SUPPLEMENTARY DISCUSSION

The value of purifying VLPs for viral metagenomic projects

When each of the 32 fecal VLP-associated viromes, sequenced to an average depth of 7.8±2.9 Mb (per sample) was used to query the 12 microbiomes sampled to an average depth of 92.2 ± 17.5 Mb, we noted that 55.8±32.4 % (mean ± SD) of viral sequences generated from the VLP preps from a given human host were detectable in that individual's sequenced fecal microbiome. When a deeply sampled VLP-virome (70.16 Mb) was used to query 0.91 Gb of pyrosequencer reads from the corresponding deeply sequenced fecal microbiome[15], the percentage of VLP-derived sequences found in the fecal community DNA sample was 76.14% (**Fig. S15).** Using the same BLAST E-value threshold cutoff, we performed a reciprocal analysis, asking what percentage of the total sequences present in each of the microbiome datasets matched to sequences present in VLP datasets generated from fecal samples collected from that human host. The results disclosed that viral reads represented 3.5±2.2% (mean ± SD) of total fecal community DNA sequences in the case of the 12 more shallowly sequenced microbiomes, and 2.5% in the case of the deeply sequenced microbiome and its corresponding deeply sequenced virome (**Fig. S15a**). These findings support a view that at the present time isolating VLPs is an efficient and direct way to characterize phage populations associated with a given (fecal) microbial community.

CRISPRs

Clusters of Interspaced Short Palindromic Repeats (CRISPR) elements are stretches of DNA composed of short palindromic repeats (23-47bp) that flank short "spacers" composed of viral DNA; their presence in a bacterial genome represents a key component of host defense against bacteriophage attack[1,2]. We used CRISPR-Finder[37] to search sequenced human gut microbial genomes; CRISPR elements were detected in 48 of 74 human gut bacterial species queried and in a prominent human gut archaeon

(*Methanobrevibacter smithii*) (see **Table S9** for the list of genomes). We identified a total of 95 different direct repeats and 2,196 different spacers. These direct repeats were subsequently used to interrogate the fecal microbiome datasets to identify reads that contained at least two copies of the same direct repeat. The spacers interposed between these repeats were subsequently extracted, and together with the spacers from the 121 sequenced human gut microbial genomes used to search for sequences with high similarity in VLP viromes (defined by Cross_match; maximum of 1 gap allowed and similarity over ≥ 90% of its length). This effort yielded 1,262 reads that were similar (≥90% identity) to a spacer sequence. Sixteen of the 38 VLP viromes (including technical replicates of samples from members of families F1-4 and the deeply sequenced virome) had hits to spacers derived from fecal microbiomes. In the 12 sequenced fecal community microbiomes for which there were corresponding VLP preparations, the only hits to the viromes were spacers represented in another individual's microbiome (**Table S11**). In this analysis of fecal microbiome datasets from a single time point, and at this depth of shotgun sequencing of the microbiome, the absence of detectable viral sequences with significant similarity to bacterial spacers in a given individual's fecal microbiome suggests that viruses to which their bacterial communities were resistant are not represented in the corresponding VLP preparation. If temperate phage dominate in the fecal microbiome, we would not expect such resistance to appear at least as judged by the representation of CRISPR spacers in viromes and microbiomes. However, additional and deeper shotgun datasets of total fecal microbial community DNA need to be generated from samples collected at all time points surveyed for each individual in order to further assess whether resistance does or does not occur.

Eukaryotic viruses represented in VLP viromes

Although 73% of the sequences in the NR_Viral_DB belong to eukaryotic viruses, none of the VLP samples yielded reads covering more than 50% of the genome of any known eukaryotic virus (tblastx, e-value < 1e-3). Eukaryotic viruses with hits throughout more than 20% of their genomes included: (i) six non-human Herpesviridae with 22–49% genome coverage; (ii) one Maculavirus (Grapevine_fleck_virus; 39% coverage of its 7,564bp genome); (iii) one Aquareovirus (Aquareovirus_A_segment_11; 26% coverage of this 783bp segment of its genome); (iv) one Parapoxivirus (Bovine_papular_stomatitis_virus; 22% coverage of its 134,431bp genome); and (v) one human Rotavirus (Human_rotavirus_G3_segment_11; 21% coverage of this 1,043bp genome segment).

Supplementary Figures and Legends

Fig. S1. Percent of pyrosequencing reads generated from VLP preparations that map to the NR_Viral_DB. Sample-by-sample distribution of the percentage of VLP-derived reads with hits (tblastx, e-value < 1e-3) to the NR_Viral_DB. See the legend to **Fig. 1** in the main text for an explanation of the nomenclature used to designate samples.

Reyes *et al.*, Supplementary Fig. 1

Fig. S2. **Correlating family-level bacteria taxa present in fecal samples with the known bacterial hosts of bacteriophage present in the NR_Viral_DB and their identified homologs in fecal VLP metagenomic datasets**. The Universal Bacterial 16S rRNA tree in Greengenes was downloaded (http://greengenes.lbl.gov/Download/Taxonomic_Outlines/), collapsed at a family level based on NCBI taxonomy, and branches colored according to their assigned phyla. The left panel corresponds to distribution and relative abundance (1.0 = most abundant) of the different samples according to 16S rRNA data. The middle panel shows the distribution and percent coverage of bacteriophage genomes from the NR_Viral_DB by VLP-derived reads; phage genomes are classified according to their host taxonomy. The right panel shows the distribution and relative abundance of the known bacterial hosts of phage present in the NR_Viral_DB. Columns are sorted by individual and time points of fecal sampling. Green arrows point to ssDNA phage from *Chlamydia* and *Bdellovibrio* known to be preferentially amplified by WGA methods[17]. For sample abbreviations see **Fig 1** in the main text.

Reyes *et al.*, Supplementary Fig. 2

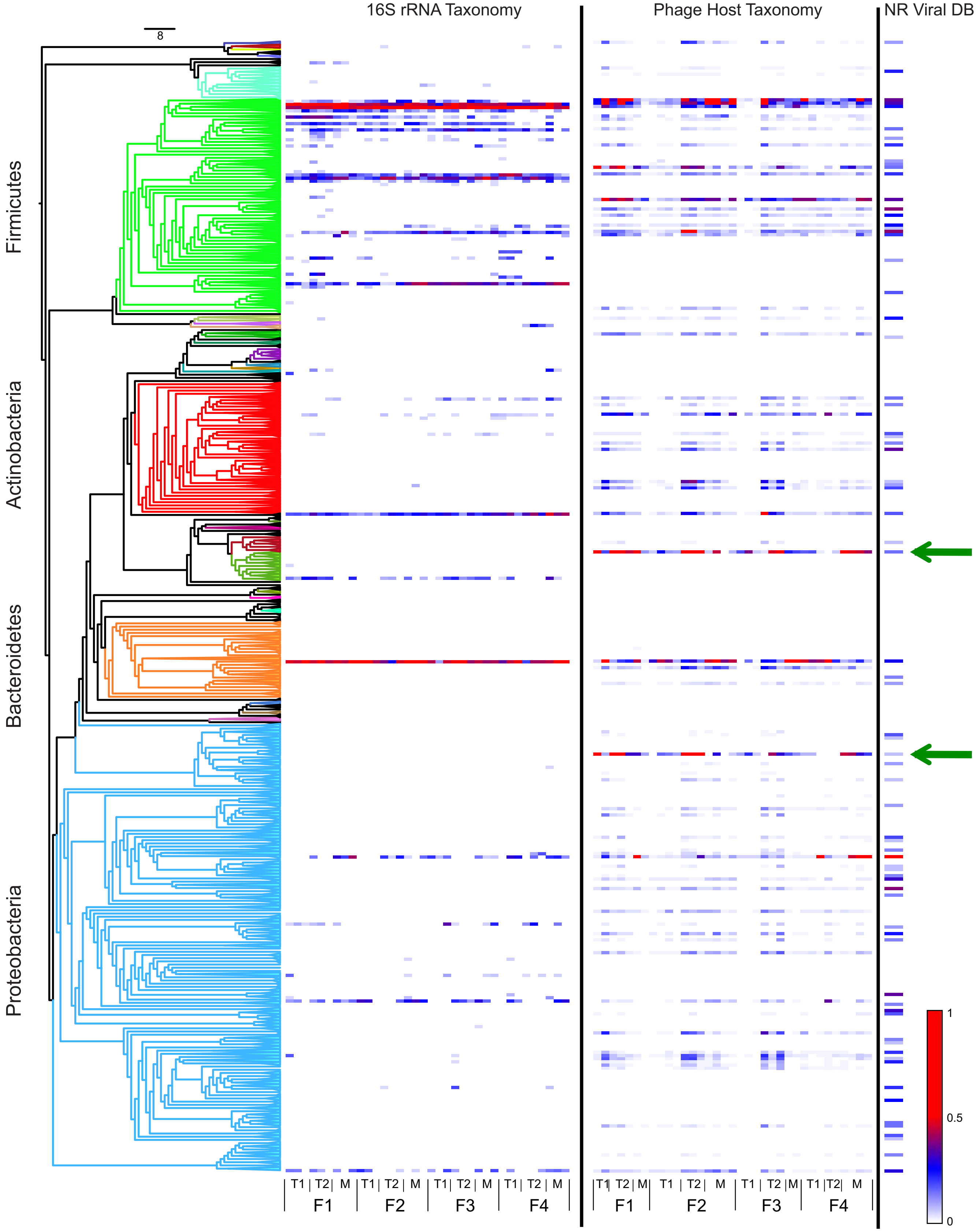


Fig. S3. Representative Monte Carlo simulations for cross contigs defining intrapersonal vs interpersonal variation in VLP DNA viromes. Monte Carlo simulations for the percent shared viral genotypes (virotypes) and percent permuted rank abundance of virotypes between pairs of fecal VLP samples. Colors indicate the likelihood score for a given position. Intra-personal variation is displayed in panels a-c: F4T3.2 vs F4T3.3 (a); F2T1.1 vs F2T1.3 (b); F1T2.2 vs F1T2.3 (c). Inter-personal variation is illustrated in panels d-f: F3M vs F3T1 (d); F2T1 vs F2T2 (e); F4T1 vs F4T2 (f).

Reyes *et al.*, Supplementary Fig. 3

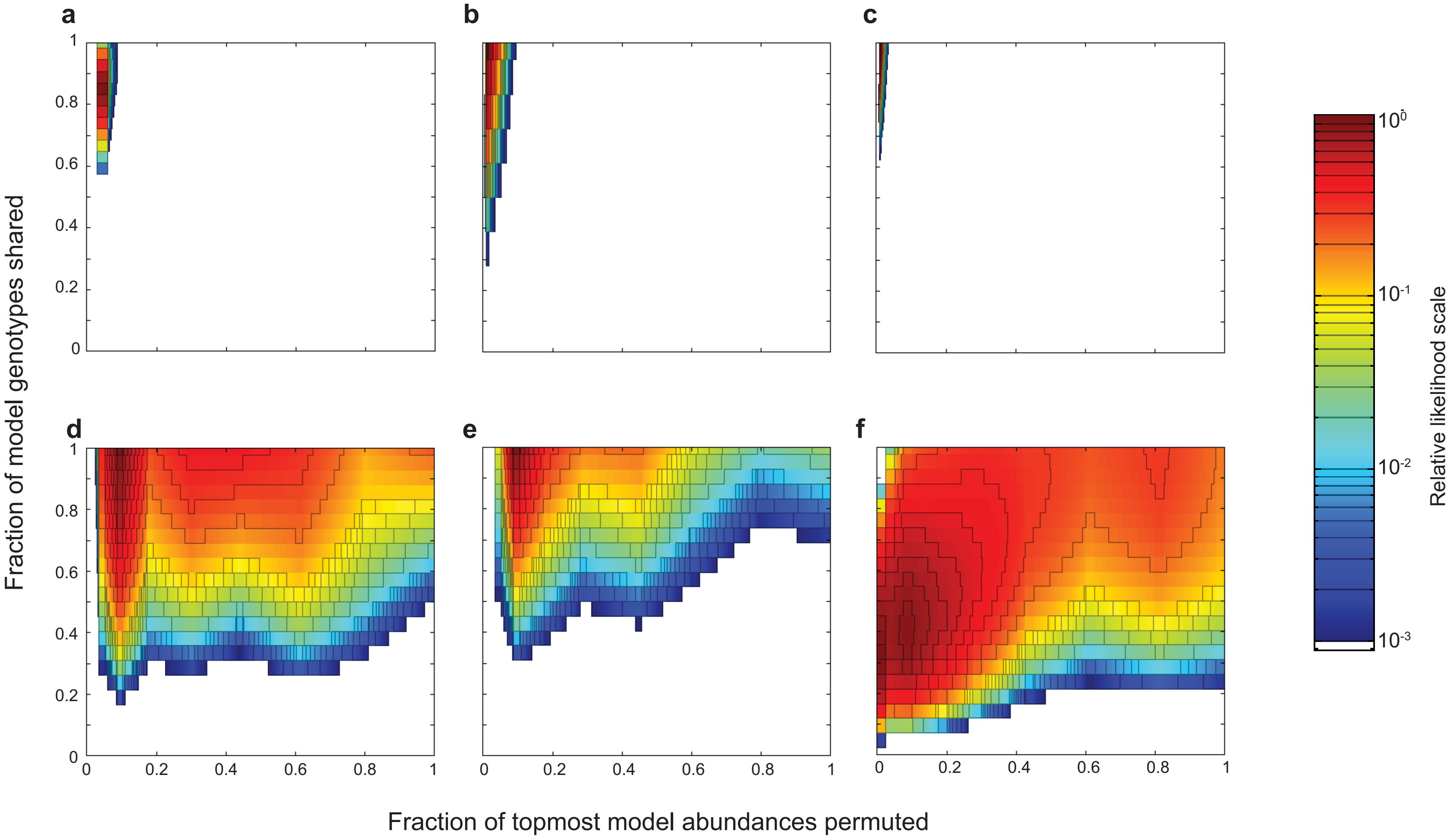


Fig. S4. Beta-diversity analysis: clustering of fecal VLP-associated viromes and bacterial 16S rRNA data. Unrooted, jack-knifed (100 iterations) consensus UPGMA trees obtained from Hellinger-based distance matrices are shown for bacterial 16S rRNA data (a) and VLP-derived viromes (b). The color key provides information about the family (F), and family member. Bars represent Hellinger distances.

Reyes *et al.*, Supplementary Fig. 4

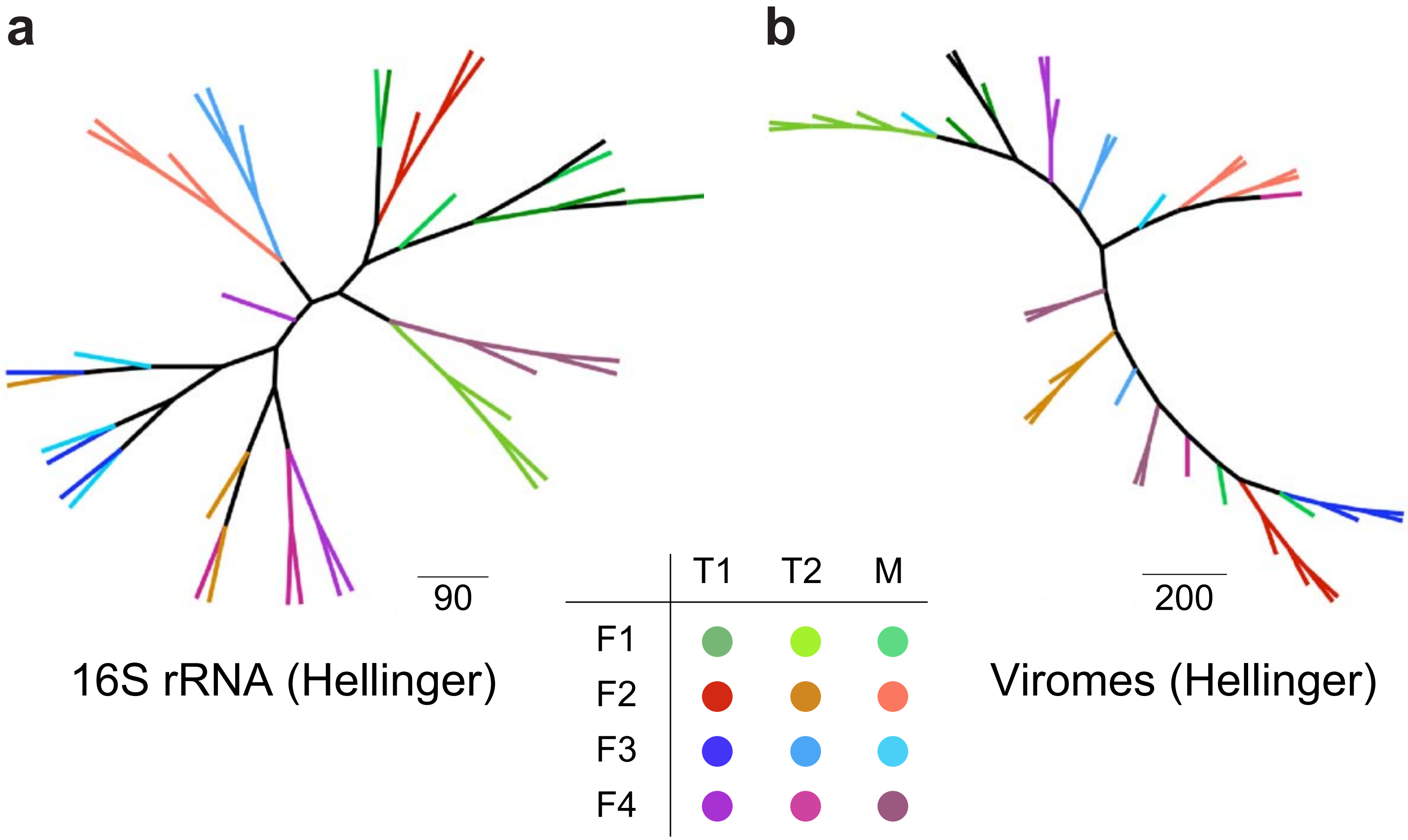


Fig. S5. Beta-diversity analysis. Branch support for the trees displayed in **Fig. S4**. Hellinger-based UPGMA trees for bacterial 16S rRNA data and VLP-derived viromes are displayed in panels a and b, respectively. The color key provides information about family (F) and the family member.

Reyes *et al.*, Supplementary Fig. 5

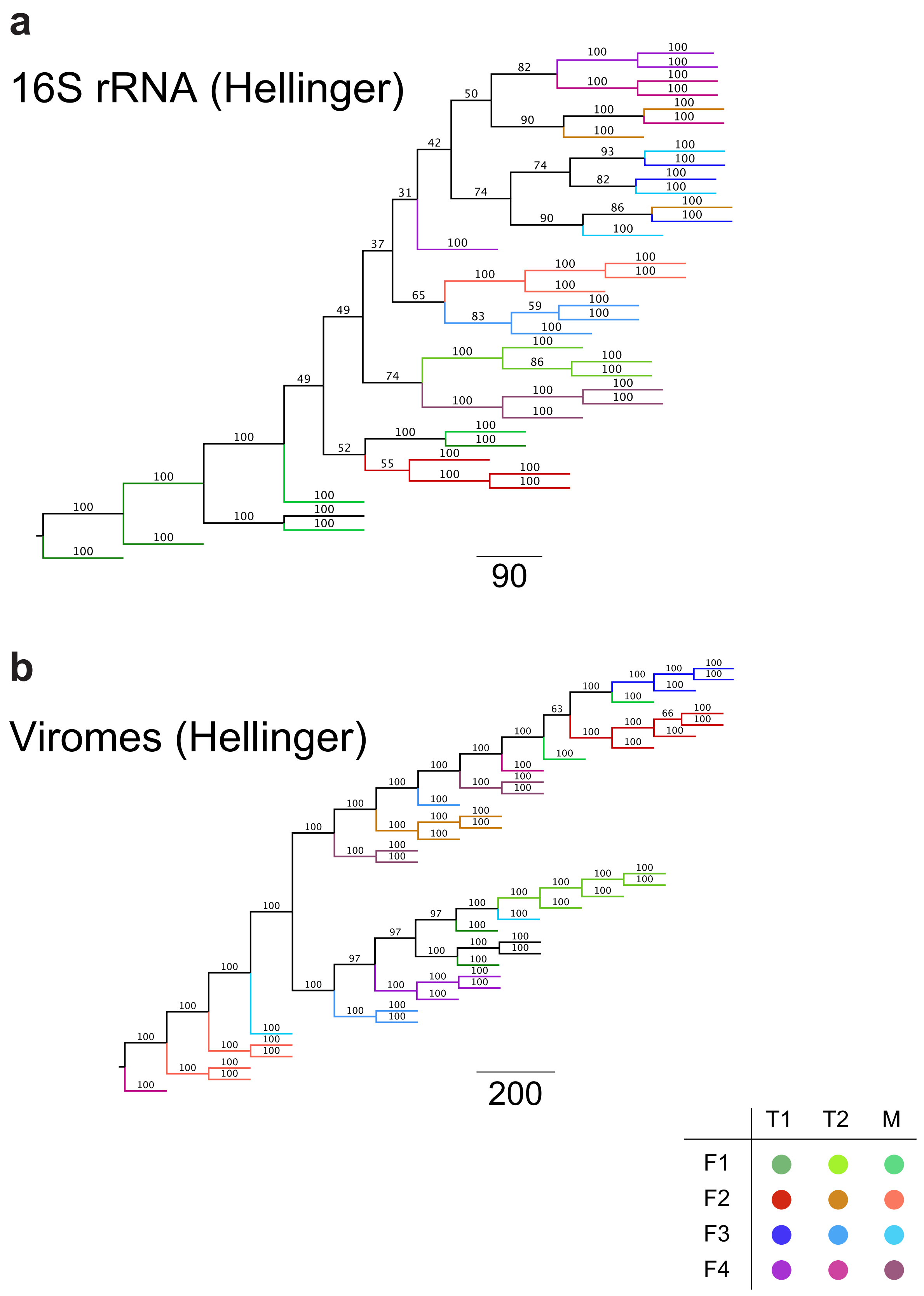
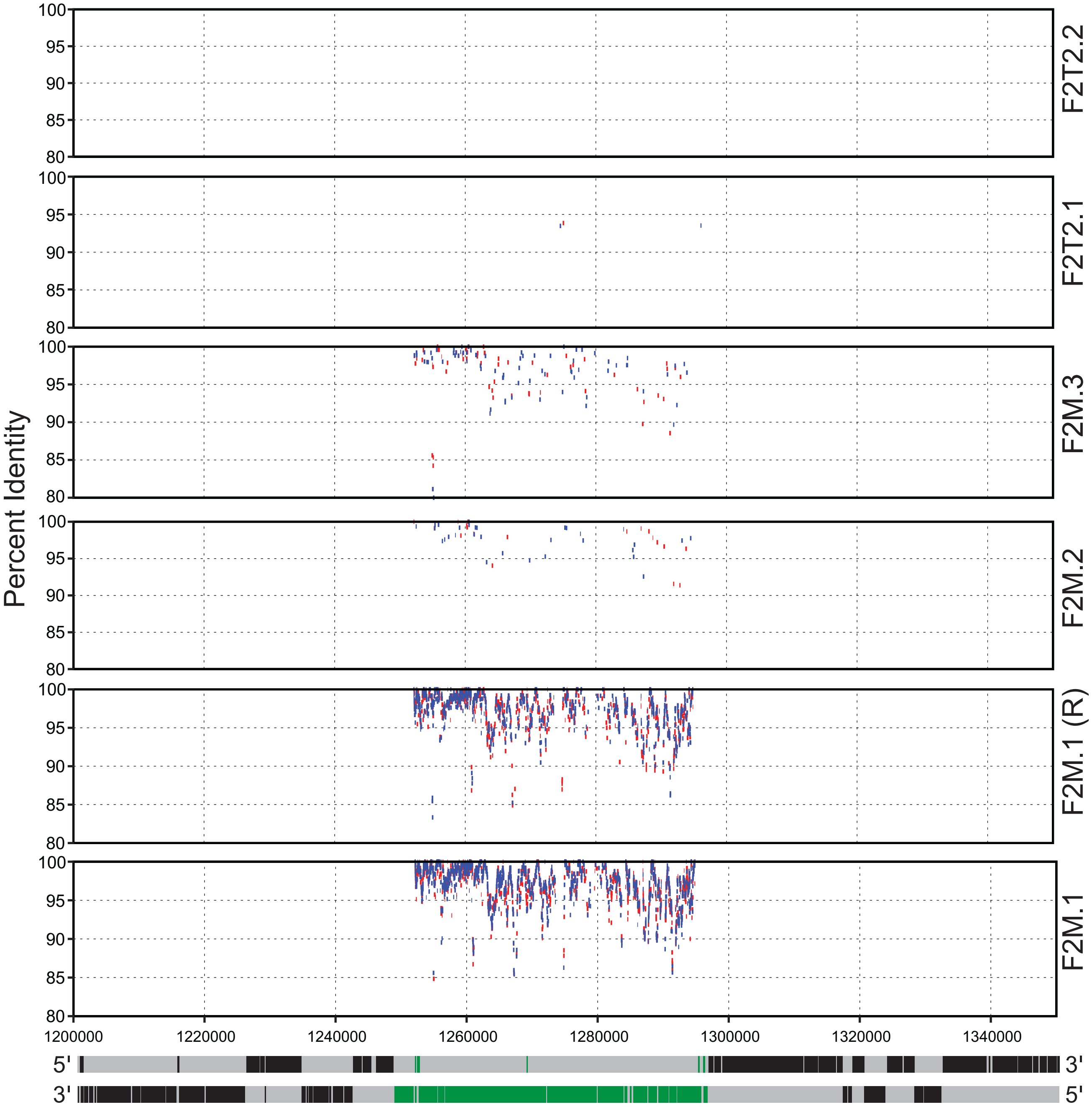


Fig. S6. Percent similarity plots of VLP virome reads mapping to a predicted prophage in *Ruminococcus torques* ATCC 27756. The genes present within the ~60 Kbp prophage are shown in green, and those present on either strand of the flanking bacterial genome are shown in black at the bottom of the figure. Pyrosequencer reads, generated from fecal VLPs, prepared at 2 or more time points from a co-twin (T2) and her mother (M) belonging to family 2 (F2) and having ≥80% identity with prophage genes, are displayed as blue dots (each dot represents a single read with a hit to the positive strand of the prophage) or red dots (negative strand hits)

Reyes *et al.*, Supplementary Fig. 6



Ruminococcus torques ATCC 27756

Fig. S7. Length distribution of viral contigs assembled from VLP-derived pyrosequencing reads. A frequency histogram of contig length is shown.

Reyes *et al.*, Supplementary Fig. 7

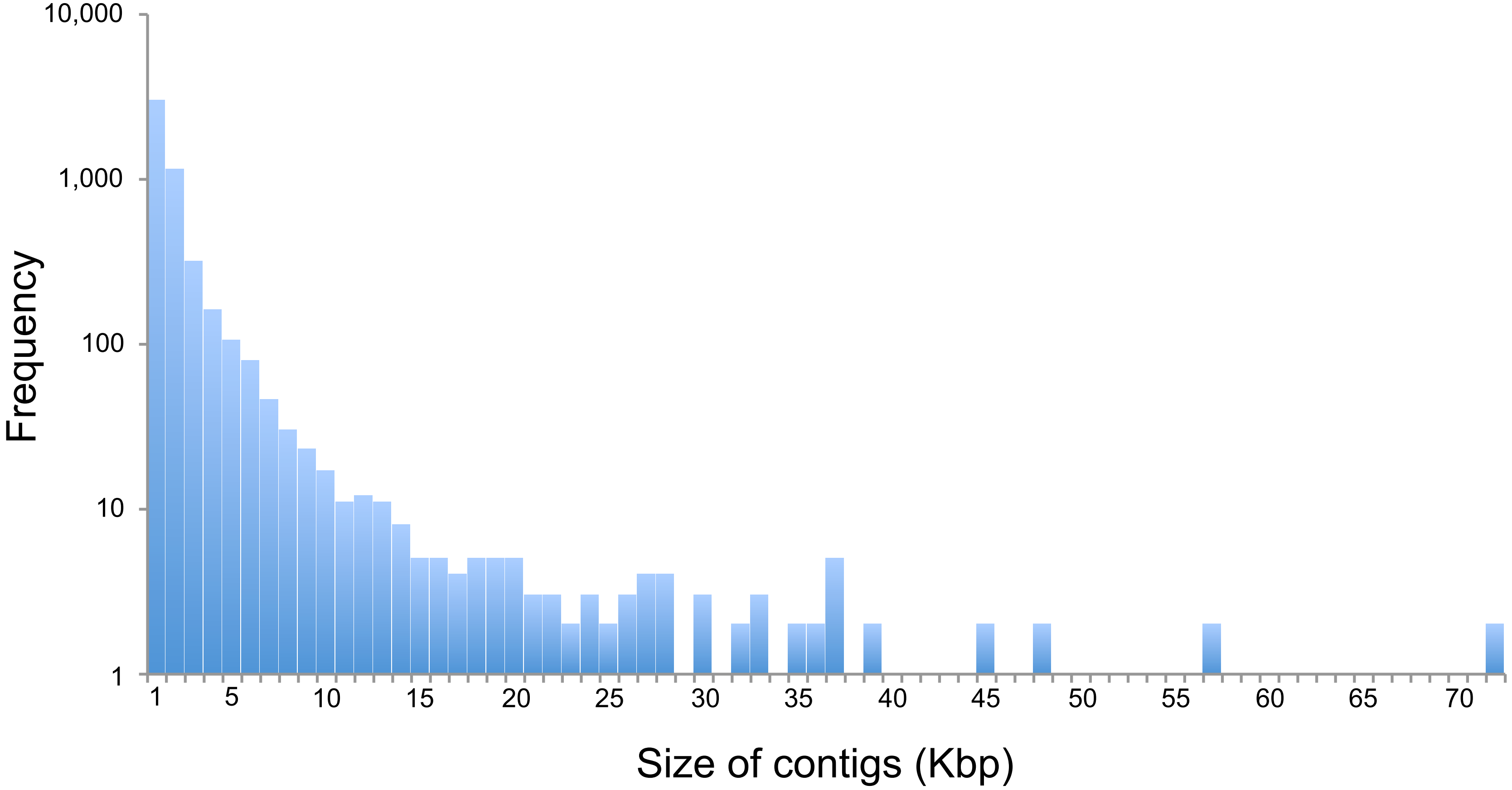
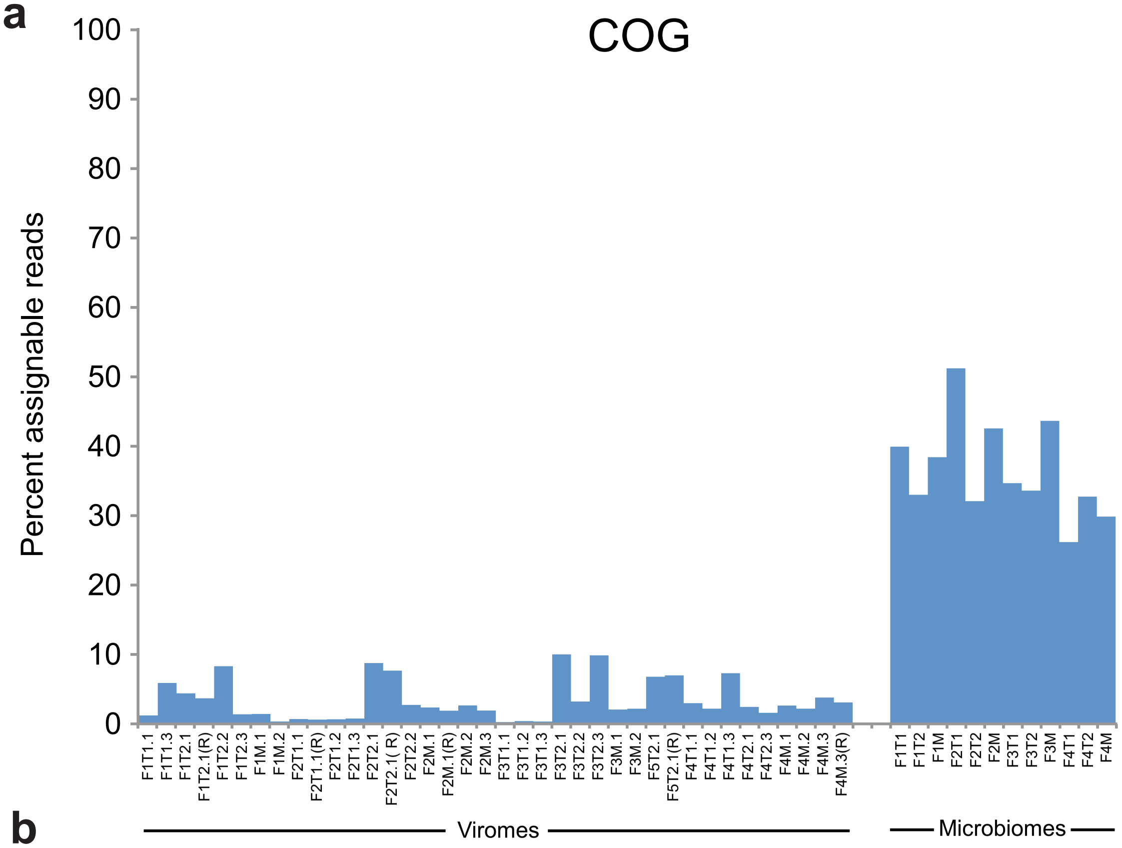


Fig. S8. Percentage of fecal virome and microbiome reads with significant hits to COG categories and KEGG second level pathways. Sample by sample percentage of reads with significant hits (blastx, e-value cutoff <1e-5) to (a) COG (STRING v7) and (b) KEGG (v44) databases.

Reyes *et al.*, Supplemental Fig. 8

a



b

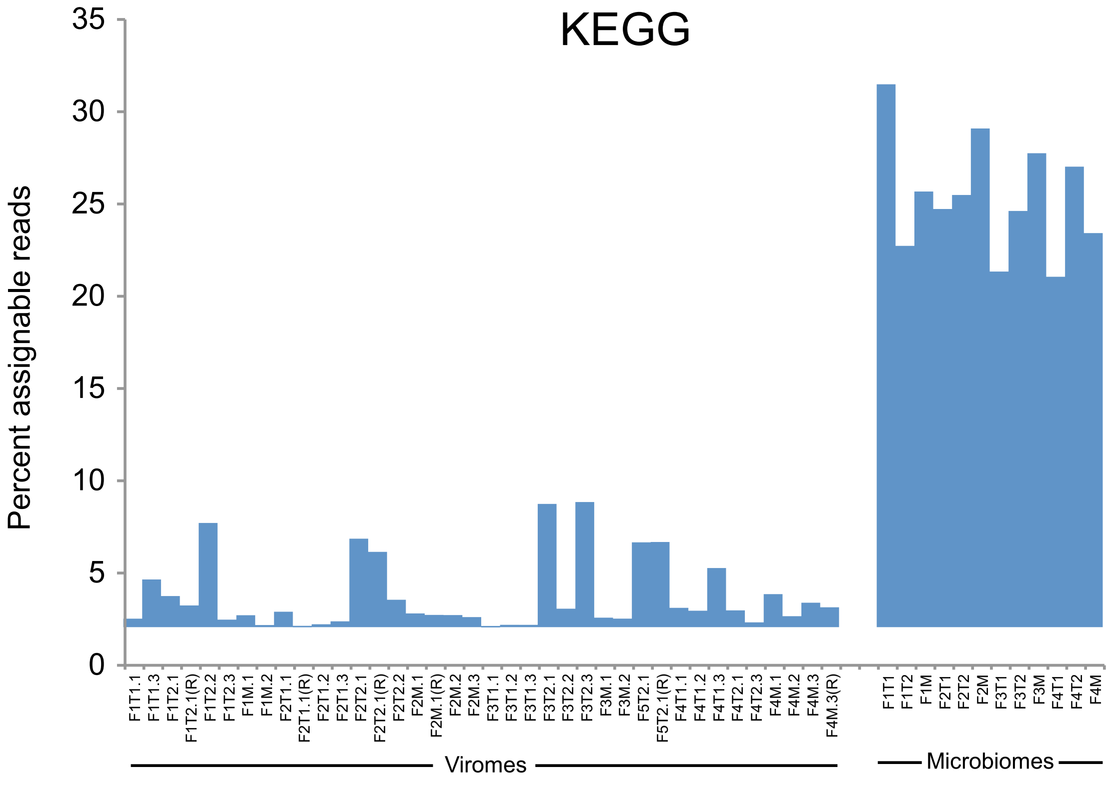


Fig. S9. KEGG and COG annotations reveal significant differences in functions between fecal VLP-associated viromes and microbiomes. Only COG-categories (panel a) and KEGG second level pathways (panel b) with significant differences in their representation between fecal microbiomes and VLP-associated viromes are shown (mean ± s.e.m plotted; $p < 0.05$; two sample t-test calculated using Metastats).

Reyes *et al.*, Supplementary Fig. 9

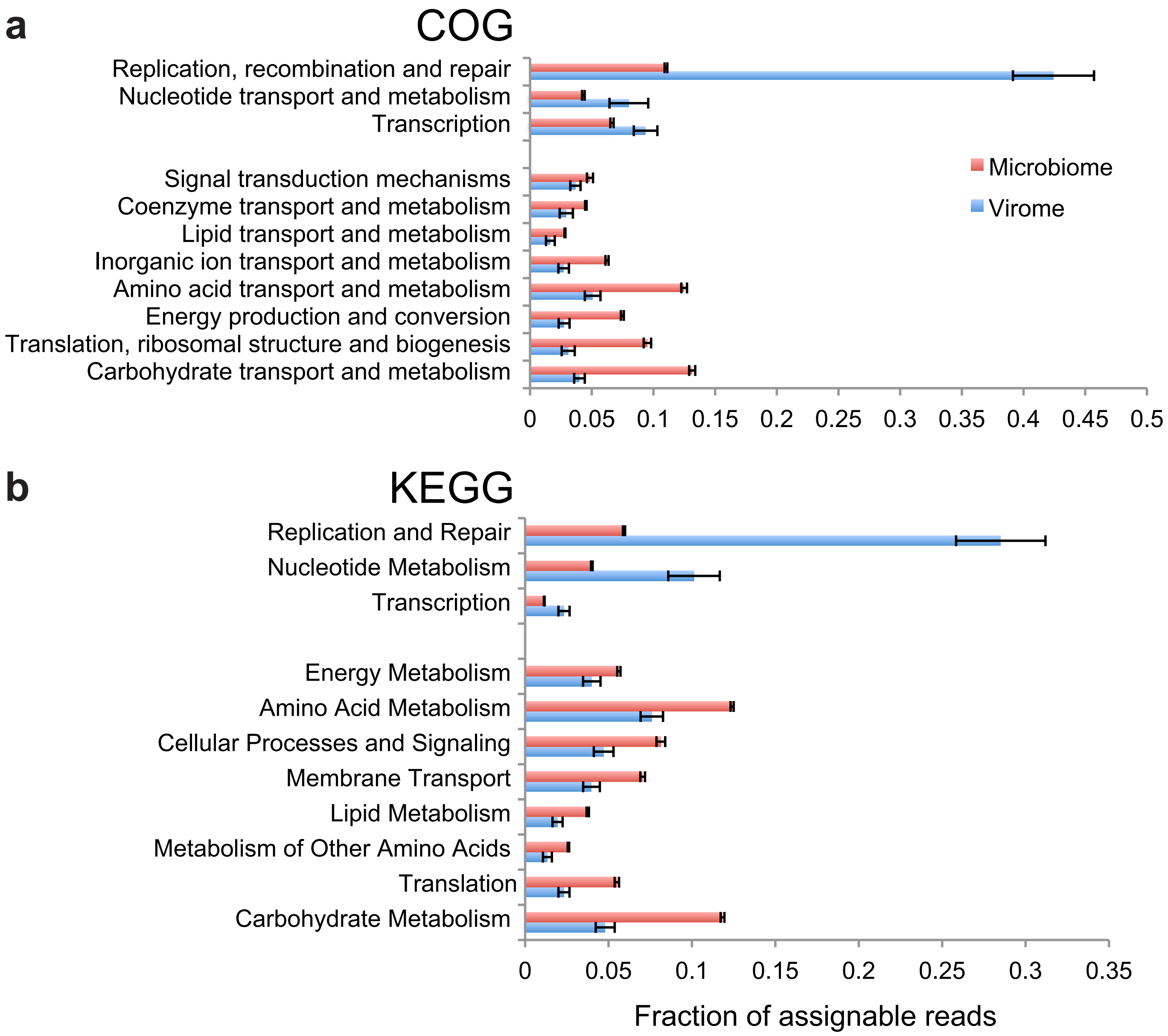


Fig. S10. A sample-by-sample view of the proportional representation of COG categories in sequenced VLP-associated viromes and gut microbiomes. Blastx assignment (e-value cutoff <1e-5) of reads to functional categories. Shown from top to bottom are proteins from viruses in the NR_Viral_DB and fecal VLP-derived viromes, plus proteins from 121 sequenced human gut-associated microbial genomes and fecal microbiomes. See **Fig. 1** for sample nomenclature.

Reyes *et al.*, Supplementary Fig. 10

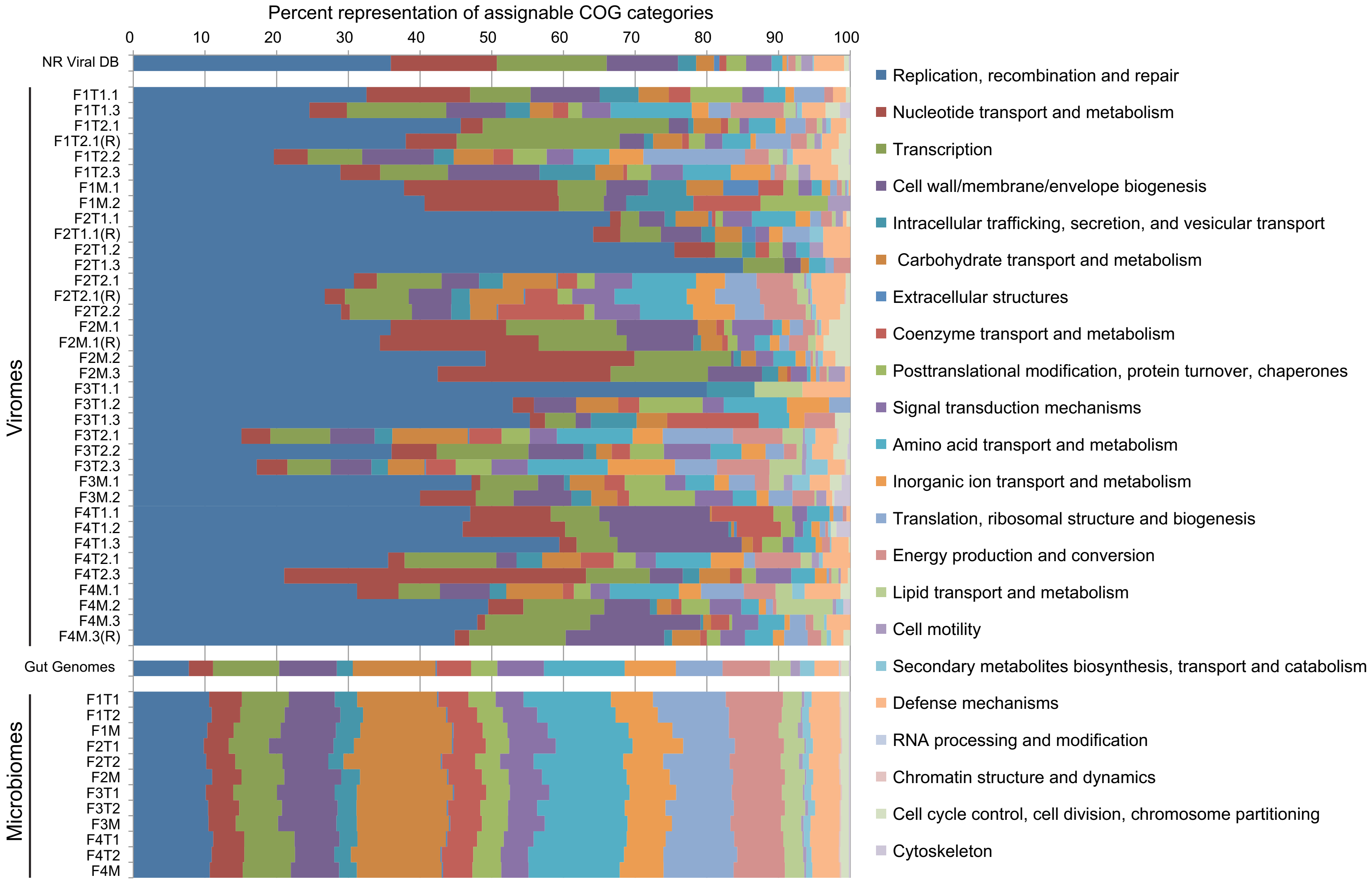


Fig. S11. Comparison of the representation of KEGG and COG groups in proteins encoded by large VLP-derived contigs and in the NR_Viral_DB. Searches (blastp, e-value <1e-5) were performed against the STRING COG database (panel a) and KEGG (results for second level pathways are shown in panel b).

Reyes *et al.*, Supplementary Fig. 11

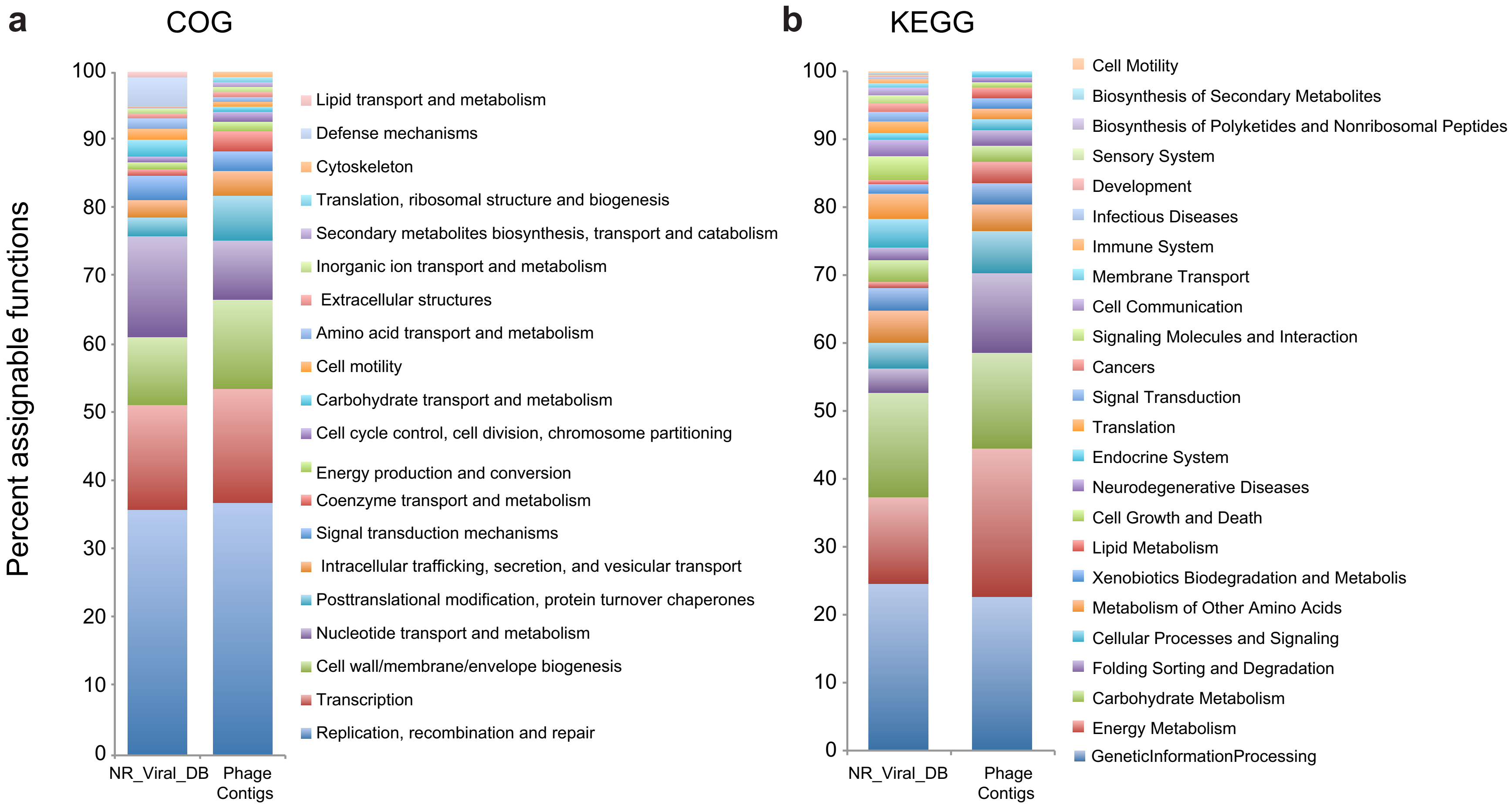
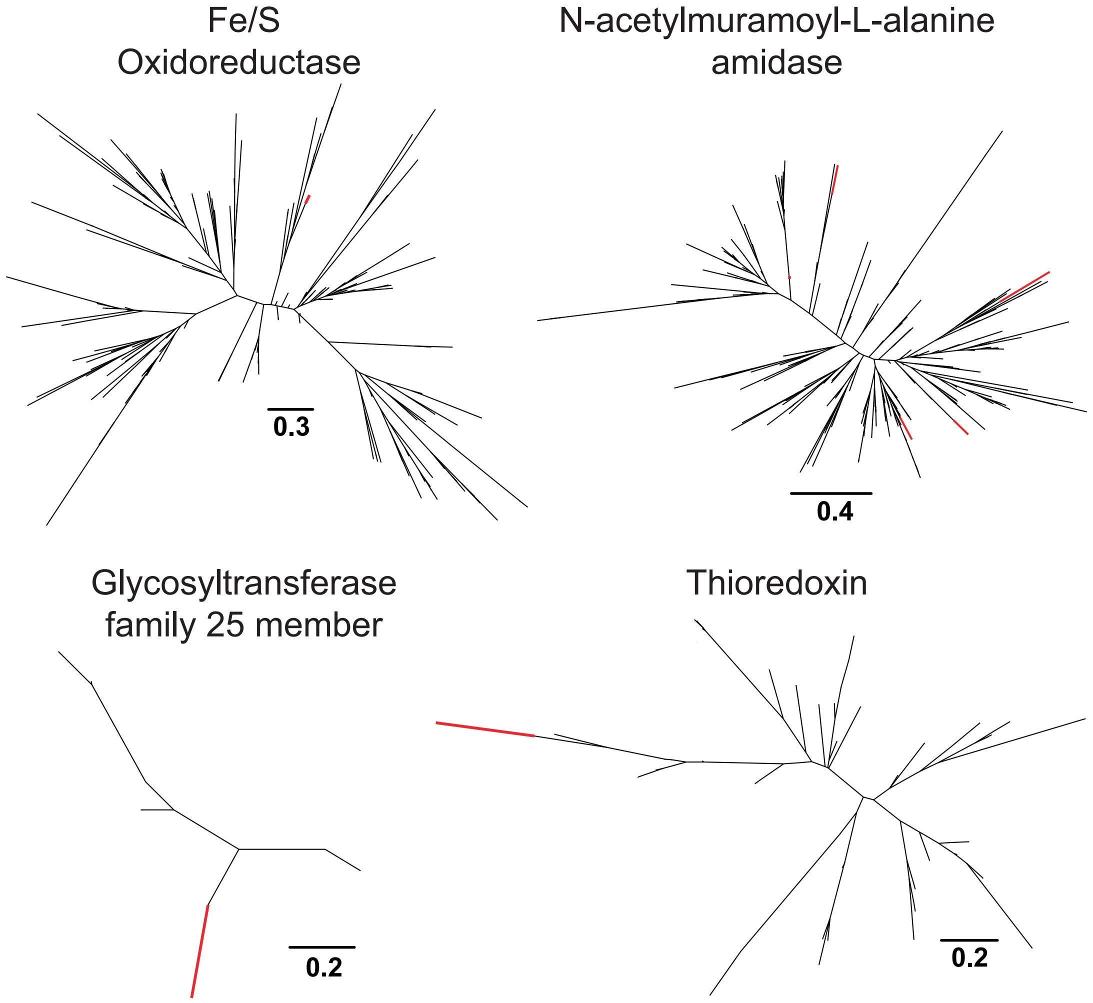


Fig. 12. Representative phylogenetic trees of bacterial proteins present in large contigs assembled from VLP-viromes with no homologs in the NR_Viral_DB. Multiple alignment of the indicated viral protein (highlighted in red) with all proteins from 121 human gut microbial genomes that harbored the same domain or motif was performed using Muscle[19]. Approximate maximum likelihood trees were generated using FastTree[20]. Bars represent the number of amino acid substitutions per position.

Reyes *et al.*, Supplementary Fig. 12



Fe/S Oxidoreductase

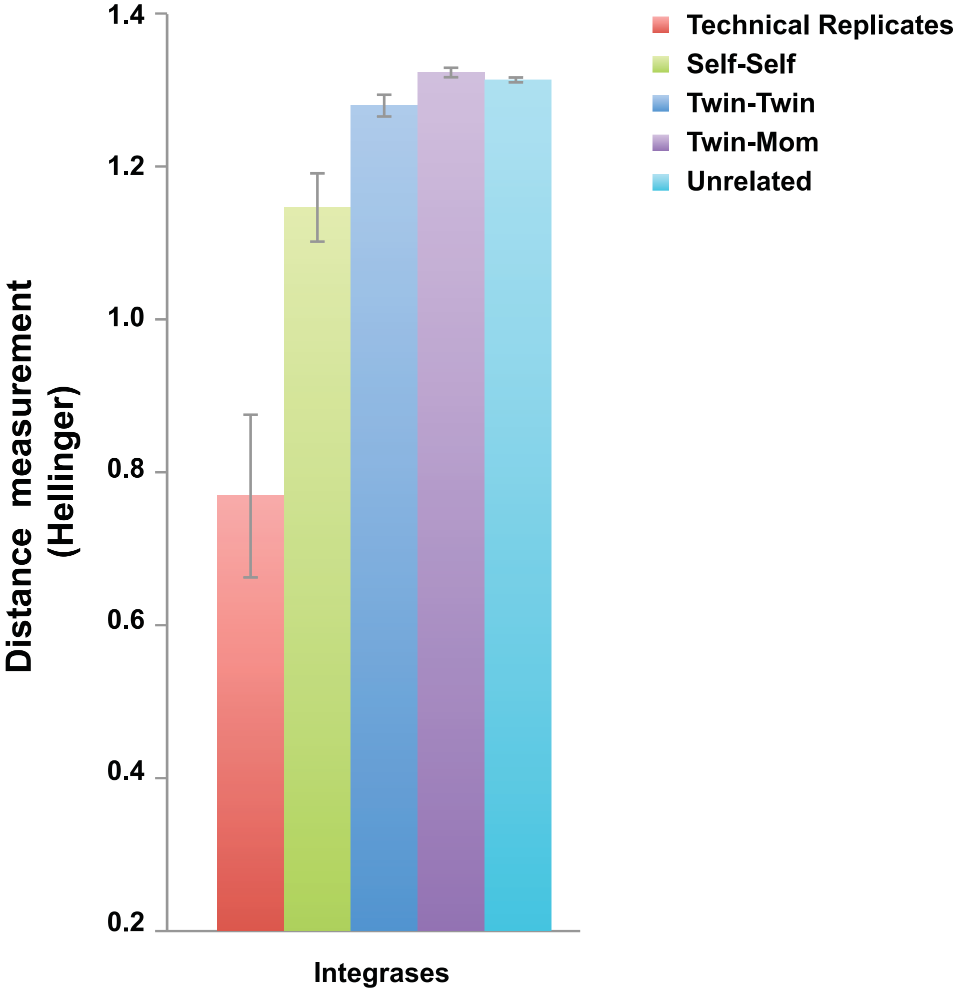
N-acetylmuramoyl-L-alanine amidase

Glycosyltransferase family 25 member

Thioredoxin

Fig. S13. Sequence diversity of integrase genes in VLP viromes. The number of pyrosequencer reads in each VLP sample with significant hits to known integrases present in the NR_Viral_DB and in prophages found in 121 human gut microbial genomes were identified and used to generate a distance matrix. Average distances among technical replicates (two shotgun datasets produced from a given VLP DNA preparation), among samples obtained from the same individual over time (intrapersonal variation), and samples obtained from co-twins, twins and their mothers or unrelated individuals, are graphed (mean ± s.e.m). The significance of differences between the groups was calculated using Student's t-test. *** $p< 0.001$; ** $p<0.01$, ns, $p>0.05$.

Reyes *et al.*, Supplementary Fig. 13



	🟦	🟪	🟦	🟩
🟥	***	***	***	***
🟩	***	***	***	
🟦	ns	**		
🟪	ns			

Fig. S14. Normalized RNA-Seq counts for predicted prophages in *Bacteroides thetaiotaomicron* VPI-5482. RNA-Seq was performed using rRNA-depleted RNA samples prepared from cecal and fecal contents harvested from gnotobiotic mice co-colonized for 2 weeks with *B. thetaiotaomicron* and *M. formatexigens* (n=3 animals). Expression levels are shown for each ORF (see color key for normalized read counts; normalization based on sequencing effort and length of each predicted ORF). Active expression is defined as a normalized read count >100. This strain of *B. thetaiotaomicron* contains two prophages. One of the prophages (labeled 1) contains a linked pair of highly expressed ORFs encoding an Xre family anti-toxin (BT4733) and a putative toxin (BT4732) while the other prophage contains a cluster of three highly expressed genes [two hypothetical proteins flanking a Xre family anti-toxin (BT4035)].

Reyes *et al.*, Supplementary Fig. 14

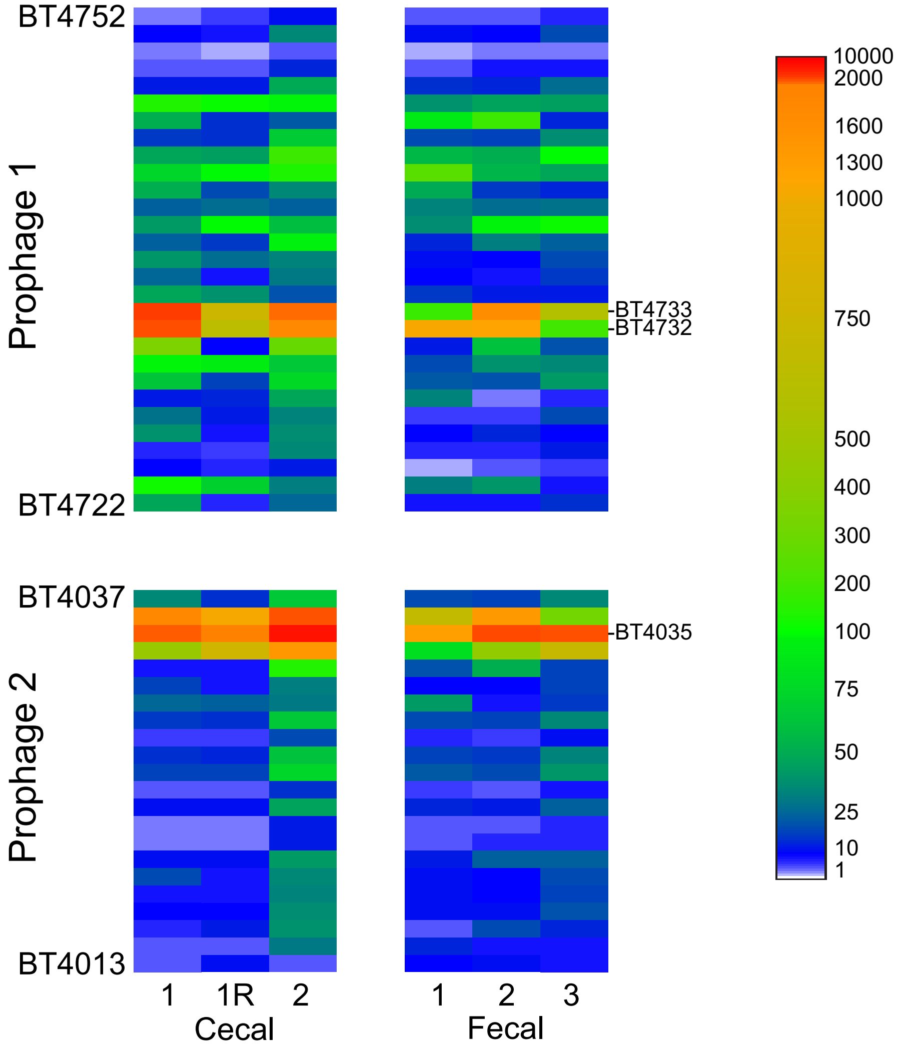
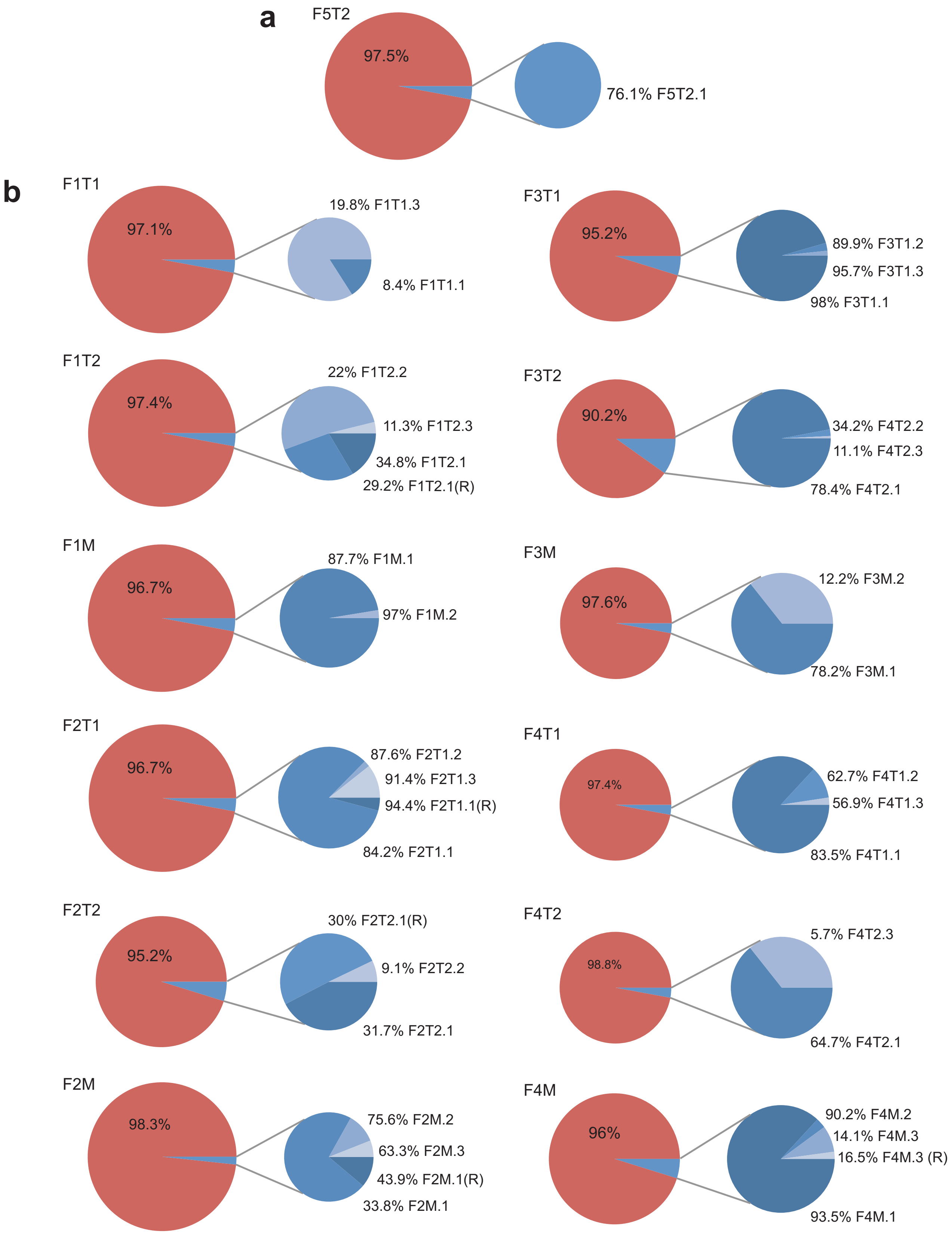


Fig. S15. Representation of VLP pyrosequencer reads in fecal microbiomes and vice versa. The percentage of reads from fecal microbiomes with significant similarity to VLP-derived reads (blastn, e-value <1e-7) is represented as a blue wedge within the red pie charts. This wedge is expanded to the right in the form a second blue pie chart that shows the percentage of reads from each of the different time point VLP preparations that have significant similarity with reads from the fecal microbiome from time point 1. (a) The percentage of shared reads between the deeply sequenced F5T2 VLP preparation and corresponding deeply sequenced fecal microbiome (293,654 and 2,579,680 reads, respectively). (b) Data derived from shallowly sequenced fecal viromes and microbiomes.

Reyes *et al.*, Supplementary Fig. 15



Supplementary Tables

Table S1. Sequencing effort for VLP preparations from 32 fecal samples obtained from 4 sets of MZ twins and their mothers. Technical replicates were performed on 6 DNA samples involving independent whole genome amplification and shotgun 454 FLX pyrosequencing. Sample F5T2.1 was subjected to deeper sequencing. NA, a fecal specimen was not available in sufficient quantity to purify VLPs at this time point.

Table S2. Sequencing effort for bacterial 16S rRNA genes present in the fecal microbiota of study participants.

Table S3. Shotgun sequencing effort for fecal community DNA (microbiome) samples.

Table S4. CD-hit cluster-based alpha diversity metrics.

Table S5. PHACCS-based alpha diversity metrics.

Table S6. Matrix of VLP samples x reference human microbial gut genomes where significant coverage to prophages in the microbial host was identified. (a) Percent of the prophage genome covered by reads from a given VLP sample. (b) Fold coverage per bp of the prophage genome, normalized to 10,000 reads. Yellow highlights instances where the prophage was covered over more than 50% of its length.

Table S7. Matrix of VLP samples x the 88 large contigs assembled from the aggregate VLP dataset showing (a) the percentage of the contig covered by reads from a given VLP sample and (b) fold-coverage per bp of each contig. Data are normalized by randomly mapping 14,000 reads per VLP sample. Yellow highlights instances where a given contig has ≥50% coverage with reads from a given VLP virome.

Table S8. Contig x VLP sample matrix of contigs present in more than one VLP sample. The percent identity between VLP-pyrosequencer reads and the corresponding contig is shown**.** Dashes indicate that no reads corresponding to the contig were present in the VLP sample **(**threshold; minimum of 80% overlap with the contig over the length of the read)

Table S9. List of 121 sequenced human gut-derived microbial genomes used for generating a reference list of KEGG second level pathways and COG assignments. The presence or absence of CRISPRs in the genomes of the 74 organisms used to search for these elements is noted.

Table S10. List of 396 sequenced microbial genomes used to identify prophage sequences for the NR_Viral_DB.

Table S11. The number of VLP reads with similarity over more than 90% of their length to CRISPR spacers present in either the fecal microbiome of the same individual or in a reference gut-associated microbial genome. Counts are normalized (10,000 reads per sample). The bottom row shows the number of spacers identified in each fecal microbiome sample (see *Supplementary Discussion* for further definition of the criteria used to select spacers from the fecal microbiome datasets to query VLP datasets).

Table S1. Shotgun pyrosequencing of DNA purified from fecal VLPs

Number of reads per dataset

Family	Family Member	Time.1	Time.2	Time.3
F1	T1	23,284	NA	43,609
	T2	43,350	31,045	28,044
	M	48,769	24,498	NA
F2	T1	71,432	40,414	26,634
	T2	40,932	36,319	NA
	M	58,379	27,647	35,561
F3	T1	18,578	19,866	30,441
	T2	53,410	33,807	28,922
	M	25,098	37,697	NA
F4	T1	24,563	20,674	14,483
	T2	27,937	NA	30,013
	M	24,980	23,731	30,207
F5	T2	48,432		

Replicates

F4	M.3	16,081
F1	T2.1	30,141
F2	T1.1	19,506
F2	T2.1	41,719
F2	M.1	34,012
F5	T2.1	245,222

Number of nucleotides per dataset

Family	Family Member	Time.1	Time.2	Time.3
F1	T1	5,480,138	NA	10,439,191
	T2	10,319,498	7,278,429	6,663,475
	M	11,874,381	6,013,335	NA
F2	T1	17,244,379	9,845,469	6,560,103
	T2	9,684,954	8,812,248	NA
	M	14,173,440	6,637,882	8,597,871
F3	T1	4,487,018	4,788,497	7,333,512
	T2	12,609,532	8,164,157	6,950,693
	M	6,019,033	9,106,849	NA
F4	T1	5,949,161	4,919,740	3,437,755
	T2	6,772,788	NA	7,115,420
	M	5,934,901	5,652,431	7,365,500
F5	T2	11,712,401		

Replicates

F4	M.3	3,931,490
F1	T2.1	7,360,504
F2	T1.1	4,747,264
F2	T2.1	10,056,492
F2	M.1	8,297,597
F5	T2.1	58,444,932

Table S2. Pyrosequencing effort for 16S rRNA V2 region amplicons

Number of reads per dataset

Family	Family Member	Time.1	Time.2	Time.3
F1	T1	6,415	1,627	40,583
	T2	15,495	1,957	2,074
	M	7,870	1,799	2,816
F2	T1	9,343	2,886	3,030
	T2	13,991	3,606	2,562
	M	7,717	4,325	3,350
F3	T1	9,837	3,953	10,392
	T2	19,586	5,045	2,417
	M	15,294	4,752	30,294
F4	T1	11,936	4,220	1,588
	T2	12,672	4,603	1,856
	M	13,789	3,284	3,071

Table S3. Shotgun pyrosequencing effort for fecal community DNA (microbiome) samples

Number of reads per dataset

Family	Family Member	Time.1
F1	T1	217,386
	T2	443,640
	M	510,972
F2	T1	414,754
	T2	490,776
	M	535,763
F3	T1	498,880
	T2	495,040
	M	413,772
F4	T1	519,072
	T2	549,700
	M	434,187
F5	T2	2,579,680

Number of nucleotides per dataset

Family	Family Member	Time.1
F1	T1	51,926,180
	T2	79,297,532
	M	103,228,389
F2	T1	95,417,867
	T2	101,090,755
	M	118,742,924
F3	T1	82,616,445
	T2	98,548,138
	M	89,199,789
F4	T1	92,506,950
	T2	112,549,303
	M	81,764,398
F5	T2	910,456,203

Table S4. CD-hit cluster-based alpha diversity metrics

		Sample	Cluster Shannon Index	Expected #Clusters	Average Genome Size	Expected #Virotypes
Family 1	Co-twin 1	1	10.94	56327	5594	403
		3	13.39	653728	34025	769
	Co-twin 2	1	11.59	108486	11863	366
		1(R)	12.24	207752	14533	572
		2	13.64	838359	12091	2773
		3	12.31	221107	19716	449
	Mother	1	12.66	315416	73725	171
		2	11.22	74731	42095	71
Family 2	Co-twin 1	1	12.00	163342	30437	215
		1(R)	11.72	122634	49931	98
		2	11.98	160020	43526	147
		3	11.91	149282	47792	125
	Co-twin 2	1	13.09	482461	9908	1948
		1(R)	13.43	682963	15585	1753
		2	12.56	285882	7552	1514
	Mother	1	13.13	503089	32162	626
		1(R)	13.22	552748	38326	577
		2	12.66	315897	49478	255
		3	12.83	374537	32373	463
Family 3	Co-twin 1	1	11.11	66702	51239	52
		2	11.27	78619	31767	99
		3	11.02	60812	43052	57
	Co-twin 2	1	13.19	536799	43499	494
		2	12.32	224460	30704	292
		3	13.60	803043	20416	1573
	Mother	1	11.82	136453	36091	151
		2	12.07	174413	27692	252
Family 4	Co-twin 1	1	12.52	274276	28033	391
		2	11.71	121210	26193	185
		3	11.66	115724	39954	116
	Co-twin 2	1	12.47	259988	31832	327
		3	11.71	121545	45812	106
	Mother	1	12.73	338213	14954	905
		2	12.22	202101	27135	298
		3	12.62	304012	31133	391
		3(R)	12.15	188697	29101	259
Family 5	Co-twin 2	1	13.49	725374	33678	862
		1(R)	13.85	1034661	33940	1219

Table S5. PHACCS-based alpha diversity metrics

		Sample	Shannon Index	Evenness	Richness
Family 1	Co-twin 1	1	3.72	0.90	62
		3	4.22	0.87	131
	Co-twin 2	1	3.68	0.91	60
		2	5.24	0.83	536
		3	3.37	0.94	37
	Mother	1	3.82	0.69	245
		2	4.09	0.66	490
Family 2	Co-twin 1	1	2.19	0.98	10
		2	2.28	0.97	11
		3	2.99	0.84	127
	Co-twin 2	1	4.51	0.92	138
		2	4.07	0.95	71
	Mother	1	3.57	0.92	50
		2	2.71	0.95	18
		3	2.85	0.95	20
Family 3	Co-twin 1	1	2.29	0.97	11
		2	2.46	0.97	13
		3	3.46	0.78	984
	Co-twin 2	1	3.72	0.85	78
		2	3.02	0.92	28
		3	4.25	0.92	102
	Mother	1	2.84	0.96	19
		2	2.93	0.96	21
Family 4	Co-twin 1	1	3.31	0.94	34
		2	3.34	0.94	35
		3	2.62	0.92	18
	Co-twin 2	1	2.83	0.96	19
		3	2.61	0.96	15
	Mother	1	3.40	0.97	33
		2	3.18	0.97	27
		3	3.23	0.91	36

Table S6a. Matrix of normalized fold-coverage per basepair of prophage genomes by VLP sample reads (10,000 reads used per VLP

		Family 1							
		Co-twin 1		Co-twin 2				Mother	
	Sample	1	3	1	1(R)	2	3	1	2
Bacterial Genomes	*Alistipes putredinis*	0.01	5.68	-	-	-	0.00	-	-
	Bacteroides caccae ATCC 43185	-	-	0.00	0.00	0.00	0.01	-	-
	Bacteroides coprocola	-	-	-	-	-	-	-	-
	Bacteroides sp.D2	0.36	-	-	-	-	-	-	-
	Bacteroides stercoris	-	0.21	-	-	-	-	0.00	0.01
	Bacteroides thetaiotaomicron 3731	0.39	-	-	-	-	-	-	-
	Blautia hansenii	-	0.01	-	-	0.01	-	-	-
	Blautia hydrogenotrophica DSM10507	-	-	-	-	-	0.00	0.00	0.01
	Clostridium leptum	0.07	0.25	0.05	0.05	0.21	0.22	-	-
	Escherichia fergusonii ATCC35469	-	-	-	-	-	-	-	-
	Holdemania filiformis	-	0.07	-	-	-	-	-	-
	Ruminococcus gnavus	-	0.01	-	-	0.01	-	-	-
	Ruminococcus torques ATCC27756	-	0.11	-	0.00	0.01	0.00	0.01	0.00

		Family 2										
		Co-twin 1				Co-twin 2			Mother			
	Sample	1	1(R)	2	3	1	1(R)	2	1	1(R)	2	3
Bacterial Genomes	*Alistipes putredinis*	-	-	0.00	-	-	-	-	-	-	-	-
	Bacteroides caccae ATCC 43185	-	-	-	-	0.30	0.27	0.09	0.03	0.02	-	0.00
	Bacteroides coprocola	-	-	-	-	-	-	0.00	0.09	0.18	3.64	-
	Bacteroides sp.D2	-	-	-	-	-	-	-	-	-	-	-
	Bacteroides stercoris	-	0.00	-	0.00	-	-	0.00	0.01	0.01	0.00	0.00
	Bacteroides thetaiotaomicron 3731	-	-	-	-	-	-	-	-	-	-	-
	Blautia hansenii	-	-	-	-	-	-	-	0.00	-	-	-
	Blautia hydrogenotrophica DSM10507	0.00	0.00	0.00	-	0.00	-	0.00	-	0.00	0.02	-
	Clostridium leptum	-	-	-	-	0.00	-	-	0.03	0.03	0.01	0.01
	Escherichia fergusonii ATCC35469	-	-	-	-	-	-	-	-	-	-	0.01
	Holdemania filiformis	-	-	-	-	0.01	0.02	0.00	0.00	-	-	-
	Ruminococcus gnavus	-	-	-	-	-	-	-	0.00	-	-	-
	Ruminococcus torques ATCC27756	-	-	-	-	0.01	0.01	-	5.23	5.00	0.16	0.42

		Family 3							
		Co-twin 1			Co-twin 2			Mother	
	Sample	1	2	3	1	2	3	1	2
Bacterial Genomes	*Alistipes putredinis*	-	-	-	0.00	0.00	-	-	-
	Bacteroides caccae ATCC 43185	-	-	-	-	-	-	-	-
	Bacteroides coprocola	-	0.00	-	0.00	0.00	0.01	-	-
	Bacteroides sp.D2	-	-	-	0.13	0.00	0.00	0.00	0.82
	Bacteroides stercoris	-	0.15	0.01	0.01	-	-	0.01	0.50
	Bacteroides thetaiotaomicron 3731	-	-	-	0.16	0.00	0.00	0.00	0.90
	Blautia hansenii	-	-	-	1.14	0.01	0.00	-	-
	Blautia hydrogenotrophica DSM10507	-	-	-	-	-	-	0.13	0.04
	Clostridium leptum	-	0.00	0.00	0.01	0.01	0.18	0.00	-
	Escherichia fergusonii ATCC35469	-	-	-	-	-	-	-	-
	Holdemania filiformis	-	-	-	0.01	0.01	-	0.00	0.00
	Ruminococcus gnavus	-	-	-	1.16	0.01	0.00	-	-
	Ruminococcus torques ATCC27756	-	-	-	0.03	-	-	-	-

		Family 4									Family 5	
		Co-twin 1			Co-twin 2		Mother				Co-twin 2	
	Sample	1	2	3	1	3	1	2	3	3(R)	1	1(R)
Bacterial Genomes	*Alistipes putredinis*	-	-	-	-	-	-	-	-	-	-	-
	Bacteroides caccae ATCC 43185	0.00	0.01	-	-	-	-	-	-	-	-	-
	Bacteroides coprocola	-	-	-	0.00	0.00	0.01	0.39	-	-	0.01	0.01
	Bacteroides sp.D2	0.00	0.00	-	0.00	-	-	0.01	0.04	0.03	-	-
	Bacteroides stercoris	0.30	0.21	0.90	-	0.01	-	0.01	-	-	0.36	0.33
	Bacteroides thetaiotaomicron 3731	0.00	0.00	-	0.00	-	-	0.01	0.04	0.03	-	-
	Blautia hansenii	0.01	-	-	-	-	0.01	-	0.01	0.02	-	-
	Blautia hydrogenotrophica DSM10507	0.16	0.00	-	-	-	-	0.00	-	-	-	0.00
	Clostridium leptum	-	-	-	-	-	-	-	-	-	0.04	0.04
	Escherichia fergusonii ATCC35469	-	-	0.01	-	0.00	-	0.30	0.00	-	-	-
	Holdemania filiformis	-	-	-	0.00	-	0.00	-	-	-	0.00	-
	Ruminococcus gnavus	0.01	-	-	-	-	0.01	-	0.01	0.02	-	-
	Ruminococcus torques ATCC27756	-	-	-	0.01	0.02	-	-	-	0.01	0.07	0.09

Table S6b. Matrix of normalized percent coverage of prophage genomes by VLP sample reads (10,000 VLP reads used per sample)

		Family 1							
		Co-twin 1		Co-twin 2				Mother	
	Sample	1	3	1	1(R)	2	3	1	2
Bacterial Genomes	*Alistipes putredinis*	1.1%	81.3%	-	-	-	0.5%	-	-
	Bacteroides caccae ATCC 43185	-	-	1.1%	0.6%	0.3%	1.4%	-	-
	Bacteroides coprocola	-	-	-	-	-	-	-	-
	Bacteroides sp.D2	52.9%	-	-	-	-	-	-	-
	Bacteroides stercoris	-	53.1%	-	-	-	-	1.3%	0.9%
	Bacteroides thetaiotaomicron 3731	55.0%	-	-	-	-	-	-	-
	Blautia hansenii	-	2.8%	-	-	0.8%	-	-	-
	Blautia hydrogenotrophica DSM10507	-	-	-	-	-	0.8%	0.8%	2.1%
	Clostridium leptum	7.2%	36.9%	6.6%	9.9%	20.2%	20.0%	-	-
	Escherichia fergusonii ATCC35469	-	-	-	-	-	-	-	-
	Holdemania filiformis	-	25.2%	-	-	-	-	-	-
	Ruminococcus gnavus	-	2.8%	-	-	0.8%	-	-	-
	Ruminococcus torques ATCC27756	-	20.6%	-	0.6%	1.2%	0.6%	2.9%	0.6%

		Family 2										
		Co-twin 1				Co-twin 2			Mother			
	Sample	1	1(R)	2	3	1	1(R)	2	1	1(R)	2	3
Bacterial Genomes	*Alistipes putredinis*	-	-	0.5%	-	-	-	-	-	-	-	-
	Bacteroides caccae ATCC 43185	-	-	-	-	9.4%	9.2%	8.3%	4.5%	4.3%	-	0.3%
	Bacteroides coprocola	-	-	-	-	-	-	0.5%	27.4%	59.9%	94.8%	-
	Bacteroides sp.D2	-	-	-	-	-	-	-	-	-	-	-
	Bacteroides stercoris	-	0.8%	-	0.6%	-	-	0.4%	1.8%	2.5%	0.4%	0.5%
	Bacteroides thetaiotaomicron 3731	-	-	-	-	-	-	-	-	-	-	-
	Blautia hansenii	-	-	-	-	-	-	-	0.7%	-	-	-
	Blautia hydrogenotrophica DSM10507	0.4%	0.4%	0.5%	-	0.4%	-	0.4%	-	2.3%	5.2%	-
	Clostridium leptum	-	-	-	-	0.4%	-	-	2.1%	4.3%	2.1%	1.2%
	Escherichia fergusonii ATCC35469	-	-	-	-	-	-	-	-	-	-	3.2%
	Holdemania filiformis	-	-	-	-	4.9%	6.7%	0.4%	0.4%	-	-	-
	Ruminococcus gnavus	-	-	-	-	-	-	-	0.7%	-	-	-
	Ruminococcus torques ATCC27756	-	-	-	-	1.1%	1.1%	-	89.1%	90.7%	18.2%	46.7%

		Family 3							
		Co-twin 1			Co-twin 2			Mother	
	Sample	1	2	3	1	2	3	1	2
Bacterial Genomes	*Alistipes putredinis*	-	-	-	1.5%	0.5%	-	-	-
	Bacteroides caccae ATCC 43185	-	-	-	-	-	-	-	-
	Bacteroides coprocola	-	0.5%	-	1.8%	0.5%	2.2%	-	-
	Bacteroides sp.D2	-	-	-	46.6%	0.5%	0.6%	0.6%	85.3%
	Bacteroides stercoris	-	22.4%	3.3%	3.8%	-	-	2.9%	72.1%
	Bacteroides thetaiotaomicron 3731	-	-	-	55.0%	0.7%	0.6%	0.6%	93.1%
	Blautia hansenii	-	-	-	90.1%	2.9%	0.7%	-	-
	Blautia hydrogenotrophica DSM10507	-	-	-	-	-	-	26.2%	12.8%
	Clostridium leptum	-	0.4%	0.4%	2.3%	1.2%	21.6%	0.5%	-
	Escherichia fergusonii ATCC35469	-	-	-	-	-	-	-	-
	Holdemania filiformis	-	-	-	6.0%	2.2%	-	0.7%	0.7%
	Ruminococcus gnavus	-	-	-	90.1%	2.9%	0.7%	-	-
	Ruminococcus torques ATCC27756	-	-	-	7.2%	-	-	-	-

		Family 4									Family 5	
		Co-twin 1			Co-twin 2		Mother				Co-twin 2	
	Sample	1	2	3	1	3	1	2	3	3(R)	1	1(R)
Bacterial Genomes	*Alistipes putredinis*	-	-	-	-	-	-	-	-	-	-	-
	Bacteroides caccae ATCC 43185	0.3%	1.3%	-	-	-	-	-	-	-	-	-
	Bacteroides coprocola	-	-	-	1.0%	1.5%	2.1%	56.0%	-	-	2.9%	5.4%
	Bacteroides sp.D2	0.6%	0.6%	-	0.6%	-	-	1.2%	10.5%	3.7%	-	-
	Bacteroides stercoris	43.6%	33.2%	50.3%	-	1.5%	-	3.1%	-	-	66.9%	72.3%
	Bacteroides thetaiotaomicron 3731	0.6%	0.6%	-	0.6%	-	-	1.2%	11.5%	3.6%	-	-
	Blautia hansenii	1.4%	-	-	-	-	1.3%	-	3.8%	3.5%	-	-
	Blautia hydrogenotrophica DSM10507	31.4%	0.9%	-	-	-	-	0.4%	-	-	-	0.8%
	Clostridium leptum	-	-	-	-	-	-	-	-	-	8.6%	13.0%
	Escherichia fergusonii ATCC35469	-	-	0.5%	-	0.6%	-	27.9%	0.6%	-	-	-
	Holdemania filiformis	-	-	-	1.0%	-	0.4%	-	-	-	0.2%	-
	Ruminococcus gnavus	1.4%	-	-	-	-	1.3%	-	3.8%	3.5%	-	-
	Ruminococcus torques ATCC27756	-	-	-	1.2%	2.3%	-	-	-	0.5%	13.8%	27.1%

Table S7a. Matrix of normalized fold-coverage per basepair of 88 assembled large contigs by VLP sample reads (14,000 reads used per VLP sample)

	Family 1								Family 2											Family 3								Family 4									Family 5	
	Co-twin 1		Co-twin 2				Mother		Co-twin 1				Co-twin 2			Mother				Co-twin 1			Co-twin 2			Mother		Co-twin 1			Co-twin 2		Mother				Co-twin 2	
Sample	1	3	1	1(R)	2	3	1	2	1	1(R)	2	3	1	1(R)	2	1	1(R)	2	3	1	2	3	1	2	3	1	2	1	2	3	1	3	1	2	3	3(R)	1	1(R)
Contig_1292	0.83	0.06	-	-	-	0.07	5.00	-	-	-	-	-	-	-	-	0.02	3.00	0.04	0.05	-	-	-	1.85	0.07	0.64	0.45	0.13	-	-	-	-	0.12	-	-	-	-	0.03	0.07
Contig_2729	-	10.76	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_507	-	5.34	-	-	-	-	-	-	-	-	-	-	-	-	-	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.17	0.28	-	-
Contig_1331	-	5.97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	-	-	-	-	-	7.00	0.04	-	-	-	-	-	-
Contig_1935	0.19	4.38	0.02	-	-	0.03	-	-	-	-	-	-	-	-	-	3.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.01
Contig_2989	-	3.95	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3580	-	3.47	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_255	-	-	8.72	4.26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9.00	-	-	-	-	-	-	-	7.41	7.94	-	-
Contig_20922	-	-	3.42	1.72	6.39	3.37	0.05	6.00	8.31	6.75	4.09	12.90	0.36	0.18	0.44	0.06	0.05	0.03	0.28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	0.03	-	-
Contig_1233	0.02	-	4.20	1.77	5.28	2.57	0.07	-	9.63	10.20	5.01	15.45	0.25	0.22	0.49	0.05	0.05	-	0.37	-	-	-	-	-	-	-	-	-	-	-	0.02	-	-	-	0.03	0.06	-	-
Contig_396	-	-	3.55	1.69	3.43	0.16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.05	-	-	-	-	-	-	-	0.01	-	-	-
Contig_389	0.01	-	3.25	1.59	3.09	0.11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_384	-	-	3.42	1.60	2.26	0.03	-	8.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_381	0.01	-	3.33	1.46	1.88	0.09	5.00	-	-	-	-	-	-	-	-	0.02	7.00	0.03	0.05	-	-	-	0.02	-	0.02	0.11	4.00	-	-	-	-	0.09	-	-	7.00	-	-	-
Contig_378	-	-	2.50	1.38	1.18	0.02	-	3.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_878	0.56	0.03	0.07	-	0.02	10.68	0.01	-	-	-	-	-	-	-	-	0.14	0.06	0.35	0.34	-	-	-	0.03	0.01	0.10	0.75	0.52	-	-	-	-	0.91	-	-	-	-	0.05	-
Contig_2660	-	-	-	-	-	6.21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_424	0.05	-	-	-	-	-	5.08	0.03	4.00	8.00	-	-	-	-	-	0.01	6.00	-	-	-	-	-	-	-	0.03	-	-	-	-	-	-	-	-	-	4.17	4.42	-	-
Contig_2194	0.10	-	7.00	7.00	8.00	0.02	4.54	0.03	0.07	0.06	0.03	0.09	-	-	3.00	0.01	4.00	-	5.00	-	-	-	-	-	0.03	-	-	-	-	-	-	-	-	-	3.71	3.37	-	-
Contig_4764	-	-	-	-	-	-	2.26	0.19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_4022	-	-	-	-	-	-	2.98	0.07	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20743	-	-	-	-	-	-	2.77	0.06	3.00	0.04	-	0.15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.31	0.07	4.95	-	-	3.90	0.01	-	-	4.00	-
Contig_3075	-	-	-	-	-	-	2.02	0.04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3756	-	-	-	-	-	-	2.14	0.06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_22692	-	-	-	-	-	-	-	-	-	-	8.76	-	-	-	3.75	-	-	-	-	-	-	-	-	-	-	-	8.00	5.00	-	-	-	-	-	-	-	-	-	-
Contig_20560	-	-	-	-	-	-	-	-	-	-	-	-	8.16	8.82	2.91	-	8.00	-	0.02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20939	-	0.44	-	-	-	-	-	-	-	-	-	-	1.10	1.61	0.45	-	0.03	-	0.04	-	-	-	-	-	-	-	0.04	0.39	0.05	-	0.09	0.04	-	-	0.02	-	-	-
Contig_2613	0.02	-	-	-	-	0.02	-	-	-	-	0.14	-	1.05	1.36	0.62	-	-	-	-	-	-	-	0.09	0.06	-	-	-	-	-	0.09	-	-	-	-	-	-	-	-
Contig_21357	-	-	-	-	4.00	0.01	-	-	-	-	-	-	6.61	5.62	2.02	8.00	4.00	-	-	-	-	-	7.00	-	-	-	0.03	0.04	0.02	4.00	0.39	0.02	-	-	-	-	-	-
Contig_2821	-	0.02	-	-	-	-	-	-	0.04	0.02	0.10	-	6.31	6.81	0.53	8.00	-	-	-	-	-	-	-	0.01	-	-	0.45	0.28	-	0.05	0.13	0.07	0.02	0.02	0.03	0.03	-	0.02
Contig_21166	-	-	-	-	-	-	-	-	-	-	-	-	7.04	6.10	2.16	6.00	6.00	-	-	-	3.00	-	0.24	0.03	-	0.02	4.00	-	-	-	0.23	-	-	-	-	-	-	-
Contig_2825	-	0.19	-	-	0.04	-	-	-	0.02	0.02	0.09	-	5.37	5.11	0.42	0.04	0.05	-	0.01	-	-	0.01	0.03	-	0.01	-	0.44	0.50	0.05	0.02	0.18	0.08	-	-	0.01	8.00	0.03	0.03
Contig_2618	7.00	0.06	-	-	0.04	-	-	-	0.03	0.04	-	-	3.19	3.92	1.27	0.01	-	-	-	-	-	-	3.00	-	-	-	-	6.00	-	-	-	-	-	-	-	-	5.00	0.01
Contig_21891	5.00	0.04	-	-	0.05	-	-	-	-	-	-	-	3.04	3.38	1.00	-	-	-	6.00	-	-	-	5.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02
Contig_3634	-	-	-	-	-	-	-	-	-	-	-	-	2.89	2.28	0.70	-	-	-	-	-	0.27	-	0.04	0.02	-	-	-	-	0.01	-	-	-	-	-	-	-	-	-
Contig_1191	-	9.00	-	-	0.02	-	-	-	0.11	0.06	0.26	-	2.27	1.84	0.46	0.02	0.03	0.03	-	-	-	-	5.00	0.07	-	-	-	0.01	-	-	-	0.01	-	-	-	-	-	-
Contig_2570	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9.02	4.43	-	0.02	-	-	0.04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_228	0.01	0.11	-	-	0.05	0.05	-	-	-	-	-	-	0.03	0.02	-	3.58	3.19	0.08	0.53	-	-	-	0.09	-	5.00	-	-	-	-	-	0.03	0.03	-	0.02	9.00	-	0.07	0.10
Contig_3228	-	0.15	-	-	-	-	-	-	-	-	3.00	-	0.03	0.02	-	1.28	1.51	0.20	0.09	-	-	-	0.03	0.03	-	-	6.00	3.00	-	-	-	-	-	4.00	-	-	-	-
Contig_19787	-	0.04	-	-	-	0.10	-	-	-	-	-	-	0.08	-	-	2.81	2.59	0.06	0.31	-	-	-	0.11	-	0.01	-	-	-	-	-	0.07	0.07	-	-	0.01	-	0.03	0.08
Contig_2285	0.03	0.03	-	9.00	0.06	0.04	0.02	-	-	-	-	-	7.00	0.02	-	1.62	1.68	0.33	1.31	-	-	-	0.02	-	0.03	-	-	-	-	-	-	-	-	0.05	-	-	-	0.01
Contig_2984	-	-	0.04	0.04	0.06	0.03	-	-	0.04	0.04	0.02	0.19	-	-	-	1.80	1.28	-	0.14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3236	-	-	0.02	-	-	0.05	-	-	-	-	-	0.01	0.02	0.02	3.00	2.19	1.33	0.04	0.05	-	-	-	-	-	-	-	-	0.15	0.10	0.04	6.00	0.05	-	-	-	0.01	9.00	-
Contig_3645	-	-	-	-	-	-	0.02	-	-	-	-	-	-	-	-	1.44	1.11	0.03	0.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1218	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23	0.32	7.50	-	-	-	-	-	0.02	-	-	-	-	-	-	-	0.03	-	0.73	-	-	0.02	-
Contig_22750	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	0.11	0.22	5.88	-	-	-	-	-	-	0.02	-	-	-	-	-	0.02	-	0.07	0.47	-	-	-	0.02
Contig_19925	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.15	0.57	2.02	0.43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1299	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.04	0.04	-	6.69	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_22223	5.00	0.02	-	-	0.01	0.02	-	-	-	-	-	-	0.01	-	-	1.26	1.12	0.22	1.62	-	-	-	0.05	-	0.02	-	-	-	-	-	8.00	0.01	-	0.10	0.03	0.02	5.00	-
Contig_9891	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.77	0.72	-	0.08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_291	0.26	0.71	0.03	-	0.03	0.05	-	-	-	-	0.02	-	-	-	0.02	-	-	0.03	0.08	-	-	-	1.09	7.13	0.33	0.77	0.02	0.91	0.05	0.03	0.02	-	0.06	-	0.17	0.09	0.06	0.01
Contig_2090	0.02	-	-	-	-	-	0.07	0.14	0.09	0.08	0.09	0.07	7.00	5.00	0.05	9.00	0.01	0.01	0.02	0.13	0.12	0.10	0.72	16.66	0.03	4.00	-	3.00	-	-	0.05	4.00	0.02	0.03	6.00	5.00	-	-
Contig_2087	0.03	-	-	-	-	-	0.04	0.07	-	-	-	-	-	-	-	5.00	7.00	-	0.01	0.06	0.04	0.08	0.39	15.42	3.00	-	-	-	-	-	0.03	-	5.00	0.03	5.00	5.00	-	-
Contig_2061	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.35	4.89	0.23	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2058	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.25	4.86	0.18	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2244	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.03	0.05	3.48	4.18	0.21	0.88	3.90	-	-	-	-	-	0.02	-	-	-	-	-
Contig_20697	0.21	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	0.03	5.00	0.08	0.08	-	-	0.01	6.67	0.47	4.54	1.24	0.03	-	-	-	-	0.19	-	-	-	-	-	0.03
Contig_2271	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.31	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1567	0.14	0.01	-	-	-	0.28	0.02	-	-	-	-	0.20	-	-	-	1.10	0.82	3.93	3.57	-	-	-	3.89	0.05	0.61	22.64	1.79	-	-	-	0.15	6.38	0.10	0.12	0.07	0.07	0.02	0.04
Contig_1553	7.00	-	-	-	-	0.11	-	-	-	-	-	0.05	-	-	-	0.03	0.04	0.14	0.21	-	-	-	0.15	0.05	0.36	24.24	0.52	-	-	-	0.02	0.47	-	-	-	-	-	-
Contig_1550	0.04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.87	0.61	2.72	2.95	-	-	-	0.42	0.02	0.06	23.67	0.22	-	-	-	0.19	6.95	-	-	-	-	-	-
Contig_329	-	-	0.02	5.00	0.03	-	-	-	-	-	-	0.01	-	-	-	0.12	0.13	0.25	0.33	-	-	-	0.12	4.00	0.05	21.15	0.17	-	-	-	0.01	0.81	-	-	-	-	-	5.00
Contig_2750	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.19	-	-	-	-	-	-	-	-	0.03	6.25	-	0.08	-	-	-	-	-	-	-	-	-	-
Contig_19814	5.00	-	-	-	-	0.25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.09	-	-	1.58	-	16.71	0.02	-	-	-	-	-	-	-	-	-	-
Contig_22976	-	-	-	-	-	-	-	-	0.02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	21.51	20.22	-	0.02	-	-	-	-	-	0.05	0.06
Contig_406	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.10	4.69	-	-	-	-	-	-	-	9.00	9.00
Contig_22975	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9.70	6.56	-	-	-	-	-	-	-	0.01	0.03
Contig_405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.07	3.84	-	-	-	-	7.00	-	-	4.00	8.00
Contig_423	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.00	-	14.70	7.62	-	-	-	-	0.03	-	-	0.05	0.01
Contig_22716	-	-	-	-	-	-	2.08	0.06	0.02	0.06	-	0.17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.49	0.09	7.12	0.02	-	5.93	0.02	-	-	-	-
Contig_1603	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.00	-	-	-	3.38	2.71	-	-	-	0.02	0.02	-	-	7.00	-
Contig_3397	9.00	-	-	-	-	0.03	-	-	-	-	-	-	0.02	0.02	-	0.03	0.01	-	-	-	-	-	9.00	-	-	-	0.03	0.01	0.01	0.03	8.80	0.19	-	-	-	-	-	-
Contig_2803	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.00	0.04	-	14.05	-	-	-	-	-	0.01	-
Contig_3434	-	-	-	-	-	-	-	-	-	-	-	-	0.15	0.15	0.03	3.00	-	-	-	-	-	-	4.00	-	-	0.16	-	-	0.01	-	6.55	0.84	-	-	-	-	-	-
Contig_3431	-	-	-	-	-	-	-	-	-	-	-	-	0.12	0.19	0.10	0.03	6.00	0.01	-	-	-	-	0.06	0.02	-	0.03	-	-	-	-	7.34	0.28	-	-	-	-	-	-
Contig_368	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	-	4.20	0.59	-	-	-	-	-	-
Contig_23558	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.00	-	-	-	-	-	-	-	8.31	0.90	-	-	-	-	-	-
Contig_2788	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.00	0.01	-	-	-	4.75	4.51	-	-	-	-
Contig_2786	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9.00	-	-	-	-	9.14	9.68	-	-	-	-
Contig_22597	4.00	-	-	-	-	0.04	-	-	-	-	-	4.00	-	-	-	0.02	8.00	0.06	0.04	-	-	-	0.08	-	0.03	0.13	-	0.02	-	-	-	0.22	9.56	11.02	-	-	-	-
Contig_251	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8.22	7.72	-	-
Contig_254	-	-	6.56	3.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10.84	11.76	-	-
Contig_1978	-	0.16	-	-	-	-	-	-	-	-	-	-	0.06	0.10	-	-	-	-	-	-	-	0.20	-	-	0.02	-	-	-	-	-	-	-	0.07	-	-	-	1.70	1.68
Contig_61	-	0.14	-	-	-	0.02	0.36	3.00	0.12	0.10	-	7.00	2.00	0.01	4.00	0.28	0.23	0.12	0.21	-	-	-	-	-	-	-	-	-	2.00	0.45	0.11	0.02	7.00	-	0.33	0.14	0.82	0.86
Contig_20385	-	-	-	-	-	-	-	-	-	-	-	-	0.14	0.08	0.29	-	-	-	-	-	-	-	0.03	0.03	-	-	-	-	-	-	-	-	-	-	-	-	0.81	0.63
Contig_19861	0.06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.43	0.62
Contig_20734	0.03	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.48	0.52
Contig_3218	-	0.04	-	-	-	-	-	-	0.02	-	0.02	0.02	0.02	0.02	4.00	-	9.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	-	-	-	0.20	0.51

Table S7b. Matrix of normalized percent coverage of 88 large viral contigs by VLP sample reads (14,000 reads used per VLP sample)

	Family 1								Family 2											Family 3								Family 4									Family 5	
	Co-twin 1		Co-twin 2				Mother		Co-twin 1				Co-twin 2			Mother				Co-twin 1			Co-twin 2			Mother		Co-twin 1			Co-twin 2		Mother				Co-twin 2	
Sample	1	3	1	1(R)	2	3	1	2	1	1(R)	2	3	1	1(R)	2	1	1(R)	2	3	1	2	3	1	2	3	1	2	1	2	3	1	3	1	2	3	3(R)	1	1(R)
Contig_1292	56.3%	5.9%	-	-	-	0.3%	0.5%	-	-	-	-	-	-	-	-	0.9%	0.3%	0.9%	0.9%	-	-	-	17.4%	6.0%	20.5%	3.5%	1.4%	-	-	-	-	0.9%	-	-	-	-	3.0%	7.1%
Contig_2729	-	100.0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_507	-	99.0%	-	-	-	-	-	-	-	-	-	-	-	-	-	1.0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	14.9%	23.2%	-	-
Contig_1331	-	98.6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.3%	-	-	-	-	-	0.7%	2.3%	-	-	-	-	-	-
Contig_1935	15.2%	97.2%	1.7%	-	-	3.0%	-	-	-	-	-	-	-	-	-	0.3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.4%
Contig_2989	-	96.9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3580	-	92.1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_255	-	-	100.0%	98.2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5%	-	-	-	-	-	-	-	76.0%	76.6%	-	-
Contig_20922	-	-	95.2%	85.0%	98.2%	95.0%	4.4%	0.6%	76.1%	74.5%	74.6%	75.8%	25.5%	16.7%	31.7%	5.7%	4.4%	2.7%	23.7%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.6%	1.3%	-	-
Contig_1233	1.6%	-	98.8%	84.9%	98.2%	91.8%	6.6%	-	91.0%	91.1%	91.5%	92.0%	22.1%	18.4%	41.1%	4.2%	4.6%	-	31.4%	-	-	-	-	-	-	-	-	-	-	-	1.7%	-	-	-	0.5%	0.5%	-	-
Contig_396	-	-	98.0%	82.4%	94.6%	15.1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.4%	-	-	-	-	-	-	-	1.1%	-	-	-
Contig_389	1.3%	-	94.2%	79.3%	89.7%	10.3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_384	-	-	98.3%	81.3%	87.3%	2.9%	-	0.8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_381	1.3%	-	95.6%	76.3%	83.3%	3.5%	0.5%	-	-	-	-	-	-	-	-	0.7%	0.3%	0.7%	0.7%	-	-	-	1.2%	-	1.2%	1.6%	0.4%	-	-	-	-	0.8%	-	-	0.7%	-	-	-
Contig_378	-	-	89.7%	74.2%	68.1%	1.9%	-	0.3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_878	42.9%	2.5%	2.6%	-	1.5%	100.0%	1.0%	-	-	-	-	-	-	-	-	6.7%	3.6%	7.9%	7.9%	-	-	-	2.0%	1.2%	6.6%	6.9%	5.5%	-	-	-	-	8.0%	-	-	-	-	4.8%	-
Contig_2660	-	-	-	-	-	98.0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_424	4.6%	-	-	-	-	-	100.0%	2.7%	0.4%	0.4%	-	-	-	-	-	0.7%	0.6%	-	-	-	-	-	-	-	3.2%	-	-	-	-	-	-	-	-	-	32.5%	32.1%	-	-
Contig_2194	8.2%	-	0.5%	0.3%	0.5%	0.5%	98.4%	3.0%	0.8%	0.8%	0.8%	0.8%	-	-	0.3%	1.1%	0.4%	-	0.3%	-	-	-	-	-	2.5%	-	-	-	-	-	-	-	-	-	30.9%	30.3%	-	-
Contig_4764	-	-	-	-	-	-	93.8%	17.6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_4022	-	-	-	-	-	-	90.4%	6.9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20743	-	-	-	-	-	-	89.7%	5.8%	0.3%	3.6%	-	12.9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	24.0%	6.6%	77.8%	-	-	77.6%	1.4%	-	-	0.4%	-
Contig_3075	-	-	-	-	-	-	88.8%	3.8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3756	-	-	-	-	-	-	88.4%	5.6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_22692	-	-	-	-	-	-	-	-	-	-	100.0%	-	-	-	95.9%	-	-	-	-	-	-	-	-	-	-	-	0.7%	0.5%	-	-	-	-	-	-	-	-	-	-
Contig_20560	-	-	-	-	-	-	-	-	-	-	-	-	100.0%	100.0%	93.9%	-	0.8%	-	1.8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20939	-	33.5%	-	-	-	-	-	-	-	-	-	-	63.9%	80.6%	34.5%	-	2.5%	-	4.4%	-	-	-	-	-	-	-	3.8%	27.5%	4.9%	-	8.6%	4.4%	-	-	2.3%	-	-	-
Contig_2613	2.1%	-	-	-	-	2.1%	-	-	-	-	12.4%	-	51.9%	65.5%	39.9%	-	-	-	-	-	-	-	7.2%	5.7%	-	-	-	-	-	0.8%	-	-	-	-	-	-	-	-
Contig_21357	-	-	-	-	0.4%	0.9%	-	-	-	-	-	-	98.9%	99.8%	83.0%	0.8%	0.4%	-	-	-	-	-	0.7%	-	-	-	3.0%	0.6%	2.3%	0.4%	9.5%	1.7%	-	-	-	-	-	-
Contig_2821	-	1.6%	-	-	-	-	-	-	3.7%	2.4%	9.5%	-	99.7%	99.6%	44.3%	0.8%	-	-	-	-	-	-	-	1.2%	-	-	33.9%	21.6%	-	5.2%	12.0%	6.5%	2.3%	2.4%	1.7%	3.4%	-	2.1%
Contig_21166	-	-	-	-	-	-	-	-	-	-	-	-	99.6%	99.5%	83.2%	0.6%	0.6%	-	-	-	0.3%	-	6.5%	2.5%	-	1.2%	0.4%	-	-	-	4.3%	-	-	-	-	-	-	-
Contig_2825	-	7.7%	-	-	2.7%	-	-	-	1.5%	1.6%	7.9%	-	99.2%	99.5%	34.5%	3.6%	4.3%	-	1.3%	-	-	1.4%	3.1%	-	1.2%	-	33.5%	38.5%	4.7%	1.9%	14.3%	8.2%	-	-	1.2%	0.8%	3.1%	2.7%
Contig_2618	0.7%	5.4%	-	-	4.0%	-	-	-	3.2%	4.2%	-	-	91.3%	97.4%	73.1%	1.3%	-	-	-	-	-	-	0.3%	-	-	-	-	0.3%	-	-	-	-	-	-	-	-	0.5%	1.1%
Contig_21891	0.5%	2.2%	-	-	3.1%	-	-	-	-	-	-	-	93.8%	91.6%	59.8%	-	-	-	0.6%	-	-	-	0.5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.8%
Contig_3634	-	-	-	-	-	-	-	-	-	-	-	-	90.2%	88.2%	52.2%	-	-	-	-	-	26.2%	-	3.7%	1.5%	-	-	-	-	0.6%	-	-	-	-	-	-	-	-	-
Contig_1191	-	0.9%	-	-	1.5%	-	-	-	7.9%	6.4%	22.3%	-	87.0%	86.4%	33.8%	1.6%	2.8%	2.5%	-	-	-	-	0.5%	5.0%	-	-	-	1.1%	-	-	-	1.2%	-	-	-	-	-	-
Contig_2570	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100.0%	97.8%	-	2.0%	-	-	3.8%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_228	1.0%	9.5%	-	-	2.5%	4.7%	-	-	-	-	-	-	2.8%	1.7%	-	94.5%	89.6%	8.1%	35.2%	-	-	-	8.1%	-	0.5%	-	-	-	-	-	3.4%	3.3%	-	1.5%	0.9%	-	7.4%	9.9%
Contig_3228	-	1.3%	-	-	-	-	-	-	-	-	0.3%	-	1.0%	1.1%	-	69.6%	78.2%	17.9%	8.7%	-	-	-	2.8%	3.2%	-	-	0.3%	0.3%	-	-	-	-	-	0.4%	-	-	-	-
Contig_19787	-	4.2%	-	-	-	4.0%	-	-	-	-	-	-	8.1%	-	-	74.4%	75.0%	6.1%	23.6%	-	-	-	10.6%	-	1.3%	-	-	-	-	-	6.5%	6.5%	-	-	1.2%	-	3.2%	7.7%
Contig_2285	3.1%	2.9%	-	0.9%	5.6%	4.0%	1.8%	-	-	-	-	-	0.7%	1.8%	-	70.5%	71.6%	30.1%	69.5%	-	-	-	2.2%	-	2.5%	-	-	-	-	-	-	-	-	5.1%	-	-	-	1.1%
Contig_2984	-	-	1.5%	1.5%	1.5%	1.5%	-	-	1.4%	1.4%	1.4%	1.4%	-	-	-	86.4%	69.8%	-	13.7%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3236	-	-	1.8%	-	-	3.8%	-	-	-	-	-	1.3%	0.3%	0.3%	0.3%	85.1%	69.2%	4.0%	4.9%	-	-	-	-	-	-	-	-	2.3%	4.2%	2.1%	0.6%	2.3%	-	-	-	1.0%	0.9%	-
Contig_3645	-	-	-	-	-	-	1.5%	-	-	-	-	-	-	-	-	76.5%	68.6%	2.7%	31.6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1218	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	22.2%	31.3%	100.0%	-	-	-	-	-	1.7%	-	-	-	-	-	-	-	2.5%	-	46.2%	-	-	2.2%	-
Contig_22750	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.7%	11.4%	19.0%	100.0%	-	-	-	-	-	-	2.0%	-	-	-	-	-	2.4%	-	5.6%	36.9%	-	-	-	2.2%
Contig_19925	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	66.7%	40.7%	88.0%	35.6%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1299	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.2%	4.0%	-	99.9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_22223	0.5%	2.4%	-	-	1.0%	1.6%	-	-	-	-	-	-	1.1%	-	-	68.6%	65.8%	21.0%	78.5%	-	-	-	4.8%	-	2.0%	-	-	-	-	-	0.8%	1.2%	-	9.7%	3.0%	2.4%	0.5%	-
Contig_9891	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	83.9%	50.7%	-	8.2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_291	24.1%	50.2%	2.8%	-	2.5%	4.8%	-	-	-	-	1.9%	-	-	-	2.1%	-	-	3.3%	7.5%	-	-	-	66.5%	100.0%	28.3%	44.6%	2.4%	60.9%	4.5%	2.5%	1.9%	-	6.1%	-	15.0%	9.3%	6.2%	1.3%
Contig_2090	1.8%	-	-	-	-	-	0.4%	0.4%	0.4%	0.4%	0.4%	1.0%	0.4%	0.3%	0.4%	0.3%	0.7%	0.3%	0.3%	0.5%	0.5%	0.7%	30.7%	100.0%	0.4%	0.3%	-	0.3%	-	-	0.4%	0.3%	0.4%	0.4%	0.1%	0.3%	-	-
Contig_2087	2.6%	-	-	-	-	-	0.3%	0.3%	-	-	-	-	-	-	-	0.3%	0.3%	-	0.3%	0.3%	0.3%	0.3%	20.9%	100.0%	0.3%	-	-	-	-	-	0.3%	-	0.3%	0.3%	0.3%	0.3%	-	-
Contig_2061	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	74.0%	98.5%	18.5%	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2058	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	70.3%	96.0%	17.4%	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2244	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.4%	5.1%	97.3%	95.8%	19.8%	48.0%	91.4%	-	-	-	-	-	1.8%	-	-	-	-	-
Contig_20697	15.8%	1.3%	-	-	-	-	-	-	-	-	-	-	-	-	-	1.5%	0.5%	1.5%	1.5%	-	-	1.3%	86.2%	37.4%	99.9%	10.0%	2.5%	-	-	-	-	1.5%	-	-	-	-	-	2.5%
Contig_2271	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	98.3%	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1567	10.7%	1.1%	-	-	-	2.7%	1.9%	-	-	-	-	12.3%	-	-	-	28.2%	26.7%	30.1%	30.1%	-	-	-	20.2%	5.0%	26.3%	100.0%	34.7%	-	-	-	12.2%	30.3%	3.0%	1.3%	1.4%	0.5%	1.6%	4.1%
Contig_1553	0.7%	-	-	-	-	0.9%	-	-	-	-	-	3.4%	-	-	-	1.8%	1.8%	3.3%	3.3%	-	-	-	7.3%	4.0%	12.8%	100.0%	20.2%	-	-	-	1.8%	3.3%	-	-	-	-	-	-
Contig_1550	2.5%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	23.2%	19.8%	23.7%	24.0%	-	-	-	4.9%	2.1%	3.4%	100.0%	18.1%	-	-	-	12.2%	24.2%	-	-	-	-	-	-
Contig_329	-	-	0.5%	0.5%	0.5%	-	-	-	-	-	-	0.7%	-	-	-	5.5%	4.7%	5.1%	5.3%	-	-	-	1.7%	0.4%	3.2%	100.0%	15.4%	-	-	-	1.2%	5.5%	-	-	-	-	-	0.5%
Contig_2750	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.8%	-	-	-	-	-	-	-	-	0.7%	100.0%	-	0.7%	-	-	-	-	-	-	-	-	-	-
Contig_19814	0.5%	-	-	-	-	3.0%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8.8%	-	-	74.8%	-	100.0%	1.5%	-	-	-	-	-	-	-	-	-	-
Contig_22976	-	-	-	-	-	-	-	-	2.1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100.0%	100.0%	-	2.2%	-	-	-	-	-	4.5%	5.8%
Contig_406	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.9%	99.6%	-	-	-	-	-	-	-	0.9%	0.9%
Contig_22975	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100.0%	98.9%	-	-	-	-	-	-	-	1.3%	2.6%
Contig_405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	98.2%	97.5%	-	-	-	-	0.7%	-	-	0.4%	0.8%
Contig_423	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.7%	-	99.3%	94.6%	-	-	-	-	2.7%	-	-	4.7%	1.4%
Contig_22716	-	-	-	-	-	-	77.6%	6.4%	1.8%	6.0%	-	15.1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	37.7%	9.4%	99.1%	1.9%	-	99.3%	1.6%	-	-	-	-
Contig_1603	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.4%	-	-	-	95.8%	91.3%	-	-	-	0.3%	0.6%	-	-	0.7%	-
Contig_3397	0.9%	-	-	-	-	1.9%	-	-	-	-	-	-	1.7%	1.7%	-	2.6%	1.1%	-	-	-	-	-	0.9%	-	-	-	1.7%	1.4%	1.1%	3.1%	100.0%	17.9%	-	-	-	-	-	-
Contig_2803	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.6%	3.9%	-	100.0%	-	-	-	-	-	1.0%	-
Contig_3434	-	-	-	-	-	-	-	-	-	-	-	-	5.0%	4.8%	2.1%	0.3%	-	-	-	-	-	-	0.4%	-	-	0.8%	-	-	1.4%	-	99.9%	22.8%	-	-	-	-	-	-
Contig_3431	-	-	-	-	-	-	-	-	-	-	-	-	6.6%	6.1%	3.7%	2.0%	0.6%	1.4%	-	-	-	-	3.0%	1.9%	-	1.9%	-	-	-	-	99.3%	23.3%	-	-	-	-	-	-
Contig_368	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.5%	-	98.6%	45.0%	-	-	-	-	-	-
Contig_23558	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.3%	-	-	-	-	-	-	-	99.3%	58.8%	-	-	-	-	-	-
Contig_2788	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.3%	0.7%	-	-	-	99.6%	99.2%	-	-	-	-
Contig_2786	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.9%	-	-	-	-	99.5%	100.0%	-	-	-	-
Contig_22597	0.4%	-	-	-	-	0.4%	-	-	-	-	-	0.4%	-	-	-	0.4%	0.4%	0.4%	0.4%	-	-	-	0.3%	-	1.4%	0.8%	-	2.3%	-	-	-	1.4%	99.9%	100.0%	-	-	-	-
Contig_251	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.3%	98.0%	-	-
Contig_254	-	-	74.7%	69.3%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	98.0%	100.0%	-	-
Contig_1978	-	15.0%	-	-	-	-	-	-	-	-	-	-	5.6%	10.3%	-	-	-	-	-	-	-	15.3%	-	-	2.1%	-	-	-	-	-	-	-	2.3%	-	-	-	82.4%	78.6%
Contig_61	-	13.5%	-	-	-	1.5%	22.7%	0.3%	7.9%	6.9%	-	0.7%	0.2%	1.0%	0.4%	13.2%	12.3%	8.6%	11.6%	-	-	-	-	-	-	-	-	-	0.2%	26.9%	10.0%	1.8%	0.7%	-	27.1%	12.6%	52.4%	56.8%
Contig_20385	-	-	-	-	-	-	-	-	-	-	-	-	2.3%	2.3%	2.3%	-	-	-	-	-	-	-	0.8%	0.8%	-	-	-	-	-	-	-	-	-	-	-	-	51.0%	45.7%
Contig_19861	5.9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	35.5%	46.0%
Contig_20734	3.1%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	38.0%	38.1%
Contig_3218	-	4.2%	-	-	-	-	-	-	1.8%	-	2.0%	2.0%	0.4%	0.4%	0.4%	-	0.9%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.9%	-	-	-	18.3%	38.3%

Table S8. Percent similarity between reads in VLP samples and 88 large viral contigs

	Family 1							Family 2								Family 3								Family 4							
	Co-twin 1		Co-twin 2			Mother		Co-twin 1			Co-twin 2		Mother			Co-twin 1			Co-twin 2			Mother		Co-twin 1			Co-twin 2		Mother		
Sample	1	3	1	2	3	1	2	1	2	3	1	2	1	2	3	1	2	3	1	2	3	1	2	1	2	3	1	3	1	2	3
Contig_255	-	-	99.75	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	89.50
Contig_381	-	-	99.81	99.78	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_389	-	-	99.85	99.82	99.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_396	-	-	99.77	99.77	99.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_378	-	-	99.81	99.79	99.66	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_384	-	-	99.86	99.80	99.93	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1233	-	-	99.82	99.82	99.83	-	-	88.16	87.96	88.03	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20922	-	-	99.83	99.81	99.83	-	-	88.09	88.06	88.05	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20743	-	-	-	-	-	99.84	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	92.87	-	-	-	-	-
Contig_424	-	-	-	-	-	99.79	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	86.32
Contig_2194	-	-	-	-	-	99.84	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	86.75
Contig_22692	-	-	-	-	-	-	-	-	99.86	-	-	99.88	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_21166	-	-	-	-	-	-	-	-	-	-	99.67	98.95	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2613	-	-	-	-	-	-	-	-	-	-	98.02	99.04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_21891	-	-	-	-	-	-	-	-	-	-	99.41	99.29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_21357	-	-	-	-	-	-	-	-	-	-	99.80	99.69	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_20560	-	-	-	-	-	-	-	-	-	-	99.70	99.74	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2618	-	-	-	-	-	-	-	-	-	-	99.84	99.77	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_3634	-	-	-	-	-	-	-	-	-	-	99.48	99.80	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1567	-	-	-	-	-	-	-	-	-	-	-	-	86.67	86.58	86.54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_1550	-	-	-	-	-	-	-	-	-	-	-	-	87.81	87.99	87.64	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_19925	-	-	-	-	-	-	-	-	-	-	-	-	99.82	99.79	99.78	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2285	-	-	-	-	-	-	-	-	-	-	-	-	99.93	100.00	99.88	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_228	-	-	-	-	-	-	-	-	-	-	-	-	99.95	-	99.94	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_22223	-	-	-	-	-	-	-	-	-	-	-	-	99.92	99.74	99.95	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Contig_2090	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.88	99.85	-	-	-	-	-	-	-	-	-	-	-
Contig_2244	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.82	99.83	99.82	-	-	-	-	-	-	-	-	-	-
Contig_291	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.82	99.81	99.84	-	-	-	-	-	-	-	-	-	-
Contig_2061	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.83	99.85	99.85	-	-	-	-	-	-	-	-	-	-
Contig_20697	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.78	99.88	99.87	-	-	-	-	-	-	-	-	-	-
Contig_2058	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.70	99.81	99.94	-	-	-	-	-	-	-	-	-	-
Contig_1567	-	-	-	-	-	-	-	-	-	-	-	-	-	86.58	86.54	-	-	-	87.30	-	-	99.81	-	-	-	-	-	86.71	-	-	-
Contig_1550	-	-	-	-	-	-	-	-	-	-	-	-	-	87.99	87.64	-	-	-	-	-	-	99.83	-	-	-	-	-	87.89	-	-	-
Contig_1603	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.76	99.73	-	-	-	-	-	-
Contig_423	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.71	99.73	-	-	-	-	-	-
Contig_22976	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.73	99.74	-	-	-	-	-	-
Contig_405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.77	99.78	-	-	-	-	-	-
Contig_406	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.78	99.78	-	-	-	-	-	-
Contig_22975	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.77	99.81	-	-	-	-	-	-
Contig_2788	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.83	99.78	-
Contig_22597	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.81	99.80	-
Contig_2786	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	99.83	99.80	-

Table S9. List of 121 reference human gut-associated microbial genomes

Species and strain designation	GenBank Accesion number	CRISPR Loci
Actinomyces odontolyticus	NZ_AAYI00000000	Present
Alistipes putredinis	NZ_ABFK00000000	Absent
Anaerococcus hydrogenalis	NZ_ABXA00000000	Present
Anaerofustis stercorihominis	NZ_ABIL00000000	Absent
Anaerostipes caccae	NZ_ABAX00000000	Present
Anaerotruncus colihominis	NZ_ABGD00000000	Absent
Bacteroides caccae ATCC 43185	NZ_AAVM00000000	Absent
Bacteroides capillosus	NZ_AAXG00000000	Present
Bacteroides cellulosilyticus	NZ_ACCH00000000	
Bacteroides coprocola	NZ_ABIY00000000	Present
Bacteroides coprophilus	NZ_ACBW00000000	
Bacteroides dorei	NZ_ABWZ00000000	Present
Bacteroides eggerthii	NZ_ABVO00000000	Absent
Bacteroides finegoldii	NZ_ABXI00000000	Absent
Bacteroides fragilis 3 1 12	NZ_ABZX00000000	Present
Bacteroides fragilis NCTC 9343	NC_006347	
Bacteroides fragilis YCH46	NC_003228	
Bacteroides intestinalis	NZ_ABJL00000000	Absent
Bacteroides ovatus ATCC 8483	NZ_AAXF00000000	Absent
Bacteroides plebeius	NZ_ABQC00000000	Present
Bacteroides sp 1 1 6	NZ_ACIC00000000	
Bacteroides sp. D1	NZ_ACAB00000000	
Bacteroides sp. D2	NZ_ACGA00000000	
Bacteroides stercoris	NZ_ABFZ00000000	Absent
Bacteroides thetaiotaomicron 7330		
Bacteroides thetaiotaomicron VPI-5482	NC_004663	Absent
Bacteroides uniformis ATCC 8492	NZ_AAYH00000000	Absent
Bacteroides vulgatus ATCC 8482	NC_009614	Absent
Bacteroides WH2		
Bacteroides xylanisolvens XB1A	ProjectID:39177	Absent
Bifidobacterium adolescentis	NZ_AAXD00000000	Present
Bifidobacterium adolescentis ATCC 15703	NC_008618	Present
Bifidobacterium angulatum	NZ_ABYS00000000	Present
Bifidobacterium animalis HN019.fa	NZ_ABOT00000000	
Bifidobacterium animalis subsp. lactis AD011	NC_011740	
Bifidobacterium breve	NZ_ACCG00000000	Absent
Bifidobacterium catenulatum	NZ_ABXY00000000	Present
Bifidobacterium dentium	NZ_ABIX00000000	Present
Bifidobacterium gallicum	NZ_ABXB00000000	Present
Bifidobacterium longum DJO10A	NC_010816	Present
Bifidobacterium longum NCC2705	NC_004307	
Bifidobacterium pseudocatenulatum	NZ_ABXX00000000	Present
Blautia hansenii	NZ_ABYU00000000	Present
Blautia hydrogenotrophica DSM 10507	NZ_ACBZ00000000	Absent
Bryantella formatexigens DSM 14469	NZ_ACCL00000000	
Butyrivibrio crossotus	NZ_ABWN00000000	Absent
Catenibacterium mitsuokai	NZ_ACCK00000000	
Citrobacter sp	NZ_ACDJ00000000	
Clostridium asparagiforme	NZ_ACCJ00000000	
Clostridium bartlettii	NZ_ABEZ00000000	Present
Clostridium bolteae	NZ_ABCC00000000	Present
Clostridium hiranonis	NZ_ABWP00000000	Present
Clostridium hylemonae	NZ_ABYI00000000	
Clostridium leptum	NZ_ABCB00000000	Absent
Clostridium methylpentosum	NZ_ACEC00000000	
Clostridium nexile	NZ_ABWO00000000	Present
Clostridium ramosum	NZ_ABFX00000000	Present
Clostridium scindens ATCC 35704	NZ_ABFY00000000	Present
Clostridium sp L2	NZ_AAYW00000000	Present
Clostridium sp M62 1	NZ_ACFX00000000	

Clostridium sp SS2	NZ_ABGC00000000	
Clostridium spiroforme DSM 1552	NZ_ABIK00000000	Present
Clostridium sporogenes	NZ_ABKW00000000	Present
Clostridium symbiosum	ProjectID:18183	Present
Collinsella aerofaciens ATCC 25986	NZ_AAVN00000000	Absent
Collinsella intestinalis	NZ_ABXH00000000	Absent
Collinsella stercoris	NZ_ABXJ00000000	Absent
Coprococcus comes	NZ_ABVR00000000	
Coprococcus eutactus	NZ_ABEY00000000	Present
Desulfovibrio piger ATCC29098	NZ_ABXU00000000	Present
Desulfovibrio piger GOR1	ProjectID:30377	
Dorea formicigenerans	NZ_AAXA00000000	Absent
Dorea longicatena DSM 13814	NZ_AAXB00000000	Present
Enterobacter cancerogenus	NZ_ABWM00000000	Present
Escherichia coli str. K-12 substr. MG1655	NC_000913	
Escherichia fergusonii ATCC 35469	NC_011835	
Eubacterium biforme	NZ_ABYT00000000	Present
Eubacterium dolichum	NZ_ABAW00000000	Present
Eubacterium eligens ATCC 27750	NC_012778	Present
Eubacterium hallii	NZ_ACEP00000000	Present
Eubacterium rectale ATCC 33656	NC_012781	Present
Eubacterium ventriosum	NZ_AAVL00000000	Present
Faecalibacterium prausnitzii M21/2	NZ_ABED00000000	Present
Fusobacterium sp. 4 1 13	NZ_ACDE00000000	
Fusobacterium varium ATCC27725	NZ_ACIE00000000	
Helicobacter pylori HPAG1	NC_008086	
Holdemania filiformis	NZ_ACCF00000000	
Lactobacillus casei ATCC 334	NC_008526	
Lactobacillus delbrueckii subsp. bulgaricus ATCC 11842	NC_008054	
Lactococcus lactis subsp. cremoris MG1363	NC_009004	
Lactococcus lactis subsp. cremoris SK11	NC_008527	
Lactococcus lactis subsp. lactis Il1403	NC_002662	
M23A		
Methanobrevibacter smithii ATCC 35061	NC_009515	Present
Methanobrevibacter smithii DSM2374	NZ_ABYV00000000	
Methanobrevibacter smithii DSM2375	NZ_ABYW00000000	
Methanosphaera stadtmanae DSM 3091	NC_007681	
Mitsuokella multacida	NZ_ABWK00000000	Present
Parabacteroides distasonis ATCC 8503	NC_009615	Present
Parabacteroides johnsonii	NZ_ABYH00000000	Absent
Parabacteroides merdae	NZ_AAXE00000000	Present
Peptostreptococcus micros	NZ_ABEE00000000	Present
Prevotella copri	NZ_ACBX00000000	Absent
Proteus penneri	NZ_ABVP00000000	
Providencia alcalifaciens	NZ_ABXW00000000	Absent
Providencia rettgeri	NZ_ACCI00000000	
Providencia rustigianii	NZ_ABXV00000000	Present
Providencia stuartii	NZ_ABJD00000000	Present
Roseburia intestinalis L1	NZ_ABYJ00000000	
Ruminococcus bromiiL263	ProjectID:39153	Absent
Ruminococcus gnavus	NZ_AAYG00000000	Present
Ruminococcus lactaris	NZ_ABOU00000000	Present
Ruminococcus obeum ATCC 29174	NZ_AAVO00000000	Absent
Ruminococcus torques ATCC 27756	NZ_AAVP00000000	Absent
Shigella sp. D9	NZ_ACDL00000000	
Streptococcus infantarius	NZ_ABJK00000000	Present
Streptococcus thermophilus CNRZ1066	NC_006448	
Streptococcus thermophilus DN001147.fa	NC_006449	
Streptococcus thermophilus LMD-9	NC_008532	
Subdoligranulum variabile	NZ_ACBY00000000	
Victivallis vadensis ATCC BAA-548	NZ_ABDE00000000	

Table S10. List of 396 microbial genomes used to search for prophages

Species and strain designation	GenBank Accession No
Acinetobacter sp. ADP1	NC_005966
Actinomyces odontolyticus	NZ_AAYI00000000
Agrobacterium tumefaciens str. C58	NC_003063
Agrobacterium tumefaciens str. C58	NC_003062
Alistipes putredinis	NZ_ABFK00000000
Anabaena variabilis ATCC 29413	NC_007413
Anaerococcus hydrogenalis	NZ_ABXA00000000
Anaerofustis stercorihominis	NZ_ABIL00000000
Anaerostipes caccae	NZ_ABAX00000000
Anaerotruncus colihominis	NZ_ABGD00000000
Anaplasma marginale str. St. Maries	NC_004842
Aquifex aeolicus VF5	NC_000918
Aster yellows witches'-broom phytoplasma AYWB	NC_007716
Azoarcus sp. EbN1	NC_006513
Bacillus anthracis str. 'Ames Ancestor'	NC_007530
Bacillus anthracis str. Ames	NC_003997
Bacillus anthracis str. Sterne	NC_005945
Bacillus cereus ATCC 10987	NC_003909
Bacillus cereus ATCC 14579	NC_004722
Bacillus cereus E33L	NC_006274
Bacillus clausii KSM-K16	NC_006582
Bacillus halodurans C-125	NC_002570
Bacillus licheniformis ATCC 14580	NC_006270
Bacillus subtilis subsp. subtilis str. 168	NC_000964
Bacillus thuringiensis serovar konkukian str. 97-27	NC_005957
Bacteroides caccae ATCC 43185	NZ_AAVM00000000
Bacteroides capillosus	NZ_AAXG00000000
Bacteroides cellulosilyticus	NZ_ACCH00000000
Bacteroides coprocola	NZ_ABIY00000000
Bacteroides coprophilus	NZ_ACBW00000000
Bacteroides dorei	NZ_ABWZ00000000
Bacteroides eggerthii	NZ_ABVO00000000
Bacteroides finegoldii	NZ_ABXI00000000
Bacteroides fragilis NCTC 9343	NC_003228
Bacteroides fragilis YCH46	NC_006347
Bacteroides intestinalis	NZ_ABJL00000000
Bacteroides ovatus ATCC 8483	NZ_AAXF00000000
Bacteroides plebeius	NZ_ABQC00000000
Bacteroides stercoris	NZ_ABFZ00000000
Bacteroides thetaiotaomicron VPI-5482	NC_004663
Bacteroides uniformis ATCC 8492	NZ_AAYH00000000
Bacteroides vulgatus ATCC 8482	NC_009614
Bartonella henselae str. Houston-1	NC_005956
Bartonella quintana str. Toulouse	NC_005955
Bdellovibrio bacteriovorus HD100	NC_005363
Bifidobacterium adolescentis	NZ_AAXD00000000
Bifidobacterium adolescentis ATCC 15703	NC_008618
Bifidobacterium angulatum	NZ_ABYS00000000
Bifidobacterium breve	NZ_ACCG00000000
Bifidobacterium catenulatum	NZ_ABXY00000000
Bifidobacterium dentium	NZ_ABIX00000000
Bifidobacterium gallicum	NZ_ABXB00000000
Bifidobacterium longum DJO10A	NC_010816
Bifidobacterium longum NCC2705	NC_004307
Bifidobacterium pseudocatenulatum	NZ_ABXX00000000
Blautia hansenii	NZ_ABYU00000000
Blautia hydrogenotrophica DSM 10507	NZ_ACBZ00000000
Bordetella pertussis Tohama I	NC_002929
Borrelia burgdorferi B31	NC_001318
Borrelia garinii PBi	NC_006156
Bradyrhizobium japonicum USDA 110	NC_004463
Brucella abortus biovar 1 str. 9-941	NC_006932
Brucella abortus biovar 1 str. 9-941	NC_006933
Brucella melitensis 16M	NC_003317
Brucella melitensis 16M	NC_003318
Brucella melitensis biovar Abortus	NC_007618
Brucella suis 1330	NC_004310
Brucella suis 1330	NC_004311
Bryantella formatexigens DSM 14469	NZ_ACCL00000000
Buchnera aphidicola str. APS (Acyrthosiphon pisum)	NC_002528
Buchnera aphidicola str. Bp (Baizongia pistaciae)	NC_004545
Buchnera aphidicola str. Sg (Schizaphis graminum)	NC_004061
Burkholderia mallei ATCC 23344	NC_006348
Burkholderia mallei ATCC 23344	NC_006349
Burkholderia pseudomallei 1710b	NC_007434
Burkholderia pseudomallei 1710b	NC_007435
Burkholderia pseudomallei K96243	NC_006350
Burkholderia pseudomallei K96243	NC_006351
Burkholderia sp. 383	NC_007509
Burkholderia sp. 383	NC_007510
Burkholderia sp. 383	NC_007511
Burkholderia thailandensis E264	NC_007650
Burkholderia thailandensis E264	NC_007651
Butyrivibrio crossotus	NZ_ABWN00000000
Campylobacter jejuni RM1221	NC_003912
Campylobacter jejuni subsp. jejuni NCTC 11168	NC_002163
Candidatus Blochmannia floridanus	NC_005061
Candidatus Blochmannia pennsylvanicus str. BPEN	NC_007292
Candidatus Pelagibacter ubique HTCC1062	NC_007205
Carboxydothermus hydrogenoformans Z-2901	NC_007503
Catenibacterium mitsuokai	NZ_ACCK00000000
Caulobacter crescentus CB15	NC_002696
Chlamydia muridarum Nigg	NC_002620
Chlamydia trachomatis A/HAR-13	NC_007429
Legionella pneumophila subsp. pneumophila str. Philadelphia	NC_002942
Leifsonia xyli subsp. xyli str. CTCB07	NC_006087
Leptospira interrogans serovar Copenhageni str. Fiocruz L1-	NC_005823
Leptospira interrogans serovar Copenhageni str. Fiocruz L1-	NC_005824
Leptospira interrogans serovar Lai str. 56601	NC_004342
Leptospira interrogans serovar Lai str. 56601	NC_004343
Listeria innocua Clip11262	NC_003212
Listeria monocytogenes EGD-e	NC_003210
Listeria monocytogenes str. 4b F2365	NC_002973
M23A	
Magnetospirillum magneticum AMB-1	NC_007626
Mannheimia succiniciproducens MBEL55E	NC_006300
Mesoplasma florum L1	NC_006055
Mesorhizobium loti MAFF303099	NC_002678
Methanobrevibacter smithii ATCC 35061	NC_009515
Methanobrevibacter smithii DSM2374	NZ_ABYV00000000
Methanobrevibacter smithii DSM2375	NZ_ABYW00000000
Methanosphaera stadtmanae DSM 3091	NC_007681
Methylococcus capsulatus str. Bath	NC_002977
Mitsuokella multacida	NZ_ABWK00000000
Moorella thermoacetica ATCC 39073	NC_007644
Mycobacterium avium subsp. paratuberculosis str. k10	NC_002944
Mycobacterium bovis AF2122/97	NC_002945
Mycobacterium leprae TN	NC_002677
Mycobacterium tuberculosis CDC1551	NC_002755
Mycobacterium tuberculosis H37Rv	NC_000962
Mycoplasma capricolum subsp. capricolum ATCC 27343	NC_007633
Mycoplasma gallisepticum R	NC_004829
Mycoplasma genitalium G-37	NC_000908
Mycoplasma hyopneumoniae 232	NC_006360
Mycoplasma hyopneumoniae 7448	NC_007332
Mycoplasma hyopneumoniae J	NC_007295
Mycoplasma mobile 163K	NC_006908
Mycoplasma mycoides subsp. mycoides SC str. PG1	NC_005364
Mycoplasma penetrans HF-2	NC_004432
Mycoplasma pneumoniae M129	NC_000912
Mycoplasma pulmonis UAB CTIP	NC_002771
Mycoplasma synoviae 53	NC_007294
Neisseria gonorrhoeae FA 1090	NC_002946
Neisseria meningitidis MC58	NC_003112
Neisseria meningitidis Z2491	NC_003116
Nitrobacter winogradskyi Nb-255	NC_007406
Nitrosococcus oceani ATCC 19707	NC_007484
Nitrosomonas europaea ATCC 19718	NC_004757
Nitrosospira multiformis ATCC 25196	NC_007614
Nocardia farcinica IFM 10152	NC_006361
Nostoc sp. PCC 7120	NC_003272
Novosphingobium aromaticivorans DSM 12444	NC_007794
Oceanobacillus iheyensis HTE831	NC_004193
Onion yellows phytoplasma OY-M	NC_005303
Parabacteroides distasonis ATCC 8503	NC_009615
Parabacteroides johnsonii	NZ_ABYH00000000
Parabacteroides merdae	NZ_AAXE00000000
Parachlamydia sp. UWE25	NC_005861
Pasteurella multocida subsp. multocida str. Pm70	NC_002663
Pelobacter carbinolicus DSM 2380	NC_007498
Pelodictyon luteolum DSM 273	NC_007512
Peptostreptococcus micros	NZ_ABEE00000000
Photobacterium profundum SS9	NC_006370
Photobacterium profundum SS9	NC_006371
Photorhabdus luminescens subsp. laumondii TTO1	NC_005126
Porphyromonas gingivalis W83	NC_002950
Prevotella copri	NZ_ACBX00000000
Prochlorococcus marinus str. MIT 9312	NC_007577
Prochlorococcus marinus str. MIT 9313	NC_005071
Prochlorococcus marinus str. NATL2A	NC_007335
Prochlorococcus marinus subsp. marinus str. CCMP1375	NC_005042
Prochlorococcus marinus subsp. pastoris str. CCMP1986	NC_005072
Propionibacterium acnes KPA171202	NC_006085
Proteus penneri	NZ_ABVP00000000
Providencia alcalifaciens	NZ_ABXW00000000
Providencia rettgeri	NZ_ACCI00000000
Providencia rustigianii	NZ_ABXV00000000
Providencia stuartii	NZ_ABJD00000000
Pseudoalteromonas haloplanktis TAC125	NC_007481
Pseudoalteromonas haloplanktis TAC125	NC_007482
Pseudomonas aeruginosa PAO1	NC_002516
Pseudomonas fluorescens Pf-5	NC_004129
Pseudomonas fluorescens PfO-1	NC_007492
Pseudomonas putida KT2440	NC_002947
Pseudomonas syringae pv. phaseolicola 1448A	NC_005773
Pseudomonas syringae pv. syringae B728a	NC_007005
Pseudomonas syringae pv. tomato str. DC3000	NC_004578
Psychrobacter arcticum 273-4	NC_007204
Ralstonia eutropha JMP134	NC_007347
Ralstonia eutropha JMP134	NC_007348
Ralstonia solanacearum GMI1000	NC_003295
Rhodobacter sphaeroides 2.4.1	NC_007493
Rhodobacter sphaeroides 2.4.1	NC_007494
Rhodopirellula baltica SH 1	NC_005027
Rhodopseudomonas palustris CGA009	NC_005296
Rhodopseudomonas palustris HaA2	NC_007778
Rhodospirillum rubrum ATCC 11170	NC_007643
Rickettsia conorii str. Malish 7	NC_003103

Chlamydia trachomatis D/UW-3/CX	NC_000117
Chlamydophila abortus S26/3	NC_004552
Chlamydophila caviae GPIC	NC_003361
Chlamydophila pneumoniae AR39	NC_002179
Chlamydophila pneumoniae CWL029	NC_000922
Chlamydophila pneumoniae J138	NC_002491
Chlamydophila pneumoniae TW-183	NC_005043
Chlorobium chlorochromatii CaD3	NC_007514
Chlorobium tepidum TLS	NC_002932
Chromobacterium violaceum ATCC 12472	NC_005085
Citrobacter sp	NZ_ACDJ00000000
Clostridium acetobutylicum ATCC 824	NC_003030
Clostridium asparagiforme	NZ_ACCJ00000000
Clostridium bartlettii	NZ_ABEZ00000000
Clostridium bolteae	NZ_ABCC00000000
Clostridium hiranonis	NZ_ABWP00000000
Clostridium hylemonae	NZ_ABYI00000000
Clostridium leptum	NZ_ABCB00000000
Clostridium methylpentosum	NZ_ACEC00000000
Clostridium nexile	NZ_ABWO00000000
Clostridium perfringens str. 13	NC_003366
Clostridium ramosum	NZ_ABFX00000000
Clostridium scindens ATCC 35704	NZ_ABFY00000000
Clostridium sp L2	NZ_AAYW00000000
Clostridium sp M62 1	NZ_ACFX00000000
Clostridium sp SS2	NZ_ABGC00000000
Clostridium spiroforme DSM 1552	NZ_ABIK00000000
Clostridium sporogenes	NZ_ABKW00000000
Clostridium symbiosum	ProjectID:18183
Clostridium tetani E88	NC_004557
Collinsella aerofaciens ATCC 25986	NZ_AAVN00000000
Collinsella intestinalis	NZ_ABXH00000000
Collinsella stercoris	NZ_ABXJ00000000
Colwellia psychrerythraea 34H	NC_003910
Coprococcus comes	NZ_ABVR00000000
Coprococcus eutactus	NZ_ABEY00000000
Corynebacterium diphtheriae NCTC 13129	NC_002935
Corynebacterium efficiens YS-314	NC_004369
Corynebacterium glutamicum ATCC 13032	NC_006958
Corynebacterium jeikeium K411	NC_007164
Coxiella burnetii RSA 493	NC_002971
Cyanobacteria bacterium Yellowstone A-Prime	NC_007775
Cyanobacteria bacterium Yellowstone B-Prime	NC_007776
Dechloromonas aromatica RCB	NC_007298
Dehalococcoides ethenogenes 195	NC_002936
Dehalococcoides sp. CBDB1	NC_007356
Deinococcus radiodurans R1	NC_001263
Deinococcus radiodurans R1	NC_001264
Desulfotalea psychrophila LSv54	NC_006138
Desulfovibrio desulfuricans G20	NC_007519
Desulfovibrio piger ATCC29098	NZ_ABXU00000000
Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough	NC_002937
Dorea formicigenerans	NZ_AAXA00000000
Dorea longicatena DSM 13814	NZ_AAXB00000000
Ehrlichia canis str. Jake	NC_007354
Ehrlichia ruminantium str. Gardel	NC_006831
Ehrlichia ruminantium str. Welgevonden	NC_005295
Ehrlichia ruminantium str. Welgevonden	NC_006832
Enterobacter cancerogenus	NZ_ABWM00000000
Enterococcus faecalis V583	NC_004668
Erwinia carotovora subsp. atroseptica SCRI1043	NC_004547
Erythrobacter litoralis HTCC2594	NC_007722
Escherichia coli CFT073	NC_004431
Escherichia coli K12	NC_000913
Escherichia coli O157:H7	NC_002695
Escherichia coli O157:H7 EDL933	NC_002655
Eubacterium biforme	NZ_ABYT00000000
Eubacterium dolichum	NZ_ABAW00000000
Eubacterium hallii	NZ_ACEP00000000
Eubacterium ventriosum	NZ_AAVL00000000
Faecalibacterium prausnitzii M21/2	NZ_ABED00000000
Francisella tularensis subsp. tularensis SCHU S4	NC_006570
Frankia sp. CcI3	NC_007777
Fusobacterium nucleatum subsp. nucleatum ATCC 25586	NC_003454
Geobacillus kaustophilus HTA426	NC_006510
Geobacter metallireducens GS-15	NC_007517
Geobacter sulfurreducens PCA	NC_002939
Gloeobacter violaceus PCC 7421	NC_005125
Gluconobacter oxydans 621H	NC_006677
Haemophilus ducreyi 35000HP	NC_002940
Haemophilus influenzae 86-028NP	NC_007146
Haemophilus influenzae Rd KW20	NC_000907
Hahella chejuensis KCTC 2396	NC_007645
Helicobacter hepaticus ATCC 51449	NC_004917
Helicobacter pylori 1039kx1	ProjectID:40515
Helicobacter pylori 1039kx2	ProjectID:40519
Helicobacter pylori 26695	NC_000915
Helicobacter pylori 345kx1	ProjectID:40545
Helicobacter pylori 345kx2	ProjectID:40547
Helicobacter pylori 98_10	NZ_ABSX00000000
Helicobacter pylori B128	NZ_ABSY00000000
Helicobacter pylori G27	NC_011333
Helicobacter pylori HPAG1	NC_008086
Helicobacter pylori J99	NC_000921
Helicobacter pylori P12	NC_011498
Holdemania filiformis	NZ_ACCF00000000
Idiomarina loihiensis L2TR	NC_006512
Rickettsia felis URRWXCal2	NC_007109
Rickettsia prowazekii str. Madrid E	NC_000963
Rickettsia typhi str. Wilmington	NC_006142
Roseburia intestinalis L1	NZ_ABYJ00000000
Ruminococcus gnavus	NZ_AAYG00000000
Ruminococcus lactaris	NZ_ABOU00000000
Ruminococcus obeum ATCC 29174	NZ_AAVO00000000
Ruminococcus torques ATCC 27756	NZ_AAVP00000000
Salinibacter ruber DSM 13855	NC_007677
Salmonella enterica subsp. enterica serovar Choleraesuis st	NC_006905
Salmonella enterica subsp. enterica serovar Paratyphi A str	NC_006511
Salmonella enterica subsp. enterica serovar Typhi str. CT18	NC_003198
Salmonella enterica subsp. enterica serovar Typhi Ty2	NC_004631
Salmonella typhimurium LT2	NC_003197
Shewanella oneidensis MR-1	NC_004347
Shigella boydii Sb227	NC_007613
Shigella dysenteriae Sd197	NC_007606
Shigella flexneri 2a str. 2457T	NC_004741
Shigella flexneri 2a str. 301	NC_004337
Shigella sonnei Ss046	NC_007384
Silicibacter pomeroyi DSS-3	NC_003911
Sinorhizobium meliloti 1021	NC_003047
Sodalis glossinidius str. 'morsitans'	NC_007712
Staphylococcus aureus RF122	NC_007622
Staphylococcus aureus subsp. aureus COL	NC_002951
Staphylococcus aureus subsp. aureus MRSA252	NC_002952
Staphylococcus aureus subsp. aureus MSSA476	NC_002953
Staphylococcus aureus subsp. aureus Mu50	NC_002758
Staphylococcus aureus subsp. aureus MW2	NC_003923
Staphylococcus aureus subsp. aureus N315	NC_002745
Staphylococcus aureus subsp. aureus NCTC 8325	NC_007795
Staphylococcus aureus subsp. aureus USA300	NC_007793
Staphylococcus epidermidis ATCC 12228	NC_004461
Staphylococcus epidermidis RP62A	NC_002976
Staphylococcus haemolyticus JCSC1435	NC_007168
Staphylococcus saprophyticus subsp. saprophyticus ATCC 1530	NC_007350
Streptococcus agalactiae 2603V/R	NC_004116
Streptococcus agalactiae A909	NC_007432
Streptococcus agalactiae NEM316	NC_004368
Streptococcus infantarius	NZ_ABJK00000000
Streptococcus mutans UA159	NC_004350
Streptococcus pneumoniae R6	NC_003098
Streptococcus pneumoniae TIGR4	NC_003028
Streptococcus pyogenes M1 GAS	NC_002737
Streptococcus pyogenes MGAS10394	NC_006086
Streptococcus pyogenes MGAS315	NC_004070
Streptococcus pyogenes MGAS5005	NC_007297
Streptococcus pyogenes MGAS6180	NC_007296
Streptococcus pyogenes MGAS8232	NC_003485
Streptococcus pyogenes SSI-1	NC_004606
Streptococcus thermophilus CNRZ1066	NC_006449
Streptococcus thermophilus LMG 18311	NC_006448
Streptomyces avermitilis MA-4680	NC_003155
Streptomyces coelicolor A3(2)	NC_003888
Subdoligranulum variabile	NZ_ACBY00000000
Symbiobacterium thermophilum IAM 14863	NC_006177
Synechococcus elongatus PCC 6301	NC_006576
Synechococcus elongatus PCC 7942	NC_007604
Synechococcus sp. CC9605	NC_007516
Synechococcus sp. CC9902	NC_007513
Synechococcus sp. WH 8102	NC_005070
Synechocystis sp. PCC 6803	NC_000911
Syntrophus aciditrophicus SB	NC_007759
Thermoanaerobacter tengcongensis MB4	NC_003869
Thermobifida fusca YX	NC_007333
Thermosynechococcus elongatus BP-1	NC_004113
Thermotoga maritima MSB8	NC_000853
Thermus thermophilus HB27	NC_005835
Thermus thermophilus HB8	NC_006461
Thiobacillus denitrificans ATCC 25259	NC_007404
Thiomicrospira crunogena XCL-2	NC_007520
Thiomicrospira denitrificans ATCC 33889	NC_007575
Treponema denticola ATCC 35405	NC_002967
Treponema pallidum subsp. pallidum str. Nichols	NC_000919
Tropheryma whipplei str. Twist	NC_004572
Tropheryma whipplei TW08/27	NC_004551
Ureaplasma parvum serovar 3 str. ATCC 700970	NC_002162
Vibrio cholerae O1 biovar eltor str. N16961	NC_002505
Vibrio cholerae O1 biovar eltor str. N16961	NC_002506
Vibrio fischeri ES114	NC_006840
Vibrio fischeri ES114	NC_006841
Vibrio parahaemolyticus RIMD 2210633	NC_004603
Vibrio parahaemolyticus RIMD 2210633	NC_004605
Vibrio vulnificus CMCP6	NC_004459
Vibrio vulnificus CMCP6	NC_004460
Vibrio vulnificus YJ016	NC_005139
Vibrio vulnificus YJ016	NC_005140
Wigglesworthia glossinidia endosymbiont of Glossina brevipa	NC_004344
Wolbachia endosymbiont of Drosophila melanogaster	NC_002978
Wolbachia endosymbiont strain TRS of Brugia malayi	NC_006833
Wolinella succinogenes DSM 1740	NC_005090
Xanthomonas axonopodis pv. citri str. 306	NC_003919
Xanthomonas campestris pv. campestris str. 8004	NC_007086
Xanthomonas campestris pv. campestris str. ATCC 33913	NC_003902
Xanthomonas campestris pv. vesicatoria str. 85-10	NC_007508
Xanthomonas oryzae pv. oryzae KACC10331	NC_006834
Xanthomonas oryzae pv. oryzae MAFF 311018	NC_007705

Lactobacillus acidophilus NCFM	NC_006814	Xylella fastidiosa 9a5c	NC_002488
Lactobacillus johnsonii NCC 533	NC_005362	Xylella fastidiosa Temecula1	NC_004556
Lactobacillus plantarum WCFS1	NC_004567	Yersinia pestis biovar Medievalis str. 91001	NC_005810
Lactobacillus sakei subsp. sakei 23K	NC_007576	Yersinia pestis CO92	NC_003143
Lactococcus lactis subsp. lactis Il1403	NC_002662	Yersinia pestis KIM	NC_004088
Legionella pneumophila str. Lens	NC_006369	Yersinia pseudotuberculosis IP 32953	NC_006155
Legionella pneumophila str. Paris	NC_006368	Zymomonas mobilis subsp. mobilis ZM4	NC_006526

Table S11. Number of VLP reads with similarity over more than 90% of their length to CRISPR spacers present in either the fecal microbiome of the same individual or a reference human gut-associated microbial genome

			Family 1			Family 2			Family 3			Family 4			Family 5	
			Co-twin 1	Co-twin 2	Mother	Co-twin 1	Co-twin 2	Mother	Co-twin 1	Co-twin 2	Mother	Co-twin 1	Co-twin 2	Mother	Co-twin 2	Ref Gut Genomes
Family 1	Co-twin 1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	8.16
		3	-	-	-	-	-	-	-	-	-	-	-	-	0.23	0.69
	Co-twin 2	1	-	-	-	-	-	-	-	-	-	-	-	-	3.65	66.35
		1(R)	-	-	-	-	-	-	-	-	-	-	-	-	7.38	80.05
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	65.07
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	1.78
	Mother	1	-	-	-	-	-	-	-	-	-	-	-	-	-	0.41
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Family 2	Co-twin 1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		1(R)	-	-	-	-	-	-	-	0.42	-	-	-	-	-	3.78
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	4.45
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Co-twin 2	1	-	-	-	-	-	-	-	-	-	-	-	-	8.39	0.24
		1(R)	-	-	-	-	-	-	-	-	-	-	-	-	5.62	-
		2	-	-	-	-	-	-	-	-	-	-	-	-	3.03	1.93
	Mother	1	-	-	-	-	-	-	-	-	-	-	-	-	-	0.29
		1(R)	-	-	-	-	-	-	-	-	-	-	-	-	1.03	0.34
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	0.36
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Family 3	Co-twin 1	1	-	-	-	-	-	-	-	-	-	-	-	-	0.54	-
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Co-twin 2	1	-	-	-	0.94	-	-	-	-	-	-	-	-	-	2.81
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	0.59
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Mother	1	-	-	-	-	-	-	-	-	-	-	-	-	-	7.17
		2	-	-	-	-	-	-	0.53	-	-	-	-	-	0.80	4.51
Family 4	Co-twin 1	1	-	-	-	-	-	-	-	-	-	-	-	-	2.04	6.11
		2	-	-	-	-	-	-	-	-	-	-	-	-	0.48	9.19
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Co-twin 2	1	-	-	-	-	-	-	-	-	-	-	-	-	8.23	46.18
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Mother	1	-	-	-	-	-	-	-	-	-	-	-	-	-	0.40
		2	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		3(R)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Family 5	Co-twin 2	1	-	-	-	-	-	-	-	-	-	-	-	-	2.68	0.21
		1(R)	-	-	-	-	-	-	-	-	-	-	-	-	3.56	1.23
No. of spacers per fecal microbiome			**9**	**24**	**52**	**13**	**36**	**17**	**26**	**14**	**22**	**21**	**57**	**49**	**241**	**2196**