Title: Complete genome of the *Bordetella pertussis* strain used to produce the Brazilian cellular vaccine

Authors Akamatsu, MA[1], Nishiyama-Jr, MY[1], Kitajima, JP[2], Morone, MSLC[1], Oliveira UC[1], Bezerra, MFB[1], Sakauchi, MA[1], Raw, I[1], Junqueira-de-Azevedo,ILM[1], Carvalho, E[1], Ho, PL[1]

Institutions: [1]Instituto Butantan, São Paulo, SP, Brazil; [2]Mendelics, São Paulo, SP, Brazil

Abstract:

Whooping cough is a respiratory disease caused by *Bordetella pertussis*. Despite the reduction in the incidence of pertussis after the introduction of vaccination, the risk of outbreaks is still a public health concern and several countries have reported an increasing incidence of this disease. The recent changes in the epidemiological diversity of the circulating *B. pertussis* suggest that the selective pressure promoted by the widespread vaccination led to changes in genome diversity, including the in virulence factors. In this scenario, it is important to characterize the strains used to produce the vaccines, to understand the selective pressure caused by the vaccines and the possible epidemiological changes related to it. We sequenced and characterized the genome of the *B. pertussis* strain Bp137, used to produce the cellular pertussis vaccine in Brazil, at Butantan Institute. The whole genome was sequenced using 454, Illumina and PacBio technologies; the assembly was performed with CLC Genomics Workbench. A single contig of 4.13Mb was assembled, 3,939 genes were annotated, being 3,703 proteins-coding sequences and 175 pseudogenes, in addition to tRNA and rRNA genes. This genome represent the 4th complete *B. pertussis* genome. Compared to the *B. pertussis* strain Tohama (the reference for pertussis genome) we found 43 positional changes, including 22 inversions. When two other strains – CS (Chinese vaccine strain) and 18323 (virulent strain) – were added to this analysis, we observed that Bp137 has the more divergent genome, especially when compared to Tohama and CS, which are mainly collinear genomes. Additionally, it was identified 1,254 single-nucleotide polymorphisms between Bp137 and Tohama. A clustering analysis, based on presence and absence of genes, also showed that Bp137 is the most dissimilar strain, compared to Tohama, CS and 18323. Indeed, some key virulence factors, such as Pertussis Toxin, Pertactin and Fimbriae 2, have different alleles when comparing Bp137 to Tohama. It is important to notice that some specific/different alleles are in generally associated with pertussis resurgence and outbreaks. We conclude that the strain used to produce the Brazilian vaccine have differences in genomic level and in virulence factors profile when compared to other vaccine strains. In a scenario in which new strains are becoming an epidemiological concern, diversity in vaccine strains could be strategically important to amplify global vaccine coverage.

Key-word: *Bordetella pertussis*, Genome, Vaccines production, whooping cough

Supported by: Fundação Butantan and FAPESP