**These relative abundance values are approximate based off Figure 1. Adult subgingival plaque**: ***Bacteroides* 12%**, *Bifidobacterium* 7%, *Streptococcus* 2%, *Prevotella* 3%, *Unclassified_Lachnospiraceae* 4%, ***Unclassified_Clostridiales* 10%**, ***Unclassified_S24-7* 12%**, *Blautia* 3%, *Faecalibacterium* 3%, *Unclassified_Ruminococcaceae* 2.5%, *Lactobacillus* 7%, *Allobaculum* 6%, others 25% **Adult tongue**: *Bacteroides* 3%, *Bifidobacterium* 2%, ***Streptococcus* 17%**, ***Prevotella* 17%**, *Unclassified_Lachnospiraceae* 2%, *Unclassified_Clostridiales* 2%, *Unclassified_S24-7 2%*, *Blautia* 0.5%, *Faecalibacterium* 0.5%, *Unclassified_Ruminococcaceae* 0.2%, *Lactobacillus* 1%, *Allobaculum* 1%, *Leptotrichia* 7%, *Neisseria* 7%, others 31%			
Li 2012									
Lloyd-Price, 2017									Temporal dynamics of species and microbial pathways at each body site were evaluated by Jaccard distance and Gaussian decomposition to evaluate variance. In the gut, Bacteroidetes species (in particular the *Bacteroides* genus) exhibited primarily inter-individual variation where Firmicutes were more temporally dynamic within individuals. **Bacteroidetes species were found to be highly personalized in the gut community.** Species abundance in the oral microbiomes (vs. gut) exhibited greater time-

									varying dynamics and biological noise overall
Schmidt 2019					**Genera predominantly fecal** (species belonging to genera in >10% of fecal but <10% of saliva): *Akkermansia, Odoribacter, Tannerella, Parabacteroides, Bacteroides, Bilophila, Desulfovibrio, Klebsiella, Eggerthella, Collinsella, Turicibacter, Coprobacillus, Clostridium, Holdemania, Oscillibacter, Ruminococcaceae bacterium, Psuedoflavonifractor, Subdoligranulum, Faecalibacterium, Anaerotruncus, Ruminococcus, butyrate-producing bacterium, Coprococcus, Marvinbryantia, Roseburia, Blautia, Ruminococcus, Lachnospiraceae bacterium 1, Lachnospiraceae bacterium 2, Lachnospiraceae bacterium 9, Dorea*.	**Genera predominantly oral** (species belonging to genera in >10% of saliva but <10% of fecal): *Treponema, Capnocytophaga,* Bacteroidetes, *Porphyromonas, Prevotella, Campylobacter, Cardiobacterium, Aggregatibacter, Lautropia, Eikenella, Kingella, Neisseria, Leptotrichia, Fusobacterium, Atopobium, Corynebacterium, Bulledia, Catonella, Selenomonas, Dialister, Filifactor, Pseudoramibacter, Abiotrophia, Lachnospiracara*	By prevalence across subjects, of 310 profiled species, 44% were predominantly fecal (in >10% of fecal but <10% of saliva). These included core members of the GM such as *Clostridium sp*., *Ruminococcus sp*., and *Bacteriodes sp*.	By prevalence across subjects, of 310 profiled species, 16% were predominantly oral (in >10% of saliva but <10% of fecal).	Subset of 46 individuals had longitudinal data (mean 79 days in between sampling intervals). Both oral and fecal strain populations were usually stable even with extended periods of time.

Segata 2012	**Group 4 Stool**: Actinobacteria 0.5%, Bacteroidetes 65.2%, Firmicutes 29.64%, Fusobacteria 0.07%, Proteobacteria 2.91%, Spirochaetes 0%, TM7 0.01%. **Stool microbiome was different from oral by high percentage of Bacteroidetes**	**Oral group 1 (buccal mucosa, keratinized gingiva, hard palate)**: Actinobacteria 4.62%, Bacteroidetes 11.55%, Firmicutes 62.26%, Fusobacteria 3.74%, Proteobacteria 17.43%, TM7 0.17%. **Oral group 2 (throat, PT, TD, saliva)**: Actinobacteria 7.26%, Bacteroidetes 19.59%, Firmicutes 42.97%, Fusobacteria 10.96%, Proteobacteria 17.24%, Spirochaetes 0.3%, TM7 1.15%, **Oral group 3 (supPlaque, subPlaque):** Actinobacteria 23.69%, Bacteroidetes 19.98%, Firmicutes 22.87%, Fusobacteria 13.39%, Proteobacteria 17.39%,	Biomarkers for stool came from Lachnospiraceae and Ruminococcaceae families (Firmicutes phylum)		**Group 4 Stool**: Actinomyces (Actinobacteria) 0.03%, Corynebacterium (Actinobacteria) 0.01%, Rothia (Actinobacteria) 0.002%, **Bacteroides (Bacteroidetes) 47.82%**, **Parabacteroides (Bacteroidetes) 4.09%**, Porphyromonas (Bacteroidetes) 0.01%, **Prevotella (Bacteroidetes) 3.16%**, **Alistipes (Bacteroidetes) 5.51%**, Capnocytophaga (Bacteroidetes) 0.003%, Gemella (Firmicutes) 0.01%, Granulicatella (Firmicutes) 0.002%, Streptococcus (Firmicutes) 0.07%, Oribacterium (Firmicutes) 0.001%, Roseburia (Firmicutes) 2.08%, **Faecalibacterium (Firmicutes) 4.58%**, Oscillibacter (Firmicutes) 2.18%, Ruminococcus (Firmicutes) 1.62%, Subdoligranulum (Firmicutes) 1.43%, Veillonella (Firmicutes) 0.07%, Fusobacterium (Fusobacteria) 0.07%, Leptotrichia (Fusobacteria) 0.002%, Kingella (Proteobacteria) 0.003%, Neisseria (Proteobacteria) 0.004%, Campylobacter (Proteobacteria) 0.004%, Actinobacillus (Proteobacteria) 0.001%, Haemophilus (Proteobacteria) 0.01%.	**Oral group 1 (buccal mucosa, keratinized gingiva, hard palate):** Actinomyces (Actinobacteria) 2.34%, Corynebacterium (Actinobacteria) 0.31%, Rothia (Actinobacteria) 1.59%, Bacteroides (Bacteroidetes) 0.23%, Parabacteroides (Bacteroidetes) 0.02%, **Porphyromonas (Bacteroidetes) 3.04%**, **Prevotella (Bacteroidetes) 3.66%,** Alistipes (Bacteroidetes) 0.01%, Capnocytophaga (Bacteroidetes) 0.61%, **Gemella (Firmicutes) 5.22%**, Granulicatella (Firmicutes) 1.8%, **Streptococcus (Firmicutes) 47.32%**, Oribacterium (Firmicutes) 0.35%, Roseburia (Firmicutes) 0.63%, Faecalibacterium (Firmicutes) 0.02%, Oscillibacter (Firmicutes) 0.01%, Ruminococcus (Firmicutes) 0.004%, Subdoligranulum (Firmicutes) 0.003%, **Veillonella (Firmicutes) 5.23%**, Fusobacterium (Fusobacteria) 2.7%, Leptotrichia (Fusobacteria) 0.97%, Kingella (Proteobacteria) 0.15%, **Neisseria (Proteobacteria) 3.49%**, Campylobacter (Proteobacteria) 0.59%, Actinobacillus (Proteobacteria) 1.1%, **Haemophilus (Proteobacteria) 4.11%**. **Oral group 2 (throat, palatine tonsils, tongue dorsum, saliva)**: Actinomyces (Actinobacteria) 5.03%, Corynebacterium (Actinobacteria) 0.22%, Rothia (Actinobacteria) 1.28%, Bacteroides (Bacteroidetes) 0.15%, Parabacteroides (Bacteroidetes) 0.02%, **Porphyromonas (Bacteroidetes) 3.78%**, **Prevotella (Bacteroidetes) 11.56%**, Alistipes (Bacteroidetes) 0.01%,			Seperated 10 sites sampled/analyzed into 4 bacterial community types based on microbial community abundance. 3 oral groups- Oral group 1 (buccal mucosa, keratinized gingiva, hard palate), Oral group 2 (throat, palatine tonsils, tongue dorsum, saliva) and Oral group 3 (supPlaque, subPlaque) and 1 stool group. Firmicutes and Bacteroidetes were dominant in all groups but at different proportions. Close oral body sites (tongue dorsum and hard palate) had very different community structure, even at the phylum level. Prevotella, Veillonella and Streptococcus are least variable across both body sites and individuals. Genera that include important human pathogens colonize the throat/tonsils: *Streptococcus pneumoniae, Streptococcus pyogenes, Neisseria meningitidis, Haemophilus influenzae* were all represented in the microbita of the upper digestive tract sites. **Authors postulated that saliva, via its impact of pH and nutrient availability, is a key driver of microbial composition in habitats above the stomach.**

		Spirochaetes 1.43%, TM7 0.62%. **Higher relative abundance of TM7, Synergistetes and SR1 in all oral groups vs. gut/stool. Firmicutes dominated the oral tissue surface and saliva microbial communities. Dental plaque taxa had more evenly distributed Firmicutes, Bacteroidetes, Acinobacteria, Proteobacteria and Fusobacteria distribution.**				Capnocytophaga(Bacteroidetes) 1.25%, Gemella (Firmicutes) 1.56%, Granulicatella (Firmicutes) 1.52%, **Streptococcus (Firmicutes) 20.36%**, Oribacterium (Firmicutes) 1.65%, Roseburia (Firmicutes) 0.01%, Faecalibacterium (Firmicutes) 0.01%, Oscillibacter (Firmicutes) 0.01%, Ruminococcus (Firmicutes) 0.01%, Subdoligranulum (Firmicutes) 0.003%, **Veillonella (Firmicutes) 10.24%**, **Fusobacterium (Fusobacteria) 7.28%**, **Leptotrichia (Fusobacteria) 3.47%**, Kingella (Proteobacteria) 0.1%, **Neisseria (Proteobacteria) 6.61%**, Campylobacter (Proteobacteria) 1.8%, Actinobacillus (Proteobacteria) 0.5%, Haemophilus (Proteobacteria) 2.25%.			

Stearns 2011	Colon and stool with predominance of Firmicutes and Bacteroidetes. Greater Firmicutes. Also Proteobacteria and Fusobacteria represented in some samples.	**Oral sites sampled: mouth plaque (L&R supra-gingival and sub-gingival, tongue**. Mouth with greater representation of bacterial phyla but smaller relative abundances: Actinobacteria- about 10%, Firmicutes and Bacteroidetes in relatively equal abundance around 20-30%, greater abundances of both Fusobacteria and Proteobacteria ~10-20%. TM7, SR1 and Spirochaetes also seen							Reported was more phylogenetic variability between subjects then there was between sample sites from within each GI location (i.e. mouth, large intestine). Very small sample size. Mouth samples had the most phylogenetically similar microbiomes across subjects.
Vasapolli 2019	Lower GI (mucosal) showed a low abundance of Actinobacteria but high abundance in feces.	Upper GI tract characterized by a high abundance of Bacteroidetes	Clostridia and Erysipelotrichia colonized the lower GI mainly.	Bacilli and Negativicutes colonized the upper GI predominantly	*Bacteroides* negligible in saliva and upper GI but increased in abundance in lower GI and feces. *Blautia, Clostridium XI , Faecalibacterium* and *Ruminococcus* (Clostridia) predominate in the lower GI tract. *Bifidobacterium* and *Sutterellaceae* higher in feces vs. lower GI tract mucosal samples.	Upper GI tract (+ saliva) characterized by a high abundance of *Fusobacteria*. Highest Prevotella in saliva-diminished in abundance down upper GI and negligible in lower GI tract. *Gemella, Streptococcus* and *Veillonella* more likely in upper GI.			

Supplemental Table 2d. Results synthesis: Shotgun

Author Year	Purpose	Sites (if different)	Database	Oral Results	Gut/Stool Results	Comparison	Intepretation/Notes
Costello 2009	Only 16S used						
Ding 2014	Only 16S used						
Huse 2012	Only 16S used						
Iwauchi 2019	Only 16S used						
Li 2012	Only 16S used						
Lloyd-Price 2017	Obtain panmicrobal (bacterial, archaeal, viral, eukaryotic) profiling, strain characterization, functional profiling and metabolic pathway analysis. Evaluated if metabolic pathways were core to body sites.	Primarily target subset of 6 body sites: anterior nares, buccal mucosa, supragingival plaque, tongue dorsum, stool and posterior fornix,	Taxonomic profiling performed using MetaPhlAn2, strain characterization using StrainPhlAn2, functional profiling performed using HUMAnN2,	Pathways significantly enriched in supragingival plaque: nitrate reduction pathway (both denitrification-pwy and PWY-6748), lipid IVA biosynthesis (naglipasyn-pwy), NAD/NADP-NADH/NADPH cytosolic interconversion (pwy-7268); Pathway significantly enriched in buccal mucosa: (Kdo)2-lipid A biosynthesis (KDO-NAGLIPASYN-PWY); Pathways significantly enriched in tongue dorsum: ADP-L-glycero-beta-D-manno-heptose biosynthesis (PWY0-1241), L-1,2-propanediol degradation (PWY-7013)	Metabolic pathways that were specifically enriched in stool were: thiazole biosynthesis, L-rhamnose degratation, beta-D-glucuronide and D-glucuronate degradation, 8-anomino-7-oxononanoate biosynthesis, D--galacuturonate degradation, L-histidine degradation, 4-deoxy-L-theo-hex-enopyranuronate degradation, mannan degradation, hexuronide and hexuronate degradation, L-glutamate and L-glutamine biosynthesis	Adenosine nucleotides de novo biosynthesis was broadly distributed across all body sites. Pyruvate fermentation to propionate enriched in both oral and gut sites.	There are metabolic pathways that are distributed across body sites, pathways that are core to multiple body areas (multicore pathways) and pathways that were core to all targeted body sites (supercore pathways). A majority of pathways sequenced were annotated to <10% of genera, multicore pathways were annotated to 48% of genera and supercore to 70%. Site-enriched pathways are suggestive of functional adaptation by the microbiota to a particular niche within the human body. Temporal abundance for metabolic pathways were calculated by Gaussian variance decomposition and pathway abundances at all body sites (except the posterior fornix) were less personalized than the taxa that encoded them--consistent with the hypothesis that community assembly is primarily mediated by functional niches rather than a requirement for a specific organism. Functional dynamics in the gut were slow, possibly related to trends in response to long-term factors like dietary patterns. Conversely, dynamics in the oral cavity sites were rapid (in particular the buccal mucosa) in accordance with the enrichment of the habitat for fast energy harvest and much greater environmental exposure.

Schmidt 2019	Whole study implemented shotgun sequencing. Metagenomic sequencing allowed investigators to track populations at the resolution of strain (vs. genus or species) to establish and quantify microbial transmission between oral and gut sites. SNVs were profiled across metagenomes as a proxy for strain populations	All samples were analyzed using shotgun metagenomics sequencing	Metagenomic reads mapped in NG less; Vertical coverage (sequencing depth) and horizontal coverage (breadth) quantified using MetaSNV (https://doi.org/10.1371/journal.pone.0182392). Phenotypes were annotated using the PATRIC database	Transmission scores were negatively correlated with genome size-- transmitted species generally had smaller genomes than non-transmitted ones. Oxygen tolerant species (aerobes and facultative anaerobes) showed 7 fold higher scores than anaerobes. No association with transmission scores with sporulation and motility.		Of the 125 species prevalent in both the mouth and the gut, 77% showed evidence of oral-fecal transmission. Strains of *Streptococcus, Veillonella, Actinomyces* and *Haemophilus* fell into this category. All members of the *Prevotella* genus were occasionally transmitted. Subset of longitudinal data studies and oral SNVs observed at baseline time period 0 were enriched among fecal/gut SNVs that newly appeared over time during follow up sampling.	Transmission score between oral to gut was created based on SNVs (proxy for bacterial strain differences). Transmission score quantified how much the similarity between oral and gut SNV profiles within an individual deviated from an inter-individual background. Longitudinal oral-gut SNV data supports the authors' oral-gut transmission hypothesis and supports that transmission is in the direction of mouth to gut and not vice versa.

Segata 2012	Examine abundance of microbial metabolic pathways from the shotgun sequencing (in subset of patients)	One body site from each of the four digestive tract groups: buccal mucosa (group 1), tongue dorsum (group 2), supragingival plaque (group 3), stool (group 4).	KEGG (enzyme families), Kos (orthologous groups), and complete metabolic modules (Kmods)	PTS transporters for small sugars were most abundant in the oral cavity: Monosaccharides-mannose transporters (M00276) & fructose transporters (M00304), as well as galactosamine transporters (M00287). Putrescine transporters (M00193, M00300) were also higher in oral cavity (anaerobiosis-related pathway). Enzymes that needed for hydrogen utilization (CoM methyltransferase, K14082) and production (hydrogenase-4, K12136) were identified in the oral cavity (potential bacterial contributers from genus *Veillonella* and *Selenomonas* or from an unclassified organism from the Pasteurellaceae*)*. Supragingival plaque had higher threhalose (M00270, M00204), alpha-glucosides (M00201, M00200) and cellobiose (M00206) transport. Protoporphyrinogen oxidase (K00231, related to iron transport) higher in oral cavity, possibly linked to *Prevotella*.	Stool microbiome had higher lactose/arabinose (M00199) and oligogalacturonide (M00202) transporters, as well as dermatan (M00076), chondroitin (M00077) and heparin-sulfate (M00078) polysaccharide degradation. Beta-glucosidase (K05347), glycolysis pathway module (M0001), ammonia production pathway (M00028-urea cycle and M00029-ornithine biosynthesis), methane production pathway (M00356 and M00357, both methanogenesis) more abundant in stool. Compared to upper digestive tract stool had higher abundance of specific multiple antibiotic resistance protein (K05595) and association with the pyruvate:ferredoxin oxidoreductase pathway. Uroporphyrinogen synthase (K01719) higher in stool (iron pathway). Gene encoding hemerithryn (K07216) was detected at multiple body sites but highly enriched in stool, hemerithryn gene associated with Clostridiales.	Oral had higher transporters for small sugars, galactosamine transporters, putrescine transporters, supragingival had higher threhalose, alpha-gulcoside and cellobiose transport. Stool had higher lactose/arabinose and oligogalacturonide transporters, dermatan polysaccharide degraders, chondroitin polysaccharide degraders, and heparin-sulfate polysaccharide degraders. Stool also more abundant in pathways related to: Beta-glucosidase cellulose degradation, glycolysis pathway module, ammonia production, and methane production. The enzymes that were present for hydrogen utilization and production that were identified in the oral cavity were almost completely absent from the gut. Cystathione-beta-lyase (K14155) and methionine-gamma-lyase (K01761), both related to hydrogen production, enriched in stool.	Difference in sugar transporters in oral vs. gut sites likely related to site-specific necessity of aerobic vs. anerobic bacterial pathways as energy sources. Beta-glucosidase: enzyme critical to pathway for cellulose breakdown to Beta-D-glucose. Glycolysis pathway module related to the Embden-Meyerhiff pathway for glucose metabolism to pyruvate in the colon--agrees with 16S data where Ruminococcus more prevalent in stool (import colonizers of plant-derived material in the gut and possess cellulytic activity. Increased abundance of genes related to ammonia and methane production is consistent with the colonic microbiome as a significant source of ammonia production. Association of stool with the pyruvate:ferredoxin oxidoreductase pathway: due to its role in conversion of metronidazole to active nitrosis form can determine sensitivity to the antibiotic. Iron transporters were widely distributers among all four body sites but specific mechanisms of iron uptake and sequestration differed as needed by body site. Hydrogen sulfide gas involved in regulation of host response (low concentrations) and in host-cell toxicity and inhibition of SCFA production (in colon) at high concentrations [refs 60-65]. Hydrogen sulfide related to cystathione-beta-lyase and methionine-gamma-lyase.
Stearns 2011	Only 16S used						
Vasapolli 2019	Only 16S used						