Genomic Survey of *Bordetella pertussis* Diversity, United States, 2000–2013

Michael R. Weigand, Margaret M. Williams, Yanhui Peng, Dane Kania, Lucia C. Pawloski, Maria L. Tondella, CDC Pertussis Working Group¹

We characterized 170 complete genome assemblies from clinical *Bordetella pertussis* isolates representing geographic and temporal diversity in the United States. These data capture genotypic shifts, including increased pertactin deficiency, occurring amid the current pertussis disease resurgence and provide a foundation for needed research to direct future public health control strategies.

Whooping cough (pertussis) remains a public health challenge in the United States where, despite high vaccine coverage, an increased number of cases have been reported since the late 1980s. This resurgence has included >48,000 cases reported in 2012 and notable recent statewide epidemics (1). Likely causes of the increase in reporting include heightened awareness, expanded surveillance, improved laboratory diagnostics, and waning protection conferred by acellular pertussis (aP) vaccine formulations (1,2).

The United States exclusively uses aP vaccines composed of inactivated *Bordetella pertussis* immunogenic proteins pertussis toxin (Pt), pertactin (Prn), and filamentous hemagglutinin (Fha), either with or without fimbria (Fim) types 2 and 3. Genetic divergence of circulating *B. pertussis* away from vaccine reference strains has led to allelic mismatch and the rapid emergence of Prn deficiency (3). Although such recent genetic changes may be ascribed to vaccine-driven immune selection (4), aP vaccines remain effective (5).

The chromosome of *B. pertussis* also undergoes frequent structural rearrangement (6) that presents unique challenges to thorough investigation of genetic contributions to disease resurgence, limiting assessment of public health strategies. Until recently, genomic data with sufficient resolution to study sequence and structural variation were available only for vaccine and laboratory reference strains. However, pathogen evolution must be explored through multinomic characterization of circulating

Author affiliation: Centers for Disease Control and Prevention, Atlanta, Georgia, USA

genotypes. To address this gap, we developed a dataset of complete, reference-quality genome sequence assemblies from isolates representing the geographic and temporal diversity of *B. pertussis* circulating in the United States during 2000–2013.

The Study

The Centers for Disease Control and Prevention (CDC) maintains a collection of *B. pertussis* isolates recovered by state public health laboratories through routine surveillance and outbreaks or the Enhanced Pertussis Surveillance/ Emerging Infections Program Network (7). We selected a subset of isolates (n = 170) to account for potential geographic diversity. We stratified all isolates in the collection by state and time period (2000-2002, 2003-2009, 2010, 2011, 2012, and 2013) chosen according to diversity indices reported previously (8), with additional emphasis on more recent sampling. We then randomly sampled the stratified collection to maximize the number of source states (n = 34)during each period with equal weighting (Figure 1, panel A, B). Most isolates were characterized by existing molecular approaches, multilocus variable-number tandem-repeat analysis (MLVA), and pulsed-field gel electrophoresis (PFGE), as described previously (9). The selected isolates included 17 MLVA types, with type 27 the most prevalent, and 33 PFGE profiles, with profile CDC013 the most prevalent (Appendix Table 1, https://wwwnc.cdc.gov/EID/ article/25/4/18-0812-App1.pdf).

We performed whole-genome shotgun sequencing and assembly as described previously (10) (Appendix). Genome assembly yielded a single circular contig for all isolates, and we performed sequence-based molecular typing (Appendix). Nearly all isolates (96%) were of the predominant type prn2-ptxP3-ptxA1 with either fimH1 or fimH2, and few harbored alternate types such as prn1ptxP1-ptxA2-fimH1 (Figure 1, panel C). Prn deficiency has been observed in ≥ 16 independent mutations to prn (6); we observed 10 deficient alleles among 57/170 isolates in our study, including missense substitutions, deletions, promoter disruption, and various IS481 insertions. The proportion of isolates with Prn-deficient alleles increased rapidly beginning in 2010 (Figure 1, panel C), consistent with a larger

¹Additional members of the CDC Pertussis Working Group are listed at the end of this article.



Figure 1. Bordetella pertussis diversity, United States, 2000–2013. A) Geographic origin of *B. pertussis* isolates selected to maximize the number of source states from each of 6 time periods. Pie chart diameter represents the number of isolates, as detailed in the key, and colors indicate time periods, as shown in panel B. B) Isolate frequency by time period. C) Relative abundance of MLST types *prn2-ptxP3-ptxA1-fimH1* (gray), *prn2-ptxP3-ptxA1-fimH2* (white), and other (black). Red line indicates frequency of pertactin-deficient alleles. D) Abundance distribution of genome structures. Black bar indicates unique structures (singletons) and gray bars the 16 colinear groups. MLST, multilocus sequence typing.

molecular survey of US isolates conducted previously that included some used in this study (3). We also determined MLVA type from genome assemblies using a custom bioinformatics pipeline (wgsMLVA) based on traditional PCR primer sequences (11) (Appendix). None of the genomes encoded known 23S ribosomal RNA mutation associated with erythromycin resistance (12).

To determine variation in chromosome structure, we performed exhaustive pairwise alignment of assembled genomes as previously described (6). Of the 170 assemblies, 129 clustered into 16 groups of ≥ 2 colinear genomes (lacking observable rearrangement or deletion >1,500 bp), whereas 41 assemblies (singletons) exhibited unique structures not shared with any others in the dataset. Observed structures largely correlated with PFGE, a proxy for chromosome structure, clustering isolates with shared PFGE profiles. The abundance of common structures reflected predominant PFGE profiles, and the largest cluster corresponded to profile CDC013 (Figure 1, panel D). Differences between many common structures could be attributed to large inversions flanked by insertions of the multicopy IS481. Select singleton structures resulted from tandem duplication of large regions (15.5–190 kbp) in the genomes of 5 isolates (D236, D665, H624, J085, and J139) that were also flanked by copies of IS481.

We reconstructed a maximum-likelihood phylogeny of the isolate genomes from 840 core variable single-nucleotide polymorphisms (SNPs) determined from the reference Tohama I (GenBank accession no. CP010964) (Appendix). The resulting tree topology revealed deep divergence of lineages bearing alleles *ptxP1* and *ptxP3*, as well as clear distinctions between clades of prn2-ptxP3-ptxA1-fimH1 and prn2-ptxP3ptxA1-fimH2 (Figure 2). Only certain prn-disrupting mutations (e.g., nonsense C1273T, promoter disruption) and chromosome structures (e.g., cluster-4, cluster-6, cluster-7, cluster-9) appeared phylogenetically linked, meaning isolates sharing them were also related according to their SNP patterns. However, each group of related isolates was recovered across multiple states and time periods, suggesting that genotypes, whether defined by gene sequence or chromosome structure, were stable enough to be widely circulated. Prn deficiency due to IS481 disruption has resulted from \geq 7 independent events among the isolates in this dataset, but related isolates with these mutations were likewise geographically and temporally distributed. These results are consistent with phylogenies of circulating *B. pertussis* reported elsewhere (6,13).

Conclusions

We have developed a representative dataset of complete genome sequence assemblies derived from *B. pertussis* clinical isolates recovered in the United States that captures

DISPATCHES



Figure 2. Phylogenetic reconstruction of all 170 isolates and the reference Tohama I (GenBank accession no. CP010964). Isolate metadata and molecular characteristics are color coded, as detailed in the key. Scale bar indicates substitutions per site. CDC, Centers for Disease Control and Prevention; fim, fimbria; fwd, forward insertion; rev, reverse insertion; PFGE, pulsed-field gel electrophoresis; prn, pertactin; ptx, pertussis toxin.

shifting population genetics concurrent with disease resurgence. We selected isolates to maximize the geographic diversity of circulating *B. pertussis* across 6 time periods during 2000–2013 and to span the time period in which Prn deficiency emerged as the predominant molecular type. Although the sparse sampling of individual states and regions prohibited detailed analyses of geographic distribution, we did recover isolates with shared SNP patterns and chromosome structures from disparate states. Our results illustrate underlying challenges to the molecular study of pertussis resurgence, including a circulating mixture of gene sequence (SNP) and chromosome structure variants.

The genomic data we provide will aid open research toward improved vaccine development and disease control strategies. Because little to no such high-quality data existed previously, the contribution of genome evolution to pertussis resurgence has not been fully appreciated. A subset of these data has already helped elucidate historical patterns of chromosome rearrangement (6). However, comparative genomics alone is not sufficient to understand the resurgence in pertussis. Further laboratory experimentation using in vitro and in vivo infection models is needed to link outcomes with novel, bioinformatically determined genetic variation, such as discrete rearrangements and tandem duplications. Potential differences in antigen expression resulting from these changes in gene organization, which may influence the burden of disease, remain untested. Our results provide needed context to guide such investigations by highlighting representative, circulating genotypes as they continue their divergence from existing laboratory and vaccine reference strains. Data such as those presented here critically establish the necessary foundation for collaborative development of advanced diagnostics, novel molecular typing methods, and improved vaccine formulations.

Additional members of the CDC Pertussis Working Group: Dhwani Batra, Katherine E. Bowden, Mark Burroughs, Pamela K. Cassiday, Jamie K. Davis, Taccara Johnson, Hong Ju, Phalasy Juieng, Kristen Knipe, Vladimir N. Loparev, Stacey W. Martin, Christine Miner, Lori A. Rowe, Tami H. Skoff, Mili Sheth, Kevin Tang.

This work was made possible through support from the Advanced Molecular Detection program at the US Centers for Disease Control and Prevention.

About the Author

Dr. Weigand is a bioinformatics research scientist in the Pertussis and Diphtheria Laboratory, Division of Bacterial Diseases, National Center for Immunization and Respiratory Diseases, CDC, Atlanta. His primary research interest is comparative genomics of bacterial pathogens, with a current focus on *Bordetella pertussis*.

References

- Clark TA. Changing pertussis epidemiology: everything old is new again. J Infect Dis. 2014;209:978–81. http://dx.doi.org/10.1093/ infdis/jiu001
- Ausiello CM, Cassone A. Acellular pertussis vaccines and pertussis resurgence: revise or replace? MBio. 2014;5:e01339–14. http://dx.doi.org/10.1128/mBio.01339–14
- Pawloski LC, Queenan AM, Cassiday PK, Lynch AS, Harrison MJ, Shang W, et al. Prevalence and molecular characterization of pertactin-deficient *Bordetella pertussis* in the United States. Clin Vaccine Immunol. 2014;21:119–25. http://dx.doi.org/10.1128/ CVI.00717-13
- Martin SW, Pawloski L, Williams M, Weening K, DeBolt C, Qin X, et al. Pertactin-negative *Bordetella pertussis* strains: evidence for a possible selective advantage. Clin Infect Dis. 2015;60:223–7. http://dx.doi.org/10.1093/cid/ciu788
- Breakwell L, Kelso P, Finley C, Schoenfeld S, Goode B, Misegades LK, et al. Pertussis vaccine effectiveness in the setting of pertactin-deficient pertussis. Pediatrics. 2016;137:e20153973. http://dx.doi.org/10.1542/peds.2015-3973

- Weigand MR, Peng Y, Loparev V, Batra D, Bowden KE, Burroughs M, et al. The history of *Bordetella pertussis* genome evolution includes structural rearrangement. J Bacteriol. 2017;199:e00806–16. http://dx.doi.org/10.1128/ JB.00806-16
- Skoff TH, Baumbach J, Cieslak PR. Tracking pertussis and evaluating control measures through enhanced pertussis surveillance, Emerging Infections Program, United States. Emerg Infect Dis. 2015;21:1568–73. http://dx.doi.org/10.3201/ eid2109.150023
- Schmidtke AJ, Boney KO, Martin SW, Skoff TH, Tondella ML, Tatti KM. Population diversity among *Bordetella pertussis* isolates, United States, 1935–2009. Emerg Infect Dis. 2012;18:1248–55. http://dx.doi.org/10.3201/eid1808.120082
- Hardwick TH, Cassiday PK, Weyant RS, Bisgard KM, Sanden GN. Changes in predominance and diversity of genomic subtypes of *Bordetella pertussis* isolated in the United States, 1935 to 1999. Emerg Infect Dis. 2002;8:44–9. http://dx.doi.org/10.3201/ eid0801.010021 article
- Bowden KE, Weigand MR, Peng Y, Cassiday PK, Sammons S, Knipe K, et al. Genome structural diversity among 31 *Bordetella pertussis* isolates from two recent U.S. whooping cough statewide epidemics. mSphere. 2016;May–Jun:e00036–16. http://dx.doi.org/ 10.1128/mSphere.00036-16
- Schouls LM, van der Heide HG, Vauterin L, Vauterin P, Mooi FR. Multiple-locus variable-number tandem repeat analysis of Dutch *Bordetella pertussis* strains reveals rapid genetic changes with clonal expansion during the late 1990s. J Bacteriol. 2004; 186:5496– 505. http://dx.doi.org/10.1128/JB.186.16.5496-5505.2004
- Bartkus JM, Juni BA, Ehresmann K, Miller CA, Sanden GN, Cassiday PK, et al. Identification of a mutation associated with erythromycin resistance in *Bordetella pertussis*: implications for surveillance of antimicrobial resistance. J Clin Microbiol. 2003;41:1167–72. http://dx.doi.org/10.1128/ JCM.41.3.1167-1172.2003
- Bart MJ, Harris SR, Advani A, Arakawa Y, Bottero D, Bouchez V, et al. Global population structure and evolution of *Bordetella pertussis* and their relationship with vaccination. MBio. 2014;5:e01074. http://dx.doi.org/10.1128/ mBio.01074-14

Address for correspondence: Michael R. Weigand, Centers for Disease Control and Prevention, 1600 Clifton Rd NE, Mailstop H18-B, Atlanta, GA 30329-4027, USA; email: mweigand@cdc.gov

Genomic Survey of *Bordetella pertussis* Diversity in the United States, 2000–2013

Appendix

Methods

We performed whole-genome shotgun sequencing using a combination of the PacBio RSII (Pacific Biosciences, http://www.pacb.com/), Illumina HiSeq/MiSeq (Illumina, http://www.illumina.com/), and Argus (OpGen; http://www.opgen.com) platforms as described previously (1). The cumulative coverage depth of PacBio and Illumina sequencing for each isolate is listed in Appendix Table 1. Genomes were assembled using HGAP v3 (Pacific Biosciences) followed by structure confirmation with restriction digest optical mapping (OpGen) and further sequence polishing by Illumina read mapping with CLC Genomics Workbench (QIAGEN, http://www.qiagen.com). Completed assemblies were submitted to the National Center for Biotechnology Information (NCBI) for annotation by the automated Prokaryotic Genome Annotation Pipeline (PGAP). Genome sequence-based molecular characterization was performed using either completed assemblies or individual sequencing reads. Alleles for common molecular typing loci (*ptxP, ptxA, ptxB, fimH,* and *prn*) were assigned by genome alignment to a curated set of wild-type and deficient alleles using high-stringency.

Molecular typing by multiple-locus variable number tandem repeat analysis (MLVA) was determined from closed genome assemblies using a custom bioinformatics pipeline (wgsMLVA). Similar to traditional PCR-based approaches, wgsMLVA uses primer sequences to identify 6 Variable Number Tandem Repeat (VNTR) sites which contain a varying set of short sequence repeats (2). The number of repeat monomers is counted at each site to build a 6-number VNTR profile summarized as an MLVA type (www.mlva.net). Traditionally, each VNTR is amplified by PCR and repeat numbers are inferred from the molecular weight; with higher molecular weights corresponding to the addition of known repeat monomers. By contrast, wgsMLVA leverages high resolution genome assemblies to directly count repeat monomers in each VNTR site using an exact-match search. This approach produces a more accurate count that does not rely on estimations calculated from VNTR length.

Of the 170 isolates characterized in this study, 128 had been characterized by traditional MLVA using PCR before whole-genome sequencing. Profiles calculated using wgsMLVA were identical to those determined by traditional MLVA for 127/128 (99.2%) isolates (Appendix Table 1). Comparison of VNTR profiles in H811 calculated by the 2 methods revealed a discrepancy of 1 repeat in a VNTR3 locus (Appendix Table 2). Traditional MLVA cannot differentiate the 2 VNTR3 loci and relies on measurable differences in electrophoretic mobility; otherwise, VNTR3b is counted as 0 ambiguously reporting it as either missing or equal to VNTR3a. Because wgsMLVA directly counts repeat monomers at each locus independently, such a discrepancy is not wholly unexpected given that VNTR3 encodes the smallest repeat monomer at 5 bp, compared with the larger 15 bp monomer of VNTR1.

The wgsMLVA pipeline is implemented in Python 2.7, free of external libraries, packages, or other dependencies. A user-supplied reference database is required to match MLVA types from calculated VNTR profiles; an updated database can be downloaded from www.mlva.net.

Single nucleotide polymorphisms in each isolate genome were determined from the reference Tohama I (CP010964) by mapping Illumina reads with snippy (https://github.com/tseemann/snippy). A maximum-likelihood phylogeny was reconstructed using RAxML (*3*) and additional tree annotation was performed using iTOL (*4*).

Source Code

The source code for calculating MLVA types from complete genome assemblies of *Bordetella pertussis* using wgsMLVA is available at https://github.com/danek90/wgsMLVA

References

 Bowden KE, Weigand MR, Peng Y, Cassiday PK, Sammons S, Knipe K, et al. Genome structural diversity among 31 *Bordetella pertussis* isolates from two recent US whooping cough statewide epidemics. mSphere. 2016; 1:e00036–16. http://dx.doi.org/10.1128/mSphere.00036-16

- 2. Schouls LM, van der Heide HG, Vauterin L, Vauterin P, Mooi FR. Multiple-locus variable-number tandem repeat analysis of Dutch *Bordetella pertussis* strains reveals rapid genetic changes with clonal expansion during the late 1990s. J Bacteriol. 2004;186:5496–505. <u>PubMed</u> <u>http://dx.doi.org/10.1128/JB.186.16.5496-5505.2004</u>
- 3. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3. <u>PubMed</u> <u>http://dx.doi.org/10.1093/bioinformatics/btu033</u>
- 4. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242–5. <u>PubMed</u> <u>http://dx.doi.org/10.1093/nar/gkw290</u>
- 5. Weigand MR, Peng Y, Loparev V, Batra D, Bowden KE, Burroughs M, et al. The history of *Bordetella pertussis* genome evolution includes structural rearrangement. J Bacteriol. 2017;199:e00806–16.
 <u>PubMed http://dx.doi.org/10.1128/JB.00806-16</u>

								Mol	ecular typing	loci				
												Accession		
ID	Year	State	EPS [†]	MLVA	wqsMLVA	PFGE	Structure	prn	ptxP	ptxA	fimH	no.	Depth [‡]	Reference
C505	2000	MI		NT	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP011687	293x	5
C549	2000	GA		36	36	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP013077	311x	5
C569	2000	IN		27	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP025347	382x	This study
C571	2000	IN		27	27	CDC010	Cluster-3	prn2	ptxP3	ntxA1	fimH1	CP011167	265x	5
C734	2000	ID		27	27	CDC002	Cluster-2	prn2	ptxP3	ntxA1	fimH1	CP013078	278x	5
C742	2000	ОH		NT	27	CDC013	Singleton	prn2	ntxP3	ntxA1	fimH2	CP011688	254x	5
C756	2000	тх		NT	27	CDC010	Cluster-10	nrn2	ptxP3	$ntx\Delta 1$	fimH1	CP025368	204x 337x	This study
C757	2000	TY		16	16	CDC013	Cluster-1	prn2 prn2	ntvP3	ptxA1	fimH2	CP013070	202v	5
C971	2000			NT	27		Singloton	prn2	ptxP2	$p i \lambda \pi I$	fimU1	CP025345	232A 112v	This study
C071	2000	M I		27	27		Cluster 1	prn2	pixr 3	$\rho tx A 1$	fimU2	CP025545	2/22	This study
C954	2000	NU		27	27	CDC013	Ciuster-1	prinz prm2	pixr 3	pixA1	1111111Z	CF010901	2404	F
C936	2001	INJ			27		Singleton	prnz prn2	pixP3	pixA1	fim 10	CP011100	349X	5 F
C975	2000				30	CDC013	Cluster-1	prnz prn2	pixP3	pixA1	fim 10	CP013000	3//X	5 F
D175	2000	CA			200	CDC013	Cluster-1	prnz	ptxP3	ptxA1	11mH2	CP011689	262X	C This study
D236	2001	01		27	27	CDC150	Singleton	prnz	ptxP3	ptxA1	timH2	CP025530	284X	i nis study
D321	2001	MO		27	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	timH2	CP011690	290x	- 5
D322	2001	MO		27	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	timH2	CP025358	337x	This study
D422	2002	CA		18	18	CDC154	Singleton	prn2	ptxP3	ptxA1	fimH2	CP016959	338x	This study
D502	2001	IL		NI	27	CDC007	Singleton	prn2	ptxP3	ptxA1	fimH1	CP011691	293x	5
D521	2000	MN		NT	27	CDC082	Cluster-5	prn2	ptxP3	ptxA1	fimH2	CP011169	261x	5
D665	2002	NV		27	27	CDC013	Singleton	prn2	ptxP3	ptxA1	fimH2	CP025526	258x	This study
D717	2002	NV		27	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP016964	384x	This study
D735	2000	OH		NT	18	CDC010	Singleton	prn2	ptxP3	ptxA1	fimH1	CP016960	399x	This study
D799	2002	ID		27	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP016963	414x	This study
D869	2002	AZ		NT	25	CDC082	Singleton	prn2	ptxP3	ptxA1	fimH2	CP025528	233x	This study
D879	2002	AZ		NT	36	CDC010	Singleton	prn2	ptxP3	ptxA1	fimH1	CP011170	234x	5
D919	2002	NY		NT	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP025355	219x	This study
D925	2002	NY		27	27	CDC082	Cluster-5	prn2	ptxP3	ptxA1	fimH2	CP016968	369x	This study
E024	2003	MD		27	27	CDC010	Cluster-10	prn2	ptxP3	ptxA1	fimH1	CP011692	236x	5
E025	2003	MD		27	27	CDC010	Cluster-10	prn2	ptxP3	ptxA1	fimH1	CP016967	320x	This study
E087	2002	MA		NT	27	CDC010	Singleton	prn2	ptxP3	ptxA1	fimH1	CP025480	295x	This study
E140	2002	MA		NT	16	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP025354	420x	This study
E150	2003	OH		32	32	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP011171	327x	5
E153	2003	SC		NT	28	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP025359	382x	This study
E191	2003	SC		27	27	CDC123	Cluster-7	prn2	ptxP3	ptxA1	fimH2	CP025478	252x	This study
E194	2003	WA		27	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP013080	353x	5
E198	2003	KY		26	26	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP025385	203x	This study
E365	2004	MO		27	27	CDC013	Cluster-1	prn2	, ptxP3	, ptxA1	fimH2	CP025387	309x	This study
E368	2004	MO		25	25	CDC013	Cluster-13	prn2	, ptxP3	, ptxA1	fimH2	CP013869	430x	5
E530	2000	MT		NT	27	CDC002	Cluster-2	prn2	, ptxP3	, ptxA1	fimH1	CP011693	200x	5
E537	2001	MT		27	27	CDC013	Cluster-1	prn2	, ptxP3	, ptxA1	fimH2	CP016958	352x	This study
E541	2003	MT		27	27	CDC010	Singleton	prn2	ptxP3	ptxA1	fimH1	CP016966	403x	This study
E555	2004	MT		NT	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP011172	332x	5
E587	2005	DF		27	27	CDC082	Singleton	prn2	ptxP3	ptxA1	fimH2	CP011173	309x	5
E602	2005	DE		218	218	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP013081	348x	5
E809	2005	A7		NT	18	CDC013	Cluster-1	prn2	ptxP3	ntxA1	fimH2	CP011174	305x	5
	2000				.0	000010		PINE		P0011		51 51 11 4	0000	5

Appendix Table 1. Detailed data about Bordetella pertussis isolates, 2000–2013*

Product Structure product product product Accession E898 2005 AZ 16 16 CDC010 Cluster-3 prn2 puAP prod.1 fimH CP016862 227.X This study E976 2005 NY NT 227 CDC020 Singleton prn1 puAP3 ptAA1 fimH CP016865 365.X This study E976 2005 NE 2.7 CDC010 Cluster-3 prn2 puAP3 ptAA1 fimH CP01177 258.K 5 F013 2005 NE 2.7 CDC0446 Cluster-3 prn2 puAP3 ptAA1 fimH 2 CP013877 330.K 5 F501 2004 NY NT 2.7 CDC0446 Cluster-6 prn2 puAP3 ptAA1 fimH2 CP013877 330.K 5 F578 2007 NG 2.7 2.7 CDC0182 Cluster-5 prn2 ptA23 p									Molec	ular typing	loci							
D Year State EPS ¹ ML/A yeFGE Structure pm pt/P ptAA fimH no. Depth ¹ Reference E986 2005 CA 70 70 CDC021 Singleton pm1 ptAP ptAA1 fimH1 CP016956 365x This study E976 2005 NE 27 27 CDC020 Singleton pm1 ptAP1 ptAA1 fimH1 CP011176 27x 5 F013 2005 NE 27 27 CDC010 Cluster-3 pm2 ptAP3 ptAA1 fimH1 CP011177 28x 5 F013 2005 CA NT 27 CDC044 Cluster-8 pm2 ptAP3 ptAA1 fimH2 CP013300 3 5 F561 2007 NC 27 27 CDC048 Sluster-5 pm2 ptAP3 ptAA1 fimH2 CP023342 505x This study F563										5 51 5			Accession					
E898 2005 AZ 16 CDC010 Cluster.3 pm2 pbkP1	ID	Year	State	EPS [†]	MLVA	wqsMLVA	PFGE	Structure	prn	ptxP	ptxA	fimH	no.	Depth [‡]	Reference			
E945 2005 CA 70 CDC021 Singleton pm1 pbtP1 pbtA1 fimH1 CP011956 365x This study F011 2005 NE 27 27 CDC020 Singleton pm1 pbtP3 pbtA1 fimH1 CP011176 276x 5 F031 2005 NE 27 27 CDC010 Cluster-3 pm2 pbtP3 pbtA1 fimH1 CP011177 288x 5 F561 2004 NY NT 27 CDC046 Cluster-8 pm2 pbtP3 pbtA1 fimH2 CP025523 370x This study F578 2007 NC 27 27 CDC046 Cluster-5 pm2 pbtA3 firmH2 CP025342 605x This study F685 2007 NC 27 27 CDC042 Cluster-5 pm2 pbtA3 firmH2 CP011173 302x 5 F676 2008 NA 365	E898	2005	AZ	_	16	16	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP016962	297x	This study			
E976 2005 NY NT 227 CDC020 Singleton pm1 pbP1 pbP2 pbA1 fmH1 CP011175 267x 5 F013 2005 NE 27 27 CDC010 Cluster-3 pm2 pbA21 fmH1 CP011976 278x 5 F013 2005 NE 27 CDC040 Cluster-3 pm2 pbA21 fmH1 CP019870 320x 5 F601 2006 GA NT 27 CDC046 Cluster-8 pm2 pbA23 fmH2 CP02337 400x This study F580 2007 NC 27 27 CDC046 Cluster-1 pm2 pbA23 fmH2 CP011871 440x 5 F652 2007 NC NT 179 CDC048 Cluster-1 pm2 pbA23 fmH2 CP011871 302x 5 F642 2008 NA 277 27 CDC046 Cluster-5	E945	2005	CA		70	70	CDC021	Singleton	prn1	ptxP1	ptxA1	fimH1	CP016956	365x	This study			
F011 2005 NE 27 27 CDC010 Cluster-3 pm2 pbP3 pbP3 fmH1 CP011176 278x 5 F013 2005 CA NT 27 CDC010 Cluster-3 pm2 pbP3 fmH1 CP011177 288x 5 F569 2006 GA NT 27 CDC046 Cluster-8 pm2 pbP3 pbA1 fmH2 CP025523 370x This study F578 2007 NC 27 27 CDC046 Cluster-6 pm2 pbP3 pbA1 fmH2 CP02537 400x This study F678 2007 NC 27 27 CDC046 Cluster-6 pm2 pbP3 pbA1 fmH2 CP025342 505x This study F678 2007 NC 27 27 CDC040 Cluster-6 pm2 pbP3 pbA1 fmH2 CP011180 332x 5 F678 2008 NC	E976	2005	NY		NT	227	CDC020	Singleton	prn1	ptxP1	ptxA2	fimH1	CP011175	267x	5			
F013 2005 NE 27 27 CDC010 Cluster-3 pm2 pbP3 pbP3 <th< td=""><td>F011</td><td>2005</td><td>NE</td><td></td><td>27</td><td>27</td><td>CDC010</td><td>Cluster-3</td><td>prn2</td><td>ptxP3</td><td>ptxA1</td><td>fimH1</td><td>CP011176</td><td>278x</td><td>5</td></th<>	F011	2005	NE		27	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP011176	278x	5			
Pi34 2005 CA NT 27 CDC/002 Singleton pn2 prp3 prp41 fimH2 CP011177 288x This F569 2006 GA NT 27 CDC/046 Cluster-8 pn2 phP3 ptx41 fimH2 CP025537 400x 5 F578 2007 NC 27 27 CDC/046 Cluster-5 pn2 ptP3 ptx41 fimH2 CP025342 505x This study F687 2007 NC 27 27 CDC/046 Cluster-5 pn2 ptP3 ptx41 fimH2 CP011178 302x 5 F687 2008 NC 36 36 CDC/022 Cluster-1 pn2 ptA1 fimH2 CP011178 302x 5 F687 2008 NC 36 36 CDC/042 Cluster-1 pn2 ptA1 fimH2 CP011181 275x 5 F934 2007 IL NT	F013	2005	NE		27	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP016965	361x	This study			
Föri 2004 NY NT 27 CDC/046 Cluster-8 pm2 ptP3 ptxA1 fimH2 CP03523 320x 5 F669 2006 GA NT 27 CDC/046 Cluster-8 pm2 ptP3 ptxA1 fimH2 CP025327 400x This study F878 2007 NC 27 CDC/046 Cluster-1 pm2 ptP3 ptxA1 fimH2 CP025342 565x This study F878 2007 NC 27 CDC/013 Cluster-1 pm2 ptP3 ptxA1 fimH2 CP011178 305x 5 F878 2008 VA 27 CDC/013 Cluster-2 pm2 ptP3 ptxA1 fimH2 CP011180 392x 5 F844 2008 VA 27 CDC/013 Singleton pm2 ptP3 ptxA1 fimH2 CP013873 372x 5 F948 2007 <il< td=""> NT 16 CDC/13</il<>	F034	2005	CA		NT	27	CDC002	Singleton	prn2	ptxP3	ptxA1	fimH1	CP011177	288x	5			
Få69 2006 GA NT 27 CDC046 Cluster-8 pm2 ptxP3 pbcA1 fimH2 CP025352 370x This study F578 2007 NC 27 CDC046 Singleton pm2 ptxP3 ptxA1 fimH2 CP025357 400x This study F657 2007 NC 27 CDC046 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP02537 400x This study F657 2008 CO 176 CDC042 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP011371 400x This study F667 2008 VA 27 CDC013 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP011180 352x 5 F647 2008 VA 27 CDC002 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP011182 35x 5 F948 2007 L NT 27 <td< td=""><td>F501</td><td>2004</td><td>NY</td><td></td><td>NT</td><td>27</td><td>CDC046</td><td>Cluster-8</td><td>prn2</td><td>ptxP3</td><td>ptxA1</td><td>fimH2</td><td>CP013870</td><td>320x</td><td>5</td></td<>	F501	2004	NY		NT	27	CDC046	Cluster-8	prn2	ptxP3	ptxA1	fimH2	CP013870	320x	5			
F778 2007 MS 27 27 CDC046 Singleton pm2 ptxP3 ptxA1 timt/2 CP025342 500x This study F580 2007 NC 27 27 CDC082 Cluster-5 pm2 ptxP3 ptxA1 timt/2 CP025342 505x This study F656 2007 CO NT 179 CDC013 Cluster-5 pm2 ptxP3 ptxA1 timt/2 CP011179 302x 5 F670 2008 NC 36 36 CDC002 Cluster-5 pm2 ptxP3 ptxA1 timt/2 CP011179 302x 5 F687 2008 VA 27 27 CDC046 Cluster-1 pm2 ptxP3 ptxA1 timt/2 CP011373 383x 5 F734 2007 LL NT 158 CDC013 Singleton pm1-signal_seq_del ptxA1 timt/2 CP013874 410x 5 5 5	F569	2006	GA		NT	27	CDC046	Cluster-8	prn2	ptxP3	ptxA1	fimH2	CP025523	370x	This study			
F80 2007 NC 27 27 CDC082 Cluster-6 pm2 ptxP3 ptxA1 fimHz CP023342 505. This study F657 2007 NC NT 179 CDC013 Cluster-5 pm2 ptxP3 ptxA1 fimHz CP011178 305.x 5 F670 2008 NA 27 CDC013 Cluster-5 pm2 ptxA3 ptxA1 fimHz CP011178 305.x 5 F687 2008 NA 27 27 CDC012 Cluster-2 pm12 ptxA3 ptxA1 fimHz CP011181 275.x 5 F684 2009 GA NT 27 CDC044 Cluster-1 pm2 ptxA3 ptxA1 firmHz CP013673 410.x 5 F948 2007 IL NT 186 CDC260 Singleton pm1-signal_seq_del ptxP3 ptxA1 firmHz CP013673 410.x 5 605 50 70	E578	2007	MS		27	27	CDC046	Singleton	prn2	ntxP3	ntxA1	fimH2	CP025357	400x	This study			
F657 2007 CO NT 179 CDC013 Cluster-1 <i>pm2 ptxP3 ptxA1 fim42</i> CP011871 440x 5 F658 2008 CO 176 176 CDC013 Cluster-5 <i>pm12 ptxP3 ptxA1 fim42</i> CP011178 302x 5 F684 2008 NC 36 36 CDC002 Cluster-5 <i>pm12 ptxP3 ptxA1 fimH2</i> CP011180 392x 5 F687 2008 VA 277 CDC002 Cluster-1 <i>pm2 ptxP3 ptxA1 fimH4</i> CP011877 440x 5 F948 2007 IL NT 158 CDC171 Singleton <i>pm2 ptxP3 ptxA1 fimH4</i> CP011827 440x 5 F944 2007 IL NT 158 CDC13 Singleton <i>pm2 ptxP3 ptxA1 fimH4</i> CP011823 48x 5 5 <	F580	2007	NC		27	27	CDC082	Cluster-5	prn2	ntxP3	ntxA1	fimH2	CP025342	505x	This study			
F658 2008 CO 176 176 CDC082 Cluster-5 pr/2 pt/2	F657	2007	CO		NT	179	CDC013	Cluster-1	prn2	ntxP3	ntxA1	fimH2	CP013871	440x	5			
Loco Loco Loco Dock 1 Loco Dock 1 ImmL2 Constraint Dock 1 ImmL2 Constraint Source Source <td>F658</td> <td>2007</td> <td>00</td> <td></td> <td>176</td> <td>176</td> <td>CDC082</td> <td>Cluster-5</td> <td>prn2</td> <td>ntvP3</td> <td>$ptx \Delta 1$</td> <td>fimH2</td> <td>CP011178</td> <td>302v</td> <td>5</td>	F658	2007	00		176	176	CDC082	Cluster-5	prn2	ntvP3	$ptx \Delta 1$	fimH2	CP011178	302v	5			
Local Local <th< td=""><td>F670</td><td>2000</td><td>W/A</td><td></td><td>27</td><td>27</td><td>CDC013</td><td>Cluster-1</td><td>prn2</td><td>ntvP3</td><td>ptxA1</td><td>fimH2</td><td>CP011179</td><td>305x</td><td>5</td></th<>	F670	2000	W/A		27	27	CDC013	Cluster-1	prn2	ntvP3	ptxA1	fimH2	CP011179	305x	5			
Los 2003 I/O 30 CDC002 Cluster-1 pm2 ptk/3 ptk/3 fim/1 CP011181 275x 5 F778 2004 OH 27 27 CDC002 Cluster-11 pm2 ptk/3 ptk/3 fim/12 CP011873 410x 5 F948 2007 IL NT 158 CDC101 Singleton pm2 ptk/3 ptk/41 fim/12 CP011873 410x 5 F948 2007 IL 18 EOC200 Singleton pm2 ptk/3 ptk/41 fim/12 CP011887 410x 5 G057 2004 MN NT 26 CDC013 Cluster-1 pm2 ptk/3 ptk/1 fim/12 CP013874 416x 5 G022 2008 MA NT 27 CDC013 Cluster-1 pm2 ptk/3 ptk/1 fim/12 CP013876 406x 5 G035 2002 MN NT </td <td>F684</td> <td>2003</td> <td>NC</td> <td></td> <td>36</td> <td>36</td> <td>CDC082</td> <td>Cluster-5</td> <td>prn2</td> <td>ntvP3</td> <td>ptxA1</td> <td>fimH2</td> <td>CP011180</td> <td>302v</td> <td>5</td>	F684	2003	NC		36	36	CDC082	Cluster-5	prn2	ntvP3	ptxA1	fimH2	CP011180	302v	5			
100 200 VA 21 CDCOUL Cluster 2 pint pint< pint< pint< pint< pint< pint<	F687	2000			27	27		Cluster-2	prn2	ntvP3	$p t x \Lambda 1$	fimH1	CP011181	275v	5			
PT/F 2004 OP 21 21 CDC046 Cluster 11 pt/L3 pt/L3 <th l3<="" th=""> <th l3<="" th=""> <th l3<="" th=""></th></th></th>	<th l3<="" th=""> <th l3<="" th=""></th></th>	<th l3<="" th=""></th>		E770	2000			27	27		Cluster-2	prn2	pixr 3	pixA1	fimU2	CF011101	213X 507v	5
P354 2009 GA N1 27 CDC013 Singleton pIIII pIIII pIIIII CP013br3 410x 5 F948 2007 IL 186 CDC111 Singleton pm1-signal_seq_del ptkP3 ptkA1 fimH1 CP013br3 338x This study G057 2004 MN NT 26 CDC013 Singleton pm2 ptkP3 ptkA1 fimH2 CP013br3 416x 5 G085 2008 MA NT 16 CDC013 Cluster-1 pm2 ptkP3 ptkA1 fimH2 CP013br3 416x 5 G102 2008 MA NT 27 CDC010 Cluster-3 pm2 ptkP3 ptkA1 fimH2 CP013br3 56 5 G965 2002 MN NT 16 CDC042 Singleton pm2 ptkP3 ptkA1 fimH2 CP013br3 55 5 54 54 54 54	F//0 E024	2004				27	CDC040	Singlaton	prn2	pixr3	ρ_{IXAI}	fimU2	CF013072	JZ7X	5			
P346 2007 IL NI 136 CDC/11 Singleton pin2 pbxA2 fimH1 CP01122 388x This study G057 2004 MN NT 26 CDC013 Cluster-1 pm2 pbxA3 pbxA1 fimH2 CP01129 369x 5 G057 2008 MA NT 26 CDC013 Cluster-1 pm2 pbxP3 pbxA1 fimH2 CP013874 416x 5 G102 2008 MA NT 27 CDC010 Cluster-1 pm2 pbxP3 pbxA1 fimH2 CP013876 406x 5 G965 2009 MA NT 27 CDC046 Cluster-5 pm2 pbxP3 pbxA1 fimH2 CP011694 246x 5 H348 2010 GA 158 158 CDC082 Cluster-5 pm2 pbxP3 pbxA1 fimH2 CP011694 246x 5 H348 2010 GA <td>F934</td> <td>2009</td> <td>GA</td> <td></td> <td></td> <td>21</td> <td>CDC013</td> <td>Singleton</td> <td>prinz</td> <td>$\mu x F S$</td> <td>pixA1</td> <td>fim 10</td> <td>CF013073</td> <td>410X</td> <td>5</td>	F934	2009	GA			21	CDC013	Singleton	prinz	$\mu x F S$	pixA1	fim 10	CF013073	410X	5			
Pisod Loof Loof Loof CDC200 Singleton pm2 ptxA1 limH2 CP012129 Saok Saok <td>F940</td> <td>2007</td> <td>1</td> <td></td> <td>106</td> <td>100</td> <td></td> <td>Singleton</td> <td>prinz</td> <td>$\rho_{\rm LXP3}$</td> <td>pixA1</td> <td>IIIII⊓Z fimLl1</td> <td>CPUTTI62</td> <td>300X</td> <td>D This study</td>	F940	2007	1		106	100		Singleton	prinz	$\rho_{\rm LXP3}$	pixA1	IIIII⊓Z fimLl1	CPUTTI62	300X	D This study			
Gos Z004 MN NT 26 CDC013 Singleton pm2 ptxA1 fimH2 CP012129 369x 5 G102 2008 MA NT 16 CDC013 Singleton pm2 ptxP3 ptxA1 fimH2 CP013874 416x 5 G102 2008 MA NT 27 CDC010 Cluster-1 pm2 ptxP3 ptxA1 fimH1 CP013874 406x 5 G965 2009 MA NT 16 CDC013 Cluster-1 pm2 ptxP3 ptxA1 fimH1 CP013874 406x 5 H344 2009 MA NT 27 CDC046 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP011384 407x 5 H348 2010 GA 158 158 CDC082 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP013874 46x 5 H348 2010 GA	F954	2007				100		Singleton	prn -signal_seq_dei	pixP1	pixA2	iiiii⊓ i fime UO	CP025366	330X				
Gobs 2008 MA NT 16 CDC013 Singleton pm2 ptxA1 timH2 CP0138/4 416x 5 G102 2008 MA NT 27 CDC010 Cluster-1 pm2 ptxA3 ptxA1 fimH2 CP02538/4 416x 5 G807 2009 MN NT 27 CDC013 Cluster-1 pm2 ptxA3 ptxA1 fimH2 CP013876 406x 5 H034 2009 MA NT 27 CDC046 Cluster-5 pm2 ptxA1 fimH2 CP013876 36x 5 H346 2010 GA 158 158 CDC082 Cluster-5 pm2 ptxA3 ptxA1 fimH2 CP0113877 366x 5 H341 2010 GA 27 27 CDC046 Cluster-7 pm2 ptxA3 ptxA1 fimH2 CP0113878 244x 5 H382 2010 CA 27	G057	2004	IVIN			26	CDC013	Cluster-1	prn2	ptxP3	ptxA1	timH2	CP012129	369X	5			
G102 2008 MA NI 27 CDC013 Cluster-1 pm2 ptxP3 ptxA1 timmL2 CP025388 248x Instudy G807 2005 MN NT 16 CDC010 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP013876 83x 5 5 H034 2009 MA NT 27 CDC042 Singleton pm2 ptxP3 ptxA1 fimH2 CP013876 83x 5 5 H346 2010 GA 158 158 CDC082 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP011694 236x 5 H348 2010 GA 27 27 CDC046 Cluster-9 pm2 ptxP3 ptxA1 fimH2 CP013878 24x 5 5 H382 2010 CA 27 27 CDC270 Singleton pm2 ptxP3 ptxA1 fimH2 CP011387 24x 5 5 5 5 5 <td>G085</td> <td>2008</td> <td>MA</td> <td></td> <td></td> <td>16</td> <td>CDC013</td> <td>Singleton</td> <td>prnz</td> <td>ptxP3</td> <td>ptxA1</td> <td>timH2</td> <td>CP013874</td> <td>416X</td> <td>5</td>	G085	2008	MA			16	CDC013	Singleton	prnz	ptxP3	ptxA1	timH2	CP013874	416X	5			
G80/r 2005 MN NI 27 CDC10 Cluster-3 pm2 ptxP3 ptxA1 timH1 CP013876 383x 5 H034 2009 MA NT 16 CDC13 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP013876 383x 5 H034 2009 MA NT 27 CDC242 Singleton pm2 ptxP3 ptxA1 fimH2 CP013876 383x 5 H346 2010 GA 158 158 CDC082 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP013877 366x 5 H348 2010 GA 27 27 CDC046 Cluster-8 pm2 ptxP3 ptxA1 fimH2 CP013878 224x 5 H382 2010 CA 27 27 CDC020 Singleton pm2 ptxP3 ptxA1 fimH2 CP013878 24x 5 H332 2006	G102	2008	MA		NI	27	CDC013	Cluster-1	prn2	ptxP3	ptxA1	timH2	CP025388	248x	This study			
G965 2002 MN NI 16 CDC013 Cluster-1 pm2 ptxP3 ptxA1 timH2 CP013876 406x 5 H034 2009 FL 27 CDC042 Singleton pm2 ptxP3 ptxA1 fimH2 CP013376 406x 5 H346 2010 GA 158 158 CDC082 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP0113376 366x 5 H348 2010 GA 27 27 CDC082 Cluster-8 pm2 ptxP3 ptxA1 fimH1 CP013878 224x 5 H381 2010 MA 27 27 CDC046 Cluster-1 pm2 ptxP3 ptxA1 fimH1 CP013878 224x 5 H382 2010 CA 27 27 CDC013 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP013878 248x 5 H437 2006 TN	G807	2005	MN		NI	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	timH1	CP013875	383x	5			
H034 2009 MA NT 27 CDC242 Singleton pm2 ptxP3 ptxA1 fimH2 CP025356 322x This study H320 2009 FL 27 27 CDC046 Cluster-8 pm2 ptxP3 ptxA1 fimH2 CP011234 407x 5 H346 2010 GA 158 158 CDC082 Cluster-9 pm2 ptxP3 ptxA1 fimH1 CP011874 266x 5 H346 2010 GA 27 27 CDC046 Cluster-8 pm2 ptxP3 ptxA1 fimH1 CP013877 366x 5 H382 2010 CA 27 27 CDC046 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP011878 24x 5 H437 2006 TN NT 77 CDC013 Cluster-7 pm2 ptxP3 ptxA1 fimH2 CP011183 249x 5 H533 2009 IN NT N/A CDC237 Cluster-7 pm2 ptxP3 ptxA1	G965	2002	MN		NT	16	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP013876	406x	5			
H320 2009 FL 27 27 CDC046 Cluster-8 pm2 ptxP3 ptxA1 fimH2 CP011234 407x 5 H346 2010 GA 158 158 CDC082 Cluster-5 prm2 ptxP3 ptxA1 fimH2 CP011694 236x 5 H348 2010 GA 27 27 CDC046 Cluster-9 prm2 ptxP3 ptxA1 fimH2 CP013878 224x 5 H382 2010 CA 27 27 CDC013 Cluster-1 prm2 ptxP3 ptxA1 fimH2 CP013878 224x 5 H437 2006 TN NT 77 CDC013 Cluster-1 prm2 ptxP3 ptxA1 fimH2 CP01195 258x 5 H533 2009 IN NT 77 CDC013 Cluster-7 prm2 ptxP3 ptxA1 fimH1 CP01895 258x 5 H540 2010 SC NT 28 CDC237 Cluster-3 prm2 ptxP3 ptxA1	H034	2009	MA		NT	27	CDC242	Singleton	prn2	ptxP3	ptxA1	fimH2	CP025356	322x	This study			
H346 2010 GA 158 158 CDC082 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP011694 236x 5 H348 2010 GA 27 27 CDC082 Cluster-9 ptr2 ptxP3 ptxA1 fimH1 CP011877 366x 5 H382 2010 CA 27 27 CDC046 Cluster-8 ptr2 ptxP3 ptxA1 fimH1 CP013878 224x 5 H382 2010 CA 27 27 CDC013 Cluster-1 ptr2 ptxP3 ptxA1 fimH2 CP011805 258x 5 H437 2006 TN NT 77 CDC013 Cluster-7 ptr2 ptxP3 ptxA1 fimH2 CP01183 249x 5 H533 2009 IN NT 27 CDC237 Cluster-3 ptr2 ptxP3 ptxA1 fimH1 CP013879 288x 5 H540 2010 SC NT 27 CDC237 Cluster-3 ptxP3 ptxA1 fimH1 <td>H320</td> <td>2009</td> <td>FL</td> <td></td> <td>27</td> <td>27</td> <td>CDC046</td> <td>Cluster-8</td> <td>prn2</td> <td>ptxP3</td> <td>ptxA1</td> <td>fimH2</td> <td>CP011234</td> <td>407x</td> <td>5</td>	H320	2009	FL		27	27	CDC046	Cluster-8	prn2	ptxP3	ptxA1	fimH2	CP011234	407x	5			
H348 2010 GA 27 27 CDC082 Cluster-9 pm2 ptxP3 ptxA1 fimH1 CP013877 366x 5 H361 2010 MA 27 27 CDC046 Cluster-8 pm2 ptxP3 ptxA1 fimH2 CP013878 224x 5 H382 2010 CA 27 27 CDC013 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP011085 258x 5 H437 2006 TN NT 77 CDC013 Cluster-7 pm2 ptxP3 ptxA1 fimH1 CP013879 288x 5 H540 2010 SC NT 28 CDC237 Cluster-4 pm2-15481-1613fwd ptxP3 ptxA1 fimH1 CP013880 292x 5 H541 2010 SC NT 28 CDC237 Cluster-3 pm2 ptxP3 ptxA1 fimH1 CP013880 292x 5 H541 2010 SC NT 27 CDC270 Singleton pm2 ptxP3 pt	H346	2010	GA		158	158	CDC082	Cluster-5	prn2	ptxP3	ptxA1	fimH2	CP011694	236x	5			
H361 2010 MA 27 27 CDC046 Cluster-8 pm2 ptxP3 ptxA1 fimH2 CP013878 224x 5 H382 2010 CA 27 27 CDC270 Singleton pm2 ptxP3 ptxA1 fimH1 CP013878 224x 5 H37 2006 TN NT 77 CDC013 Cluster-1 pm2 ptxP3 ptxA1 fimH1 CP013878 224x 5 H520 2009 IN NT 77 CDC013 Cluster-7 pm2 ptxP3 ptxA1 fimH2 CP011183 249x 5 H533 2009 IN NT N/A CDC217 Cluster-4 pm2-16481-1613fwd ptxP3 ptxA1 fimH1 CP013879 288x 5 H541 2010 SC NT 27 CDC237 Cluster-3 pm2 ptxP3 ptxA1 fimH1 CP013880 292x 5 H541 2010 SC NT 27 CDC237 Cluster-3 pm2 ptxP3 pt	H348	2010	GA		27	27	CDC082	Cluster-9	prn2	ptxP3	ptxA1	fimH1	CP013877	366x	5			
H382 2010 CA 27 27 CDC270 Singleton pm2 ptxP3 ptxA1 fimH1 CP013082 334x 5 H437 2006 TN NT 77 CDC013 Cluster-1 pm2 ptxP3 ptxA1 fimH2 CP011095 258x 5 H520 2009 IN NT 27 CDC013 Cluster-7 pm2 ptxP3 ptxA1 fimH2 CP011832 249x 5 H533 2009 IN NT V/A CDC217 Cluster-9 pm2 ptxP3 ptxA1 fimH1 CP013879 28x 5 H540 2010 SC NT 28 CDC237 Cluster-4 pm2-1S481-1613fwd ptxP3 ptxA1 fimH1 CP013880 292x 5 H541 2010 SC NT 27 CDC237 Cluster-13 pm2 ptxP3 ptxA1 fimH1 CP013881 249x 5 H636 2010 TN 16 6 CDC278 Singleton pm2 ptxP3 pt	H361	2010	MA		27	27	CDC046	Cluster-8	prn2	ptxP3	ptxA1	fimH2	CP013878	224x	5			
H4372006TNNT77CDC013Cluster-1pm2ptxP3ptxA1fimH2CP011695258x5H5202009INNT27CDC013Cluster-7pm2ptxP3ptxA1fimH2CP011183249x5H5332009INNTN7CDC217Cluster-9pm2ptxP3ptxA1fimH1CP013879288x5H5402010SCNT28CDC237Cluster-4pm2-lS481-1613fwdptxP3ptxA1fimH1CP025373266xThis studyH5412010SCNT27CDC237Cluster-13pm2ptxP3ptxA1fimH1CP025373266xThis studyH5792010PA128128CDC013Cluster-13pm2ptxP3ptxA1fimH1CP025529391xThis studyH6362010TN1616CDC278Singletonpm2-wt-C638TptxP3ptxA1fimH2CP011185270x5H6372010TNNT27CDC123Cluster-7pm2-stop-C739TptxP3ptxA1fimH2CP011185270x5H6392010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP01130209x5H6392010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP01130209x5H64020	H382	2010	CA		27	27	CDC270	Singleton	prn2	ptxP3	ptxA1	fimH1	CP013082	334x	5			
H5202009INNT27CDC013Cluster-7pm2ptxP3ptxA1fimH2CP011183249x5H5332009INNTNTN/ACDC217Cluster-9pm2ptxP3ptxA1fimH1CP013879288x5H5402010SCNT28CDC237Cluster-4pm2-IS481-1613fwdptxP3ptxA1fimH1CP013880292x5H5412010SCNT27CDC237Cluster-3pm2ptxP3ptxA1fimH1CP025373266xThis studyH5792010PA128128CDC013Cluster-13pm2ptxP3ptxA1fimH1CP025529391xThis studyH6362010TN1616CDC278Singletonpm2-wt-C638TptxP3ptxA1fimH2CP011185270x5H6372010TNNT27CDC123Cluster-7pm2-Stop-C739TptxP3ptxA1fimH2CP011185270x5H6392010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP02130209x5H6402010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP025360427xThis studyH6422010MI3636CDC046Cluster-11pm9ptxP3ptxA1fimH2CP025360427xThis study<	H437	2006	ΤN		NT	77	CDC013	Cluster-1	prn2	ptxP3	ptxA1	fimH2	CP011695	258x	5			
H5332009INNTN/ACDC217Cluster-9pm2ptxP3ptxP1fimH1CP013879288x5H5402010SCNT28CDC237Cluster-4pm2-IS481-1613fwdptxP3ptxA1fimH1CP013880292x5H5412010SCNT27CDC237Cluster-3pm2ptxP3ptxA1fimH1CP013880292x5H5412010SCNT27CDC237Cluster-3pm2ptxP3ptxA1fimH1CP025373266xThis studyH5792010PA128128CDC013Cluster-13pm2ptxP3ptxA1fimH1CP02529391xThis studyH6362010TN1616CDC278Singletonpm2-wt-C638TptxP3ptxA1fimH2CP011881249x5H6372010TNNT27CDC123Cluster-7pm2-Stop-C739TptxP3ptxA1fimH2CP011185270x5H6392010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP025371406xThis studyH6402010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP025371406xThis studyH6422010CA2727CDC013Singletonpm9ptxP3ptxA1fimH2CP025360427xThis study	H520	2009	IN		NT	27	CDC013	Cluster-7	prn2	ptxP3	ptxA1	fimH2	CP011183	249x	5			
H5402010SCNT28CDC237Cluster-4prn2-IS481-1613fwdptxP3ptxA1fimH1CP013880292x5H5412010SCNT27CDC237Cluster-3prn2ptxP3ptxA1fimH1CP025373266xThis studyH5792010PA128128CDC013Cluster-13prn2ptxP3ptxA1fimH1CP025373266xThis studyH6242010ORY2727CDC270Singletonprn2ptxP3ptxA1fimH1CP025529391xThis studyH6362010TN1616CDC278Singletonprn2-wt-C638TptxP3ptxA1fimH2CP011881249x5H6372010TNNT27CDC123Cluster-7prn2-Stop-C739TptxP3ptxA1fimH2CP01185270x5H6392010MI3636CDC046Cluster-11prn2ptxP3ptxA1fimH2CP012130209x5H6402010MI3636CDC046Cluster-11prn2ptxP3ptxA1fimH2CP025371406xThis studyH6422010CA2727CDC013Singletonprn9ptxP3ptxA1fimH2CP025360427xThis studyH6452010MAY2727CDC013Cluster-1prn9ptxP3ptxA1fimH1CP011186 <td< td=""><td>H533</td><td>2009</td><td>IN</td><td></td><td>NT</td><td>N/A</td><td>CDC217</td><td>Cluster-9</td><td>prn2</td><td>ptxP3</td><td>ptxA1</td><td>fimH1</td><td>CP013879</td><td>288x</td><td>5</td></td<>	H533	2009	IN		NT	N/A	CDC217	Cluster-9	prn2	ptxP3	ptxA1	fimH1	CP013879	288x	5			
H5412010SCNT27CDC237Cluster-3pm2ptxP3ptxP1fimH1CP025373266xThis studyH5792010PA128128CDC013Cluster-13pm2ptxP3ptxA1fimH2CP011184336x5H6242010ORY2727CDC270Singletonpm2ptxP3ptxA1fimH1CP025529391xThis studyH6362010TN1616CDC278Singletonpm2-wt-C638TptxP3ptxA1fimH2CP01185249x5H6372010TNNT27CDC123Cluster-7pm2-stop-C739TptxP3ptxA1fimH2CP011185270x5H6392010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP012130209x5H6402010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP025371406xThis studyH6422010CA2727CDC013Singletonpm9ptxP3ptxA1fimH2CP025360427xThis studyH6552010MAY2727CDC013Cluster-1pm2ptxP3ptxA1fimH1CP011186342x5H6722010MAY2727CDC013Cluster-1pm2ptxP3ptxA1fimH1CP011186342x5 <td>H540</td> <td>2010</td> <td>SC</td> <td></td> <td>NT</td> <td>28</td> <td>CDC237</td> <td>Cluster-4</td> <td>prn2-IS481-1613fwd</td> <td>ptxP3</td> <td>ptxA1</td> <td>fimH1</td> <td>CP013880</td> <td>292x</td> <td>5</td>	H540	2010	SC		NT	28	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP013880	292x	5			
H579 2010 PA 128 128 CDC013 Cluster-13 prn2 ptxP3 ptxA1 fimH2 CP011184 336x 5 H624 2010 OR Y 27 27 CDC270 Singleton prn2 ptxP3 ptxA1 fimH1 CP025529 391x This study H636 2010 TN 16 16 CDC278 Singleton prn2-wt-C638T ptxP3 ptxA1 fimH2 CP011881 249x 5 H637 2010 TN NT 27 CDC123 Cluster-7 prn2-Stop-C739T ptxP3 ptxA1 fimH2 CP011185 270x 5 H639 2010 MI 36 36 CDC046 Cluster-11 prn2 ptxP3 ptxA1 fimH2 CP012130 209x 5 H640 2010 MI 36 36 CDC046 Cluster-11 prn2 ptxP3 ptxA1 fimH2 CP025371 406x This study H642 2010 CA 27 27 CDC013 Singleton	H541	2010	SC		NT	27	CDC237	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP025373	266x	This study			
H624 2010 OR Y 27 27 CDC270 Singleton prn2 ptxP3 ptxA1 fimH1 CP025529 391x This study H636 2010 TN 16 16 CDC278 Singleton prn2-wt-C638T ptxP3 ptxA1 fimH2 CP013881 249x 5 H637 2010 TN NT 27 CDC123 Cluster-7 prn2-stop-C739T ptxP3 ptxA1 fimH2 CP013881 249x 5 H639 2010 MI 36 36 CDC046 Cluster-11 prn2 ptxP3 ptxA1 fimH2 CP012130 209x 5 H640 2010 MI 36 36 CDC046 Cluster-11 prn2 ptxP3 ptxA1 fimH2 CP012130 209x 5 H642 2010 CA 27 27 CDC013 Singleton prn9 ptxP3 ptxA1 fimH2 CP025360 427x This study H645 2010 MA Y 27 27 CDC013 C	H579	2010	PA		128	128	CDC013	Cluster-13	prn2	ptxP3	ptxA1	fimH2	CP011184	336x	5			
H636 2010 TN 16 16 CDC278 Singleton pm2-wt-C638T ptxP3 ptxA1 fimH2 CP013881 249x 5 H637 2010 TN NT 27 CDC123 Cluster-7 pm2-Stop-C739T ptxP3 ptxA1 fimH2 CP013881 249x 5 H639 2010 MI 36 36 CDC046 Cluster-11 pm2 ptxP3 ptxA1 fimH2 CP012130 209x 5 H640 2010 MI 36 36 CDC046 Cluster-11 pm2 ptxP3 ptxA1 fimH2 CP025371 406x This study H642 2010 CA 27 27 CDC013 Singleton pm9 ptxP3 ptxA1 fimH2 CP025360 427x This study H665 2010 MA Y 27 27 CDC013 Cluster-4 pm2-IS481-1613fwd ptxP3 ptxA1 fimH1 CP011186 342x 5 H672 2010 MN Y 27 27 CDC013	H624	2010	OR	Y	27	27	CDC270	Singleton	prn2	ptxP3	ptxA1	fimH1	CP025529	391x	This study			
H6372010TNNT27CDC123Cluster-7pm2-Stop-C739TptxP3ptxA1fimH2CP011185270x5H6392010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP012130209x5H6402010MI3636CDC046Cluster-11pm2ptxP3ptxA1fimH2CP025371406xThis studyH6422010CA2727CDC013Singletonpm9ptxP3ptxA1fimH2CP025360427xThis studyH6652010MAY2727CDC013Cluster-4pm2-IS481-1613fwdptxP3ptxA1fimH1CP011186342x5H6722010MNY2727CDC013Cluster-1pm2ptxP3ptxA1fimH2CP025349443xThis studyH6772010OR2727CDC013Cluster-1pm14UNKptxA1fimH2CP025367417xThis study	H636	2010	ΤN		16	16	CDC278	Singleton	prn2-wt-C638T	, ptxP3	ptxA1	fimH2	CP013881	249x	5			
H6392010MI3636CDC046Cluster-11prn2ptxP3ptxA1fimH2CP012130209x5H6402010MI3636CDC046Cluster-11prn2ptxP3ptxA1fimH2CP025371406xThis studyH6422010CA2727CDC013Singletonprn9ptxP3ptxA1fimH2CP025360427xThis studyH6652010MAY2727CDC013Cluster-4prn2-IS481-1613fwdptxP3ptxA1fimH1CP011186342x5H6722010MNY2727CDC013Cluster-1prn2ptxP3ptxA1fimH2CP025349443xThis studyH6772010OR2727CDC013Cluster-1prn14UNKptxA1fimH2CP025367417xThis study	H637	2010	ΤN		NT	27	CDC123	Cluster-7	prn2-Stop-C739T	, ptxP3	, ptxA1	fimH2	CP011185	270x	5			
H640 2010 MI 36 36 CDC046 Cluster-11 prn2 ptxP3 ptxA1 fimH2 CP025371 406x This study H642 2010 CA 27 27 CDC013 Singleton prn9 ptxP3 ptxA1 fimH2 CP025371 406x This study H655 2010 MA Y 27 27 CDC237 Cluster-4 prn2-IS481-1613fwd ptxP3 ptxA1 fimH1 CP011186 342x 5 H672 2010 MN Y 27 27 CDC013 Cluster-1 prn2 ptxP3 ptxA1 fimH2 CP025349 443x This study H677 2010 OR 27 27 CDC013 Cluster-1 prn14 UNK ptxA1 fimH2 CP025367 417x This study	H639	2010	MI		36	36	CDC046	Cluster-11	prn2	ptxP3	ptxA1	fimH2	CP012130	209x	5			
H642 2010 CA 27 27 CDC013 Singleton prn9 ptxP3 ptxA1 fimH2 CP025360 427x This study H665 2010 MA Y 27 27 CDC237 Cluster-4 prn2-IS481-1613fwd ptxP3 ptxA1 fimH2 CP025360 427x This study H672 2010 MN Y 27 27 CDC013 Cluster-1 prn2 ptxP3 ptxA1 fimH2 CP025349 443x This study H677 2010 OR 27 27 CDC013 Cluster-1 prn14 UNK ptxA1 fimH2 CP025367 417x This study	H640	2010	MI		36	36	CDC046	Cluster-11	prn2	ptxP3	ptxA1	fimH2	CP025371	406x	This study			
H665 2010 MA Y 27 27 CDC237 Cluster-4 prn2-IS481-1613fwd ptxP3 ptxA1 fimH1 CP011186 342x 5 H672 2010 MN Y 27 27 CDC013 Cluster-1 prn2 ptxP3 ptxA1 fimH2 CP025349 443x This study H677 2010 OR 27 27 CDC013 Cluster-1 prn14 UNK ptxA1 fimH2 CP025367 417x This study	H642	2010	CA		27	27	CDC013	Singleton	prn <u>9</u>	ntxP3	ntxA1	fimH2	CP025360	427x	This study			
H672 2010 MN Y 27 27 CDC013 Cluster-1 prn2 ptxP3 ptxA1 fimH2 CP025349 443x This study H677 2010 OR 27 27 CDC013 Cluster-1 prn2 ptxP3 ptxA1 fimH2 CP025349 443x This study H677 2010 OR 27 27 CDC013 Cluster-1 prn14 UNK ptxA1 fimH2 CP025367 417x This study	H665	2010	MA	Y	27	27	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP011186	342x	5			
H677 2010 OR 27 27 CDC013 Cluster-1 prn14 UNK ptxA1 fimH2 CP025367 417x This study	H672	2010	MN	Ŷ	27	27	CDC013	Cluster-1	prn2	ntxP3	ntxA1	fimH2	CP025349	443x	This study			
	H677	2010	OR		27	27	CDC013	Cluster-1	prn14	UNK	ptxA1	fimH2	CP025367	417x	This study			
H681 2009 PA NT 27 CDC013 Cluster-1 prp2-IS481-1613fwd ptyP3 ptyA1 fimH2 CP012078 335y 5	H681	2009	PA		NT	27	CDC013	Cluster-1	prn2-IS481-1613fwd	ntxP3	ntx∆1	fimH2	CP012078	335x	5			
H682 2009 PA NT 18 CDC125 Singleton prn2 ptxP3 ptxA1 fimH1 CP013083 260x 5	H682	2009	PA		NT	18	CDC125	Singleton	prn2	ptxP3	ptxA1	fimH1	CP013083	260x	5			

Accession	
ID Year State EPS ^T MLVA wasMLVA PFGE Structure prn ptxP ptxA fimH no. D	epth [‡] Reference
H697 2011 PA NT 27 CDC002 Cluster-2 prn2-Stop-C1273T ptxP3 ptxA1 fimH1 CP025365	75x This study
H698 2010 PA 27 27 CDC002 Cluster-2 prn2-Stop-C1273T ptxP3 ptxA1 fimH1 CP013084	66x 5
H703 2010 MN 27 27 CDC013 Singleton prn2 ptxP3 ptxA1 fimH2 CP011187	20x 5
H706 2010 CO 27 27 CDC010 Singleton prn2 ptxP3 ptxA1 fimH1 CP013085	72x 5
H707 2010 CO 120 120 CDC082 Cluster-5 pm2 ptxP3 ptxA1 fimH2 CP011188	79x 5
H709 2010 NY 16 16 CDC037 Cluster-12 prn2 ptxP3 ptxA1 fimH2 CP025364	70x This study
H710 2010 NY 16 16 CDC082 Cluster-5 prn2-wt-C638T ptxP3 ptxA1 fimH2 CP011236	29x 5
H729 2011 MA 27 27 CDC217 Cluster-9 prn2 ptxP3 ptxA1 fimH1 CP011189	39x 5
H730 2011 MA 22 22 CDC202 Singleton prn2 ptxP3 ptxA1 fimH2 CP013086	65x 5
H740 2011 GA NT 186 CDC266 Singleton prn1-signal seg del ptxP1 ptxA2 fimH1 CP011190	88x 5
H742 2011 FL 27 27 CDC002 Cluster-2 pm2 ptxP3 ptxA1 fimH1 CP025346	95x This study
H754 2011 PA NT 27 CDC010 Singleton pm2 ptxP3 ptxA1 fimH1 CP011191 '	84x 5
H765 2011 NY 27 27 CDC002 Cluster-2 prn2-Stop-C1273T ptxP3 ptxA1 fimH1 CP011192	59x 5
H771 2011 CA 27 27 CDC013 Cluster-1 prn2 ptxP3 ptxA1 fimH2 CP013087	07x 5
H778 2011 OR Y 27 27 CDC013 Cluster-1 prn2-IS481-1613rev ptxP3 ptxA1 fimH2 CP025362 '	07x This study
H784 2011 OR Y 27 27 CDC273 Singleton prp2 ptxP3 ptxA1 fimH2 CP011193	30x 5
H800 2011 MO 27 27 CDC253 Cluster-6 pm2 ptxP3 ptxA1 fimH1 CP011194	68x 5
H806 2011 EL 27 27 CDC010 Cluster-3 pro2- ptxP3 ptxA1 fimH1 CP011195	04x 5
promoter disrupt	•
H810 2011 CT Y 27 27 CDC237 Cluster-4 prr2-IS481-1613fwd ptxP3 ptxA1 fimH1 CP011196	35x 5
H811 2011 CT Y 27 18 CDC002 Cluster-2 pro2-IS481-1613rev ptxP3 ptxA1 fimH1 CP025361	53x This study
H812 2011 NM Y 27 27 CDC269 Cluster-14 pro2-IS481-240rey ptxP3 ptxA1 fimH1 CP011197 '	56x 5
H813 2011 NM Y 27 27 CDC269 Cluster-14 pm2-IS481-240rev ptxP3 ptxA1 fimH1 CP025351	71x This study
H814 2007 NM 27 27 CDC013 Cluster-1 pr/2 ptxP3 ptxA1 fimH2 CP025374	12x This study
H834 2009 NM 27 27 CDC013 Cluster-7 pm2 ptxP3 ptxA1 fimH2 CP011235 (05x 5
H851 2011 CA 27 27 CDC253 Cluster-6 prn2-IS481-240rev ptxP3 ptxA1 fimH1 CP011237	50x 5
H852 2011 MN Y 27 27 CDC024 Cluster-15 pm2 ptxP3 ptxP3 ptxA1 fimH1 CP012079	62x 5
H877 2012 QR Y 32 32 CDC046 Singleton pm2 ptxP3 ptxA1 fimH2 CP025382	88x This study
Hence 2011 NY 36 36 CDC217 Cluster-9 pm^2 $ptrP3 ptrA1 fimH1 CP025363$	34x This study
H911 2012 GA 27 27 CDC237 Cluster-4 pr2-IS481-1613fwd ptrP3 ptxA1 fimH1 CP011238	60x 5
H915 2011 WA 27 27 CDCD46 Cluster-8 nr/2-IS481-2735fwd ptr/3 ntx 41 fimH2 CP011239	73x 5
Here 2π 2π 2π 2π 2π 2π 2π 2π	33x This study
1075 2012 QR Y 27 27 CDC326 Cluster-15 pm2 ptr/2 ptxP3 ptxA1 fimH1 CP011240	22x 5
1112 2012 CO Y 27 27 CDC237 Cluster-4 pro2-IS481-1613fwd ptrP3 ptrA1 fimH1 CP011241	48x 5
112 2012 NY Y 27 27 CDC002 Cluster-2 $prizeStor-C1273T$ $ptxP3$ $ptxP1$ $fmH1$ CP025370	77x This study
1182 2012 WA 158 158 CDC002 Cluster-2 pro2-15481-1613rev ptrP3 ptx/1 fimH1 CP026996	76x This study
1187 2012 A7 27 27 CDC253 Cluster-6 priz-15/481-240rey ptr/23 ptr/41 fimH1 CP012132	41x 5
188 2012 GA 27 27 CDC002 Singleton $nn2$ $ntxP3$ $ntxA1$ $fimH1$ CP025379	93x This study
1223 2012 El 27 27 CDC123 Cluster-7 pm2 ptrP3 ptrA1 fimH2 CP025369	36x This study
1220 2012 FL 27 27 CDC046 Singleton nr_2 (S481.1613rev ntr_2 ntr_4 1 $fimH_2$ CP011198	34y 5
1238 2012 CA 27 27 CDC010 Cluster-3 <i>priz</i> Co-romoter dis <i>ptxP3 ptxP1 fimH2</i> CP011190	48x 5
250 2012 NY Y 27 27 CDC002 Cluster-2 m^2 Stor-C1273T $ntrP3$ $ntrA1$ $fimH1$ CP012133	40X 5
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	65x This study
1372 2012 NM Y 27 27 CDC082 Cluster-5 nrn2/15/8/11/613ray ntv23 ntv41 fimH2 CD055372	58x This study
1373 2012 NM Y 27 27 CDC253 Cluster-6 pm2.1040100 ptr 3 ptr 4 fimHz CP011200 $^{\circ}$	92x 5
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	25x 5
1420 2012 WA 27 27 CDCO02 Cluster-2 pm2 of the part of parts pixed in min of the pixed of the pi	03x This study

								Moleci	ular typing	loci				
												Accession		
ID	Year	State	EPS [†]	MLVA	wgsMLVA	PFGE	Structure	prn	ptxP	ptxA	fimH	no.	Depth [‡]	Reference
1439	2012	CO	Y	27	27	CDC300	Cluster-16	<i>prn2</i> -IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP025375	441x	This study
I462	2012	СТ	Y	27	27	CDC253	Cluster-6	<i>prn2</i> -IS <i>4</i> 81-240rev	ptxP3	ptxA1	fimH1	CP025376	432x	This study
1598	2013	WA		27	27	CDC010	Cluster-3	<i>prn2</i> -promoter_dis	ptxP3	ptxA1	fimH1	CP025380	744x	This study
1602	2013	GA		36	36	CDC237	Cluster-4	<i>prn2</i> -IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP011202	289x	5
1623	2012	VT		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP025386	457x	This study
1692	2011	VT		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP025378	394x	This study
1705	2011	VT		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP025524	113x	This study
1730	2013	CO	Y	27	27	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP011203	291x	5
1752	2013	СТ	Y	313	313	CDC074	Singleton	prn2	ptxP3	ptxA1	fimH1	CP011204	441x	5
1763	2013	VT		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP011205	438x	5
1859	2012	VT		27	27	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP025477	197x	This study
1892	2007	VT		16	16	CDC037	Cluster-12	prn2	ptxP3	ptxA1	fimH2	CP025479	239x	This study
1896	2007	VT		16	16	CDC170	Cluster-12	prn2	ptxP3	ptxA1	fimH2	CP025381	410x	This study
1915	2010	VT		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP011206	293x	5
1944	2013	CA		27	27	CDC104	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP011207	176x	5
1945	2013	GA		27	27	CDC237	Singleton	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP025384	206x	This study
1955	2013	CA		27	27	CDC237	Cluster-4	<i>prn2</i> -IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP025531	394x	This study
1958	2013	NM	Y	27	27	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP025350	371x	This study
1975	2013	NY	Y	27	27	CDC253	Cluster-6	<i>prn2-IS481-240rev</i>	ptxP3	ptxA1	fimH1	CP011242	151x	5
1998	2013	WA		27	27	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP011243	695x	5
J018	2013	MN	Y	27	27	CDC010	Cluster-3	prn2-Stop-C223T	ptxP3	ptxA1	fimH1	CP011208	199x	5
J022	2013	OR	Y	27	27	CDC253	Cluster-6	prn2-IS481-240rev	ptxP3	ptxA1	fimH1	CP011244	338x	5
J024	2013	NY	Y	27	27	CDC010	Cluster-3	prn2	ptxP3	ptxA1	fimH1	CP025353	406x	This study
J038	2013	AZ		27	27	CDC237	Cluster-4	<i>prn9</i> -IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP012087	105x	5
J066	2013	СТ	Y	27	27	CDC237	Cluster-4	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP026998	329x	This study
J076	2013	NM	Y	27	27	CDC253	Singleton	<i>prn2-IS481-240rev</i>	ptxP3	ptxA1	fimH1	CP011762	348x	5
J077	2013	MN	Y	27	27	CDC300	Cluster-16	prn2-IS481-1613fwd	ptxP3	ptxA1	fimH1	CP025344	363x	This study
J085	2013	CO	Y	27	27	CDC375	Singleton	<i>prn2</i> -IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP026997	313x	This study
J139	2013	ТΧ		27	27	CDC377	Singleton	<i>prn2</i> -IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP025527	289x	This study
J184	2012	IN		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP025383	377x	This study
J185	2013	IN		28	28	CDC237	Cluster-4	<i>prn</i> 2-IS <i>4</i> 81-1613fwd	ptxP3	ptxA1	fimH1	CP025343	344x	This study
J234	2013	VT		27	27	CDC002	Cluster-2	prn2-Stop-C1273T	ptxP3	ptxA1	fimH1	CP025348	424x	This study

*EPS, Enhanced Pertussis Surveillance/Emerging Infections Program Network; MLVA, multiple-locus variable number tandem repeat analysis; PFGE, pulsed-field gel electrophoresis; wgs, whole genome sequence.

†Average coverage depth of all PacBio and Illumina sequencing data combined.

Appendix Table 2. Comparison of variable number tandem repeat profiles for H811 isolates of	Bordetella pertussis	. 2000–2013*
---	----------------------	--------------

pponaix		Valiable Hallibel	tanaoni ropoat p		blatte el Beraete	na pontacolo, Eco	2010
Method	MLVA type	VNTR1	VNTR3a	VNTR3b	VNTR4	VNTR5	VNTR6
MLVA	18	8	7	0†	7	6	7
wgsMLVA	27	8	6	7	7	6	7

*MLVA, multiple-locus variable number tandem repeat analysis; VNTR, variable number tandem repeat profile; wgs, whole-genome sequence. †Isolates collected through the EPS/EIP Network are marked with Y in this column.

‡If 2 discrete bands cannot be observed, traditional MLVA cannot differentiate the 2 VNTR3 loci and VNTR3b is reported as 0.