Clostridium perfringens Whole Genome Sequencing improves outbreak source attribution

A case for including C. perfringens in the GenomeTrakr database
Acknowledgments

The folks that did the work
Jaclyn Carey
Jocelyn Cole

WGS enterics group
Samantha Wirth
Michelle Dickinson

GalaxyTrakr
Jimmy Sanders
Justin Payne
Hugh Rand
Dan Benisatto
Charles Strittmatter

Sequencing Core
Matt Shudt supervisor
Nathalie Boucher
Melissa Leisner
Helen Ling
Joshua Williams
Zhen Zhang

Bacteriology Laboratory
Kimberlee Musser, Director
Nellie Dumas, Associate Director
Lisa Mingle Enterics Supervisor
Meet Priscilla Perfringens

- 2nd most common cause of bacterial food borne illness
  - About 1,000,000 cases/yr in the US
  - Poor food handling

- Widely distributed in the environment

- About 40% of peoples are asymptomatic carriers

- No ongoing surveillance
  - Investigate outbreaks at request of epi. or clinicians
At the Wadsworth we characterize suspect food and stool samples

1) Screen primary samples by rt-PCR
2) Enumerate to confirm etiology
   I. Food $10^5$ cells/gr
   II. Stool $10^6$ spores/gr
3) Isolate
4) Confirm ID with Maldi-TOF

THERE IS NO TYPING BELOW THE SPECIES LEVEL.

5) Source attribution can be challenging
   I. Wide environmental distribution
   II. Asymptomatic carriers
A retrospective analysis of a small cohort of outbreak and sporadic samples

*C. perfringens* samples collected between 2010 - 2016

- 37 clinical
- 15 food
- 7 distinct foodborne outbreaks

- Extracted on QIAcube & Sequenced on MiSeq
- Reference based hqSNP analysis using the CFSAN SNP Pipeline
- Implemented in GalaxyTrakr
Seven outbreaks from 2010 to 2016

- 28 isolates from 7 distinct foodborne outbreaks
- 24 sporadic isolates
- Outbreaks from
  I. 3 restaurants
  II. 2 jails
  III. 1 wedding
  IV. 1 cafeteria
WGS confirms genetic relatedness of outbreaks

- Company cafeteria and prison outbreaks

NY82540082  Chicken
NY82539841  Pulled Pork
1 SNP

NY82540655  Corn in Cup
NY82540577  Corn in Tray
0 SNPS
WGS confirms genetic relatedness of outbreaks

Restaurant associated

NY83906307 Stool
NY83906252 Stool
NY82540943 Stool
NY83906155 Prime Rib

0 - 1 SNPS
But WGS can show us something new

- Patients that are linked epidemiologically may not have the same genotype
- Are these folks carriers?
- Or was the source polyclonal?

Wedding outbreak

- Stool
- Roast Beef
- NY83899388
- NY82540322
- 0 SNPS
- 570 SNPS

Another jail

- Stool
- Mexican Corn
- NY83889968
- NY82539716
- NY83890009
- 3 SNPS
- 108 - 110 SNPS
A very large roast beef from the restaurant Easter dinner in 2016
Even a single food sample may harbor multiple genotypes

- One large roast beef (RB) cut into 4 sections
- Harbors 2 genotypes
Again a member of the outbreak cohort has a different genotype

- Thanksgiving restaurant outbreak 2016

**NY84139692** Stool

NY83909353

| NY82540655 | Stool |
| NY82540577 | Stool |
| NY83889968 | Stool |
| NY82539716 | Stool |
| NY83890009 | Stool |
| NY82541080 | Stool |

| NY84139278 | Stool |
| NY84139156 | Stool |
| NY84140474 | Stool |
| NY84141319 | Stool |
| NY84141151 | Stool |
| NY84139829 | Stool |
| NY84139376 | Stool |
| NY82581450 | Gravy |
| NY82541376 | MP |

0 - 7 SNPS

829 - 897 SNPS

MP = Mashed Potatoes
WGS shows relatedness regardless of spore count

- The same patient sampled on consecutive days

<table>
<thead>
<tr>
<th>Spore count</th>
<th>Date</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insufficient specimen</td>
<td>11/25/16</td>
<td></td>
</tr>
<tr>
<td>94,000,000 CFU/g</td>
<td>11/26/16</td>
<td></td>
</tr>
</tbody>
</table>

MP = Mashed Potatoes
What we have learned so far

• WGS clustering analysis is largely concordant with epidemiological clusters.
  • Using the CFSAN pipeline related samples are generally 0-3 SNP apart
  • But outliers exist
• WGS confirms relatedness even with low enumeration counts
  • These samples may not be considered part of the outbreak
• Multiple strains can be detected in a single source (giant roast beef)
How much genetic diversity can we detect in a primary sample?

• Multiple picks have been saved from the original enumeration
• We sequenced picks from RB#3 and patient stool
• Do the other picks match the second strain?
Multiple picks reveal genetic diversity within a sample

- Multiple picks generally have the same genotype, but not always

- What if we had only sequenced pick 3 from the stool?

SNPs:
- 0 - 3 SNPs
- 1 - 4 SNPs
- 767 - 781 SNPs
- 868 - 872 SNPs
Analysis of Picks from the Thanksgiving Outbreak

- Sequence picks from gravy and outlier stool
- Do the other picks match the second strain?

MP = Mashed Potatoes
Analysis of Picks from Thanksgiving Outbreak

- Same patient

Gravy picks:
- NY82581450
- NY84141151
- NY87455557
- NY84139278
- NY87455515
- NY84140474
- NY84141319
- NY87455467
- NY84139156
- NY84139829
- NY84139376
- NY87455425
- NY82541376

Stool picks:
- NY87454849
- NY87454897
- NY84139692
- NY87454804
- NY87454761

MP = Mashed Potatoes

- Multiple picks did not bring outlier into outbreak cluster

0 - 8 SNPS

890 - 975 SNPS

0 - 1 SNPS
How can WGS on *C. perfringens* be helpful at the Wadsworth Center?

- Refine epidemiological investigations
  - Since WGS is not used for surveillance, it won’t provoke an investigation
- Detect multiple strains in a single sample
  - Metagenomics?
- Show relatedness with insufficient specimen
Should *C. perfringens* be included in the GT database?

In support:
- Very common source of food borne illness
- Not much genomic data available, so would inform population structure etc.
- Likely to identify new vehicles or mechanisms for outbreaks
- No national data base available

The current approach is adequate:
- Disease is generally not severe and is self-limiting
- No national surveillance in place