

# Comparative Analysis of 40 Genomes provides new insights into biomass degradation by *Clostridia*

## Background

- The genus *Clostridium* contains the most widely studied biomass-degrading organisms and includes *Clostridium thermocellum*, a potential biofuel-producing organism.

## Approach

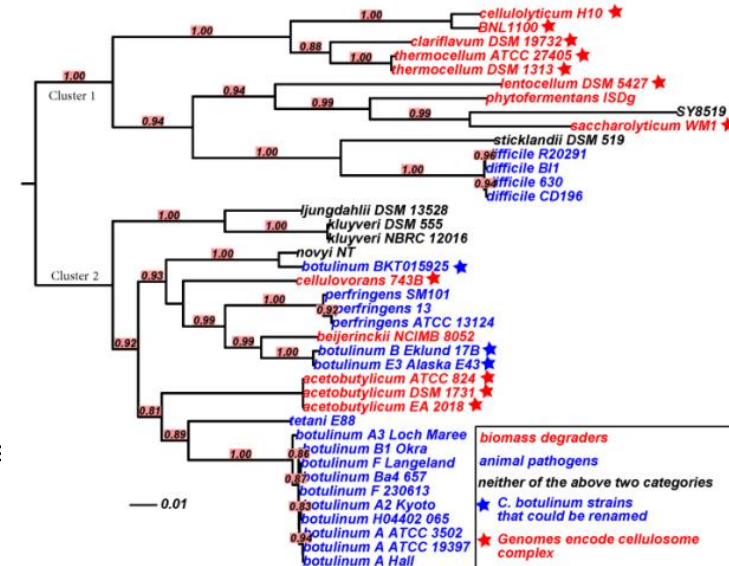
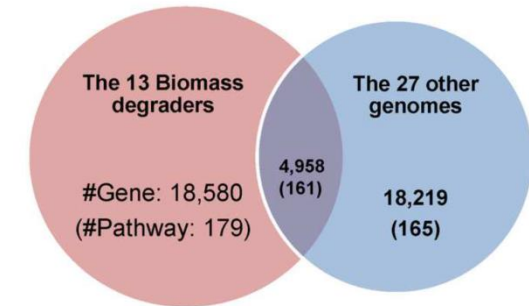
- 40 sequenced clostridial genomes available from NCBI as of April 2012 were selected for the comparative study including 13 biomass degraders, 21 pathogens and six others.

## Outcomes

- Comparative analyses revealed that biomass degraders tend to have larger genome sizes, more genes, higher GC content and high percentage of secretory proteins than other clostridial strains. They also have more % genes that are CAZymes and cellulosomal genes.
- The biomass degraders also have larger pan-genomes and encode more metabolic pathways than all other sequenced *Clostridia* combined.
- The study suggests that three of the *Clostridium botulinum* strains should be renamed according to the 16S rRNA phylogeny since these three strains are separated in the phylogeny tree from the other *C. botulinum* strains and differ in genome size, GC content, and the number of genes.

## Significance

- New insights provided by this analysis could provide useful information to mechanistic studies of biomass degradation and the range of metabolic pathways in *Clostridia*. This range of pathways may aid to designing or improving more efficient biomass degraders.
  - Note: Over 100 clostridial genomes have now been sequenced (2014); therefore these types of analyses will need to be repeated on a regular basis as more genomes become available.



16S rRNA tree of the 40 *Clostridium* genomes: biomass degraders (red), animal pathogens (blue), other (black)