Journal of Infection (2016) xx, 1–10

ELSEVIER

BIA British Infection Association

www.elsevierhealth.com/journals/jinf

Genomic dissection of Australian *Bordetella pertussis* isolates from the 2008–2012 epidemic

Azadeh Safarchi [a], Sophie Octavia [a], Sunny Z. Wu [a], Sandeep Kaur [a], Vitali Sintchenko [b,c], Gwendolyn L. Gilbert [b,c], Nicholas Wood [b], Peter McIntyre [b], Helen Marshall [d], Anthony D. Keil [e], Ruiting Lan [a,*]

[a] *School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, New South Wales, Australia*
[b] *Centre for Infectious Diseases and Microbiology – Public Health, Institute of Clinical Pathology and Medical Research, Pathology West, Westmead Hospital, New South Wales, Australia*
[c] *Marie Bashir Institute for Infectious Diseases and Biosecurity, The University of Sydney, New South Wales, Australia*
[d] *Vaccinology and Immunology Research Trials Unit, Women's and Children's Hospital and School of Medicine and Robinson Research Institute, University of Adelaide, South Australia, Australia*
[e] *Department of Microbiology, PathWest Laboratory Medicine WA, Princess Margaret Hospital for Children, Perth, Australia*

Accepted 14 January 2016
Available online ■ ■ ■

KEYWORDS
Bordetella pertussis;
Molecular epidemiology;
Whole genome sequencing;
Microevolution;
Single nucleotide polymorphism

Summary *Objectives:* Despite high pertussis vaccination coverage, Australia experienced a prolonged epidemic in 2008–2012. The predominant *Bordetella pertussis* genotype harboured pertussis toxin promoter allele, *ptxP3*, and pertactin gene allele, *prn2*. The emergence and expansion of non-expressing *prn* isolates (Prn negative), were also observed. We aimed to investigate the microevolution and genomic diversity of epidemic *B. pertussis* isolates.
Methods: We sequenced 22 *B. pertussis* isolates collected in 2008–2012 from two states of Australia which are geographically widely separated. Ten of the 22 were Prn negative isolates with three different modes of silencing of *prn* (*prn*::IS*481*F, *prn*::IS*481*R and *prn*::IS*1002*). Five pre-epidemic isolates were also sequenced for comparison.
Results: Five single nucleotide polymorphisms were common in the epidemic isolates and differentiated them from pre-epidemic isolates. The Australian epidemic isolates can be divided into five lineages (EL1–EL5) with EL1 containing only Prn negative isolates. Comparison

* Corresponding author. Tel.: +61 2 9385 2095; fax: +61 2 9385 1483.
E-mail address: r.lan@unsw.edu.au (R. Lan).

http://dx.doi.org/10.1016/j.jinf.2016.01.005

Please cite this article in press as: Safarchi A, et al., Genomic dissection of Australian *Bordetella pertussis* isolates from the 2008–2012 epidemic, J Infect (2016), http://dx.doi.org/10.1016/j.jinf.2016.01.005

with global isolates showed that three lineages remained geographically and temporally distinct whereas two lineages mixed with isolates from 2012 UK outbreak.

Conclusion: Our results suggest significant diversification and the microevolution of *B. pertussis* within the 2008–2012 Australian epidemic.

Introduction

Bordetella pertussis, a small Gram-negative bacterium, is the causative agent of the respiratory infection, pertussis. The infection can be life threatening in infants and young children. Introduction of whole cell vaccine (WCV) during the 1950s significantly reduced the morbidity and mortality of pertussis globally. However, due to the side effect profiles of WCV, acellular vaccine (ACV) was developed in the 1980s.[1] In Australia, ACV replaced the WCV initially as a booster in 1997 and then for all scheduled doses by 2000.[2] Australia predominantly uses 3-component ACVs from GlaxoSmithKline containing detoxified pertussis toxin (Ptx), pertactin (Prn), filamentous haemagglutinin (FHA). The 5-component ACV from Sanofi – Aventis containing additionally 2 types of fimbriae (Fim2 and Fim3) has also been used in Australia.[3,4]

Re-emergence of pertussis has been reported in many countries with high vaccination coverage, including the US, Canada, Japan, European countries and Australia.[5–9] Despite high pertussis vaccine uptake in Australia, (91–92%) pertussis incidence is still the highest amongst all vaccine-preventable diseases. The latest pertussis epidemic commenced in 2008 and reached its peak in 2011.[10,11] In the state of New South Wales (NSW), the pertussis notification and hospitalization rate were 2.7 and 3.9 times higher, respectively, as compared to the previous five-year average. In addition, there was a significant increase in notification and hospitalisation rates for infants aged less than one-year-old.[10]

Although it has been suggested that improvement in diagnostic laboratory techniques and increased awareness by general practitioners may explain the high rate of pertussis in developed countries, there is evidence that organism adaptation and antigenic drift may increase the incidence of pertussis in the highly immunised population.[12–14] Strain variation and pathogen adaptation have been reported in different countries as evidenced by polymorphisms in several virulence associated genes and their promoters, including those included in the ACV: *ptx* genes and its promoter *ptxP*, *prn*, *fhaB* and *fim*[14–16]. Recent studies in several countries have reported the emergence and increasing circulation of isolates that do not express Prn.[17–22] Comparative genomics of pre-vaccination and modern *B. pertussis* strains showed that single nucleotide polymorphism (SNP) in important genes including virulence-associated genes or regulatory genes may have helped the organism to survive under vaccine selection pressure.[23–25]

Previously, SNP typing classified 208 Australian *B. pertussis* isolates collected since the 1970s into 30 SNP profiles (SPs) which are further grouped into five SNP clusters.[26] An increase in prevalence of SNP cluster I was documented after the introduction of ACV.[26] Australia experienced a prolonged outbreak from 2008 to 2012. Typing of 194 *B. pertussis* isolated collected from the epidemic found three predominant SPs with SP13, SP14 and SP16 representing 49%, 24% and 12% of the total isolates genotyped.[26] During the epidemic, Prn negative isolates emerged in 2008 and increased to 78% by 2012.[27]

In this study, we used whole genome sequencing to investigate the genetic diversity of *B. pertussis* isolates associated with the Australian pertussis epidemic of 2008–2012 and to determine evolutionary characteristics of Australian epidemic *B. pertussis*. We sequenced 22 SP13 *B. pertussis* isolates collected in 2008–2012 from two states of Australia, NSW and Western Australia (WA) which are geographically widely separated. Ten of the 22 were Prn negative isolates which allowed us to determine the origin and relationships of these Prn negative isolates.

Materials and methods

Bacterial strains

A total of 27 *B. pertussis* SP13 isolates were selected based on year and state of isolation and the inactivation mechanism of *prn* gene. Amongst the 27 selected *B. pertussis* SP13 isolates, 22 of them represented isolates from the Australian 2008–2012 epidemic and are referred to henceforth as epidemic isolates. The remaining five isolates represented isolates prior to the 2008–2012 epidemic and are referred to as pre-epidemic isolates. Bacterial isolates were maintained on Bordet-Gengou agar (Oxoid) supplemented with 10% horse blood (Oxoid) at 37 °C for 3–5 days. Genomic DNA was extracted and purified from pure culture using the phenol-chloroform method as described previously.[28]

DNA sequencing and assembly

DNA libraries were constructed with the insert size of 250 bp paired-end using NexteraXT kit (Illumina) and sequenced on the MiSeq (Illumina). Genome sequencing was done in a multiplex of 24. *De novo* assembly was performed for all sequencing data using Velvet version 1.2.08[29] to combine reads into contigs. These contigs were aligned to the reference *B. pertussis* strain Tohama I (GeneBank accession number BX470248) using progressiveMauve (version 2.3.1).[30] The strain Tohama I was used as a reference to generate comparable data as all previous genomic studies used Tohama I as a reference. Some studies revealed that there are some genomic regions present in *B. pertussis* strains but are not found in Tohama I.[31,32] Therefore, the genome of *B. pertussis* strain CS,[33] which is fully sequenced, was also used to investigate the genomic diversity among these regions of differences. The

Please cite this article in press as: Safarchi A, et al., Genomic dissection of Australian *Bordetella pertussis* isolates from the 2008–2012 epidemic, J Infect (2016), http://dx.doi.org/10.1016/j.jinf.2016.01.005