**Appendix e-1:**

**Guillain-Barré syndrome Outbreak in Peru 2019 associated with *Campylobacter jejuni* infection**

Ramos AP et. al.

**Methods**

***Campylobacter jejuni* sequencing and genomic analysis**

**Isolates culture**

Vials of frozen isolates were thawed and plated to both a Chocolate agar plate and a Campylobacter CSM (Charcoal based Selective Medium) plate. Each plate was placed in a GasPak (EZ Campy Gas Generating Pouch System) and sealed it to create a microaerophilic environment needed for the growth of*Campylobacter jejuni*. The atmosphere in each bag was approximately 5% oxygen, 10% carbon dioxide, and 85% nitrogen. The sealed bags were placed in an incubator set at 42o C for 24 hours. After 24 hours, identification and confirmation of each isolate was performed using a MALDI-TOF (Bruker Biotyper) instrument. Isolates were subcultured once more under the same conditions as above before DNA isolation using standard procedures.

**Sequencing**

Genomic DNA was extracted from pure cultures using the DNeasy PowerBiofilm Kit (Qiagen, Hilden, Germany). Long-read genomic sequencing was performed using the third generation Oxford Nanopore GridION X5, Oxford, England) sequencing instrument. Each Nanopore sequencing library was prepared using 1 µg of DNA with the Rapid Barcoding Kit (SQK-RBK004, Oxford Nanopore Technologies) using R9.4 flowcells (FLO-MIN106). The four isolates were run per flowcell. MinKNOW software was used to collect sequencing data.

Whole genome sequencing (WGS) was conducted using Illumina MiSeq short-read sequencing (Illumina, San Diego, California). Approximately 100-500 ng of gDNA was used to prepare sequencing libraries using the Nextera Flex Kit. Each Illumina library was then sequenced using v3 2x300 MiSeq Reagent Kit on an Illumina Miseq.

**Tree construction methods**

83 C. jejuni genome assemblies were downloaded from NCBI. All 16 genomes with reported GBS associations in the NCBI metadata were selected and 67 additional genomes were selected to represent a wide range of the collection locations, dates, and studies available (See table below). These 83 genomes, the C. jejuni reference genome (NCTC11168)1,2, and the 4 Peruvian strains (NCBI Genebank accession numbers SAMN15508151- SAMN15508154) assembled here were then used to construct a phylogenic tree from the sequence of the hypervariable lipo-oligosaccharide region using the Nextstrain-Augur pipeline3.

The LOS region was first extracted from each genome by aligning each assembly to the C. jejuni reference (NCTC11168) with minimap2 (-c --cs -x asm5)4. The LOS region coordinates from the reference were taken to be from Cj1131c (galE) to Cj1151c (rfaD)5. The coordinate range (AL111168.1:1064895-1084736) was then lifted over using the minimap2 utility paftools liftover, and the specified region was extracted from each genome with samtools6.

Augur was then used to perform a multiple alignment of the 88 resultant LOS loci. A well described strain, ICDCCJ070017, was specified as the reference. This strain was chosen as the reference as the cst*II* locus is of particular interest in GBS association, and cst*II* is not present in strain NCTC11168, which has a class C LOS locus8, but is present in ICDCCJ07001. augur tree was then run on the resulting multiple alignment, and the tree was refined with collection date metadata (when available) and the augur refine --timetree option, which re-roots the tree at an inferred ancestor rather than rooting at the specified reference.

augur ancestral and augur translate were run to determine the nucleotide and amino acid changes along each tree branch, based on the generated tree and sequence alignments. Visualizations were produced with auspice using metadata of country, geographic region, and GBS status.

Of the 20 GBS associated genomes, 15, including the Peruvian isolates, have fairly closely related LOS regions to one another. Sample collection regions do not appear to define clades, with strains from countries with numerous samples spread throughout the tree. All four Peruvian C. jejuni isolates were found to have Asn at cst*II* codon 51, based on the alignments to ICDCCJ07001.

**Table. *Campylobacter jejuni* genome assemblies information**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genebank accession** | **Strain** | **Year** | **Country** | **GBS**  **status** |  |
| GCA\_000686265.1 | 255 | 20XX | Pakistan | No | South Asia |
| GCA\_000686285.1 | 10227 | 20XX | Vietnam | No | Southeast Asia |
| GCA\_000685605.1 | 30286 | 20XX | Vietnam | No | Southeast Asia |
| GCA\_000686305.1 | 30318 | 20XX | Vietnam | No | Southeast Asia |
| GCA\_000430385.1 | 32488 | 2004 | USA | No | North America |
| GCA\_007923345.1 | 221091 | 2017 | United Kingdom | No | Europe |
| GCA\_000835365.1 | 00-0949 | 2000 | Canada | No | North America |
| GCA\_000466075.2 | 00-2544 | 2000 | Canada | No | North America |
| GCA\_000835345.1 | 01-1512 | 2001 | Canada | No | North America |
| GCA\_009939705.1 | 15AR0919 | 2015 | New Zealand | No | Oceania |
| GCA\_009939645.1 | 15AR1555 | 2014 | New Zealand | No | Oceania |
| GCA\_006508225.1 | 2006-119 | 2006 | Hungary | No | Europe |
| GCA\_010998535.1 | 2011\_0198H | 2011 | France | No | Europe |
| GCA\_008014135.1 | 2011D-8695 | 2011 | USA | No | North America |
| GCA\_010997615.1 | 2013\_0529H | 2013 | France | No | Europe |
| GCA\_010996855.1 | 2014\_0889H | 2014 | France | No | Europe |
| GCA\_010996275.1 | 2015\_0505H | 2015 | France | No | Europe |
| GCA\_010994855.1 | 2016\_3439 | 2016 | France | No | Europe |
| GCA\_010994695.1 | 2016\_3451 | 2016 | France | No | Europe |
| GCA\_010994645.1 | 2016\_3453 | 2016 | France | No | Europe |
| GCA\_009270295.1 | ABH\_016 | 2006 | Canada | No | North America |
| GCA\_009268395.1 | ABH\_121 | 2008 | Canada | No | North America |
| GCA\_003048165.1 | BD-39 | 2006 | Bangladesh | GBS | South Asia |
| GCA\_003048115.1 | BD-67 | 2007 | Bangladesh | GBS | South Asia |
| GCA\_003048185.1 | BD-75 | 2007 | Bangladesh | GBS | South Asia |
| GCA\_000440015.1 | BJ-CJD101 | 2003 | China | No | China |
| GCA\_001547605.1 | BJ-CJD120 | 2003 | China | No | China |
| GCA\_001547625.1 | BJ-CJD39 | 2003 | China | No | China |
| GCA\_001547595.1 | BJ-CJD70 | 2003 | China | No | China |
| GCA\_001516495.1 | BJ-CJGB95377 | 1995 | China | GBS | China |
| GCA\_001516505.1 | BJ-CJGB96114 | 1996 | China | GBS | China |
| GCA\_001516515.1 | BJ-CJGB96299 | 1996 | China | GBS | China |
| GCA\_001516485.1 | BJ-CJGB96G25 | 1996 | China | GBS | China |
| GCA\_007845945.1 | CAM19-008 | 2018 | China | No | China |
| GCA\_007845785.1 | CAM19-028 | 2018 | China | No | China |
| GCA\_007845875.1 | CAM19-037 | 2018 | China | No | China |
| GCA\_002118115.1 | CAM970 | 2016 | Japan | No | Japan Korea |
| GCA\_002805275.1 | CFSAN065294 | 2002 | Brazil | No | South America |
| GCA\_002806375.1 | CFSAN065310 | 2006 | Brazil | No | South America |
| GCA\_002806535.1 | CFSAN065312 | 2007 | Brazil | No | South America |
| GCA\_002804635.1 | CFSAN065403 | 2010 | Brazil | No | South America |
| GCA\_002834105.1 | CFSAN065415 | 2016 | USA | No | North America |
| GCA\_010226215.2 | CFSAN096304 | 2018 | Chile | No | South America |
| GCA\_010353465.2 | CFSAN096324 | 2019 | Chile | No | South America |
| GCA\_011067715.1 | CFSAN096343 | 2018 | Chile | No | South America |
| GCA\_010359425.2 | CFSAN096345 | 201 | Chile | No | South America |
| GCA\_003368105.1 | CJ017CCUA | 2001 | Finland | No | Europe |
| GCA\_003368225.1 | CJ018CCUA | 2002 | Finland | No | Europe |
| GCA\_002110625.1 | CJ035CC21 | 2003 | Finland | No | Europe |
| GCA\_003368185.1 | CJ071CC464 | 1999 | Finland | No | Europe |
| GCA\_000686345.1 | Cj1 | 20XX | Thailand | No | Southeast Asia |
| GCA\_000686365.1 | Cj2 | 20XX | Thailand | No | Southeast Asia |
| GCA\_000686385.1 | Cj3 | 20XX | Thailand | No | Southeast Asia |
| GCA\_000686325.1 | Cj5 | 20XX | Thailand | No | Southeast Asia |
| GCA\_001412295.1 | CJM1cam | 20XX | United Kingdom | No | Europe |
| GCA\_002209045.1 | FDAARGOS\_265 | 20XX | USA | No | North America |
| GCA\_002983725.1 | FDAARGOS\_295 | 1997 | South Africa | No | Africa |
| GCA\_002587105.1 | FDAARGOS\_422 | 20XX | USA | No | North America |
| GCA\_002214785.1 | FORC\_046 | 2016 | South Korea | No | Japan Korea |
| GCA\_000759575.1 | G1 | 20XX | ? | GBS | ? |
| GCA\_001516565.1 | HB-CJGB-LC | 20XX | China | GBS | China |
| GCA\_001516415.1 | HB-CJGB-LL | 1993 | China | GBS | China |
| GCA\_001516425.1 | HB-CJGB-LXC | 20XX | China | GBS | China |
| GCA\_001516405.1 | HB-CJGB-QYT | 20XX | China | GBS | China |
| GCA\_001516435.1 | HB-CJGB-XWM | 20XX | China | GBS | China |
| GCA\_003971585.1 | huA17 | 2000 | Germany | No | Europe |
| GCA\_000184085.1 | ICDCCJ07001 | 2007 | China | GBS | China |
| GCA\_000355825.1 | ICDCCJ07002 | 2007 | China | No | China |
| GCA\_000686405.1 | k1 | 20XX | Pakistan | No | South Asia |
| GCA\_003716305.1 | KF017 | 2010 | Poland | No | Europe |
| GCA\_003716565.1 | KF070 | 2010 | Poland | No | Europe |
| GCA\_001570705.1 | NC05-27 | 2005 | Japan | No | Japan Korea |
| GCA\_000009085.1 | NCTC11168 | 20XX | ? | No | ? |
| GCA\_900474415.1 | NCTC11925 | 1988 | Australia | No | Oceania |
| GCA\_001237545.1 | OXC6361 | 2011 | United Kingdom | No | Europe |
| GCA\_001224485.1 | OXC6622 | 2011 | United Kingdom | No | Europe |
| GCA\_007844205.1 | PUCV-1 | 2018 | Chile | No | South America |
| GCA\_007844175.1 | PUCV-2 | 2018 | Chile | No | South America |
| GCA\_007844165.1 | PUCV-3 | 2017 | Chile | No | South America |
| GCA\_001563565.1 | RM3194 | 1994 | South Africa | No | Africa |
| GCA\_001299595.1 | RM3196 | 1994 | South Africa | GBS | Africa |
| GCA\_001299565.1 | RM3197 | 1995 | South Africa | GBS | Africa |
| GCA\_006370335.1 | SP17-196 | 2017 | USA | No | North America |
| GCA\_012692385.1 | ST122 (1902M28538) | 2019 | Denmark | No | Europe |
| SAMN15508151 | OBT12377\_S20 | 2019 | Peru | GBS | South America |
| SAMN15508152 | OBT12386\_S24 | 2019 | Peru | GBS | South America |
| SAMN15508153 | OBT12390\_S28 | 2019 | Peru | GBS | South America |
| SAMN15508154 | OBT12393\_S32 | 2019 | Peru | GBS | South America |

**References**

1 Parkhill, J. *et al.* The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. *Nature* **403**, 665-668, doi:10.1038/35001088 (2000).

2 Gundogdu, O. *et al.* Re-annotation and re-analysis of the Campylobacter jejuni NCTC11168 genome sequence. *BMC Genomics* **8**, 162, doi:10.1186/1471-2164-8-162 (2007).

3 Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121-4123, doi:10.1093/bioinformatics/bty407 (2018).

4 Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100, doi:10.1093/bioinformatics/bty191 (2018).

5 Parker, C. T. *et al.* Comparison of Campylobacter jejuni lipooligosaccharide biosynthesis loci from a variety of sources. *J Clin Microbiol* **43**, 2771-2781, doi:10.1128/JCM.43.6.2771-2781.2005 (2005).

6 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

7 Zhang, M. *et al.* Association of Anti-GT1a Antibodies with an Outbreak of Guillain-Barre Syndrome and Analysis of Ganglioside Mimicry in an Associated Campylobacter jejuni Strain. *PLoS One* **10**, e0131730, doi:10.1371/journal.pone.0131730 (2015).

8 Gilbert, M. *et al.* The genetic bases for the variation in the lipo-oligosaccharide of the mucosal pathogen, Campylobacter jejuni. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide. *J Biol Chem* **277**, 327-337, doi:10.1074/jbc.M108452200 (2002).