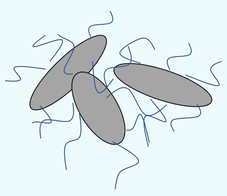
The genomic epidemiology of novel virulent *Clostridium difficile* ribotype 251 strains in Australia

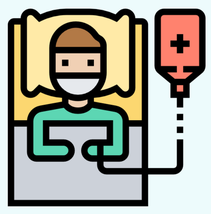
Stacey Hong[1], Korakrit Imwattana[1], Daniel R Knight[1, 2], Thomas V Riley[1, 2, 3, 4]

[1] School of Biomedical Sciences, The University of Western Australia, Crawley, Australia; [2] Medical, Molecular and Forensic Sciences, Murdoch University, Murdoch, Australia; [3] School of Medical and Health Sciences, Edith Cowan University, Joondalup, Australia; [4] PathWest Laboratory Medicine, Queen Elizabeth II Medical Centre, Nedlands, Australia.

BACKGROUND

In the last 20 years, *Clostridium difficile* infection (CDI) has reached epidemic proportions across continents, with significant increases in incidence and severity of disease following the emergence and global dissemination of the "hyper-virulent" *C. difficile* ribotype (RT) 027 strain, initially in North America and then Europe[1]. This strain never established in Australia and there is a different repertoire of circulating *C. difficile* strains causing human infection[2]. One such strain that has increased in prevalence Australia-wide is RT251[3]. Herein, comparative genomics and high-resolution core genome phylogenetics were used to examine the genomic diversity, evolutionary history and virulence potential of *C. difficile* RT251 strains in Australia.



MATERIALS & METHODS

A total of 48 *C. difficile* RT251 strains were sourced from Australia (n=28), USA (n=11), Slovenia (n=4), New Zealand (n=3) and Canada (n=1).

Genetic characterisation was performed using a variety of techniques including; toxin gene profiling[2], whole genome sequencing (WGS), *in silico* multi-locus sequencing (MLST), core-genome single nucleotide polymorphism (cgSNP) analysis, and detection of antimicrobial resistance genes[4].

Using the *C. difficile* molecular clock (within-host mutation rate), a cut-off of 0 – 2 cgSNPs between strains represents a recent clonal transmission event[4].

RESULTS

Table 1. *C. difficile* RT251 sequence type (ST) sub-lineages and allelic profiles. *RT027 was included as a comparator strain.

RT	Toxin profile	Clade	ST (n)	Housekeeping genes/Allelic profiles						
				adk	*atpA*	*dxr*	*glyA*	*recA*	*sodA*	*tpi*
251	$A^+B^+CDT^+$	2	188 (5)	1	1	9	9	14	5 (C381G)	2
		2	230 (1)	1	1	9	9	14	3	1 (A403G)
		2	231 (40)	1	1	9	9	14	3	2
		2	365 (2)	1	1	9	23 (C19A)	14	3	2
027*	$A^+B^+CDT^+$	2	1	1	1	1	10	1	3	5

Key epidemiological questions answered using WGS:

- Where did Australian *C. difficile* RT251 strains come from?
- What is the genetic relatedness between RT251 strains isolated in CDI patients in Australia?

MAJOR FINDINGS

- *C. difficile* RT251 strains clustered with RT027 in clade 2, also known as the "hyper-virulent" clade[5].
- The *C. difficile* RT251 lineage was differentiated into four sub-lineages by MLST (STs 188, 230, 231 and 365) [Table 1].
- All Australian *C. difficile* RT251 strains belonged to ST 231, suggesting a highly clonal population.
- The methyltransferase gene *ermB* was found in 3/48 (6.3%) RT251 strains, all of which belonged to the ST231 sub-lineage. This genotype is associated with the MLS_B (macrolide-lincosamide-streptogramin B-resistant) phenotype[4].
- A cgSNP maximum-likelihood phylogeny of 40 RT251/ST231 strains revealed a highly clonal population, consisting of five clonal groups (CGs 1 – 5) [**Figure 1**].
- CGs 1 – 4 consisted of genetically indistinguishable strains (≤2cgSNPs difference) isolated across five States in Australia separated by thousands of kilometres without evidence of geographical or temporal clustering [**Figure 1**]. This finding is consistent with a point source of contamination likely associated with the national food chain.
- One strain (US48★) that originated from Virginia, USA, in 2009 shared ≤2 cgSNP difference with RT251 strains isolated in Australia (CG4), suggesting a cross-continental transmission event and shared evolutionary ancestry.

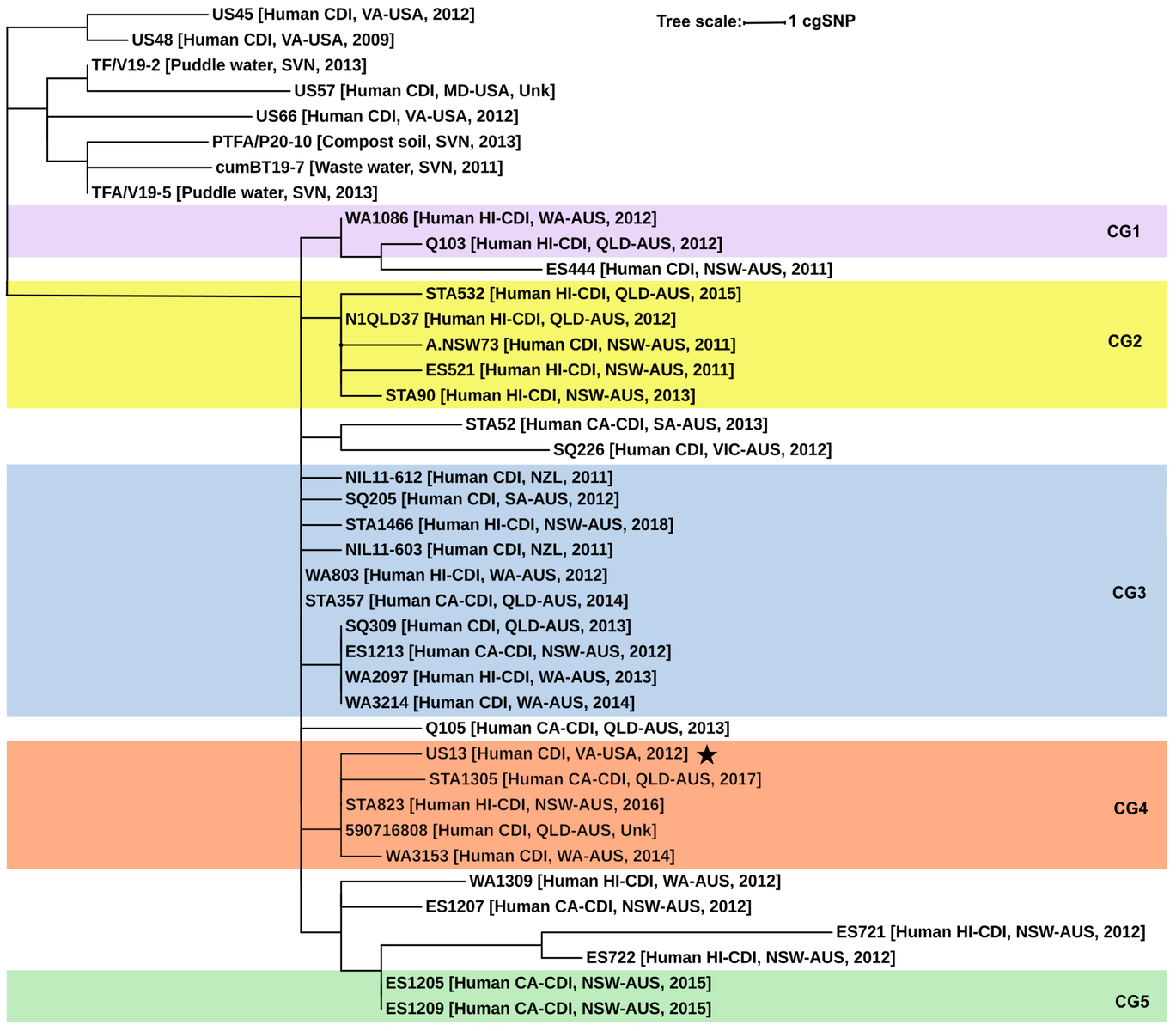


Figure 1. Population structure of 40 *C. difficile* RT251/ST231 strains based on cgSNPs.

Maximum-likelihood phylogeny based on non-repetitive, non-recombinant SNPs (n=86) identified after mapping all sequence reads against the R20291 (RT027) reference genome (accession number FN545816; 4,191,339 bp). RAxML tree is mid-point rooted and is supported by 100 nonparametric bootstrap replicates. Colour range based on CGs where all isolates differ by no more than four SNPs in their core genome.

REFERENCES

1. Freeman J *et al.* The changing epidemiology of *Clostridium difficile* infections. Clin Microbiol Rev. 2010. 23:529-49.
2. Hong S *et al.* Laboratory-based surveillance of *Clostridium difficile* Infection in Australian health care and community settings, 2013 to 2018. J Clin Microbiol. 2020. 21;58:e01552-20.
3. Wehrhahn MC *et al.* A series of three cases of severe *Clostridium difficile* infection in Australia associated with a binary toxin producing clade 2 ribotype 251 strain. Anaerobe. 2019. 55:117-123.
4. Knight DR *et al.* Evolutionary and genomic insights into *Clostridioides difficile* sequence type 11: a diverse zoonotic and antimicrobial-resistant lineage of global one health importance. *mBio*. 2019. 10:e00446-19.
5. Dingle KE et al. Clinical *Clostridium difficile*: clonality and pathogenicity locus diversity. PLoS One. 2011. 6:e19993.

UWA PERTH·AUSTRALIA | Murdoch UNIVERSITY PERTH WESTERN AUSTRALIA | AUSTRALIA ECU EDITH COWAN UNIVERSITY | PathWest LABORATORY MEDICINE WA

@iamstaceyhong