Genome characteristics of Bordetella pertussis isolates from Tunisia

Ikram Ben Fraj, Valérie Bouchez, Hanen Smaoui, Amel Kechrid, Sylvain Brisse

Genome characteristics of *Bordetella pertussis* isolates from Tunisia

Ikram Ben Fraj[1], Valérie Bouchez[2,3], Hanen Smaoui[1], Amel Kechrid[1] and Sylvain Brisse[2,3,*]

[1] University of Tunis El Manar, Children's Hospital of Tunis, Laboratory of Microbiology, UR12ES01, Tunis, Tunisia

[2]Institut Pasteur, Biodiversity and Epidemiology of Bacterial Pathogens, Paris, France

[3]National Reference Center for Whooping Cough and other Bordetella infections, Paris, France

*** Corresponding author**: Sylvain Brisse. Biodiversity and Epidemiology of Bacterial Pathogens, Institut Pasteur, 25 rue du Docteur Roux, F-75724 Paris, France.

E-mail: sylvain.brisse@pasteur.fr; Phone +33 1 45 68 83 34

Running title: *Bordetella pertussis* isolates from Tunisia

Keywords: *Bordetella pertussis*, Tunisia, genomic epidemiology, phylogeny

Accession numbers

The genomic sequence data generated in this work were submitted to the European Nucleotide Archive and are available from the International Nucleotide Sequence Database Collaboration (NCBI/ENA/DDBJ) databases under project accession number PRJEB27412 and run data accession numbers ERS2572942 to ERS2572951.

Abstract

Purpose: The recent increase in pertussis cases observed in some countries may have several causes, including the evolution of *Bordetella pertussis* populations towards escape of vaccine-induced immunity. Most genomic studies of *B. pertussis* isolates performed so far are from countries that use acellular vaccines. The objective was to analyze genomic sequences of isolates collected during the 2014 whooping cough epidemic in Tunisia, a country where whole cell vaccines are used. **Methodology**: Ten Tunisian isolates and four vaccine strains were sequenced and compared to 169 isolates from countries where acellular vaccines are used. **Results**: Phylogenetic analysis showed that Tunisian isolates are diverse, demonstrating a multi-strain 2014 epidemic peak, and are intermixed with those circulating in other world regions, showing inter-country transmission. Consistently, Tunisian isolates have antigen variant composition observed in other world regions. No pertactin-deficient strain was observed. **Conclusion**: The Tunisian *B. pertussis* population appears to be largely connected with populations from other countries.

Introduction

Pertussis is a highly contagious infectious disease, caused by the bacterium *Bordetella pertussis* and more rarely by *B. parapertussis*. Although the infection is largely controlled by vaccination, a recent increase in pertussis incidence has been observed in several countries [1]. Possible explanations of this increase include sub-optimal vaccine coverage, improved surveillance and diagnosis of the disease, waning of vaccine-induced immunity, and genetic variations in circulating *B. pertussis* isolates that escape vaccine-induced immunity [2–4].

In Tunisia, pertussis is still present as a cyclical disease despite high (98%) primo-vaccination coverage [5], and two epidemic peaks were observed in 2009 and 2014. A 9-year surveillance study (2007-2016) among 1,844 infants and children reported 306 (16.6%) *Bordetella* infection that were confirmed by real time PCR [6, 7]. While initial vaccination of Tunisian children is performed at 2, 3 and 6 months of age, a booster dose is received at the age of 18 months. Since April 2011, a pentavalent vaccine (Pentavac, Serum Institute of India), which includes a whole cell pertussis component, is used in Tunisia.

B. pertussis is an highly monomorphic bacterium [4, 8]. Therefore, molecular typing of *B. pertussis* isolates requires the full resolution of whole genome sequencing (WGS) for accurate phylogenetic comparisons and to follow the spread of sublineages, as other methods largely fail to resolve relationships at strain level [4, 9–11]. Inter-country transmission of *B. pertussis* isolates was demonstrated based on genomic-scale genotyping of a global dataset of isolates collected before 2011 [4] and of isolates collected more recently in countries using acellular vaccines [9, 11, 12]. However, only few data are currently available on the genomic sublineages of *B. pertussis* isolates collected in countries using whole cell vaccines [13, 14]. Therefore, whether transmission occurs between these countries and those using acellular vaccines is not enough documented, limiting our understanding of the drivers of *B. pertussis* diversity across countries with distinct vaccination strategies.

This study aimed to determine the genomic characteristics of Tunisian *B. pertussis* clinical isolates and to compare them with isolates from other world regions and with strains used to prepare the whooping cough vaccine used in Tunisia.

Methods

***B. pertussis* isolates and genomes.** Tunisian *Bordetella pertussis* clinical isolates were recovered during the epidemic peak of 2014. Their phenotypic characteristics and antigen genotyping of genes *ptx*P, *prn*, *ptx*A, *fim*2 and *fim*3 were described earlier [7]. Genomic DNA was isolated and purified from subcultured isolates [7] using the MagNA Pure 96 DNA and Viral NA Small Volume Kit (Roche, Germany) according to the manufacturer's instructions. Sequencing libraries were prepared using the Nextera XT DNA Sample Preparation Kit (Illumina, USA). Whole genome sequencing was performed on the Illumina® NextSeq® 500 system (Illumina, USA) using a 2×150 paired-end protocol by the Mutualized Platform for Microbiology at Institut Pasteur of Paris. AlienTrimmer v0.4.0, Musket v1.1 and KhmerStream v1.1 software packages were used for paired-end reads clipping, trimming and sequencing errors correction, respectively. SPAdes/3.9.0 was used for genome assembly. Genome sequences were deposited in the European Nucleotide Archive and their accession numbers are available in Table 1.

cgMLST and phylogenetic analysis. Genome sequence data were analyzed using a gene-by-gene approach known as core genome MLST (cgMLST) [11, 15]. cgMLST was performed using the BIGSdb platform of Institut Pasteur (http://bigsdb.pasteur.fr/bordetella) as previously described [11]. Briefly, genome assemblies were compared using BLASTN to the reference alleles of 2,038 predefined gene loci. When novel alleles were discovered in the

genomes of the Tunisian isolates, they were imported into the reference allele database and an allelic number was assigned. Isolates from other world regions were investigated in a previous study and corresponded to isolates from France, UK and USA [11]. Genomic sequences of vaccine strains from Serum Institute of India J445 (GCA_001831395), J446 (GCA_001831415), J447 (GCA_001831435) and J448 (GCA_001831455) were also included. We used IQ-TREE v1.5.4 to infer a maximum likelihood phylogenetic tree based on concatenated alignments of the sequences from each of the 2,038 cgMLST loci.

Results and Discussion

Ten isolates collected during the epidemic of 2014 at the Children's Hospital of Tunis, the only pediatric university hospital in Tunisia, were analyzed by genome sequencing. The characteristics of genomes are summarized in **Table 1**. Assemblies characteristics were similar for the ten isolates: about 300 contigs were obtained for each genome, with a mean length of 3,867,263 bp. **Figure 1** presents the genomic sequence-based phylogenetic tree of Tunisian *B. pertussis* isolates and their positioning with respect to isolates from other world regions. Tunisian isolates were all distributed in the *ptxP3* clade but were genetically diverse, with eight of them falling within the *fim3-2* clade and two within the *fim3-1* clade. TN0003 and TN0005 were closely related, with only 4 different cgMLST genes among them, whereas they were more distant from other isolates (11 to 14 different loci). Likewise, isolates TN0007, TN0001, TN0006 and TN0011 were closely related to each other within the *fim3-2* branch, displaying only 1 or 2 different loci among themselves. These two groups of genetically very similar Tunisian isolates may correspond to local chains of transmission [11]. Nevertheless, the 10 Tunisian isolates were mixed phylogenetically with isolates from other countries, indicating several independent transmissions between Tunisia and other countries. The closest neighbors of the Tunisian strains in this dataset were either from France, UK or

the USA, and diverged from them by 3 to 8 loci (**Table 2)**. This close genetic relatedness between Tunisian isolates and isolates from other geographic origins is consistent with the high global genetic homogeneity of *B. pertussis* and is likely favored by the increased international exchanges with Tunisia. Hence, despite the monomorphic nature of *B. pertussis,* the genome sequencing approach shows genetic heterogeneity among the Tunisian isolates included in this study, demonstrating that the epidemic peak of 2014 was not due to the spread of a single strain but instead, was caused by the simultaneous infection by various strains as observed previously [9, 11].

B. pertussis isolates can be characterized by their allelic profile, which recapitulates the allelic variants at virulence and antigen genes loci. As previously observed [7], the allelic profile of all Tunisian isolates is *ptxP*3-*ptxA*1-*prn*2-*fim*2-1-*fim*3-2. Notably, all Tunisian isolates had alleles *ptxA1* for the subunit 1 of pertussis toxin, and pertactin allele *prn2*, irrespective of the *fim3* clade they belong to, as observed for the majority of current *B. pertussis* isolates collected in countries using acellular *B. pertussis* vaccine [4, 11, 12]. An increase in reported *B. pertussis* isolates not producing pertactin has been observed in many countries using the acellular pertussis vaccine, including the USA [16], Japan [12] and European countries [17]. All Tunisian *B. pertussis* isolates were shown to produce pertactin [7], and the genomic sequence confirmed that they all had an intact *prn* gene. The low number of isolates from Tunisia that could be analyzed does not allow to exclude the presence of pertactin non-producing isolates in this country, but the results indicate that they are much less frequent than pertactin-producing ones. This situation therefore contrasts with countries where acellular vaccines are used since nearly two decades such as Japan, USA or some EU countries [17–19]. Therefore, the inter-country transmission of *B. pertussis* sublineages is not intense enough not erase the specificities of *B. pertussis* populations in countries with distinct vaccination strategies.

The four vaccine strains (J445 to J448) incorporated in the vaccine used in Tunisia appear to belong to early-branching clades (**Figure 1**), as does the Tohama reference strain previously shown to be distant from most *B. pertussis* isolates that currently cause infection [4, 20]. Genetic distances of the four vaccine strains to Tunisian isolates were higher for strain J446 (117 to 124 cgMLST gene mismatches) than for strain J445 (28 to 31 mismatches) and for strains J447 and J448 (17 to 24 mismatches). These genetic distances and the phylogenetic position of vaccine strains are consistent with the promoter sequence or antigenic variants of vaccine strains (J446: *ptxP*2/*ptxA4*/*prn7*/*fim*2-2/*fim*3-1; J445: *ptxP*1/*ptxA2*/*prn1*/*fim*2-1/*fim*3-1; J447 and J448: *ptxP1*/*ptxA1/prn1/fim2-1/fim3-1*), which do not match entirely with the profiles of most circulating isolates (*ptxP*3/*ptxA*1/*prn*2/*fim*2-1/*fim*3-2). Vaccine efficacy might be improved by the use of vaccine strains that have a better phylogenetic and antigenic match with the circulating isolates.

In conclusion, although based on a limited collection of isolates, this study provides a snapshot of the genomic make-up of Tunisian isolates. Despite the use of whole cell vaccines in Tunisia, *B. pertussis* isolates from this country are phylogenetically and antigenically closely related to those circulating in countries using acellular vaccines, whereas they are less related to vaccine strains. However, our sample of *B. pertussis* isolates did not reveal the presence of pertactin-deficient strain in Tunisia. The relative contributions of long-range transmission and local selection by vaccine-induced immunity in shaping *B. pertussis* populations is an important subject for future research. Surveillance of pertussis infections in Tunisia and other countries using whole-cell vaccines should be reinforced to inform control measures and vaccination policy.

Author statements

Author contributions: Conceptualization: IBF, AK, HS and SB; Experiments: IBF, VB; Data analysis: IBF, VB; Original draft preparation: IBF, VB, HS, SB; Review and editing: SB; Supervision: HS, SB; Funding: SB.

Conflict of Interest: The authors declare that there are no conflicts of interest.

Funding information: This research received no specific grant. It was supported financially by Institut Pasteur support to the research unit Biodiversity and Epidemiology of Bacterial Pathogens.

Ethical approval: not relevant

Acknowledgements: We thank Annie Landier and Sophie Guillot (National Reference Center for Whooping Cough and other Bordetella infections, Institut Pasteur) for their continuous implication in data collection, strain characterization and genomic sequencing.

References

1. **Domenech de Cellès M, Magpantay FMG, King AA, Rohani P**. The pertussis enigma: reconciling epidemiology, immunology and evolution. *Proc R Soc B Biol Sci*;283. Epub ahead of print 13 January 2016. DOI: 10.1098/rspb.2015.2309.

2. **Mooi FR**. Bordetella pertussis and vaccination: the persistence of a genetically monomorphic pathogen. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2010;10:36–49.

3. **Cherry JD**. Pertussis: Challenges Today and for the Future. *PLOS Pathog* 2013;9:e1003418.

4. **Bart MJ, Harris SR, Advani A, Arakawa Y, Bottero D, *et al.*** Global Population Structure and Evolution of Bordetella pertussis and Their Relationship with Vaccination. *mBio* 2014;5:e01074-14.

5. **World Health Organization**. WHO and UNICEF estimates of immunization coverage, Tunisia: 2016 revision. https://data.unicef.org/wp-content/uploads/country_profiles/Tunisia/immunization_country_profiles/immunization_tun.pdf (2016, accessed 28 April 2018).

6. **Zouari A, Smaoui H, Brun D, Njamkepo E, Sghaier S, *et al.*** Prevalence of Bordetella pertussis and Bordetella parapertussis infections in Tunisian hospitalized infants: results of a 4-year prospective study. *Diagn Microbiol Infect Dis* 2012;72:303–317.

7. **Ben Fraj I, Kechrid A, Guillot S, Bouchez V, Brisse S, *et al.*** Pertussis epidemiology in Tunisian infants and children and characterization of Bordetella pertussis isolates: results of a 9-year surveillance study, 2007 to 2016. *Journal of Medical Microbiology*. Epub ahead of print 2018. DOI: 10.1099/jmm.0.000892.

8. **Diavatopoulos DA, Cummings CA, Schouls LM, Brinig MM, Relman DA, *et al.*** Bordetella pertussis, the causative agent of whooping cough, evolved from a distinct, human-associated lineage of B. bronchiseptica. *PLoS Pathog* 2005;1:e45.

9. **Sealey KL, Harris SR, Fry NK, Hurst LD, Gorringe AR, *et al.*** Genomic Analysis of Isolates From the United Kingdom 2012 Pertussis Outbreak Reveals That Vaccine Antigen Genes Are Unusually Fast Evolving. *J Infect Dis* 2015;212:294–301.

10. **Xu Y, Liu B, Gröndahl-Yli-Hannuksila K, Tan Y, Feng L, *et al.*** Whole-genome sequencing reveals the effect of vaccination on the evolution of Bordetella pertussis. *Sci Rep* 2015;5:12888.

11. **Bouchez V, Guglielmini J, Dazas M, Landier A, Toubiana J, *et al.*** Genomic Sequencing of Bordetella pertussis for Epidemiology and Global Surveillance of Whooping Cough. *Emerg Infect Dis* 2018;24:988–994.

12. **Zomer A, Otsuka N, Hiramatsu Y, Kamachi K, Nishimura N, *et al.*** Bordetella pertussis population dynamics and phylogeny in Japan after adoption of acellular pertussis

vaccines. *Microb Genomics*. Epub ahead of print 17 May 2018. DOI: 10.1099/mgen.0.000180.

13. **Mosiej E, Krysztopa-Grzybowska K, Polak M, Prygiel M, Lutyńska A**. Multi-locus variable-number tandem repeat analysis of Bordetella pertussis isolates circulating in Poland in the period 1959-2013. *J Med Microbiol* 2017;66:753–761.

14. **Dienstbier A, Pouchnik D, Wildung M, Amman F, Hofacker IL, *et al.*** Comparative genomics of Czech vaccine strains of Bordetella pertussis. *Pathog Dis*;76. Epub ahead of print 1 October 2018. DOI: 10.1093/femspd/fty071.

15. **Maiden MCJ, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, *et al.*** MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol* 2013;11:728–736.

16. **Martin SW, Pawloski L, Williams M, Weening K, DeBolt C, *et al.*** Pertactin-negative Bordetella pertussis strains: evidence for a possible selective advantage. *Clin Infect Dis Off Publ Infect Dis Soc Am* 2015;60:223–227.

17. **Zeddeman A, van Gent M, Heuvelman CJ, van der Heide HG, Bart MJ, *et al.*** Investigations into the emergence of pertactin-deficient Bordetella pertussis isolates in six European countries, 1996 to 2012. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull*;19.

18. **Otsuka N, Han H-J, Toyoizumi-Ajisaka H, Nakamura Y, Arakawa Y, *et al.*** Prevalence and Genetic Characterization of Pertactin-Deficient Bordetella pertussis in Japan. *PLOS ONE* 2012;7:e31985.

19. **Breakwell L, Kelso P, Finley C, Schoenfeld S, Goode B, *et al.*** Pertussis Vaccine Effectiveness in the Setting of Pertactin-Deficient Pertussis. *Pediatrics*;137. Epub ahead of print May 2016. DOI: 10.1542/peds.2015-3973.

20. **Caro V, Elomaa A, Brun D, Mertsola J, He Q, *et al.*** Bordetella pertussis, Finland and France. *Emerg Infect Dis* 2006;12:987–989.

Table 1: Characteristics of Tunisian *Bordetella pertussis* genomes used in this study

Isolate Name	**Month of collection**	**Age of patient**	**Number of contigs**	**Genome size (nt)**	**N50 length (nt)**	**Genotype**	**Accession number (PRJEB27412)**
TN0001	April 2014	2 months	300	3,872,771	20,131	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572942
TN0002	April 2014	3 weeks	293	3,864,234	20,328	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572943
TN0003	April 2014	6 weeks	298	3,863,302	20,108	*ptxP3 ptxA1 prn2 fim2-1 fim3-1*	ERS2572944
TN0004	April 2014	3 weeks	301	3,878,136	19,862	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572945
TN0005	April 2014	1 month	290	3,868,375	20,544	*ptxP3 ptxA1 prn3 fim2-1 fim3-1*	ERS2572946
TN0006	May 2014	6 weeks	283	3,868,089	20,787	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572947
TN0007	May 2014	2 months	288	3,865,925	20,544	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572948
TN0008	May 2014	2 months	298	3,862,980	19,926	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572949
TN0009	May 2014	2 months	293	3,865,015	20,346	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572950
TN0011	May 2014	3 months	289	3,863,801	20,541	*ptxP3 ptxA1 prn2 fim2-1 fim3-2*	ERS2572951

236 **Table 2**: Genetically closest isolate(s) for each Tunisian *Bordetella pertussis* isolate

Isolate Name	Closest isolate(s) by cgMLST distance	No. of cgMLST mismatches
TN001	FR3903; FR3916	2
TN002	FR3903; FR3916	3
TN003	FR5438; FR5439	5
TN004	FR5791; P2M	3
TN005	I538; I539; I656	2
TN006	FR5859; FR5860; FR5942; FR6031; H379; H788; H918	2
TN007	FR3903; FR3916	2
TN008	FR3903; FR3916; FR5943; FR5793	2
TN009	FR5942; FR6031; H379; H788; H918; FR5859; FR5860	2
TN011	FR3903; FR3916	2

237

238

Figure 1: Maximum-likelihood phylogenetic tree of Tunisian *Bordetella pertussis* isolates and isolates from other geographical origins based on the concatenated alignments of 2,038 cgMLST gene sequences. The tree was rooted with strain J446. Black branches correspond to the *fim3-1* clade and green branches to the *fim3-2* clade. The external circle indicates the geographic origin of isolates (long red bars, Tunisia; long blue bars: vaccine strains of Serum Institute of India; light green, France; light blue, United Kingdom; light pink, USA). The scale bar indicates the number of nucleotide substitutions per site.