



# Conserved Patterns of Symmetric Inversion in the Genome Evolution of *Bordetella* Respiratory Pathogens

Michael R. Weigand,<sup>a</sup> Yanhui Peng,<sup>a</sup> Dhvani Batra,<sup>b</sup> Mark Burroughs,<sup>b</sup> Jamie K. Davis,<sup>b</sup> Kristen Knipe,<sup>b</sup> Vladimir N. Loparev,<sup>b</sup> Taccara Johnson,<sup>a</sup> Phalasy Juieng,<sup>b</sup> Lori A. Rowe,<sup>b</sup> Mili Sheth,<sup>b</sup> Kevin Tang,<sup>b</sup> Yvette Unoarumhi,<sup>b</sup> Margaret M. Williams,<sup>a</sup> M. Lucia Tondella<sup>a</sup>

<sup>a</sup>Division of Bacterial Disease, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

<sup>b</sup>Division of Scientific Resources, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

**ABSTRACT** Whooping cough (pertussis), primarily caused by *Bordetella pertussis*, has resurged in the United States, and circulating strains exhibit considerable chromosome structural fluidity in the form of rearrangement and deletion. The genus *Bordetella* includes additional pathogenic species infecting various animals, some even causing pertussis-like respiratory disease in humans; however, investigation of their genome evolution has been limited. We studied chromosome structure in complete genome sequences from 167 *Bordetella* species isolates, as well as 469 *B. pertussis* isolates, to gain a generalized understanding of rearrangement patterns among these related pathogens. Observed changes in gene order primarily resulted from large inversions and were only detected in species with genomes harboring multicopy insertion sequence (IS) elements, most notably *B. holmesii* and *B. parapertussis*. While genomes of *B. pertussis* contain >240 copies of IS481, IS elements appear less numerous in other species and yield less chromosome structural diversity through rearrangement. These data were further used to predict all possible rearrangements between IS element copies present in *Bordetella* genomes, revealing that only a subset is observed among circulating strains. Therefore, while it appears that rearrangement occurs less frequently in other species than in *B. pertussis*, these clinically relevant respiratory pathogens likely experience similar mutation of gene order. The resulting chromosome structural fluidity presents both challenges and opportunity for the study of *Bordetella* respiratory pathogens.

**IMPORTANCE** *Bordetella pertussis* is the primary agent of whooping cough (pertussis). The *Bordetella* genus includes additional pathogens of animals and humans, including some that cause pertussis-like respiratory illness. The chromosome of *B. pertussis* has previously been shown to exhibit considerable structural rearrangement, but insufficient data have prevented comparable investigation in related species. In this study, we analyze chromosome structure variation in several *Bordetella* species to gain a generalized understanding of rearrangement patterns in this genus. Just as in *B. pertussis*, we observed inversions in other species that likely result from common mutational processes. We used these data to further predict additional, unobserved inversions, suggesting that specific genome structures may be preferred in each species.

**KEYWORDS** *Bordetella*, pertussis, whooping cough, evolution, genomics, insertion sequence, rearrangement


The genus *Bordetella* presently includes 16 named species, comprising human and animal pathogens, as well as environmental colonizers. The closely related *Bordetella* species *B. pertussis*, *B. parapertussis*, and *B. bronchiseptica* are collectively known as the “classic bordetellae” due to their pathogenic ecologies and were previously re-

**Citation** Weigand MR, Peng Y, Batra D, Burroughs M, Davis JK, Knipe K, Loparev VN, Johnson T, Juieng P, Rowe LA, Sheth M, Tang K, Unoarumhi Y, Williams MM, Tondella ML. 2019. Conserved patterns of symmetric inversion in the genome evolution of *Bordetella* respiratory pathogens. *mSystems* 4:e00702-19. <https://doi.org/10.1128/mSystems.00702-19>.

**Editor** Jack A. Gilbert, University of California San Diego

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Michael R. Weigand, [mweigand@cdc.gov](mailto:mweigand@cdc.gov).

 Clinical isolates of *Bordetella* respiratory pathogens have varied chromosome structures thanks to frequent inversion between copies of insertion sequence elements.

**Received** 24 October 2019

**Accepted** 1 November 2019

**Published** 19 November 2019

garded as subspecies differentiated by host adaptation (1, 2). Most notably, *B. pertussis* is the primary causative agent of whooping cough (pertussis), a respiratory disease with high morbidity and mortality in infants too young to be vaccinated (i.e., those <2 months of age). *B. bronchiseptica* causes respiratory disease in a range of mammals, such as kennel cough in dogs, but also various infections in humans (3–6). Divergent lineages of *B. paraptussis* infect the respiratory tracts of sheep and humans (7–9). The remaining “nonclassic” species are phylogenetically more distant and have been recovered from varied hosts and pathologies, in particular, *B. holmesii* and *B. hinzii* from humans, which is reflected in their discrete virulence-associated-gene compositions (10).

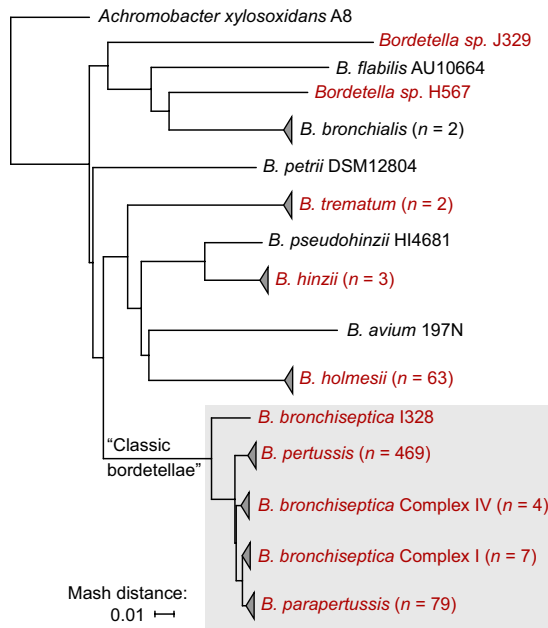
Considerable recent attention has been given to *B. pertussis* as the number of reported pertussis cases in many developed countries—including the United States—has risen over the last 2 decades, despite high or increasing coverage with pertussis-containing vaccines (11). Increased disease reporting likely results from many factors, including heightened awareness, expanded surveillance, improved laboratory diagnostic testing, and allelic mismatch between circulating and vaccine reference strains (11–13). Moreover, the protection conferred by acellular vaccine formulations—which replaced whole-cell preparations in the United States during the 1990s—waned over time, resulting in increased disease rates among vaccinated individuals (11, 14–16).

*B. holmesii* and *B. paraptussis* cause similar cough illnesses in humans that, while generally regarded as less severe, can be clinically indistinguishable from *B. pertussis* infection and meet the pertussis clinical case definition (17–24). Cocirculation and coinfection have been reported with no discernible difference in symptomology between etiologies (23–28). It is unclear whether increased detection of *B. holmesii* or *B. paraptussis* respiratory infection results from recent genetic changes or, rather, heightened awareness and improved diagnostics (24, 29, 30). Nonetheless, many diagnostic PCR assays cannot differentiate *B. pertussis* and *B. holmesii*, and retrospective testing of specimens sometimes reveals misdiagnosed cases of *B. holmesii* (26, 27, 31–33). For these reasons, the true burden of pertussis-like respiratory disease caused by *B. holmesii* and *B. paraptussis* remains understudied and largely unknown (19, 34).

Recent advances in high-throughput sequencing and bioinformatics have yielded a large collection of complete genome assemblies from *B. pertussis* clinical isolates, and comparative analyses have described considerable rearrangement fluidity (35, 36). In contrast, much lower amounts of genomic data exist for non-*pertussis* *Bordetella* species, including other human pathogens, limiting opportunities to research these species. To address these knowledge gaps, we assembled and compared closed genome sequences from 167 *Bordetella* species isolates for analyses of genome evolution and gene order rearrangement within this genus. Together with a comparison of 469 *B. pertussis* closed genome sequences, these data reveal conserved patterns of inversion within species whose genomes harbor multicopy insertion sequence (IS) elements. A linear model of observed rearrangements predicted additional potential inversions based on IS element content. These results indicated that *Bordetella* species are subject to common rearrangement processes and that each exhibits a preferred gene order.

## RESULTS

**Rearrangement in *Bordetella* spp.** Complete genome assemblies were prepared or publicly available for 11 named or provisional species in the genus *Bordetella* and were compared in the present study to identify rearrangements of gene order. A neighbor-joining phylogenetic reconstruction of these species, plus two distantly related *Bordetella* isolates (H567 and J329), rooted with outgroup *Achromobacter xylosoxidans*, clustered the classic bordetellae apart from the otherwise disparate species branches (Fig. 1), consistent with previous studies (10). The number of single nucleotide polymorphisms (SNPs) observed in each species varied. Isolates of *B. bronchiseptica* derived primarily from two distinct clades, complex I and complex IV, which differ by ~35,000 SNPs (Fig. S1A and B in the supplemental material). The distant *B. bronchiseptica* strain

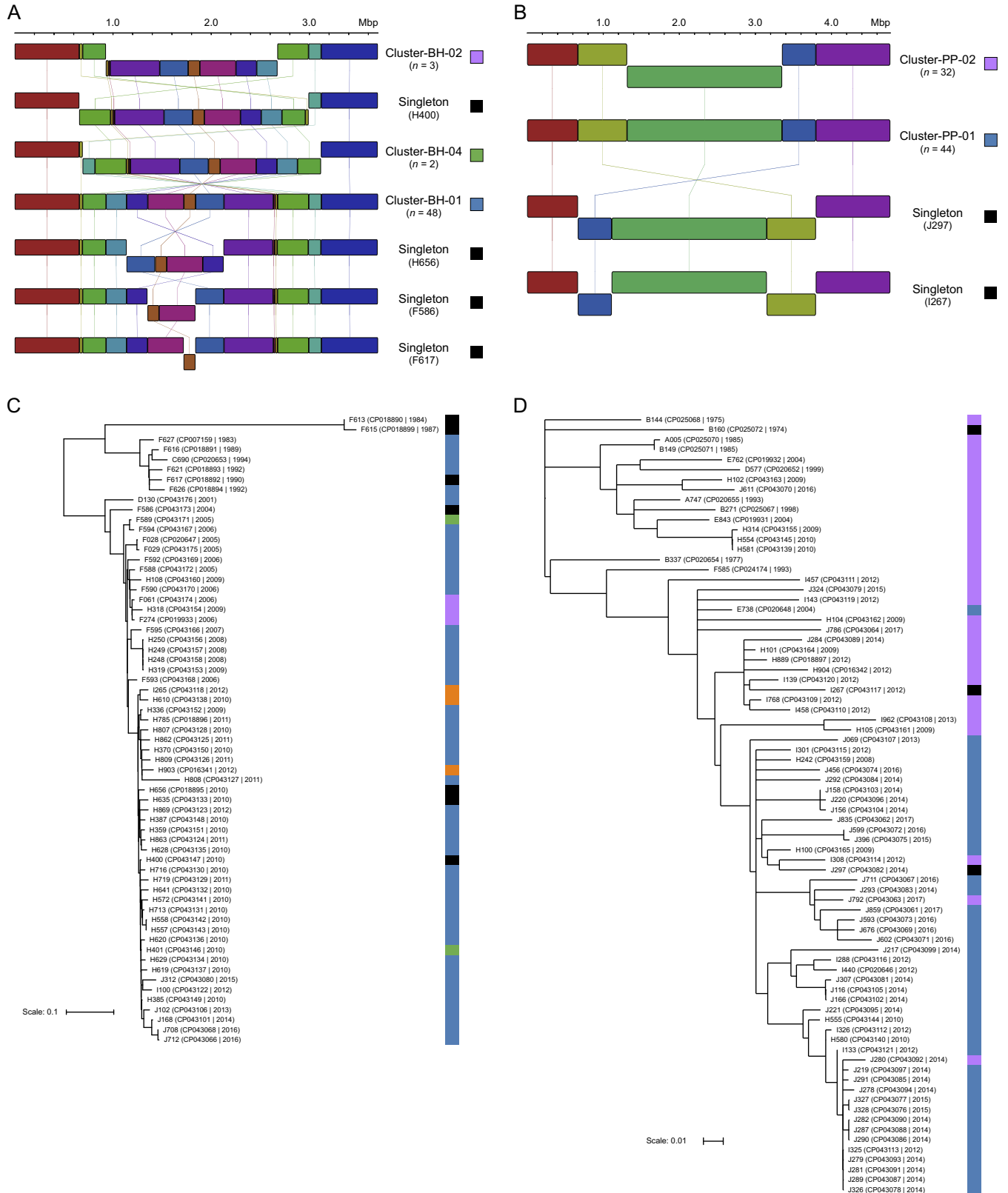


**FIG 1** Phylogeny of the genus *Bordetella*. Neighbor-joining tree of named and provisional *Bordetella* species with available complete-genome sequences based on pairwise mash distances and rooted with *Achromobacter xylosoxidans*. Species sequenced at the CDC are listed in red. Triangles indicate collapsed nodes.

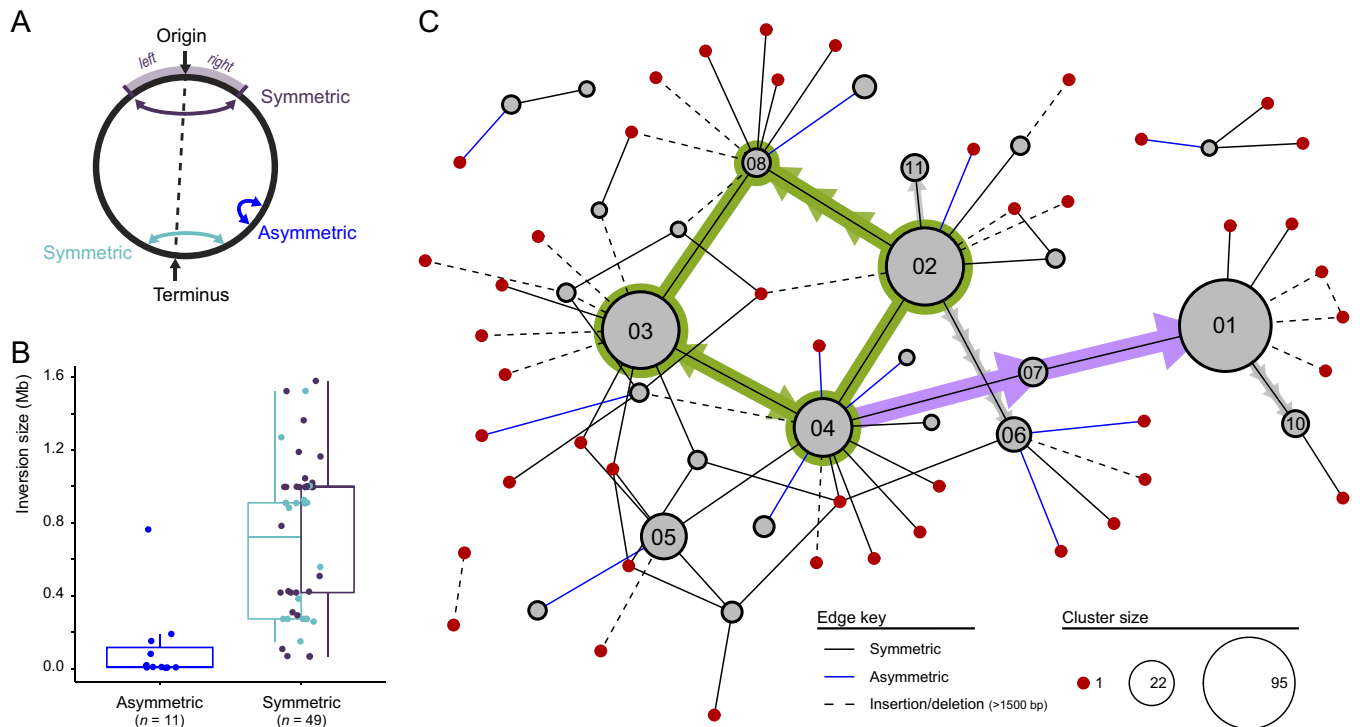
1328 (37) did not belong to either complex (Fig. 1) and was separated from other *B. bronchiseptica* isolates by  $\sim 60,000$  SNPs. Isolates of *B. pertussis* (complex II), *B. parapertussis* (complex III), and *B. holmesii* were much more similar, exhibiting  $< 100$  SNP differences between most isolates in each (Fig. S1C). Complete genomes from isolates of *B. hinzii* ( $n = 3$ ) and *Bordetella trematum* ( $n = 2$ ) differed by  $\sim 9,600$  and  $\sim 11,000$  SNPs, respectively.

Among the various *Bordetella* species, only *B. holmesii* and *B. parapertussis* displayed signs of chromosome structure variation in alignments of the isolate genomes analyzed here (Fig. 2). Genomes of *B. holmesii* ( $n = 63$ ) comprised 11 discrete chromosome structures (Fig. S2A). The most abundant structure was observed in 48 isolates (cluster BH-01), and six other structures each differed from it by a single, unique inversion (Fig. 2A). All observed inversions were symmetric, encompassing the replication origin-terminus axis, and flanked by opposing copies of IS elements. Other structures differed by small deletions, not rearrangements, including cluster BH-03 (Data Set S1). Genomes of *B. parapertussis* ( $n = 79$ ) exhibited two predominant chromosome structures of comparable levels of abundance, varying by a symmetric inversion between copies of IS1001, which largely followed phylogenetic boundaries (Fig. 2B). In both species, unique structures were also observed in only one isolate (Fig. S2B, singletons). Predicted protein-coding gene contents adjacent to select inversion breakpoints observed in *B. holmesii* and *B. parapertussis* are listed in Data Set S1. These results indicated that inversions like those frequently observed in *B. pertussis* also occur in at least a few other *Bordetella* species.

**Chromosome structural diversity of *B. pertussis*.** The chromosome of *B. pertussis* displays structural fluidity (36), and the 469 complete genomes of *B. pertussis* included here were aligned in 109,746 pairwise comparisons to identify rearrangements and group isolates with colinear chromosomes. These genomes exhibited 107 unique chromosome structures, including 32 clusters of two or more isolates (e.g., cluster BP-01 and cluster BP-02) and 75 singletons observed in only one isolate (Fig. S2C). The abundances of unique structures varied, with the largest cluster including 95/469 (20%) of the isolates; the distribution suggested that structural diversity remains under-sampled, as observed previously (36). The data set here sampled years 2000 to 2016



**FIG 2** Genome rearrangements in *Bordetella* species. (A) Eleven unique chromosome structures that included six symmetric inversions were observed among *B. holmesii* isolate genomes. (B) Two predominant structures were observed in *B. paraptentussis*. (C, D) Phylogenetic placement of structures observed in *B. holmesii* (C) and *B. paraptentussis* (D) isolates was reconstructed using maximum parsimony from 1,496 and 677 core SNPs, respectively.



**FIG 3** Genome rearrangement in *B. pertussis*. (A) Observed inversions were either symmetric, encompassing either replication origin (purple) or terminus (light blue), or asymmetric (blue). (B) Single inversions observed in pairwise alignments varied in size and were predominantly symmetric. (C) Undirected network constructed from single inversions or insertion/deletions observed between 71 unique chromosome structures. Node diameter and edge line type indicate cluster size and rearrangement, respectively, according to the key. Nodes with the highest degrees of centrality (Fig. S5A) and their connections at the network core are highlighted (green). Shaded arrows indicate directionality inferred from the SNP phylogeny (Fig. S7), including divergence toward the predominant cluster BP-01 (pink), as described in the text.

( $n = 370$ ) most densely, and a rarefaction curve of unique structures recovered during this time period did not approach saturation, as expected (Fig. S2B). Links between clusters of unique structures reported here and those reported previously are listed in Table S2.

Circulating isolates of *B. pertussis* exhibited little gene sequence diversity (13, 38), but the distribution of pairwise SNP distances between isolates within the largest clusters varied. While cluster BP-01 and cluster BP-05 appeared relatively clonal, exhibiting  $<15$  SNPs between isolates, SNP distances within others, such as cluster BP-02, cluster BP-03, cluster BP-04, cluster BP-06, and cluster BP-08, measured up to 32 SNPs (Fig. S3A). These differences largely reflected the temporal spread of isolates within a cluster. Isolates in clusters with fewer pairwise SNPs were recovered during narrower time frames than those with more SNPs and wider temporal distributions (Fig. S3B). These data illustrate that certain chromosome structures predominate due to their stability (e.g., cluster BP-02), while the abundances of others can be attributed to recent expansion (e.g., cluster BP-01).

**Rearrangement network mapping.** Pairwise alignments between a subset of genomes representing each of the unique structures in *B. pertussis*, *B. holmesii*, and *B. parapertussis* were investigated further to identify pairs that differed by either a single inversion (Fig. 3A) or insertion/deletion (i.e., a gap of  $>1,500$  bp) but were otherwise colinear. Eighty-two pairs were identified among the 107 *B. pertussis* structures, of which 49 differed by symmetric inversion, 11 by asymmetric inversion, and 22 by insertion/deletion. Symmetric inversions of varied sizes were observed near both the replication origin (*oriC*) and terminus (*dif*) and were larger than asymmetric inversions on average (Fig. 3B).

To investigate relationships among discrete *B. pertussis* chromosome structures, the identified pairs were used to construct an undirected network that connected 71

unique structures (nodes) by inferred single inversion or insertion/deletion events (Fig. 3C, edges). Sixty-two structures were connected to a single graph, representing 407 isolates (87%) in the data set. The structures in vaccine and laboratory reference strains, such as strain 134 (B202 [accession no. CP016338]), were not connected to the larger network of recent clinical isolates but instead formed small groups of 2 to 4 related structures. The remaining 36 structures, representing 49 isolates that included Tohama I, could not be connected to the network or each other.

Some of the largest structural clusters were located near the network core and connected through loops of parallel, symmetric inversions. Four of these clusters also exhibited the highest degrees of centrality, indicating their significant connectivity to other structures in the network (Fig. S4A). In contrast, cluster BP-01, which was the most abundant, appeared closer to the periphery, following a path of successive rearrangements that diverged from the network core, and therefore, it deviated from many other abundant clusters. Singletons also formed peripheral nodes, often crowded around an abundant cluster, suggesting they represent rare aberrations from predominant chromosome structures. Network edges representing insertion/deletion and asymmetric inversion were also largely responsible for connecting small, peripheral nodes to the network, while edges at the core comprised symmetric inversions. Similar patterns were observed in smaller networks constructed of chromosome structures observed in *B. holmesii* and *B. paraptussis* (Fig. S5). Taken together, these results demonstrated the extent to which symmetric inversion has shaped and continues to shape the genome compared to the relative rarity of asymmetric and insertion/deletion events.

**Rearrangement divergence dating.** The phylogenetic history of discrete chromosome structures observed in *B. pertussis* was further explored to infer the divergence of specific structures and assign putative directionality to network edges shown in Fig. 3C, where possible. A time-scaled Bayesian reconstruction of 438 isolates from the *ptxP3-prn2-ptxA1* background was calculated using 908 variable, core nucleotides (Fig. S6) with an estimated mutation rate of  $1.30 \times 10^{-7}$  substitutions per site per year. The resulting tree topology separated isolates into two large clades corresponding to the *fimH1* (*fim3-1*) and *fimH2* (*fim3-2*) alleles, consistent with previous studies (13, 36), and large structural clusters were restricted to a single clade.

Within the *fimH1* clade, there was insufficient sampling to discriminate the ancestral order of cluster BP-03 and cluster BP-04. Genomes exhibiting both structures appeared near the tree root and in isolates dating back to 2000 (Fig. S6; Table S1). However, inspection of more contemporary subclades revealed evidence for multiple inversion events between the two clusters, consistent with the wide range of pairwise SNPs in each (Fig. S3). The divergence dates of some chromosome structures could be estimated from the phylogeny (e.g., cluster BP-01 and cluster BP-07), while multiple events were predicted for others (e.g., cluster BP-10) (Fig. S6). Likewise, the *fimH2* allele background included the large central cluster BP-02, and multiple inversion events were predicted that gave rise to unrelated isolates with colinear chromosomes (e.g., cluster BP-06 and cluster BP-08) (Fig. S6). While some inversions have occurred repeatedly or reversibly, most chromosome structures exhibited stable phylogenetic linkage, suggesting that rearrangement events are either rare or under selection.

**Repeat content.** Observed inversions among genomes of *B. pertussis*, *B. paraptussis*, and *B. holmesii* were primarily flanked by IS elements, which appeared in multiple copies throughout each genome. All repetitive sequence content was quantified independently of gene annotation by counting unique 15-bp sequences (15-mers) in a representative isolate genome from each of the species analyzed here. Genomes of *B. pertussis* and *B. holmesii* each included a set of highly abundant 15-mers, while *B. paraptussis*, *B. petrii*, *B. bronchialis*, *Bordetella* sp. strain H567, *B. bronchiseptica* I328, *B. pseudohinzii*, and *B. flabilis* included decreased numbers of abundant 15-mers (Fig. S7A). Representatives from all other species included in Fig. 1 did not exhibit any abundant 15-mers. A similar pattern resulted from comparison of predicted protein-coding gene sequences (Fig. S7B), and abundant genes corresponded to transposases



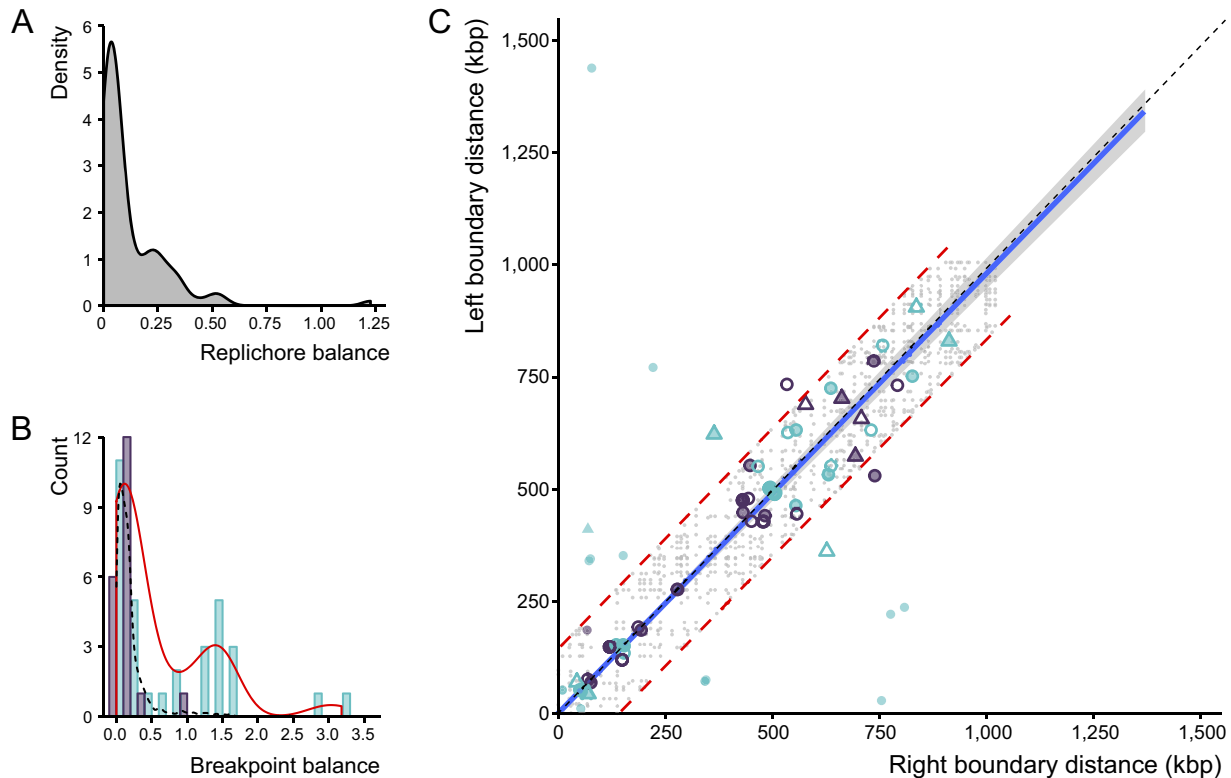
**TABLE 1** IS element contents in representative isolates of *Bordetella* species

Species (isolate)	IS element	Copy no.	No. of inversions predicted
<i>B. pertussis</i> (J549)	IS481	249	918
	IS1663	16	1
	IS1002	6	0
<i>B. holmesii</i> (C690)	IS1001-like	68	106
	IS481	48	30
	IS407	48	59
	IS21 family	4	0
<i>B. parapertussis</i> (B271)	IS1001	22	12
	IS1002	9	0
<i>B. petrii</i> (DSM1204)	IS3 family (RS00975–RS00980)	25	6
	IS3 family (RS01055–RS01060)	10	1
	IS3 family (RS20025)	6	0
	IS3 family (RS20230)	2	0
	IS110 family	6	1
	ISBope1	4	0
<i>B. bronchialis</i> (AU171976)	IS481	5	0
	IS256 family	3	0
<i>Bordetella</i> sp. H567	IS256 family	10	2
<i>B. bronchiseptica</i> I328	IS1002-like	5	0
<i>B. pseudohinzii</i> (H14681)	IS21 family	4	0
<i>B. flabilis</i> (AU10664)	IS3 family	2	0

encoded in various IS elements in *B. pertussis*, *B. holmesii*, *B. parapertussis*, *B. petrii*, *B. bronchialis*, *Bordetella* sp. H567, *B. bronchiseptica* I328, *B. pseudohinzii*, and *B. flabilis* (Table 1). The complements and copy numbers of the IS elements detected varied, with *B. pertussis* carrying the greatest number. While all species considered here included a few duplicated protein-coding genes of other functions, those absent from Table 1 did not encode any, including transposases, with >2 copies. These results support the expected correlation between genome repeat contents, specifically IS elements, and the observation of rearrangement.

**Inversion boundary prediction.** Most observed inversions were symmetric, and their “balance” was assessed by calculating the ratio of breakpoint distances relative to either the replication origin (*oriC*) or terminus (*dif*), whichever was closer, according to equation 1. The size of replichores, oppositely replicated chromosome halves, in most of the 107 unique *B. pertussis* chromosome structures were balanced, as the natural logs of their ratios were close to zero (Fig. 4A), particularly among those representing the largest clusters, and most of the deviations were observed in small clusters or singletons (Fig. S4B). Replichore sizes were similarly balanced in *B. holmesii* (Fig. S4C). Although replichore sizes in *B. parapertussis* appeared less balanced, the two predominant structures produced similar results (Fig. S4D). Many of the 49 symmetric inversions observed in *B. pertussis* were also balanced (Fig. 4B), maintaining similar relative replichore sizes before and after inversion. Some unbalanced inversions were also observed, and these appeared more common near the replication terminus than near the origin.

Linear regression of the right and left relative breakpoint distances in 55 symmetric inversions observed in *B. pertussis* and *B. holmesii* further illustrated the bias toward balance (Fig. 4C). The resulting linear model was used to predict all possible symmetric inversions in a representative isolate genome from each species based on the positions of detected IS elements (Table 1). As expected, the number of predicted inversions increased with IS element copy number and was greater than the number observed in the current data set of *B. pertussis*, *B. holmesii*, and *B. parapertussis* genomes (Fig. 5). Additional breakpoints were determined from alignment of *B. pertussis* J549 (cluster BP-04) with representatives of all 106 other structures but were still fewer than those derived from predicted inversions (Fig. 5A). Many individual copies of IS481 in J549 were predicted to engage in multiple inversions, pairing with as many as 16 other insertions, reflecting the particularly high density of this IS element and varied balance points of observed inversions. Only one genome sequence each was available for *B.*



**FIG 4** Symmetric inversion balance and prediction. (A) Density of relative replicore size balance observed in each of the 107 unique *B. pertussis* genome structures. (B) Histogram of breakpoint balance observed in symmetric inversions encompassing the replication origin (purple) or terminus (light blue). The combined densities of observed (solid line) and predicted (dotted line) symmetric inversions were scaled to 10. (C) Linear regression of right and left breakpoint distances observed in *B. pertussis* (circle) and *B. holmesii* (triangle) nearest to the replication origin (purple) or terminus (light blue). Open symbols represent duplicated points inverted for model calculation. Blue line indicates linear fit, and shading represents the 95% confidence interval. Dotted red lines denote boundaries of the model's 95% prediction interval, and gray points correspond to all predicted inversions between copies of IS elements in *B. pertussis* J549.

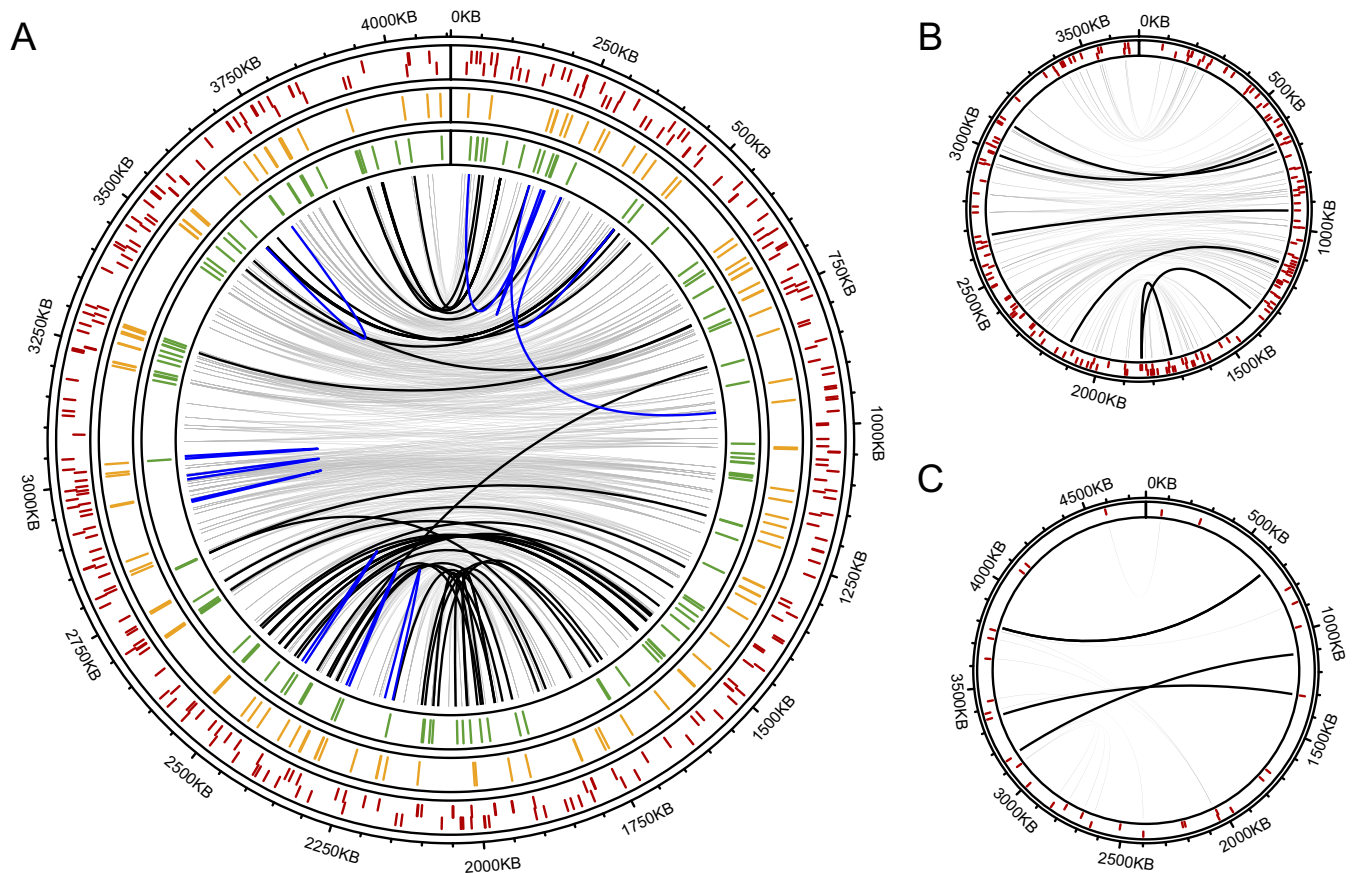
*petrii* and *Bordetella* sp. H567, but symmetric inversions were predicted based on the locations of their IS elements (Table 1). No inversions were predicted between IS elements in genomes of *B. bronchiseptica* I328, *B. bronchialis*, *B. flabilis*, or *B. pseudohinzii*, as suitable pairs could not be matched within the prediction window. This generalized model of symmetric inversion revealed considerable rearrangement potential, much of which remains unobserved. The model also illustrated that inversion potential does not simply depend on IS element copy number but on relative position and orientation as well.

## DISCUSSION

In this study, we extended analyses of chromosome structure variation in *B. pertussis* to include diverse *Bordetella* species to gain a comprehensive understanding of rearrangement patterns in this genus. Those species colonized by IS elements, in particular *B. holmesii* and *B. paraptussis*, also exhibited similar patterns of chromosome inversion. Just as in *B. pertussis*, inversions observed in these species were biased toward symmetry and maintenance of replicore balance between predominant—and thus potentially favorable—gene arrangements. The results here reveal the breadth of chromosome rearrangement in the continued genome evolution of a genus that encompasses human and other animal pathogens.

The chromosome of *B. pertussis* experiences ongoing structural fluidity, and rearrangements supply a notable source of mutational diversity within the circulating population (36, 39), first investigated before the availability of assembled genomes (40, 41). Study of complete genome assemblies from other *Bordetella* species revealed that such patterns are not a unique feature of *B. pertussis*. IS element colonization and





**FIG 5** Inversion potential of *Bordetella* species genomes based on IS element content. Chromosome maps of *B. pertussis* J549 (A), *B. holmesii* C690 (B), and *B. parapertussis* B271 (C). Tracks of red lines indicate the locations of IS elements in either the forward or reverse orientation (above or below center, respectively). Connecting lines link breakpoints of observed symmetric (black), observed asymmetric (blue), and predicted symmetric (gray) inversions. Additional tracks in panel A denote all breakpoints observed in pairwise alignments between J549 and genomes of the *ptxP1-ptxP2* (orange) or *ptxP3* (green) clades.

subsequent copy number proliferation generate regions of homology throughout genomes, measured here as both 15-mers and protein-coding genes, that supply substrates for homologous-recombination-mediated rearrangement (42). There are >240 copies of IS481 in genomes of circulating *B. pertussis* strains, which has significantly impacted the evolution of this species (39). Likewise, the genomes of *B. holmesii* and *B. parapertussis* each harbor a collection of different IS elements that have contributed to their speciation through genome reduction (10, 43, 44), as well as mediated the rearrangements observed here.

*B. holmesii* has evolved independently of the classic bordetellae, reiterating the influence of IS element colonization, not phylogenetic history, on genome structural fluidity in the genus. However, *B. holmesii* may have first acquired IS481 from *B. pertussis* via horizontal gene transfer (HGT) (43). Not all *Bordetella* species included in this study encoded IS elements, and while sample sizes for direct comparison were low for some (e.g., *B. hinzii*), these are unlikely to exhibit rearrangement, based on the results presented here. Publicly available genome sequences of *B. bronchiseptica* and ovine-specific *B. parapertussis* show differences in gene order, with putative rearrangements flanked by rRNA operons and IS elements, but were excluded from the present study because their assemblies lacked the independent structural validation employed here (2, 9, 44, 45).

Bacterial chromosomes are highly organized with respect to both the placement of protein-coding genes (46, 47) and the distribution of noncoding sequences (48, 49). Accordingly, this organization imposes limits on genome rearrangement (50, 51), and gene location can directly influence expression (52–55). Consistent with this under-

standing, some abundant *B. pertussis* chromosome structures have persisted across SNP and temporal distances. A time-scaled phylogeny also revealed a degree of flexibility in the form of repeated, or reversible, symmetric inversion among similar, abundant structures. Many structures have given rise to other rare clusters or singletons, depicted here as peripheral nodes in an undirected rearrangement network, through random exploration of mutation space. Analogous to nucleotide sequence variation, beneficial structural mutations are expected to persist, suggesting structures that are low in abundance may comprise less favorable gene order arrangements. Although sampling of *B. holmesii* and *B. parapertussis* genomes was considerably lower, each similarly exhibited a predominant chromosome structure, or two, which had given rise to rare variants through individual rearrangements. Phenotypes derived from chromosome structure variation in these species have not yet been observed experimentally, and therefore, whether specific gene arrangements are indeed favorable remains a hypothesis.

Based on observations of IS element involvement and proclivity for symmetry around the replication origin-terminus axis, additional inversions could be predicted, even in *Bordetella* species in which rearrangement could not be observed here. Not surprisingly, the number of predicted inversions was proportional to the IS element load, with *B. pertussis* exhibiting the most due to the high copy number of IS481. The model here considered not only IS element copy number, but position and orientation relative to the replication origin-terminus axis as well. However, limitations to the model likely include an oversimplification of balance, according to the predominant structures seen in *B. parapertussis*, and disregard for asymmetric inversion. As a result, predicted inversions were likely underestimated, particularly in species with sporadic IS element insertions. In species where IS element-mediated rearrangement was observed, the high number of predicted possible inversions reinforced the idea that specific gene order arrangements are likely more favorable than others. The *B. pertussis* IS481 copy number continues to slowly increase (36), and additional insertion target sites are predicted in *B. pertussis* (36) and *B. holmesii* (M. R. Weigand, unpublished data) that could facilitate even more potential inversions than those predicted here.

The present genomic study of the genus *Bordetella* highlights patterns of chromosome structural variation in two species, *B. parapertussis* and *B. holmesii*. Although *B. pertussis* is the primary agent, *B. parapertussis* also causes clinical whooping cough (18, 28, 56), and increasing reports have attributed pertussis-like cough illness to *B. holmesii* (19–22). Current acellular pertussis vaccines contain immunogenic proteins purified from *B. pertussis* and provide no cross-protection against *B. holmesii* (57, 58) and debated protection against *B. parapertussis* (59–66). However, widespread application of acellular pertussis vaccines may have influenced the circulating population structure of all three species (19, 29, 67, 68). Acellular pertussis vaccination and *B. pertussis* infection each appear to enhance *B. parapertussis* colonization (61, 69), and interactions among the three species have even been suggested to underlie the periodicity of pertussis cycles (56, 70). *B. parapertussis* and *B. holmesii* are likely endemic to the United States, and their detection through improved diagnostics warrants additional attention. While chromosome structural fluidity challenges existing approaches to the genomic surveillance of bacterial pathogens, it may hold valuable information about the emergence of these related etiologies of respiratory disease.

## MATERIALS AND METHODS

**Strain selection.** The Centers for Disease Control and Prevention's (CDC's) collection includes U.S. *B. pertussis* isolates collected through surveillance and during outbreaks. In total, 469 assembled genomes were included in the current study based on availability, and most were selected for sequencing as part of previous studies (Table S1 in the supplemental material). One set ( $n = 170$ ) was selected to capture potential geographic diversity among 28 U.S. states in 2000 to 2013 (71). Another selection ( $n = 108$ ) focused on isolates obtained through the Enhanced Pertussis Surveillance/Emerging Infection Program Network (EPS) (72) in 2011 to 2014, which prioritized isolates from hospitalized cases and forced sampling of every possible combination of year, state, vaccination status, and age group. Additionally, 133 isolates were selected prospectively from sporadic or EPS submission in 2014 to 2016; 31 epidemic isolates were sequenced previously (35); and 27 others, such as vaccine reference strains and novel

immunogen-deficient genotypes, were also included (73–76). U.S. isolates of other *Bordetella* species were selected from the CDC collection for sequencing across as many years as available, but with additional emphasis on those recovered after 2000 from patients presenting respiratory illness. A few isolates sequenced previously were also included (37, 77, 78). The present study featured new genome assemblies for 125 *B. pertussis*, 11 *B. bronchiseptica*, 60 *B. holmesii*, 78 *B. parapertussis*, 2 *B. trematum*, 1 *B. hinzii*, and 1 other *Bordetella* species isolate as outlined in Table S1.

**Genomic DNA preparation, sequencing, and assembly.** Isolates were cultured on Regan-Lowe agar without cephalixin for 72 h at 37°C. Genomic DNA isolation and purification was performed using the Genra Puregene yeast/bacteria kit (Qiagen, Valencia, CA) with slight modifications. Briefly, two aliquots of approximately  $1 \times 10^9$  bacterial cells were harvested and resuspended in 500  $\mu$ l of 0.85% sterile saline and then pelleted by centrifugation for 1 min at  $16,000 \times g$ . Recovered genomic DNA was resuspended in 100  $\mu$ l of DNA hydration solution. Aliquots were quantified using a Nanodrop 2000 (Thermo Fisher Scientific, Inc.; Wilmington, DE). Additional chloroform purification was performed to remove polysaccharides from preparations of all species except *B. pertussis* (79).

Whole-genome shotgun sequencing was performed using a combination of the PacBio RSII (Pacific Biosciences, Menlo Park, CA), Illumina HiSeq/MiSeq (Illumina, San Diego, CA), and Argus (OpGen, Gaithersburg, MA) platforms as described previously (35). Genomic DNA libraries were prepared for PacBio sequencing using the SMRTbell template preparation kit 1.0 and polymerase binding kit P4 or P6, while Illumina libraries were prepared using the NEB ultra library preparation kit (New England Biolabs, Ipswich, MA). *De novo* assembly was performed using either the Hierarchical Genome Assembly Process (HGAP) (version 3; Pacific Biosciences) (80) or FLYE (version 2.4.2) (81). The resulting consensus sequences were circularized using either circlator (version 1.5.1) (82) or gepard (version 1.30) (83). Where necessary, hybrid *de novo* assembly of PacBio and Illumina sequences was performed using Unicycler (version 0.4.0) (84). All assemblies were reordered to start at the coding region for glucose-inhibited cell division protein A (*gidA*). Assemblies were confirmed by comparison to restriction digest optical maps using the Argus system (OpGen) with MapSolver (version 2.1.1; OpGen) and further polished by mapping Illumina reads using CLC Genomics Workbench (version 10) (CLC bio, Boston, MA). Assemblies were annotated using the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP). Genome sequence-based allele characterization of *B. pertussis* molecular typing loci (*prn*, *ptxP*, *ptxA*, *ptxB*, and *fimH*) was assigned by genome alignment to a curated set of wild-type and deficient alleles using high stringency.

**Phylogenetic reconstruction.** A neighbor-joining phylogenetic reconstruction of the genus *Bordetella* was determined from pairwise Mash (version 2.0) (85) distances using mashtree (version 0.29) (86). Further phylogenetic reconstruction within species *B. holmesii* and *B. parapertussis* was calculated using kSNP3 (87), with a kmer size of 23 bp. Maximum-parsimony trees and pairwise SNP distances were calculated from core variable positions. Trees were annotated with iTOL (version 4) (88). The phylogeny of *B. pertussis* isolates from the predominant *ptxP3-prn2-ptxA1* background was reconstructed from SNPs determined relative to the reference strain C734 (accession number CP013078) using snippy (version 3.1) (<https://github.com/tseemann/snippy>) after masking all IS elements with N's. Isolates with <75% coverage breadth were removed ( $n = 6$ ). A maximum-likelihood phylogeny was estimated from the core SNP alignment using RAxML (version 8.1.16) (89), and a temporal correlation of 0.5231 was determined using TempEst (version 1.5.1) (90). Bayesian divergence time estimation from the core alignment was performed using BEAST (version 1.8.3) (91) with an HKY substitution model, four-category gamma site heterogeneity, and strict clock. Markov chain length was 100,000,000 with parameters sampled every 1,000 states. Model convergence and parameter expected sample sizes were evaluated with tracer (version 1.6) (92), and a maximum clade credibility tree was calculated after a burn-in of 10,000,000 states using TreeAnnotator (version 1.10.0) (91) with common ancestor heights. Tree visualization and annotation were performed with the ggtree R package (version 1.10.5) (93).

**Repeat content.** Repeat content in each genome was quantified by counting the abundance of unique 15-mers with jellyfish (version 2.2.6) (94). Protein-coding genes within each genome were clustered using CD-HIT (version 4.6) (95) with default parameters, except for match cutoffs, which were set at 95% nucleotide sequence identity and 90% length difference. Detected IS elements were annotated using ISfinder (96), and copy numbers were quantified by BLASTn query of a representative genome from each species.

**Genome structural variation.** Genomes from each species were aligned in all pairwise combinations using progressiveMauve (97) with optimized parameters ( $-\text{seed-weight} = 16$ ,  $-\text{hmm-identity} = 0.85$ ) and clustered according to their structures as described previously (36). To improve the accuracy of *B. pertussis* whole-genome alignments, adjacent insertions of IS481 with the same orientation and separated only by their 6-bp target sequences were first collapsed to a single insertion using a custom Perl script. A rarefaction curve of structures observed among *B. pertussis* isolates collected during 2010 to 2016 was calculated using the vegan R package (version 2.3.0) (98). Alignments were further analyzed using custom Perl scripts to identify pairs of *B. pertussis* genome clusters differing by only a single inversion or single insertion/deletion of >1,500 bp, set to exceed the length of individual IS elements, and to characterize inversion symmetry measured relative to the positions of *oriC* and *dif* (99, 100). The resulting list was used to construct an undirected network of rearrangement and insertion/deletion events (edges) between genome clusters (nodes) using the network (version 1.13.0) (101) and sna (version 2.4) (102) R packages. The centralization of each node in the network was computed as degree of centrality using sna.

The balance of symmetric inversions was calculated from the left and right breakpoint distances, measured from the nearest replication origin or terminus as follows:

$$\text{balance} = |\ln(\text{left}/\text{right})| \quad (1)$$

The replichore balance of discrete chromosome structures was determined by the same equation using the distances between the replication origin and terminus. Symmetric inversions observed among 49 *B. pertussis* alignment pairs were used to fit a linear model to the balance of breakpoint distances in the right and left replichores. The distribution of natural log ratios (balance) was used to prune extremely unbalanced inversions (outliers). The remaining 35 breakpoint distances were used to fit a linear model using the stats R package (version 3.4.4). The resulting model was used to predict all possible inversions between IS elements located within the corresponding 95% prediction intervals in a representative genome from each species. Observed and predicted inversions were visualized using the circlize R package (version 0.4.4) (103).

**Data availability.** Source code for custom scripts developed under the present study is available at [https://github.com/mikeyweigand/Bordetella\\_species](https://github.com/mikeyweigand/Bordetella_species). The whole-genome shotgun sequences have been deposited at DDBJ/EMBL/GenBank under the accession numbers listed in Table S1. The versions described in this paper are the first versions. Raw sequence data are available from the NCBI Sequence Read Archive, organized under BioProject accession numbers [PRJNA279196](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA279196) and [PRJNA287884](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA287884).

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSystems.00702-19>.

**FIG S1**, PDF file, 0.7 MB.

**FIG S2**, PDF file, 0.1 MB.

**FIG S3**, PDF file, 0.9 MB.

**FIG S4**, PDF file, 0.2 MB.

**FIG S5**, PDF file, 0.1 MB.

**FIG S6**, PDF file, 0.3 MB.

**FIG S7**, PDF file, 0.1 MB.

**TABLE S1**, XLSX file, 0.1 MB.

**TABLE S2**, DOCX file, 0.01 MB.

**DATA SET S1**, XLSX file, 0.1 MB.

## ACKNOWLEDGMENTS

We thank Pam Cassidy, Nicholas Cook, and Wen Li for technical assistance, Sandeep Joseph, Reagan Kelly, and Christine Miner for informatics assistance, Tami Skoff for insightful discussions that contributed to the development of this work, and the Enhanced Pertussis Surveillance/Emerging Infections Program Network sites and other state public health departments for contributing isolates.

This work was made possible through support from CDC's Advanced Molecular Detection (AMD) program.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

## REFERENCES

1. von Wintzingerode F, Gerlach G, Schneider B, Gross R. 2002. Phylogenetic relationships and virulence evolution in the genus *Bordetella*. *Curr Top Microbiol Immunol* 264:177–199.
2. Park J, Zhang Y, Buboltz AM, Zhang X, Schuster SC, Ahuja U, Liu M, Miller JF, Sebahia M, Bentley SD, Parkhill J, Harvill ET. 2012. Comparative genomics of the classical *Bordetella* subspecies: the evolution and exchange of virulence-associated diversity amongst closely related pathogens. *BMC Genomics* 13:545. <https://doi.org/10.1186/1471-2164-13-545>.
3. Bemis DA, Greisen HA, Appel M. 1977. Pathogenesis of canine bordetellosis. *J Infect Dis* 135:753–762. <https://doi.org/10.1093/infdis/135.5.753>.
4. Magyar T, Chanter N, Lax AJ, Rutter JM, Hall GA. 1988. The pathogenesis of turbinate atrophy in pigs caused by *Bordetella bronchiseptica*. *Vet Microbiol* 18:135–146. [https://doi.org/10.1016/0378-1135\(88\)90059-4](https://doi.org/10.1016/0378-1135(88)90059-4).
5. Szvalb AD, Rolston KV, Mori N, Tarrand JJ, Mulanovich VE. 2019. Infections with the agent of 'kennel cough' in patients with cancer. *J Infect* 78:48–53. <https://doi.org/10.1016/j.jinf.2018.07.010>.
6. McNulty MC, Shibib DR, Steinbeck JL, Mullane K, Pisano J, Matushek S, Beavis KG, Tesic V, Pitrak D. 2018. Misdiagnosis of *Bordetella bronchiseptica* respiratory infection as *Bordetella pertussis* by multiplex molecular assay. *Clin Infect Dis* 67:1919–1921. <https://doi.org/10.1093/cid/ciy469>.
7. Yuk MH, Heininger U, Martinez de Tejada G, Miller JF. 1998. Human but not ovine isolates of *Bordetella parapertussis* are highly clonal as determined by PCR-based RAPD fingerprinting. *Infection* 26:270–273. <https://doi.org/10.1007/bf02962245>.
8. Hester SE, Goodfield LL, Park J, Feaga HA, Ivanov YV, Bendor L, Taylor DL, Harvill ET. 2015. Host specificity of ovine *Bordetella parapertussis* and the role of complement. *PLoS One* 10:e0130964. <https://doi.org/10.1371/journal.pone.0130964>.
9. Brinig MM, Register KB, Ackermann MR, Relman DA. 2006. Genomic features of *Bordetella parapertussis* clades with distinct host species specificity. *Genome Biol* 7:R81. <https://doi.org/10.1186/gb-2006-7-9-r81>.
10. Linz B, Ivanov YV, Preston A, Brinkac L, Parkhill J, Kim M, Harris SR, Goodfield LL, Fry NK, Gorringer AR, Nicholson TL, Register KB, Losada L, Harvill ET. 2016. Acquisition and loss of virulence-associated factors during genome evolution and speciation in three clades of *Bordetella* species. *BMC Genomics* 17:767. <https://doi.org/10.1186/s12864-016-3112-5>.



11. Clark TA. 2014. Changing pertussis epidemiology: everything old is new again. *J Infect Dis* 209:978–981. <https://doi.org/10.1093/infdis/jiu001>.
12. Ausiello CM, Cassone A. 2014. Acellular pertussis vaccines and pertussis resurgence: revise or replace? *mBio* 5:e01339-14. <https://doi.org/10.1128/mBio.01339-14>.
13. Bart MJ, Harris SR, Advani A, Arakawa Y, Bottero D, Bouchez V, Cassiday PK, Chiang C-S, Dalby T, Fry NK, Gaillard ME, van Gent M, Guiso N, Hallander HO, Harvill ET, He Q, van der Heide HGJ, Heuvelman K, Hozbor DF, Kamachi K, Karataev GI, Lan R, Lutyńska A, Maharjan RP, Mertsola J, Miyamura T, Octavia S, Preston A, Quail MA, Sintchenko V, Stefanelli P, Tondella ML, Tsang RSW, Xu Y, Yao S-M, Zhang S, Parkhill J, Mooi FR. 2014. Global population structure and evolution of *Bordetella pertussis* and their relationship with vaccination. *mBio* 5:e01074-14. <https://doi.org/10.1128/mBio.01074-14>.
14. Klein NP, Bartlett J, Rowhani-Rahbar A, Fireman B, Baxter R. 2012. Waning protection after fifth dose of acellular pertussis vaccine in children. *N Engl J Med* 367:1012–1019. <https://doi.org/10.1056/NEJMoa1200850>.
15. Misegades LK, Winter K, Harriman K, Talarico J, Messonnier NE, Clark TA, Martin SW. 2012. Association of childhood pertussis with receipt of 5 doses of pertussis vaccine by time since last vaccine dose, California, 2010. *JAMA* 308:2126–2132. <https://doi.org/10.1001/jama.2012.14939>.
16. Warfel JM, Edwards KM. 2015. Pertussis vaccines and the challenge of inducing durable immunity. *Curr Opin Immunol* 35:48–54. <https://doi.org/10.1016/j.coi.2015.05.008>.
17. Heininger U, Schlassa D. 2016. Two distinct episodes of whooping cough caused by consecutive *Bordetella pertussis* and *Bordetella parapertussis* infections in a fully immunized healthy boy. *Pediatr Infect Dis J* 35:1275–1276. <https://doi.org/10.1097/INF.0000000000001295>.
18. He Q, Viljanen MK, Arvilommi H, Aittanen B, Mertsola J. 1998. Whooping cough caused by *Bordetella pertussis* and *Bordetella parapertussis* in an immunized population. *JAMA* 280:635–637. <https://doi.org/10.1001/jama.280.7.635>.
19. Pittet LF, Emonet S, Schrenzel J, Siegrist CA, Posfay-Barbe KM. 2014. *Bordetella holmesii*: an under-recognised *Bordetella* species. *Lancet Infect Dis* 14:510–519. [https://doi.org/10.1016/S1473-3099\(14\)70021-0](https://doi.org/10.1016/S1473-3099(14)70021-0).
20. Alba M-C, Gema C, Martín-Gómez MT, Anna F, Xavier M, Mireia J, Diego Van E, Thais C, Carlos R, Magda C, Tomàs P, Juan José G-L. 2017. Emergence of *Bordetella holmesii* as a causative agent of whooping cough, Barcelona, Spain. *Emerg Infect Dis* 23:1856. <https://doi.org/10.3201/eid2311.170960>.
21. Bottero D, Griffith MM, Lara C, Flores D, Pianciola L, Gaillard ME, Mazzeo M, Zamboni MI, Spoletti MJ, Anchart E, Ruggeri D, Sorhouet C, Fiori S, Galas M, Tondella ML, Hozbor DF. 2013. *Bordetella holmesii* in children suspected of pertussis in Argentina. *Epidemiol Infect* 141:714–717. <https://doi.org/10.1017/S095026881200132X>.
22. Mooi FR, Bruisten S, Linde I, Reubsaet F, Heuvelman K, van der Lee S, King AJ. 2012. Characterization of *Bordetella holmesii* isolates from patients with pertussis-like illness in the Netherlands. *FEMS Immunol Med Microbiol* 64:289–291. <https://doi.org/10.1111/j.1574-695X.2011.00911.x>.
23. Spicer KB, Salamon D, Cummins C, Leber A, Rodgers LE, Marcon MJ. 2014. Occurrence of 3 *Bordetella* species during an outbreak of cough illness in Ohio: epidemiology, clinical features, laboratory findings and antimicrobial susceptibility. *Pediatr Infect Dis J* 33:e162–e167. <https://doi.org/10.1097/INF.0000000000000262>.
24. Koepke R, Bartholomew ML, Eickhoff JC, Ayele RA, Rodd D, Kuennen J, Rosekrans J, Warshauer DM, Conway JH, Davis JP. 2015. Widespread *Bordetella parapertussis* infections—Wisconsin, 2011–2012: clinical and epidemiologic features and antibiotic use for treatment and prevention. *Clin Infect Dis* 61:1421–1431. <https://doi.org/10.1093/cid/civ514>.
25. Kamiya H, Otsuka N, Ando Y, Odaira F, Yoshino S, Kawano K, Takahashi H, Nishida T, Hidaka Y, Toyozumi-Ajisaka H, Shibayama K, Kamachi K, Sunagawa T, Taniguchi K, Okabe N. 2012. Transmission of *Bordetella holmesii* during pertussis outbreak, Japan. *Emerg Infect Dis* 18:1166–1169. <https://doi.org/10.3201/eid1807.120130>.
26. Rodgers L, Martin SW, Cohn A, Budd J, Marcon M, Terranella A, Mandal S, Salamon D, Leber A, Tondella M-L, Tatti K, Spicer K, Emanuel A, Koch E, McGlone L, Pawloski L, LeMaile-Williams M, Tucker N, Iyer R, Clark TA, DiOrio M. 2013. Epidemiologic and laboratory features of a large outbreak of pertussis-like illnesses associated with cocirculating *Bordetella holmesii* and *Bordetella pertussis*—Ohio, 2010–2011. *Clin Infect Dis* 56:322–331. <https://doi.org/10.1093/cid/cis888>.
27. Fong W, Timms V, Holmes N, Sintchenko V. 2018. Detection and incidence of *Bordetella holmesii* in respiratory specimens from patients with pertussis-like symptoms in New South Wales, Australia. *Pathology* 50:322–326. <https://doi.org/10.1016/j.pathol.2017.10.014>.
28. Cherry JD, Seaton BL. 2012. Patterns of *Bordetella parapertussis* respiratory illnesses: 2008–2010. *Clin Infect Dis* 54:534–537. <https://doi.org/10.1093/cid/cir860>.
29. Guiso N, Hegerle N. 2014. Other bordetellas, lessons for and from pertussis vaccines. *Expert Rev Vaccines* 13:1125–1133. <https://doi.org/10.1586/14760584.2014.942221>.
30. Williams MM, Taylor TH, Warshauer DM, Martin MD, Valley AM, Tondella ML. 2015. Harmonization of *Bordetella pertussis* real-time PCR diagnostics in the United States in 2012. *J Clin Microbiol* 53:118. <https://doi.org/10.1128/JCM.02368-14>.
31. Martini H, Detemmerman L, Soetens O, Yusuf E, Piérard D. 2017. Improving specificity of *Bordetella pertussis* detection using a four target real-time PCR. *PLoS One* 12:e0175587. <https://doi.org/10.1371/journal.pone.0175587>.
32. Tatti KM, Sparks KN, Boney KO, Tondella ML. 2011. Novel multitarget real-time PCR assay for rapid detection of *Bordetella* species in clinical specimens. *J Clin Microbiol* 49:4059–4066. <https://doi.org/10.1128/JCM.00601-11>.
33. Burgos-Rivera B, Lee AD, Bowden KE, Faulkner AE, Seaton BL, Lembke BD, Cartwright CP, Martin SW, Tondella ML. 2015. Evaluation of level of agreement in *Bordetella* species identification in three U.S. laboratories during a period of increased pertussis. *J Clin Microbiol* 53:1842–1847. <https://doi.org/10.1128/JCM.03567-14>.
34. Pittet LF, Posfay-Barbe KM. 2016. *Bordetella holmesii*: still emerging and elusive 20 years on. *Microbiol Spectr* 4:E110-0003-2015. <https://doi.org/10.1128/microbiolspec.E110-0003-2015>.
35. Bowden KE, Weigand MR, Peng Y, Cassiday PK, Sammons S, Knipe K, Rowe LA, Loparev V, Sheth M, Weening K, Tondella ML, Williams MM. 2016. Genome structural diversity among 31 *Bordetella pertussis* isolates from two recent U.S. whooping cough statewide epidemics. *mSphere* 1:e00036-16. <https://doi.org/10.1128/mSphere.00036-16>.
36. Weigand MR, Peng Y, Loparev V, Batra D, Bowden KE, Burroughs M, Cassiday PK, Davis JK, Johnson T, Juieng P, Knipe K, Mathis MH, Pruitt AM, Rowe L, Sheth M, Tondella ML, Williams MM. 2017. The history of *Bordetella pertussis* genome evolution includes structural rearrangement. *J Bacteriol* 199:e00806-16. <https://doi.org/10.1128/JB.00806-16>.
37. Weigand MR, Peng Y, Loparev V, Batra D, Bowden KE, Cassiday PK, Davis JK, Johnson T, Juieng P, Miner CE, Rowe L, Sheth M, Tondella ML, Williams MM. 2016. Complete genome sequences of four different *Bordetella* sp. isolates causing human respiratory infections. *Genome Announc* 4:e01080-16. <https://doi.org/10.1128/genomeA.01080-16>.
38. Mooi FR. 2010. *Bordetella pertussis* and vaccination: the persistence of a genetically monomorphic pathogen. *Infect Genet Evol* 10:36–49. <https://doi.org/10.1016/j.meegid.2009.10.007>.
39. Weigand MR, Williams MM, Otero G. 2018. Chapter 9. Temporal patterns of *Bordetella pertussis* genome sequence and structural evolution. *In* Rohai P, Scarpino SV (ed), *Pertussis: epidemiology, immunology, and evolution*. Oxford University Press, Oxford, United Kingdom.
40. Stibitz S, Yang MS. 1997. Genomic fluidity of *Bordetella pertussis* assessed by a new method for chromosomal mapping. *J Bacteriol* 179:5820–5826. <https://doi.org/10.1128/jb.179.18.5820-5826.1997>.
41. Stibitz S, Yang MS. 1999. Genomic plasticity in natural populations of *Bordetella pertussis*. *J Bacteriol* 181:5512–5515.
42. Lee H, Doak TG, Popodi E, Foster PL, Tang H. 2016. Insertion sequence-caused large-scale rearrangements in the genome of *Escherichia coli*. *Nucleic Acids Res* 44:7109–7119. <https://doi.org/10.1093/nar/gkw647>.
43. Diavatopoulos DA, Cummings CA, van der Heide HGJ, van Gent M, Liew S, Relman DA, Mooi FR. 2006. Characterization of a highly conserved island in the otherwise divergent *Bordetella holmesii* and *Bordetella pertussis* genomes. *J Bacteriol* 188:8385–8394. <https://doi.org/10.1128/JB.01081-06>.
44. Parkhill J, Sebaihia M, Preston A, Murphy LD, Thomson N, Harris DE, Holden MTG, Churcher CM, Bentley SD, Mungall KL, Cerdeño-Tárraga AM, Temple L, James K, Harris B, Quail MA, Achtman M, Atkin R, Baker S, Basham D, Bason N, Cherevach I, Chillingworth T, Collins M, Cronin A, Davis P, Doggett J, Feltwell T, Goble A, Hamlin N, Hauser H, Holroyd S, Jagels K, Leather S, Moule S, Norberczak H, O'Neill S, Ormond D, Price C, Rabinowitz E, Rutter S, Sanders M, Saunders D, Seeger K, Sharp S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Unwin L, Whitehead S, Barrell BG, Maskell DJ. 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and

- Bordetella bronchiseptica*. Nat Genet 35:32–40. <https://doi.org/10.1038/ng1227>.
45. Okada K, Ogura Y, Hayashi T, Abe A, Kuwae A, Horiguchi Y, Abe H. 2014. Complete genome sequence of *Bordetella bronchiseptica* S798, an isolate from a pig with atrophic rhinitis. Genome Announc 2:e00436-14. <https://doi.org/10.1128/genomeA.00436-14>.
  46. Slager J, Veening JW. 2016. Hard-wired control of bacterial processes by chromosomal gene location. Trends Microbiol 24:788–800. <https://doi.org/10.1016/j.tim.2016.06.003>.
  47. Kang Y, Gu C, Yuan L, Wang Y, Zhu Y, Li X, Luo Q, Xiao J, Jiang D, Qian M, Ahmed Khan A, Chen F, Zhang Z, Yu J. 2014. Flexibility and symmetry of prokaryotic genome rearrangement reveal lineage-associated core-gene-defined genome organizational frameworks. mBio 5:e01867-14. <https://doi.org/10.1128/mBio.01867-14>.
  48. Hendrickson H, Lawrence JG. 2006. Selection for chromosome architecture in bacteria. J Mol Evol 62:615–629. <https://doi.org/10.1007/s00239-005-0192-2>.
  49. Hendrickson HL, Barbeau D, Ceschin R, Lawrence JG. 2018. Chromosome architecture constrains horizontal gene transfer in bacteria. PLoS Genet 14:e1007421. <https://doi.org/10.1371/journal.pgen.1007421>.
  50. Campo N, Dias MJ, Daveran-Mingot ML, Ritzenthaler P, Le Bourgeois P. 2004. Chromosomal constraints in Gram-positive bacteria revealed by artificial inversions. Mol Microbiol 51:511–522. <https://doi.org/10.1046/j.1365-2958.2003.03847.x>.
  51. Valens M, Penaud S, Rossignol M, Cornet F, Boccard F. 2004. Macrodomain organization of the *Escherichia coli* chromosome. EMBO J 23:4330–4341. <https://doi.org/10.1038/sj.emboj.7600434>.
  52. Pannier L, Merino E, Marchal K, Collado-Vides J. 2017. Effect of genomic distance on coexpression of coregulated genes in *E. coli*. PLoS One 12:e0174887. <https://doi.org/10.1371/journal.pone.0174887>.
  53. Gerganova V, Berger M, Zaldastanishvili E, Sobetzko P, Lafon C, Mourez M, Travers A, Muskhelishvili G. 2015. Chromosomal position shift of a regulatory gene alters the bacterial phenotype. Nucleic Acids Res 43:8215–8226. <https://doi.org/10.1093/nar/gkv709>.
  54. Brambilla E, Sclavi B. 2015. Gene regulation by H-NS as a function of growth conditions depends on chromosomal position in *Escherichia coli*. G3 (Bethesda) 5:605–614. <https://doi.org/10.1534/g3.114.016139>.
  55. Sobetzko P, Travers A, Muskhelishvili G. 2012. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. Proc Natl Acad Sci U S A 109:E42–E50. <https://doi.org/10.1073/pnas.1108229109>.
  56. Bouchez V, Guiso N. 2015. *Bordetella pertussis*, *B. parapertussis*, vaccines and cycles of whooping cough. Pathog Dis 73:ftv055. <https://doi.org/10.1093/femspd/ftv055>.
  57. Zhang X, Weyrich LS, Lavine JS, Karanikas AT, Harvill ET. 2012. Lack of cross-protection against *Bordetella holmesii* after pertussis vaccination. Emerg Infect Dis 18:1771–1779. <https://doi.org/10.3201/eid1811.111544>.
  58. Saito M, Odanaka K, Otsuka N, Kamachi K, Watanabe M. 2016. Development of vaccines against pertussis caused by *Bordetella holmesii* using a mouse intranasal challenge model. Microbiol Immunol 60:599–608. <https://doi.org/10.1111/1348-0421.12409>.
  59. Zhang X, Rodriguez ME, Harvill ET. 2009. O antigen allows *B. parapertussis* to evade *B. pertussis* vaccine-induced immunity by blocking binding and functions of cross-reactive antibodies. PLoS One 4:e6989. <https://doi.org/10.1371/journal.pone.0006989>.
  60. Liko J, Robison SG, Cieslak PR. 2017. Do pertussis vaccines protect against *Bordetella parapertussis*? Clin Infect Dis 64:1795–1797. <https://doi.org/10.1093/cid/cix221>.
  61. Long GH, Karanikas AT, Harvill ET, Read AF, Hudson PJ. 2010. Acellular pertussis vaccination facilitates *Bordetella parapertussis* infection in a rodent model of bordetellosis. Proc Biol Sci 277:2017–2025. <https://doi.org/10.1098/rspb.2010.0010>.
  62. Heining U, Stehr K, Christenson P, Cherry JD. 1999. Evidence of efficacy of the Lederle/Takeda acellular pertussis component diphtheria and tetanus toxoids and pertussis vaccine but not the Lederle whole-cell component diphtheria and tetanus toxoids and pertussis vaccine against *Bordetella parapertussis* infection. Clin Infect Dis 28:602–604. <https://doi.org/10.1086/515154>.
  63. Stehr K, Cherry JD, Heining U, Schmitt-Grohé S, Überall M, Laussucq S, Eckhardt T, Meyer M, Engelhardt R, Christenson P. 1998. A comparative efficacy trial in Germany in infants who received either the Lederle/Takeda acellular pertussis component DTP (DTaP) vaccine, the Lederle whole-cell component DTP vaccine, or DT vaccine. Pediatrics 101:1–11. <https://doi.org/10.1542/peds.101.1.1>.
  64. David S, van Furth R, Mooi FR. 2004. Efficacies of whole cell and acellular pertussis vaccines against *Bordetella parapertussis* in a mouse model. Vaccine 22:1892–1898. <https://doi.org/10.1016/j.vaccine.2003.11.005>.
  65. Komatsu E, Yamaguchi F, Eguchi M, Watanabe M. 2010. Protective effects of vaccines against *Bordetella parapertussis* in a mouse intranasal challenge model. Vaccine 28:4362–4368. <https://doi.org/10.1016/j.vaccine.2010.04.053>.
  66. Zawadka M, Polak M, Rabczenko D, Mosiej E, Augustynowicz E, Lutyńska A. 2011. Effectiveness of the whole-cell pertussis vaccine produced in Poland against different *Bordetella parapertussis* isolates in the mouse intranasal challenge model. Vaccine 29:5488–5494. <https://doi.org/10.1016/j.vaccine.2011.05.053>.
  67. Bouchez V, Brun D, Dore G, Njamkepo E, Guiso N. 2011. *Bordetella parapertussis* isolates not expressing pertactin circulating in France. Clin Microbiol Infect 17:675–682. <https://doi.org/10.1111/j.1469-0691.2010.03303.x>.
  68. Hegerle N, Paris AS, Brun D, Dore G, Njamkepo E, Guillot S, Guiso N. 2012. Evolution of French *Bordetella pertussis* and *Bordetella parapertussis* isolates: increase of bordetellae not expressing pertactin. Clin Microbiol Infect 18:E340–E346. <https://doi.org/10.1111/j.1469-0691.2012.03925.x>.
  69. Worthington ZEV, Van Rooijen N, Carbonetti NH. 2011. Enhancement of *Bordetella parapertussis* infection by *Bordetella pertussis* in mixed infection of the respiratory tract. FEMS Immunol Med Microbiol 63:119–128. <https://doi.org/10.1111/j.1574-695X.2011.00836.x>.
  70. Bhattacharyya S, Ferrari MJ, Bjornstad ON. 2018. Species interactions may help explain the erratic periodicity of whooping cough dynamics. Epidemics 23:64–70. <https://doi.org/10.1016/j.epidem.2017.12.005>.
  71. Weigand MR, Williams MM, Peng Y, Kania D, Pawloski LC, Tondella ML, CDC Pertussis Working Group. 2019. Genomic survey of *Bordetella pertussis* diversity, United States, 2000–2013. Emerg Infect Dis 25:780–783. <https://doi.org/10.3201/eid2504.180812>.
  72. Skoff TH, Baumbach J, Cieslak PR. 2015. Tracking pertussis and evaluating control measures through Enhanced Pertussis Surveillance, Emerging Infections Program, United States. Emerg Infect Dis 21:1568–1573. <https://doi.org/10.3201/eid2109.150023>.
  73. Weigand MR, Peng Y, Loparev V, Batra D, Burroughs M, Johnson T, Juieng P, Rowe L, Tondella ML, Williams MM. 2016. Complete genome sequences of *Bordetella pertussis* vaccine reference strains 134 and 10536. Genome Announc 4:e00979-16. <https://doi.org/10.1128/genomeA.00979-16>.
  74. Weigand MR, Peng Y, Loparev V, Johnson T, Juieng P, Gairola S, Kumar R, Shaligram U, Gowrishankar R, Moura H, Rees J, Schieltz DM, Williamson Y, Woolfitt A, Barr J, Tondella ML, Williams MM. 2016. Complete genome sequences of four *Bordetella pertussis* vaccine reference strains from Serum Institute of India. Genome Announc 4:e01404-16. <https://doi.org/10.1128/genomeA.01404-16>.
  75. Williams MM, Sen K, Weigand MR, Skoff TH, Cunningham VA, Halse TA, Tondella ML, CDC Pertussis Working Group. 2016. *Bordetella pertussis* strain lacking pertactin and pertussis toxin. Emerg Infect Dis 22:319–322. <https://doi.org/10.3201/eid2202.151332>.
  76. Weigand MR, Peng Y, Cassidy PK, Loparev VN, Johnson T, Juieng P, Nazarian EJ, Weening K, Tondella ML, Williams MM. 2017. Complete genome sequences of *Bordetella pertussis* isolates with novel pertactin-deficient deletions. Genome Announc 5:e00973-17. <https://doi.org/10.1128/genomeA.00973-17>.
  77. Weigand MR, Changayil S, Kulasekarapandian Y, Batra D, Loparev V, Juieng P, Rowe L, Sheth M, Davis JK, Tondella ML. 2015. Complete genome sequences of two *Bordetella hinzii* strains isolated from humans. Genome Announc 3:e00965-15. <https://doi.org/10.1128/genomeA.00965-15>.
  78. Tatti KM, Loparev VN, Ranganathanakammal S, Changayil S, Frace M, Weil MR, Sammons S, Maccannell D, Mayer LW, Tondella ML. 2013. Draft genome sequences of *Bordetella holmesii* strains from blood (F627) and nasopharynx (H558). Genome Announc 1:e00056-13. <https://doi.org/10.1128/genomeA.00056-13>.
  79. Green MR, Sambrook J, Sambrook J. 2012. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
  80. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read



- SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
81. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
  82. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
  83. Krumsiek J, Arnold R, Rattei T. 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 23:1026–1028. <https://doi.org/10.1093/bioinformatics/btm039>.
  84. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
  85. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 17:132. <https://doi.org/10.1186/s13059-016-0997-x>.
  86. Katz LS, Griswold T, Carleton HA. 2017. Generating WGS trees with mashtree, poster 27. *Abstr ASM Conf Rapid Appl Microb Next Gener Seq Bioinformatic Pipelines*. American Society for Microbiology, Washington, DC.
  87. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31:2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>.
  88. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242. <https://doi.org/10.1093/nar/gkw290>.
  89. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
  90. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2:vew007. <https://doi.org/10.1093/ve/vew007>.
  91. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 4:vey016. <https://doi.org/10.1093/ve/vey016>.
  92. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904. <https://doi.org/10.1093/sysbio/syy032>.
  93. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 8:28–36. <https://doi.org/10.1111/2041-210X.12628>.
  94. Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27:764–770. <https://doi.org/10.1093/bioinformatics/btr011>.
  95. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
  96. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34:D32–D36. <https://doi.org/10.1093/nar/gkj014>.
  97. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. <https://doi.org/10.1371/journal.pone.0011147>.
  98. Oksanen J, Guillaume Blanchet F, Friendly M, Kindt R, Legendre P, McGinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Wagner H. 2018. vegan: Community Ecology Package. R package version 2.5-2. <https://CRAN.R-project.org/package=vegan>.
  99. Gao F, Luo H, Zhang C-T. 2013. DoriC 5.0: an updated database of *oriC* regions in both bacterial and archaeal genomes. *Nucleic Acids Res* 41:D90–D93. <https://doi.org/10.1093/nar/gks990>.
  100. Kono N, Arakawa K, Tomita M. 2011. Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. *BMC Genomics* 12:19. <https://doi.org/10.1186/1471-2164-12-19>.
  101. Butts CT. 2008. network: a package for managing relational data in R. *J Stat Softw* 24:v024i02. <https://doi.org/10.18637/jss.v024.i02>.
  102. Butts CT. 2008. Social network analysis with sna. *J Stat Softw* 24:v024i06. <https://doi.org/10.18637/jss.v024.i06>.
  103. Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. circlize implements and enhances circular visualization in R. *Bioinformatics* 30:2811–2812. <https://doi.org/10.1093/bioinformatics/btu393>.