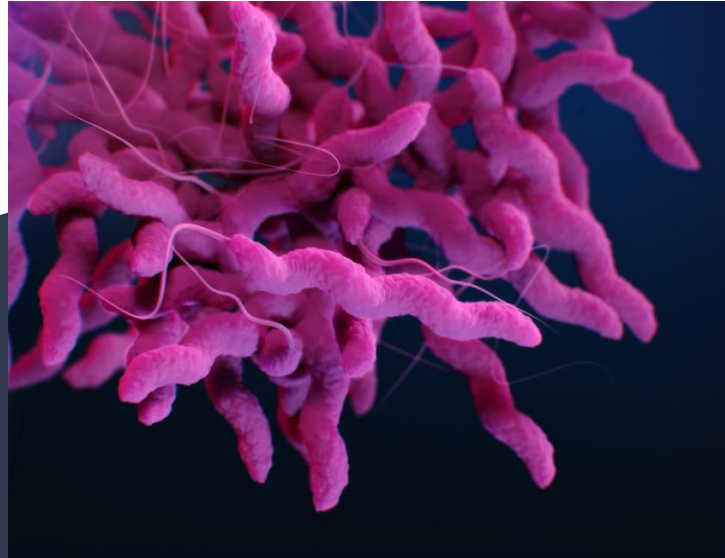


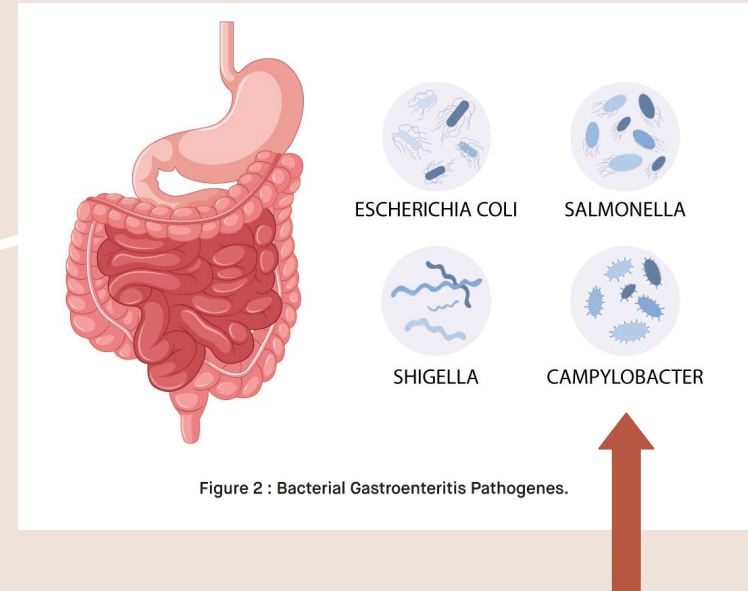
Whole genome sequencing of *C. jejuni* and *C. coli* from rhesus monkey with & without intestinal disease

Andrea Villanueva



Campylobacter jejuni and *Campylobacter coli*

- Gram negative, spiral shaped bacteria
- Leading cause of bacterial gastroenteritis that can develop into post-infectious IBS (PI-IBS) in humans
- Can cause acute campylobacteriosis in rhesus monkeys (*Macaca mulatta*), with PI-IBS-like syndromes



Purpose

- Small-scale genetic differences in *Campylobacter* strains that produce intestinal disease and those that do not
- Limited information available on whole genome sequences from *Campylobacter* species isolated from rhesus monkeys



Methods: Genome assembly using ABySS

- Parallel, paired-end genome assembly done by short sequencing
- More useful when analyzing large genomes

ABySS statistics

	n	N50 (bp)	Sum (complete genome size)
unitig	283	24641	1673549
contig	278	24641	1673925
scaffold	275	24641	1673906

Methods: Genome assembly using SPAdes

- Genome assembly for regular and single cell projects
- Better algorithm designed for smaller genomes

- Contigs
- Scaffolds
- Regular assembly graph
- Assembly graph with scaffolds

Genome Quality Assessment (QUAST)

- Compares assembly with reference genome of *C. jejuni* (as identified from SPADes scaffolds file, on NCBI)
- **BUSCO** measures how complete the genome appears

	ABySS	SPAdes
BUSCO value (%)	80.41	82.43

Genome Annotation (prokka and RAST)

Identifies genes and other key features in the assembled genome

- **Prokka:** rapid annotation designed for prokaryotic and viral genomes
- **RAST:** rapid annotation; identifies protein-encoding genes

RAST Seed Viewer

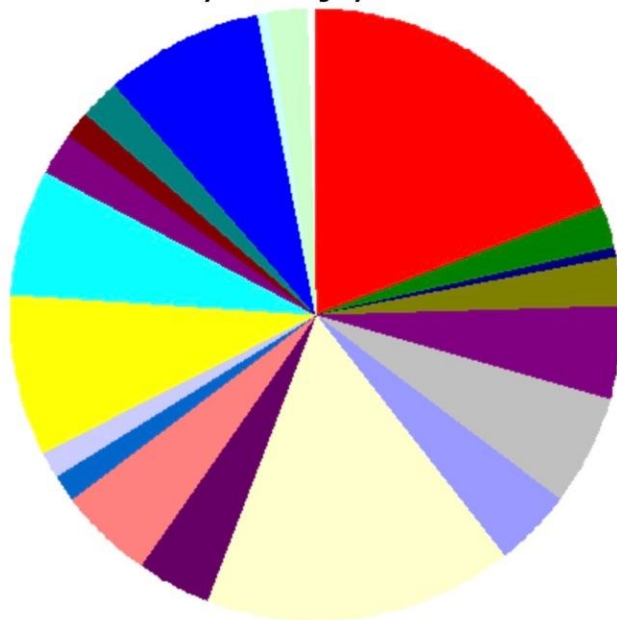
Subsystem Statistics

Features in Subsystems

Subsystem Coverage



Subsystem Category Distribution



Subsystem Feature Counts

+	Respiration (71)
+	Cell Wall and Capsule (17)
+	Motility and Chemotaxis (55)
+	Dormancy and Sporulation (1)
+	Iron acquisition and metabolism (13)
+	Virulence, Disease and Defense (19)
+	Cofactors, Vitamins, Prosthetic Groups, Pigments (68)
+	Photosynthesis (0)
+	Regulation and Cell signaling (2)
+	Miscellaneous (0)
+	Sulfur Metabolism (5)
+	Phosphorus Metabolism (15)
+	Amino Acids and Derivatives (160)
+	Stress Response (19)
+	Secondary Metabolism (4)
+	Fatty Acids, Lipids, and Isoprenoids (21)
+	Nucleosides and Nucleotides (41)
+	Cell Division and Cell Cycle (0)
+	Carbohydrates (35)
+	Phages, Prophages, Transposable elements, Plasmids (0)
+	Membrane Transport (48)
+	RNA Metabolism (31)
+	DNA Metabolism (41)
+	Nitrogen Metabolism (13)
+	Potassium metabolism (11)
+	Metabolism of Aromatic Compounds (1)
+	Protein Metabolism (136)

Results/Conclusions

TABLE 1 Characteristics and accession numbers of genome sequences of *Campylobacter jejuni* and *C. coli* isolates from rhesus macaques with and without intestinal disease

Sample	Organism	Host intestinal disease	No. of contigs	N_{50} (bp)	Coverage (%)	Genome size (bp)	G + C content (%)	GenBank accession no.
2	<i>C. jejuni</i>	Yes	20	178,959	36.33324	1,676,018	30.46	SAMN38909928
10	<i>C. jejuni</i>	No	37	175,354	46.84625	1,787,441	30.18	SAMN38909936
52	<i>C. jejuni</i>	No	14	322,944	47.28748	1,690,736	30.44	SAMN38909977
56	<i>C. jejuni</i>	No	37	198,078	34.58232	1,788,868	30.18	SAMN38909981
69	<i>C. jejuni</i>	No	32	198,078	30.19519	1,782,502	30.19	SAMN38909994
74	<i>C. jejuni</i>	No	153	20,821	31.58797	1,801,038	30.18	SAMN38909999
104	<i>C. jejuni</i>	Yes	425	5,796	23.45778	1,651,546	30.54	SAMN38910028
108	<i>C. jejuni</i>	No	203	15,826	28.62104	1,803,331	30.19	SAMN38910032
1	<i>C. coli</i>	No	27	171,129	31.53036	1,690,046	31.34	SAMN38909927
3	<i>C. coli</i>	No	34	174,774	35.09052	1,736,591	31.3	SAMN38909929
4	<i>C. coli</i>	No	32	126,735	41.7259	1,760,530	31.17	SAMN38909930
5	<i>C. coli</i>	No	34	174,774	51.48681	1,736,183	31.3	SAMN38909931
6	<i>C. coli</i>	No	39	137,763	37.23116	1,735,726	31.3	SAMN38909932
7	<i>C. coli</i>	No	22	138,429	39.28135	1,678,825	31.35	SAMN38909933
8	<i>C. coli</i>	No	29	223,285	38.24534	1,719,861	31.32	SAMN38909934
9	<i>C. coli</i>	Yes	16	163,417	42.63004	1,640,223	31.33	SAMN38909935
11	<i>C. coli</i>	No	17	162,890	45.64465	1,641,468	31.33	SAMN38909937
12	<i>C. coli</i>	No	28	162,934	50.17786	1,756,281	31.25	SAMN38909938
13	<i>C. coli</i>	No	55	219,887	49.73176	1,724,146	31.27	SAMN38909939

References

- <https://journals.asm.org/doi/pdf/10.1128/mra.00018-24>
- <https://rast.nmpdr.org/seedviewer.cgi?page=Organism&organism=197.26665>
-

GitHub link (annotation/assembly files)

- <https://github.com/villanuevandrea/Bioinformatics>