

# Application of whole genome sequencing to fully characterise campylobacter isolates from IID1 and IID2 studies

Area of research interest: [Foodborne pathogens](#)

Study duration: 2014-01-01

Planned completion: 1 February 2015

Project code: FS101072

Conducted by: University of Liverpool

[Back to top](#)

## Background

Campylobacter is the most common cause of acute bacterial gastroenteritis worldwide. In the UK alone it causes an estimated 500,000 infections each year. There have been two large studies of Infectious Intestinal Disease in the UK community (IID1 in the mid 1990s and IID2 in 2008-2009). Campylobacter was identified as the most common bacterial pathogen amongst patients presenting to primary care. Although there was little variation in the burden of illness between the two studies, the molecular epidemiology of the campylobacter isolates from these studies has not been investigated. MLST and (WGS) comparative analyses can be used to understand not only the epidemiology and any variations between the two survey periods, but also the potential sources for transmission of the pathogen to humans, by comparison with isolates from the environment, wildlife and farm animals.

[Back to top](#)

## Research Approach

The overall aims of this study were to:

- fully characterise UK campylobacter strains associated with human campylobacteriosis by whole genome sequencing using Next Generation Sequencing (NGS) technologies
- identify markers to assist with source attribution by integration of the data from IID study strains with published data obtained from non-human sources and genome data being derived from current funded projects at Liverpool

Genomic DNA from all of the campylobacter isolates from the IID1 and IID2 studies were isolated. All isolate genomes were then subjected to sequencing of paired-end libraries using the Illumina MiSeq platform. After this initial run, the genomes of a sub-set of isolates (approximately 25) were improved using a PacBio approach. Using this combination approach we constructed a comprehensive genome sequence data-set from the IID isolates, and these data were submitted to a publically accessible database. We also extracted and analysed MLST data from the IID isolates, in order to place the collection in the context of previous studies. Both MLST and genomic data were used to compare between the IID1 and IID2 collections.

The major outcomes of the study were:

- a comprehensive genome sequence dataset from the IID isolates, submitted to a publically accessible database
- analysis of MLST data from the IID isolates to place the collection in the context of previous studies based on MLST
- genome-wide phylogenetic analysis of the IID strains compared to others available in the wider database (and our parallel studies involving isolates from the environment, wildlife and farm animals)

In addition, we considerably enhanced the community's knowledge on what constitutes the core genome of campylobacter, especially in relation to isolates associated with human infections, with the potential to link variations between strains (either in accessory genome content, or in SNP variations within the core genome) with other factors such as putative source or, potentially, clinical severity, as well as other important phenotypes, such as survival characteristics in the environment or during food processing.

[Back to top](#)

## Results

WGS of all available campylobacter isolates from the IID1 and IID2 studies was carried out using the Illumina platform. From the 504 samples received, WGS data was obtained for 470 campylobacter isolates, comprising 351 from IID1 and 119 from IID2. Of these 416 were *C. jejuni* and 46 were *C. coli*. We also obtained WGS data from five *C. upsaliensis*, one *C. fetus*, one *Arcobacter butzleri* and one *Arcobacter* spp.

Analysis of MLST data extracted from the WGS data indicated that the most common clonal complexes found amongst the IID1 and IID2 strains reflected their abundance amongst the wider Campylobacter population. There were no clear variations of note between the IID1 and IID2 isolate collections with respect to breakdown according to MLST clonal complex (CC). The use of Single Nucleotide Polymorphism (SNP) phylogeny to cluster *C. jejuni* genomes based on either ribosomal loci (rMLST), or larger sets of core genes, confirmed the broad distribution of IID1 and IID2 isolates amongst the wider Campylobacter population, but highlighted some sub-divisions within MLST-based clonal complexes, and some small clusters specific to either IID1 or IID2. SNP-based phylogenetic analysis of *C. coli* isolates from IID1 and IID2 indicated that they all cluster within Clade 3, a clade associated previously with clinical isolates and agricultural sources.

Using PacBio sequencing, a further 17 high quality reference genomes were added to the general database, eleven of which (ten *C. jejuni* and one *C. coli*) assembled as a single chromosome. From 14 high quality PacBio genomes and three previously available complete genomes, we defined a *C. jejuni* core genome of 1,261 genes.

These data provide an excellent baseline for monitoring shifts in the UK population of campylobacter associated with gastrointestinal infections. By combining survey data of this nature with analyses of other isolate collections from non-human sources, it will be possible to identify changing trends and shifts in the relative importance of potential sources of transmission.

[Back to top](#)

## Research report

## **England, Northern Ireland and Wales**

PDF

[View Application of whole genome sequencing to fully characterise Campylobacter isolates from the UK Infectious Intestinal Disease as PDF\(Open in a new window\)](#) (1.18 MB)

## **England, Northern Ireland and Wales**

EXCEL

[View FS101072 - Appendix 1 as Excel\(Open in a new window\)](#) (107.35 KB)

## **England, Northern Ireland and Wales**

PDF

[View FS101072 - Appendix 2 as PDF\(Open in a new window\)](#) (483.32 KB)

## **England, Northern Ireland and Wales**

PDF

[View FS101072 - Appendix 3 as PDF\(Open in a new window\)](#) (96.74 KB)

## **England, Northern Ireland and Wales**

PDF

[View FS101072 - Appendix 4 as PDF\(Open in a new window\)](#) (14.7 KB)

## **England, Northern Ireland and Wales**

PDF

[View FS101072 - Appendix 5 as PDF\(Open in a new window\)](#) (18.79 KB)

## **England, Northern Ireland and Wales**

PPT

[View FS101072 - Appendix 6 as PPT\(Open in a new window\)](#) (567.75 KB)