Complete Genome Sequences of Two *Bordetella hinzii* Strains Isolated from Humans

Michael R. Weigand, Shankar Changayil, Yasvanth Kulasekarapandian, M. Lucia Tondella

Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Bordetella hinzii is primarily recovered from poultry but can also colonize mammalian hosts and immunocompromised humans. Here, we report the first complete genome sequences of *B. hinzii* in two isolates recovered from humans. The availability of these sequences will hopefully aid in identifying host-specific determinants variably present within this species.

Address correspondence to Michael R. Weigand, mweigand@cdc.gov.

*B*ordetella species are associated with a variety of hosts, in which they are often the etiologic agents of disease. *Bordetella hinzii* frequently colonizes poultry and causes respiratory disease in turkeys (1). However, *B. hinzii* can also opportunistically infect immunocompromised humans, leading to conditions such as respiratory disease, septicemia, and cholangitis (2–5). Rodents have been proposed to serve as reservoirs for *B. hinzii* (6, 7), suggesting that this species may be an emerging zoonotic pathogen. Therefore, comparative genomics of *B. hinzii* isolated from different sources is needed to elucidate any virulence mechanisms defining the host range in this species. Here, we report the complete genome sequences of two clinically derived *B. hinzii* strains from the culture collection at the Centers for Disease Control and Prevention. *B. hinzii* F582, a clone of ATCC 51730, was isolated from a blood sample from a male AIDS patient (8), and *B. hinzii* H568 was isolated in 2010 from a sputum sample from a male in Arkansas. Although at least eight draft genomes for *B. hinzii* exist in public databases (9, 10), including that of ATCC 51730, these are the first closed genome sequences for this species.

Genome sequencing was performed using a combination of the PacBio RSII (Pacific Biosciences, Menlo Park, CA) and HiSeq PE-100 (Illumina, Hayward, CA) platforms. Genomic DNA was isolated with the Gentra Puregene yeast/bacteria kit (Qiagen, Valencia, CA). Libraries were prepared for PacBio sequencing using the SMRTbell template prep kit 1.0 and polymerase binding kit P4, while HiSeq libraries were prepared using the NEBNext Ultra library prep kit (New England BioLabs, Ipswich, MA). PacBio sequencing reads were filtered and assembled *de novo* using the Hierarchical Genome Assembly Process algorithm (HGAP) version 3. The initial assemblies were improved by unambiguously mapping reads with BLASR (version 1) at 482× and 184× coverage for F582 and H568, respectively. The resulting consensus sequences were determined with Quiver (version 1) and manually checked for circularity. The assemblies were validated by comparison to restriction digest optical maps generated with the Argus platform (OpGen, Gaithersburg, MA) using MapSolver (version 2.1.1). Finally, the genome sequences were further polished by mapping HiSeq reads at 184× and 81× coverage for F582 and

H568, respectively. Raw reads were first quality trimmed, clipped of adapter contamination, and filtered using cutadapt (version 1.9) (11). The reads were mapped against the HGAP assembly using the basic variant detection module in CLC Genomics Workbench (version 7.5) to identify errors, which were then corrected using a custom Perl script.

The average G+C content of both genomes is 67.1%, and the genome sizes were 4,972,977 and 4,859,884 bp for F582 and H568, respectively. Genome annotation with Prokka (version 1.11) (12) identified 4,559 protein-coding genes and 64 tRNAs in F582 and 4,467 protein-coding genes and 66 tRNAs in H568. Both genomes contain three rRNA operons. A comparison of the two genomes by read mapping in CLC Genomics Workbench or *k*-mer alignment with kSNP (version 3.0) (13) revealed 8,926 to 9,688 core single-nucleotide polymorphisms (SNPs), depending on the method of identification. The two sequences also include several strain-specific genomic islands, some of which include unique prophage insertions.

Nucleotide sequence accession numbers. The complete genome sequences have been deposited at DDBJ/EMBL/GenBank under the accession numbers CP012076 and CP012077 for *B. hinzii* strains F582 and H568, respectively. The versions described in this paper are the first versions.

REFERENCES

1. **Register KB, Kunkle RA**. 2009. Strain-specific virulence of *Bordetella hinzii* in poultry. Avian Dis **53:**50–54. http://dx.doi.org/10.1637/8388-070108-Reg.1.

2. **Spilker T, Liwienski AA, LiPuma JJ**. 2008. Identification of *Bordetella* spp. in respiratory specimens from individuals with cystic fibrosis. Clin Microbiol Infect **14**:504–506. http://dx.doi.org/10.1111/j.1469-0691.2008.01968.x.

3. **Arvand M, Feldhues R, Mieth M, Kraus T, Vandamme P**. 2004. Chronic cholangitis caused by *Bordetella hinzii* in a liver transplant recipient. J Clin Microbiol **42**:2335–2337. http://dx.doi.org/10.1128/JCM.42.5.2335 -2337.2004.

4. **Gadea I, Cuenca-Estrella M, Benito N, Blanco A, Fernández-Guerrero ML, Valero-Guillén PL, Soriano F**. 2000. *Bordetella hinzii*, a "new" opportunistic pathogen to think about. J Infect **40**:298–299. http://dx.doi.org/10.1053/jinf.2000.0646.

5. **Funke G, Hess T, von Graevenitz A, Vandamme P**. 1996. Characteristics of *Bordetella hinzii* strains isolated from a cystic fibrosis patient over a 3-year period. J Clin Microbiol **34**:966–969.

6. **Jiyipong T, Morand S, Jittapalapong S, Raoult D, Rolain JM**. 2013. *Bordetella hinzii* in rodents, southeast Asia. Emerg Infect Dis **19**:502–503. http://dx.doi.org/10.3201/eid1903.120987.

7. **Hayashimoto N, Morita H, Ishida T, Yasuda M, Kameda S, Takakura A, Itoh T**. 2012. Prevalence of *Bordetella hinzii* in mice in experimental facilities in Japan. Res Vet Sci **93**:624–626. http://dx.doi.org/10.1016/j.rvsc.2011.11.001.

8. **Cookson BT, Vandamme P, Carlson LC, Larson AM, Sheffield JV, Kersters K, Spach DH**. 1994. Bacteremia caused by a novel *Bordetella* species, "*B. hinzii*." J Clin Microbiol **32**:2569–2571.

9. **Shah NR, Moksa M, Novikov A, Perry MB, Hirst M, Caroff M, Fernandez RC**. 2013. Draft genome sequences of *Bordetella hinzii* and *Bordetella trematum*. Genome Announc **1**(5):e00838-13. http://dx.doi.org/10.1128/genomeA.00838-13.

10. **Register KB, Ivanov YV, Harvill ET, Brinkac L, Kim M, Losada L**. 2015. Draft genome sequences of six *Bordetella hinzii* isolates acquired from avian and mammalian hosts. Genome Announc **3**(2):e00081-15. http://dx.doi.org/10.1128/genomeA.00081-15.

11. **Martin M**. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**:10–12. http://dx.doi.org/10.14806/ej.17.1.200.

12. **Seemann T**. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**:2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.

13. **Gardner SN, Hall BG**. 2013. When whole-genome alignments just won't work: kSNP v2 software for alignment-free SNP discovery and phylogenetics of hundreds of microbial genomes. PLoS One **8**:e81760. http://dx.doi.org/10.1371/journal.pone.0081760.