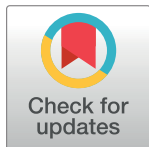


## RESEARCH ARTICLE

Genomic diversity and molecular epidemiology of *Pasteurella multocida*Emily Smith<sup>1\*</sup>, Elizabeth Miller<sup>1</sup>, Jeannette Munoz Aguayo<sup>2</sup>, Cristian Flores Figueroa<sup>2</sup>, Jill Nezworski<sup>3</sup>, Marissa Studniski<sup>4</sup>, Ben Wileman<sup>4</sup>, Timothy Johnson<sup>1,2</sup>

**1** Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, MN, United States of America, **2** Mid-Central Research and Outreach Center, University of Minnesota, Willmar, Minnesota, United States of America, **3** Blue House Veterinary LLC, Buffalo Lake, Minnesota, United States of America, **4** Select Genetics, Willmar, MN, United States of America

\* [smi8380@umn.edu](mailto:smi8380@umn.edu)

## OPEN ACCESS

**Citation:** Smith E, Miller E, Aguayo JM, Figueroa CF, Nezworski J, Studniski M, et al. (2021) Genomic diversity and molecular epidemiology of *Pasteurella multocida*. PLoS ONE 16(4): e0249138. <https://doi.org/10.1371/journal.pone.0249138>

**Editor:** Yung-Fu Chang, Cornell University, UNITED STATES

**Received:** October 2, 2020

**Accepted:** March 11, 2021

**Published:** April 6, 2021

**Copyright:** © 2021 Smith et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Genome sequences - NCBI Short Read Archive (BioProject PRJNA666906) Databases - <https://github.com/JohnsonSingerLab/Pmult-databases>.

**Funding:** This work was supported by Grant no. 2018-68003-27464 from the USDA National Institute of Food and Agriculture. This work was supported by Grant no. 2018-68003-27464 from the USDA National Institute of Food and Agriculture (<https://nifa.usda.gov>), awarded to TJ.

**Competing interests:** JN is employed by Blue House Veterinary LLC. MS and BW are employed

## Abstract

*Pasteurella multocida* is a bacterial pathogen with the ability to infect a multitude of hosts including humans, companion animals, livestock, and wildlife. This study used bioinformatic approaches to explore the genomic diversity of 656 *P. multocida* isolates and epidemiological associations between host factors and specific genotypes. Isolates included in this study originated from a variety of hosts, including poultry, cattle, swine, rabbits, rodents, and humans, from five different continents. Multi-locus sequence typing identified 69 different sequence types. *In-silico* methodology for determining capsular serogroup was developed, validated, and applied to all genome sequences, whereby capsular serogroups A, B, D, and F were found. Whole genome phylogeny was constructed from 237,670 core single nucleotide variants (SNVs) and demonstrated an overall lack of host or capsular serogroup specificity, with the exception of isolates from bovine sources. Specific SNVs within the *srIB* gene were identified in *P. multocida* subsp. *septica* genomes, representing specific mutations that may be useful for differentiating one of the three known subspecies. Significant associations were identified between capsular serogroup and virulence factors, including capsular serogroup A and OmpH1, OmpH3, PlpE, and PfhB1; capsular serogroup B and HgbA and PtfA; and capsular serogroup F and PtfA and PlpP. Various mobile genetic elements were identified including those similar to ICE*Pmu1*, ICE*hin1056*, and IncQ1 plasmids, all of which harbored multiple antimicrobial resistance-encoding genes. Additional analyses were performed on a subset of 99 isolates obtained from turkeys during fowl cholera outbreaks from a single company which revealed that multiple strains of *P. multocida* were circulating during the outbreak, instead of a single, highly virulent clone. This study further demonstrates the extensive genomic diversity of *P. multocida*, provides epidemiological context to the various genotyping schemes that have traditionally been used for differentiating isolates, and introduces additional tools for *P. multocida* molecular typing.

by Select Genetics. There are no patents, products in development or marketed products to declare. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

## Introduction

*Pasteurella multocida* is a zoonotic pathogen with transboundary dissemination across a multitude of hosts and geographic landscapes. Infections with this Gram-negative bacterium are responsible for significant morbidity and mortality in both humans and animals, as well as economic losses in the livestock industry. Common disease manifestations due to *P. multocida* include fowl cholera in poultry, hemorrhagic septicemia in cattle, and atrophic rhinitis in swine [1,2]. Human infections are typically the result of bites from cats or dogs, most of which result in mild wound infections, but complications such as osteomyelitis or septicemia can occur [3–6].

Previous studies have demonstrated that poor biosecurity or management practices have contributed to *P. multocida* outbreaks in poultry, cattle, and swine [7–9]. Other risk factors, such as co-infection with another pathogenic organism, age, and vaccination status of the animals have also been reported [10–14]. In commercial turkey production, it is mostly unknown whether outbreaks of fowl cholera are attributable to a highly virulent single clone, or whether multiple *P. multocida* strains are circulating simultaneously. The identification of multiple strains circulating at the same time may further emphasize the role of biosecurity in prevention of fowl cholera, however, there has been little effort to determine specific strains responsible for outbreaks at the whole-genome level.

To date, typing schemes that target specific genotypes or phenotypes are primarily utilized to classify *P. multocida* strains. Specifically, *P. multocida* can be classified into 5 capsular serogroups (A, B, D, E, and F) based on the antigenicity of the outer capsule [15–17]. This capsule plays a major role in the overall virulence of *P. multocida*, as isolates lacking this capsule demonstrate reduced capacity to invade and adhere to host cells and resist phagocytosis [18–20]. *P. multocida* can be further divided into 16 somatic serovars using the Heddlestone serotyping scheme based on the structure of lipopolysaccharide antigens [21,22]. Although this serotyping scheme is widely used, many isolates are non-typable with this method and results are often ambiguous due to cross-reactivity during the gel diffusion precipitin test [23–26]. More recently, a multiplex PCR assay has been developed that is able to distinguish Heddlestone serotypes of *P. multocida* based on the genetic organization of the LPS outer core biosynthesis loci with greater accuracy than the traditional gel diffusion precipitin test. [23,27]. Overall, serotyping can provide useful information about the cell-surface antigens of *P. multocida*. However, it is not an indicator of genetic diversity between isolates and has limited capacity to differentiate related strains.

*P. multocida* isolates can also be classified into one of three subspecies, subsp. *multocida*, *septica*, or *gallicida* based on the ability to ferment sorbitol and dulcitol [28]. Some previous studies have stated that *P. multocida* subsp. *multocida* and *septica* are associated with more severe clinical presentations, but these studies have been limited in scope and sample size [29,30]. Phenotypic methods have traditionally been used to determine subspecies, but these methods are time consuming and resource intensive. A PCR assay has been developed to distinguish subspecies based on the 16S rRNA gene, however, specific genes associated with sorbitol or dulcitol fermentation that may differentiate subsp. *multocida*, *septica*, and *gallicida* have not yet been identified [31]. Determining these genetic components that differentiate subspecies would allow for more rapid identification and lay a foundation for studies on epidemiological factors that may be associated with specific subspecies.

Multi-locus sequence typing (MLST) is another common method used to differentiate *P. multocida* strains and is based on the allelic profiles of seven housekeeping genes [32]. There are two different schemes available for MLST of *P. multocida*: the Rural Industries Research and Development Organization (RIRDC) scheme, which currently includes 373 different sequence types, and the Multi-host scheme, which currently consists of 127 different sequence types [33–35]. Both schemes have been widely used to sequence type isolates from various hosts, collection

dates, and geographic regions [36–41]. The use of MLST alone excludes the possibility of discovering variation that exists outside seven specific genes. Therefore, it is often used in combination with other genotypic methods, such as the identification of virulence factors or antimicrobial resistance genes, to provide a more complete picture of the *P. multocida* study population [36,42–45]. Multiple plasmids containing antimicrobial resistance genes have also been identified in *P. multocida* isolates from multiple hosts and geographic regions, some of which are highly similar to those identified in other microbial species [46–48].

While the various typing schemes and assays for differentiating *P. multocida* isolates have enhanced the understanding of underlying molecular mechanisms and genetic diversity of *P. multocida*, they alone have little value in the absence of epidemiological data. Additionally, the associations between these different typing methods and relationships with host and geography have not been fully explored. Therefore, the purpose of this study was to characterize all publicly available *P. multocida* genomes and their associated raw reads using bioinformatic and epidemiologic approaches to provide insight into the diversity of the pathogen population with respect to host, geography, and genetic indicators of virulence and antimicrobial resistance. This study also employed the application of genomic tools in outbreak investigations, using fowl cholera outbreaks as an example to highlight the genetic diversity of *P. multocida* among farms from a single commercial turkey producer.

## Results

There were a total of 656 *P. multocida* whole genome sequences included in this study. These sequences were extracted from the NCBI Short Read Archive database ( $n = 338$ ), NCBI Assembly database ( $n = 193$ ), and from archived isolates at the University of Minnesota Mid-Central Research and Outreach Center ( $n = 125$ ). The genomes were sourced from isolates from a variety of hosts, a majority of which were avian in origin, followed by bovine, rabbit, and swine species (Table 1). Most genomes came from isolates from hosts in North America ( $n = 390$ ),

**Table 1. Hosts from which *P. multocida* was isolated ( $n = 656$ ).**

Host	No.	%	Specific host
Avian	373	56.3	Chicken (130); Duck (10); Goose (3); Gull (1); Pine siskin (1); Turkey (225)*; Unknown (3)
Bovine	155	23.5	Bison (3); Buffalo (30); Cattle (123)
Camelid	2	0.3	Alpaca (2)
Canine	2	0.3	Dog (1); Unknown (1)
Caprine	2	0.3	Goat (2)
Cervid	3	0.5	Deer (3)
Feline	2	0.3	Cat (2)
Human	16	2.4	Human (16)
Lagomorph	35	5.3	Rabbit (35)
Marsupial	14	2.1	Rufous rat-kangaroo (1); Squirrel glider (11); Woylie (2)
Otariid	6	0.9	Australasian fur seal (4); Subantarctic fur seal (2)
Ovine	4	0.6	Sheep (4)
Rodent	2	0.3	Unknown (2)
Swine	21	3.2	Boar (5); Pig (16)
Unknown	19	2.9	Unknown (19)

\*125 isolates were provided by the University of Minnesota Mid-Central Research and Outreach Center for the current study.

<https://doi.org/10.1371/journal.pone.0249138.t001>

but isolates from Asia ( $n = 98$ ), Europe ( $n = 42$ ), Australia ( $n = 95$ ), and South America ( $n = 6$ ) were also present. Collection dates ranged from 1936 to 2019. The Genbank assembly accession numbers and SRA numbers of genomes used in this study and accompanying metadata can be found in [S1 Dataset](#).

### Sequence-typing demonstrates broad host range and presence of novel sequence types

Amongst the *P. multocida* genomes, 69 different sequence types (STs) were identified. Some STs appeared to be host-specific, such as the 75 genomes identified as ST20 that were exclusively from chickens in Australia. Similarly, ST79 included 99 genomes of bovine origin from North America, Asia, and Europe. However, many STs were not comprised of genomes from a single host species. For example, 36 genomes were identified as ST9, which has a broad host range amongst the genomes in this study, including avian, bovine, feline, rabbit, rodent, and swine sources. Of the 16 *P. multocida* genomes from humans, 13 fell into STs that included other hosts, such as ST6, which also included genomes from poultry or canine sources, and ST25, which included 4 genomes from human wounds and abscesses, but also included two genomes from turkeys. Additionally, ST242 contained 3 *P. multocida* genomes from human septicemia patients, but also included genomes from chickens and turkeys. There were 114 genomes that were unable to be sequence-typed. Of these, 110 (96.5%) genomes had all seven alleles identified, so it is likely that these represent novel sequence types that did not exist under the current scheme. The collection dates for genomes unable to be sequence-typed ranged from 1970 to 2018.

### *In-silico* capsular serotyping identifies minimal differences in host range and geographic distribution

An *in-silico* capsular serogrouping method was developed to identify the presence of genes encoding *P. multocida* capsular antigens A, B, D, and F within *P. multocida* genomes. There were 55 *P. multocida* genomes that had capsular serogroup data available in NCBI. The *in-silico* capsular serogrouping method correctly determined the capsular serogroup for 54 (98.1%) of these genomes, demonstrating that the *in-silico* methodology developed in this study may be a useful tool for capsular serogroup prediction when other methods are unavailable. The one genome for which the *in-silico* method did not correctly identify the capsular serogroup was designated as capsular serogroup A in the NCBI Assembly database but was identified as serogroup F using the *in-silico* methodology. When applied to all 656 genomes, capsular serogroups A ( $n = 516$ ), B ( $n = 45$ ), D ( $n = 33$ ), and F ( $n = 62$ ) were identified. Among the genomes in this study, 482 (73.5%) had >50% coverage of the *cap* gene used to distinguish capsular serogroups. The proportion of *P. multocida* genomes designated as different capsular serogroups within each host category were identified ([Table 2](#)). Capsular serogroup E was not identified from any genomes in this study. Capsular serogroup A was found in every host category with the exception of caprine hosts. All four capsular serogroups were found in isolates of avian origin, but capsular serogroup A (85.0%) was most common. For *P. multocida* isolates of bovine origin, capsular serogroup A (71.0%) was the most common, followed by capsular serogroups B (27.1%) and F (1.9%). Similarly, for isolates of swine origin, capsular serogroup A (66.7%) was the most common, followed by capsular serogroups D (19.0%), F (9.5%), and B (4.8%). All *P. multocida* isolates from human isolates were from serogroup A, with the exception of one which was of capsular serogroup F. The A and F capsular serogroups were also the only serogroups identified in cats and dogs. There did not appear to be major differences in the geographic distribution of *P. multocida* capsular serogroups. Capsular serogroups A and F

Table 2. *Pasteurella multocida* capsular serotypes by host category ( $n = 656$ ).

Species	A ( $n = 516$ )	B ( $n = 45$ )	D ( $n = 33$ )	F ( $n = 62$ )
Avian	317 (85.0)	2 (0.5)	24 (6.4)	30 (8.0)
Bovine	110 (71.0)	42 (27.1)	0 (0.0)	3 (1.9)
Camelid	2 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Canine	1 (50.0)	0 (0.0)	0 (0.0)	1 (50.0)
Caprine	0 (0.0)	0 (0.0)	2 (100.0)	0 (0.0)
Cervid	3 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Feline	1 (50.0)	0 (0.0)	0 (0.0)	1 (50.0)
Human	15 (93.8)	0 (0.0)	0 (0.0)	1 (6.3)
Lagomorph	19 (54.3)	0 (0.0)	2 (5.7)	14 (40.0)
Marsupial	13 (92.9)	0 (0.0)	0 (0.0)	1 (7.1)
Otariid	6 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)
Ovine	2 (50.0)	0 (0.0)	0 (0.0)	2 (50.0)
Rodent	1 (50.0)	0 (0.0)	0 (0.0)	1 (50.0)
Swine	14 (66.7)	1 (4.8)	4 (19.0)	2 (9.5)
Unknown	12 (62.3)	0 (0.0)	1 (5.3)	6 (31.6)

The numbers in parentheses reflect the percentage of genomes from the corresponding host type that belong to that particular serogroup.

<https://doi.org/10.1371/journal.pone.0249138.t002>

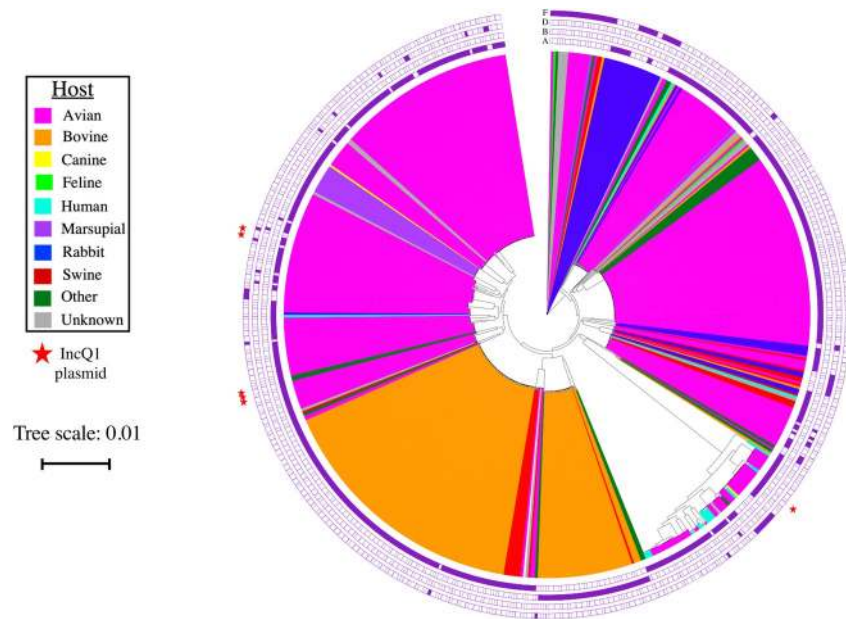
were identified in isolates from all continents. Capsular serogroups B and D were only found in isolates from Asia, Europe, and North America, but this is likely reflective of the small number of isolates that were identified as these two capsular serogroups.

### SNV-based phylogeny demonstrates apparent bovine host specificity, but lack of host specificity for other sources

There were 237,670 core single nucleotide variants (SNVs) identified from 653 genome sequences. One genome (SRR2907037) was removed from the SNV alignment because no SNVs were identified despite 98.1% of nucleotides aligning to the reference genome (GCA\_000754275.1), and three genome sequences were automatically removed from the SNV alignment by Gubbins due to a high proportion of missing data (SRR5256406, PP64, PP74). The transversion model of DNA substitution with empirical base frequencies and three rate categories was selected as the best fit model by IQ-TREE and was used to construct the phylogenetic tree (Fig 1).

Almost all of the isolates from bovine sources fell into two phylogenetic clades (Fig 1), potentially indicating some degree of host-specificity. This includes all of the ST79 ( $n = 99$ ) and ST122 ( $n = 42$ ) isolates of primarily bovine origin. The majority of ST79 genomes belonged to capsular serogroup A, while all of the ST122 genomes belonged to capsular serogroup B. Four of the ST122 isolates were bovine vaccine strains of *P. multocida*, while the rest were recovered from sources including the nasal cavity, throat, lungs, blood, and bone marrow of buffalo and cattle from different continents, with the exception of one isolate of swine origin. In contrast, the isolates of avian origin were dispersed throughout the tree. This may potentially indicate that birds are susceptible to a wider variety of *P. multocida* strains compared to bovine hosts. While there is some evidence of clustering of the isolates from rabbits and swine, these clusters were also dispersed throughout the tree. There did not appear to be any major patterns on the phylogenetic tree by collection year.

Interestingly, there was one clade with a much longer branch length compared to the rest of the clades in the phylogenetic tree that represents more than 6,000 SNV differences from all



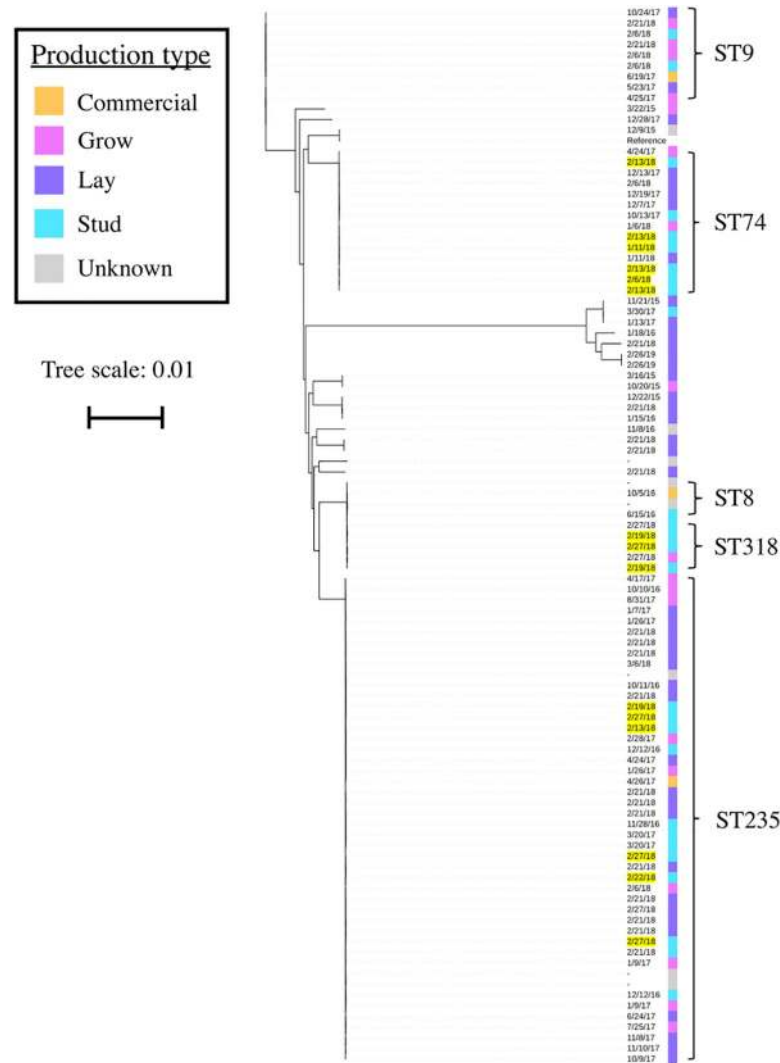
**Fig 1. SNV-based phylogeny of *Pasteurella multocida*.** Phylogenetic tree constructed from 237,670 core SNVs identified from 653 whole-genome sequences. The colors inside the tree represent different host sources (see host legend), the outer rings represent the corresponding capsular serogroup (filled = present and white = absent), and the star represents the presence of the IncQ1 plasmid (star = present and none = absent). The scale bar represents a branch length of 0.01 (approximately 2,377 SNVs).

<https://doi.org/10.1371/journal.pone.0249138.g001>

other clades. This outer clade contains 13 of the 16 *P. multocida* isolates from humans, as well as isolates of avian, canine, and rodent origin. Capsular serogroups A, D, and F were found in the isolates from this clade. Isolates from this clade originated from Asia, Europe, and North America. Subspecies information was only available for five of these isolates, two of which were subsp. *multocida*, and three of which were subsp. *septica*. Upon further investigation of what may differentiate this clade from other genomes, 1,911 genes and gene groups were identified with significantly increased or decreased odds of being present in this specific clade compared to the rest of the *P. multocida* genomes. Genomes in this clade had significantly increased odds of containing *esiB* (OR = 6.1;  $p = 0.022$ ), annotated as a secretory immunoglobulin A-binding gene, which has been shown to impair host neutrophil activation, and toxin *cdiA* (OR = 15.8;  $p = 0.009$ ), which has been shown to demonstrate contact-dependent growth inhibition of surrounding bacteria [49,50]. The *oatA* gene encoding an O-acetyltransferase, which allows the bacteria to evade an initial host immune response, was also significantly associated with outer clade (OR = infinite;  $p = 0.003$ ) [51]. Genes associated with antimicrobial resistance were also found at higher prevalence in this clade compared to all other genomes, including *etk* (OR = 14.7;  $p = 0.002$ ), a putative tyrosine-protein kinase that has been associated with resistance to polymyxin, as well as *tetD* (OR = 100.5;  $p < 0.0001$ ), a tetracycline efflux gene [52]. Additionally, the outer clade had significantly decreased odds of containing *wzc* (OR = 0.09;  $p = 0.02$ ), a putative tyrosine-protein kinase associated with susceptibility to macrolide antibiotics [53].

### SNV-based phylogeny shows persistence of multiple strains involved in fowl cholera outbreaks

A collection of 99 *P. multocida* isolates involved in outbreaks of fowl cholera within a single vertically integrated commercial turkey production company occurring over five years was



**Fig 2. SNV-based phylogeny of *P. multocida* isolates involved in recurrent fowl cholera outbreaks within the same commercial turkey company.** Phylogenetic tree constructed from 100,023 core SNVs of 99 *P. multocida* genomes. These genomes originated from clinical isolates collected during multiple fowl cholera outbreaks within a single company. The different colors represent different stages of turkey production (see legend). Collection dates are listed, and those highlighted in yellow represent successive samplings from a single indoor flock during an outbreak of fowl cholera in 2018. Clusters of specific sequence types are indicated with brackets. The scale bar represents a branch length of 0.01 (approximately 1,000 SNVs).

<https://doi.org/10.1371/journal.pone.0249138.g002>

investigated. This subset includes genomes from all four capsular serogroups identified *in-silico*. A total of 100,023 SNVs were identified amongst the genomes of these isolates, indicating that multiple clones were implicated in this persistent and recurring disease issue. The core SNV phylogeny did not demonstrate associations between different stages of production, indicating that birds of all ages were likely susceptible to the same strains of *P. multocida* (Fig 2). Although collection dates ranged from 2015 to 2019, there were no distinct patterns in collection date on the phylogenetic tree, demonstrating that similar strains of *P. multocida* have persisted on farms from the same company over time, or were re-introduced from an external source on multiple occasions. Additionally, isolates of 3 different sequence types were collected from a single flock that experienced a fowl cholera outbreak in early 2018, indicating that multiple strains of *P. multocida* may be involved in the same fowl cholera outbreak within a flock.

## Pangenome analyses reveals key genes associated with *P. multocida* subspecies

There were 68 genomes for which subspecies information was available on NCBI, which included subsp. *multocida* ( $n = 48$ ), subsp. *septica* ( $n = 16$ ), and subsp. *gallicida* ( $n = 4$ ). The genomes in each of these subspecies were comprised of multiple host sources, capsular serogroups, and sequence types. A search of the *P. multocida* pangenome, consisting of 37,854 genes, revealed the presence of sorbitol-specific PTS enzymes encoded by the *srl* operon, *srlAEBD*, which is homologous to the *gut* operon in *E. coli*, in 95.4% of *P. multocida* genomes [54]. This included 13 genomes of the *septica* subspecies, which are distinguished based on an inability to ferment sorbitol [54,55]. As no other genes associated with sorbitol fermentation were identified, this operon was further investigated in a SNV analysis of the 68 genomes of known subspecies. Interestingly, 12 subsp. *septica* sequences contained the specific SNV C→T at position 520 in the *srlB* gene, that resulted in a premature stop codon. The one subsp. *septica* sequence that did not contain this SNV had a missense mutation, C→A at position 226, in the same gene. Therefore, it is likely that the *srl* operon is the gene associated with sorbitol fermentation in *P. multocida*, and specific SNVs are responsible for the lack of sorbitol fermentation in subsp. *septica* isolates. Further investigation is needed on a larger set of genomes of known subspecies. However, this serves as the first step towards development of an *in-silico* method for distinguishing *P. multocida* subspecies.

The *P. multocida* pangenome was also searched for genes associated with dulcitol (galactitol) fermentation, and the galactitol-specific PTS enzymes *gatA*, *gatB*, and *gatC*, were identified. Three of the four genomes of the *gallicida* subspecies contained *gatABC*, compared to just one of 63 *multocida* and *septica* genomes. However, this finding was not significant due to the small sample size of *gallicida* genomes compared to the other subspecies (OR = 63.0;  $p = 0.88$ ). The association between *gatABC* and the *gallicida* subspecies warrants further analysis with a larger sample size and may eventually be used in combination with *srlAEBD* or genes not yet identified to distinguish all three *P. multocida* subspecies *in-silico*.

## Acquired antimicrobial resistance genes have integrated into *P. multocida* chromosome through mobile genetic elements

Acquired AMR-encoding genes conferring resistance to 11 different antimicrobial drug classes were identified among the *P. multocida* genomes (Table 3). The most common drug classes with resistance genes identified included aminoglycosides, with 120 (18.1%) *P. multocida* genomes containing an aminoglycoside resistance gene, followed by tetracyclines (17.5%), streptogramins (10.3%), macrolides (10.1%), and lincosamides (10.0%). Drug classes with fewer AMR genes identified included sulfonamides (9.2%), phenicols (7.6%), and beta-lactams (2.9%). Less than one percent of genomes contained genes conferring resistance to rifamycin (0.5%), fosfomycin (0.3%), and trimethoprim antibiotics (0.3%). Many of these genes were associated with mobile genetic elements determined using PlasmidFinder and ICEberg 2.0 databases [56,57].

For example, the integrative and conjugative element (ICE), *ICEPmu1*, was identified in 73 isolates of bovine origin from the United States belonging to ST79 and one isolate of bovine origin with an unknown ST and geographic location. All genomes containing *ICEPmu1* aligned to the *ICEPmu1* reference sequence with >90% identity and >50% coverage. Interestingly, *ICEPmu1* was also identified in one *P. multocida* isolate from a turkey with an unknown sequence type. All genomes containing *ICEPmu1* belonged to serotype A. Notably, *ICEPmu1* contains the genes *aph(3'')-lb*, *aph(3'')-la*, *aph(6)-ld*, conferring resistance to aminoglycoside



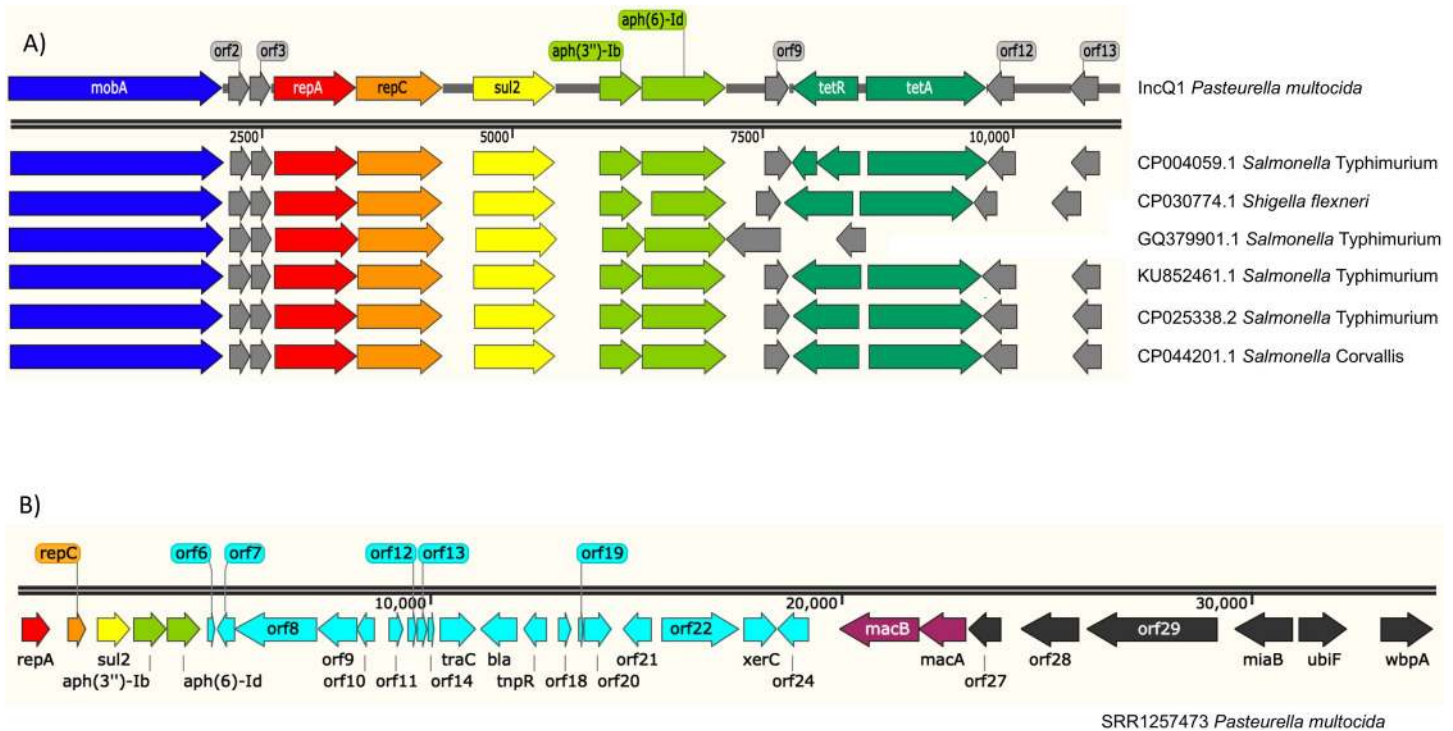
**Table 3. Prevalence of antimicrobial resistance genes by drug class across 656 *Pasteurella multocida* genomes used in this study.**

Class	No.	%	Genes
Aminoglycosides	120	18.1	<i>aadA1</i> ; <i>aadA25</i> ; <i>aadA31</i> ; <i>aadK</i> ; <i>ant(2'')-Ia</i> ; <i>aph(3'')-Ib</i> ; <i>aph(3'')-Ia</i> ; <i>aph(6)-IId</i>
Beta-lactams	19	2.9	<i>bcII</i> ; <i>bla1</i> ; <i>blaA</i> ; <i>bla<sub>ORR-1</sub></i> ; <i>bla<sub>OXA-1</sub></i> ; <i>bla<sub>OXA-2</sub></i> ; <i>blaP</i> ; <i>bla<sub>ROB-1</sub></i> ; <i>bla<sub>TEM-1</sub></i> ; <i>bla<sub>VEB-18</sub></i>
Fosfomycin	2	0.3	<i>fosA</i> ; <i>fosB</i>
Lincosamides	66	10.0	<i>msrE</i> ; <i>erm42</i>
Macrolides	67	10.1	<i>mefB</i> ; <i>mphE</i> ; <i>msrE</i> ; <i>erm42</i>
Phenicol	50	7.6	<i>catA2</i> ; <i>catA3</i> ; <i>catD</i> ; <i>cmlA5</i> ; <i>floR</i> ; <i>msrE</i>
Rifamycin	3	0.5	<i>arr-2</i> ; <i>rphC</i>
Streptogramins	68	10.3	<i>erm42</i> ; <i>satA</i> ; <i>satB</i> ; <i>msrE</i>
Sulfonamides	61	9.2	<i>sul2</i> ; <i>sul3</i>
Tetracyclines	116	17.5	<i>msrE</i> ; <i>tetA</i> ; <i>tetB</i> ; <i>tetH</i> ; <i>tetQ</i> ; <i>tetY</i>
Trimethoprim	2	0.3	<i>dfrA1</i> ; <i>dfrA32</i>

<https://doi.org/10.1371/journal.pone.0249138.t003>

drugs, *erm(42)*, conferring resistance to macrolide drugs, *tetH*, conferring resistance to tetracycline, and *sul2*, conferring resistance to sulfonamides.

IncQ1 plasmids were identified in six *P. multocida* isolates of avian origin from North America. This plasmid also contains *aph(3'')-Ib*, *aph(6)-IId*, and *sul2*, and additionally includes two tetracycline genes, *tetA* and *tetR*. These IncQ1 plasmids shared >99% nucleotide identity with IncQ1 plasmids identified in other pathogenic bacterial species, including *Salmonella enterica* subsp. Typhimurium and *Shigella flexneri* (Fig 3A). IncQ1 replicons were also identified in six isolates of bovine origin from Asia. However, upon further inspection, only a



**Fig 3. IncQ1 plasmid maps and chromosomal insertion.** A) Map of IncQ1 plasmids from *P. multocida* compared to IncQ1 plasmids of other pathogenic bacterial species. B) Partial insertion of IncQ1 plasmid into the *P. multocida* chromosome in an isolate of bovine origin, adjacent to ICE*hin1056*. Dark blue = mobilization gene; red and orange = replication genes; yellow = sulfonamide resistance gene; light green = aminoglycoside resistance genes; dark green = tetracycline resistance genes; light blue = ICE*hin1056*; mauve = macrolide resistance genes.

<https://doi.org/10.1371/journal.pone.0249138.g003>

fragment of the entire plasmid was identified in each of these genomes, and it appeared to be inserted into the chromosome at similar locations in all cases. Adjacent to the insertion was an ICE which had >99% nucleotide identity with ICE<sub>Hin1056</sub> of *Haemophilus influenzae*. Annotation of ICE<sub>Hin1056</sub> revealed the presence of additional antimicrobial resistance genes including *bla*<sub>TEM-1</sub>, which confers resistance to beta-lactam antibiotics (Fig 3B). All genomes that included a partial or whole IncQ1 plasmid were the same sequence type, ST122, which suggests that IncQ1 was potentially acquired in a single event, and possibly a second event integrated this element into the chromosome.

### Distribution of virulence factors demonstrates association with specific capsular serogroups

Genomes were searched for the presence of 15 known *P. multocida* virulence factors (Table 4). All 656 genomes contained Omp16, TbpA, and PlpB. More than 99% of genomes also contained HgbB, OmpH2, ComE, PlpD, and PfhB2, while less than 99% of genomes contained HgbA, OmpH1, OmpH3, PtfA, PlpE, PlpP, and PfhB1. These genes were further investigated for statistically significant associations with capsular serogroups A, B, D, and F (Table 5). Capsular serogroup B genomes had significantly increased odds (OR = 29.07;  $p < 0.00001$ ) of containing HgbA, Hemoglobin-binding protein A, compared to genomes of all other capsular serogroups. Capsular serogroup A genomes had significantly increased odds of containing the outer membrane proteins OmpH1 (OR = 9.09;  $p < 0.00001$ ) and OmpH3 (OR = 2.56;  $p = 0.00003$ ) compared to all other capsular serogroups. PtfA, a Type IV Fimbriae, was significantly associated with capsular serogroups B (OR = infinite;  $p = 0.00272$ ) and F (OR = 7.9;  $p = 0.00734$ ). PlpE and PlpP, protective outer membrane lipoproteins, were significantly associated with capsular serogroups A (OR = 4.25;  $p < 0.00001$ ) and F (OR = 8.96;  $p < 0.0001$ ), respectively. Lastly, capsular serogroup A genomes had statistically increased odds of containing, PfhB1, a filamentous hemagglutinin protein, compared to all other genomes (OR = 3.64;  $p < 0.00001$ ).

### Discussion

This large-scale bioinformatics study provides a comprehensive overview of the molecular epidemiology of *P. multocida* using publicly available whole-genome sequence data. Despite the

**Table 4. Prevalence of select virulence factors across 656 *Pasteurella multocida* genomes in this study.**

Protein	Acc. No.	Description	No.	%
HgbA	AAQ14873.1	Hemoglobin-binding protein A	240	36.59
HgbB	APX53047.1	Hemoglobin-binding protein B	650	99.09
Omp16	AHW46109.1	Outer membrane protein 16	656	100.00
OmpH1	SNV59447.1	Outer membrane protein H1	113	17.23
OmpH2	VEE38334.1	Outer membrane protein H2	652	99.39
OmpH3	AMK08231.1	Outer membrane protein H3	386	58.84
TbpA	AAK02460.1	Transferrin binding protein A	656	100.00
PtfA	AKO69808.1	Type IV fimbriae	530	80.79
ComE	AFF25459.1	Response regulator	650	99.09
PlpE	AAK03601.1	Protective outer membrane lipoprotein E	453	69.05
PlpP	AAK03602.1	Protective outer membrane lipoprotein P	180	27.44
PlpB	AAK03814.1	Protective outer membrane lipoprotein B	656	100.00
PlpD	VEE38139.1	Protective outer membrane lipoprotein D	651	99.24
PfhB1	AAK02141.1	Filamentous hemagglutinin protein B1	482	73.48
PfhB2	AAK02143.1	Filamentous hemagglutinin protein B2	650	99.09

<https://doi.org/10.1371/journal.pone.0249138.t004>

**Table 5. Association between *Pasteurella multocida* virulence factors and capsular serogroups in this study (OR = odds ratio, CI = confidence interval, Inf = infinity).**

	OR	95% CI	p-value
<b>HgbA</b>			
A	0.68	(0.46, 1.03)	0.060
B	29.07	(9.11, 148.51)	<0.001
D	0.37	(0.12, 0.93)	0.720
F	0.43	(0.21, 0.83)	0.226
<b>OmpH1</b>			
A	9.09	(3.35, 34.59)	<0.001
B	0.00	(0.00, 0.38)	0.008
D	0.14	(0.00, 0.87)	0.832
F	0.22	(0.04, 0.71)	0.119
<b>OmpH3</b>			
A	2.56	(1.72, 3.83)	<0.001
B	0.00	(0.00, 0.06)	<0.001
D	0.73	(0.34, 1.58)	1.00
F	1.30	(0.73, 2.36)	1.00
<b>PtfA</b>			
A	0.09	(0.03, 0.26)	<0.001
B	Inf	(2.98, inf)	0.003
D	3.85	(0.96, 33.61)	1.00
F	7.90	(2.04, 67.76)	0.007
<b>PlpE</b>			
A	4.25	(2.83, 6.42)	<0.001
B	0.02	(0.00, 0.06)	<0.001
D	0.15	(0.06, 0.35)	0.001
F	1.77	(0.92, 3.65)	1.00
<b>PlpP</b>			
A	0.75	(0.49, 1.15)	1.00
B	0.00	(0.00, 0.21)	0.00001
D	0.00	(0.00, 0.29)	0.001
F	8.96	(4.84, 17.28)	<0.001
<b>PfhB1</b>			
A	3.64	(2.41, 5.52)	<0.001
B	0.01	(0.00, 0.05)	<0.001
D	2.71	(0.93, 10.80)	1.00
F	0.68	(0.38, 1.25)	1.00

<https://doi.org/10.1371/journal.pone.0249138.t005>

broad host range of isolates in this study, there is little evidence for host predilection within *P. multocida*. This is apparent in MLST results, as well as the core SNV-based phylogenetic tree, as isolates from different hosts were found to belong to the same sequence type, and host types were distributed throughout the phylogenetic tree. This finding is similar to the results of other genotypic studies of *P. multocida* that compared different host types, but only assessed the seven MLST genes [24,37,45]. There does seem to be an exception, however, with isolates from bovine hosts. Previous studies have demonstrated a clonal population structure of *P. multocida* isolates associated with hemorrhagic septicemia in cattle [38,58,59]. Although there was limited information available on disease status for the isolates in this study, it is evident that almost all of the 155 isolates from bovine sources fell into specific phylogenetic clades that

were clustered together on the tree. Interestingly, this specificity is demonstrated further when looking at the individual bovine hosts, as 28 of the 30 isolates from buffalo belong to ST122, with collection dates ranging from 2003 to 2016. It should be acknowledged, however, that there are two different MLST schemes for *P. multocida*, limiting direct comparisons between studies that used different schemes. Additionally, the lack of host-specificity identified in this study may be limited by the data available in the NCBI SRA and Assembly databases. For example, in this study, ST20 genomes were exclusively from chicken isolates, but this sequence type has also been identified in a *P. multocida* isolate from a pig in Australia [24].

A lack of host-specificity was also noted within *P. multocida* capsular serogroups, as all but three host categories listed in Table 2 had genomes fall into multiple capsular serogroups. Four different capsular serogroups were identified in isolates of avian and swine origin, and three different capsular serogroups were identified in isolates of bovine and rabbit origin. This is contradictory with some reports in the literature that suggest *P. multocida* serogroups display host predilection [2,60–62]. There was also no evidence of a geographic distribution of specific capsular serogroups. Previous studies have claimed that certain geographic regions have increased prevalence certain serogroups, however, these studies were limited in scope and sample size [2,63,64]. The results of this study indicate that, outside of bovine sources, there is an overall lack of host specificity in relation to capsular serogroup and sequence type, as well as a lack of geographical associations with certain capsular serogroups.

Genomes of different capsular serogroups were dispersed throughout the phylogenetic tree. While genomes within particular clades mostly belonged to the same capsular serogroup, the overall distribution of those genomes on the tree indicates that differences in capsular serogroups may not always correspond to differences at the whole-genome level. One possible explanation could be homologous recombination of the capsule genes. During the SNV analysis, there were 82 recombination events detected in the region of the genome containing the capsule genes that were removed from the SNV alignment. Therefore, SNVs identified in the capsule genes that determine serogroup were not incorporated into the resulting phylogenetic tree. While this study did not include somatic serotyping, it has been shown that somatic serotyping is also a poor indicator of the genetic relatedness of *P. multocida* isolates [65].

Specific virulence genes demonstrated significant associations with particular capsular serogroups. For example, while HgbA was found in genomes of all serotypes, the odds of HgbA being present in capsular serogroup B genomes was 29 times higher compared to all other capsular serogroups. This contrasts with previous studies that found HgbA is almost always present among *P. multocida* isolates regardless of capsular serogroup or host [66–70]. PtfA, a Type IV fimbriae, was also associated with capsular serogroup B, and has previously shown to be an important factor in disease outcomes in cattle [67]. This protein was also significantly associated with capsular serogroup F. OmpH1, an outer membrane protein, was not identified in any genomes from capsular serogroup B, and was significantly associated with capsular serogroup A. However, a previous study has indicated that the *ompH1* genes of *P. multocida* have undergone extensive horizontal gene transfer, as well as intragenic and assortative recombination [71]. In this study, it appears that different alleles of the gene encoding OmpH1 exist in other capsular serogroups with enough sequence variation that they did not meet the criteria for containing OmpH1, as many genomes had blast identity scores in the 70–90% range. OmpH3 was not identified in any capsular serogroup B genomes and was significantly associated with capsular serogroup A. The protective outer membrane lipoprotein PlpE was also significantly associated with capsular serogroup A. However, PlpP was significantly associated with capsular serogroup F. In fact, capsular serogroup F was the only capsular serogroup with an odds ratio greater than one. These findings warrant further investigation into the diversity and prevalence of OmpH and Plp proteins in *P. multocida*, as these surface antigens are often

used in vaccine studies that attempt to elicit protective immunity against *P. multocida* infection [72–76].

The SNV analysis on a subset of isolates involved in fowl cholera outbreaks from a single turkey producer revealed that multiple *P. multocida* strains may be identified in a single outbreak, which also has implications for vaccination. There are live-attenuated vaccines available for the prevention of fowl cholera in turkeys, as well as adjuvant or autogenous bacterins. The choice of vaccine type is typically dependent on bird age, stage of production, and which strains are currently circulating in the flocks of interest. Traditional *P. multocida* serotype designations are very broad, whether it be the Carter capsular serogroups, Heddleston somatic serovars, or the combination of both, so serological data alone is often not enough to distinguish specific strains to use for vaccination. This emphasizes the importance of more discriminatory methods, such as whole-genome sequencing, when making decisions about vaccine strategy. However, whole-genome sequencing may not be readily available or economically feasible in many cases. Alternative typing approaches have been developed in recent years, including PFGE and REP-PCR, which may provide a more granular approach to classifying *P. multocida* isolates [77–79].

This study is the first to provide genotypic evidence for differentiation between *P. multocida* subspecies with the identification of specific mutations in the *srlB* gene in subsp. *septica* genomes. While these mutations could distinguish subsp. *septica* from *multocida* and *gallicida*, subsp. *multocida* and *gallicida* could not be distinguished from one another through the genomic approaches utilized in this study. A previous study suggested the possibility of two distinct lineages of *P. multocida* based on subspecies: subsp. *multocida* and *gallicida* as one, and subsp. *septica* as the other [39]. However, in the present study, different subspecies, including *septica*, were identified as the same sequence type, such as with ST9 and ST129, which included both *multocida* and *septica* genomes. Further, genomes of the same subspecies were found throughout the phylogenetic tree instead of in clusters as one would expect of distinct lineages. This needs further investigation, however, as this study could readily distinguish *septica* genomes with SNVs in the *srlB* gene but were unable to distinguish subsp. *multocida* and *gallicida* using similar methodology. These findings also call into question the overall value of the information provided by subspecies distinction, as different hosts, capsular serogroups, and sequence types have been found within the same subspecies, and subspecies did not correspond to whole-genome differences in this study. Although subspecies differentiation may have provided early insight into the microbiology of *P. multocida* before other typing methods were developed, the identification of two sugar utilization systems has limited application in the context of genomic or epidemiological studies.

Several mobile genetic elements were identified within the *P. multocida* genomes in this study, some of which were highly similar to those from other bacterial species. This was particularly evident with the IncQ1 plasmids identified, as almost identical plasmids were found in strains of *Salmonella* Typhimurium, *Salmonella* Corvallis and *Shigella flexneri*. The match with the highest percent identity to *P. multocida* IncQ1 plasmids was a plasmid of *Salmonella* Typhimurium U288 (CP004059.1), a significant pathogen of pigs in the United Kingdom [80]. Other similar IncQ1 plasmids were from a human *Salmonella* Typhimurium (KU852461.1) clinical isolate from Southern Italy, a human *Shigella flexneri* (CP030774.1) clinical isolate from Bangladesh, and a *Salmonella* Corvallis isolate (CP044201.1) from an unknown source in the United States [81]. In this study, the *P. multocida* genomes containing the entire IncQ1 plasmid were from poultry isolates in the United States, demonstrating that the dissemination of IncQ1 plasmids has not been hindered by host or geography. This is notably concerning due to the presence of antimicrobial resistance genes, as well as the potential to integrate into the bacterial chromosome. Additionally, the identification of this plasmid in isolates from

poultry is also concerning given that this IncQ1 plasmid contains *sul2*, *tetA*, and *tetR*, conferring resistance to sulfonamide and tetracycline drugs, both of which have been used for prevention and treatment of fowl cholera in poultry in the United States [82].

The identification of antimicrobial resistance genes, virulence genes, and mobile genetic elements in this study provide important epidemiological context to the various typing methods often used in *P. multocida* diagnostics and research. The use of whole-genome sequence data has enhanced the overall understanding of the relationships, or lack of relationships, between host, geography, specific genotypes, and phylogeny in *P. multocida*. This study relied on sequences and metadata available in NCBI databases, which allowed for the large sample size necessary to make appropriate statistical associations in most cases, but was not without limitations. Information on host, geographic location, and collection date were not available for all genomes, and only a limited number of genomes had information on capsular serogroup or subspecies. Additionally, there were several *P. multocida* genomes with low coverage of the *cap* genes used to distinguish capsular serogroups. A larger set of *P. multocida* genomes of known capsular serogroups are needed to explore this further and validate the *in-silico* methodology developed in this study. It was assumed that all of the information that accompanied the genome sequences in the NCBI databases was correct. The overrepresentation of isolates of avian origin, as well as capsular serogroup A genomes, and isolates from the United States, should be acknowledged. The addition of 125 genomes from the University of Minnesota Midwest Central Research and Outreach Center contributed to this overrepresentation, as did batch uploads within the NCBI databases that allowed for the submission of multiple sequences at the same time by a single user. Therefore, the genomic diversity demonstrated in this study may actually underrepresent the true genomic diversity of *P. multocida*. However, this is the nature of large-scale bioinformatic studies and does not undermine the considerable insight gained from using publicly available data. This study expanded the current understanding of *P. multocida* molecular epidemiology through comparative genomics of 656 whole-genome sequences that varied in host and geographic origins. The genomic landscape of *P. multocida* demonstrates ample genetic diversity and has elucidated patterns in the potential for virulence and antimicrobial resistance that can be used to inform researchers and clinicians about factors that may influence the course of *P. multocida* infections.

## Methods

### Genome download, assembly, and multi-locus sequence typing

All available *P. multocida* whole-genome sequences and associated metadata were downloaded from the NCBI Short Read Archive and NCBI Assembly databases using the SRA toolkit v. 2.9.1 and NCBI-genome-download v. 0.2.7 [83,84]. These genomes were combined with a set of 125 *P. multocida* whole-genome sequences from the University of Minnesota Mid-Central Research and Outreach Center (Willmar, MN, USA) (BioProject PRJNA666906). The *P. multocida* genomes generated during this study came from clinical isolates of fowl cholera from turkeys within a single turkey production company. DNA was extracted using the Qiagen DNeasy blood and tissue kit (Valencia, CA). The genomic library was produced using the Nextera XT library prep kit and Nextera XT index kit v. 2 (Illumina, San Diego, CA) according to the manufacturer's instructions. Sequencing was performed on an Illumina MiSeq system using a 2 x 250 dual-index approach. Raw Illumina reads were trimmed of adaptor content and filtered for quality using Trimmomatic v. 0.33 using default settings [85]. For genomes that were already assembled from the NCBI Assembly database, raw Illumina reads were simulated using Wgsim v. 1.0.2 and then trimmed using the same methods (<https://github.com/lh3/wgsim>). All genomes were assembled with Spades v. 3.10.0 using default settings [86].

Genome assemblies were assessed for quality using QUAST v. 4.3 with reference genome ATCC 43137 (GCA\_000754275.1) [87]. Genomes were excluded from subsequent analysis if the genome fraction (i.e. proportion of aligned bases from the reference) was less than 80%, which would indicate that the genome may be of a species other than *P. multocida*. Genomes were also excluded if the number of unaligned bases in the assembly was greater than 2 million, which would be indicative of possible contamination by a species other than *P. multocida*. Multi-locus sequence typing was performed on the remaining *P. multocida* genomes using the RIRDC scheme available in the PubMLST database (<https://github.com/tseemann/mlst>) [32,33,88]. This scheme was selected because it contains a larger number of sequence types that would ideally reduce the number of non-typable genomes in this study.

### ***In-silico* identification of *P. multocida* capsular serogroups**

There are genes unique to specific capsular serogroups (*hyaD*, *bcbD*, *dcbF*, *ecbJ*, and *fcuD*) within the capsular biosynthetic loci of *P. multocida*. These genes have previously been used in a multiplex PCR assay that was successful in distinguishing *P. multocida* capsular serogroups [89]. The nucleotide sequences for these genes unique to capsular serogroups A (AF067175), B (AF169324), D (AF302465), E (AF302466), and F (AF302467) were downloaded from GenBank and used to construct a custom blast database using BLAST+ v. 2.8.1 [90,91]. This database was used in a nucleotide blast of all genome assemblies in order to identify top hits to the genes encoding the capsular antigens. Hits for each genome with a bitscore of less than 1,000 were removed. Of the remaining top hits, the one with the highest percent identity was selected as the capsular serogroup for that genome. This methodology was validated by comparing the *in-silico* capsular serogroup results to the capsular serogroup information available in the metadata downloaded from NCBI. All custom databases generated in this study are available on GitHub (<https://github.com/JohnsonSingerLab/Pmult-databases>).

### **Single nucleotide variant (SNV) analysis**

Snippy v. 4.6.0 was used to call core SNVs from the trimmed reads of each genome and generate a core SNV alignment using default parameters with reference genome ATCC 43137 (GCA\_000754275.1) [92]. Gubbins v. 2.3.4 was used to remove SNVs found in recombinant regions from the alignment [93]. The ModelFinder tool within IQ-TREE was used to identify the optimal model of DNA substitution for estimating maximum-likelihood phylogeny [94,95]. A phylogenetic tree was constructed using the selected model of DNA substitution with a correction for constant sites determined using maskrc-svg (<https://github.com/kwongji/maskrc-svg>) and snp-sites [96]. An Ultrafast bootstrap approximation with 1,000 replications was performed [97]. The resulting tree was annotated using iTOL [98].

### ***P. multocida* pangenome and differences in gene content**

Genomes were annotated using Prokka v. 1.13, and Roary v. 3.10.2 was used to construct the *P. multocida* pangenome using default parameters [99,100]. Scoary v. 1.6.16 was used to identify significant differences in gene content between phylogenetic clades with Fisher's exact tests and the Benjamini-Hochburg correction to control the false discovery rate [101,102]. Odds ratios and *p*-values are reported, with a *p*-value of  $\leq 0.05$  considered significant.

The pangenome was also used to search for genes that may be associated with the ability to ferment sorbitol and dulcitol in order to identify any associations with *P. multocida* subspecies. Subspecies information was only available for a limited number of the *P. multocida* genomes in this study ( $n = 68$ ). SNVs were identified using the same methods described previously, with the reference genome consisting of the gene sequences extracted from the pangenome

reference file. SNVs within each alignment were assessed for missense or nonsense mutations that could potentially alter the function of the gene.

### Identification of antimicrobial resistance genes, mobile genetic elements, and virulence factors

Antimicrobial resistance genes were identified from the genome assemblies using Abricate v. 0.9.8 with the NCBI AMRFinder database [103,104]. Genes with coverage of less than 50% were excluded. Plasmid replicons were identified from the genome assemblies using Abricate v. 0.9.8 with the PlasmidFinder database [56,103]. When specific replicons were identified, a nucleotide blast search was performed to determine if the whole plasmid was present. Plasmid sequences were circularized, oriented to the same start site, aligned with Mauve, and visualized using SnapGene Viewer (from GSL biotech; available at [snapgene.com](http://snapgene.com)) [105].

To identify genomes with the known integrative conjugative element associated with *P. multocida*, ICEPmu1, a blast search was conducted on genome assemblies using the ICEPmu1 reference sequence from ICEberg 2.0 [57,91,106]. Genomes that aligned to ICEPmu1 with at least 50% coverage and 90% identity were considered to contain ICEPmu1. This included genomes with alignments on multiple contigs, provided that 50% of the reference sequence was still covered.

A blast search was conducted for a set of virulence factors that differ in prevalence among virulent strains of *P. multocida* as determined in a prior study [68,91]. The annotated amino acid sequences of each genome were searched for these virulence-associated proteins. Sequences producing alignments with greater than 50% coverage and 90% identity were considered to contain the virulence factor. To test whether these virulence factors were associated with specific capsular serogroups, Fisher's exact tests with the Benjamini-Hochburg correction were performed in R v. 1.2.5042 [102,107]. Odds ratios, confidence intervals, and *p*-values are reported.

## Supporting information

### S1 Dataset.

(XLSX)

## Acknowledgments

We thank the turkey producers of the United States for their willingness to collaborate in this study. Bioinformatics was supported using tools available from the Minnesota Supercomputing Institute. Sequencing reagents for this study were donated by the Mid-Central Research and Outreach Center, Willmar, MN, USA.

## Author Contributions

**Conceptualization:** Emily Smith, Timothy Johnson.

**Data curation:** Emily Smith, Jill Nezworski, Marissa Studniski, Ben Wileman.

**Formal analysis:** Emily Smith.

**Funding acquisition:** Timothy Johnson.

**Investigation:** Emily Smith, Timothy Johnson.

**Methodology:** Emily Smith, Elizabeth Miller, Jeannette Munoz Aguayo, Cristian Flores Figueroa, Timothy Johnson.



**Project administration:** Emily Smith.

**Supervision:** Timothy Johnson.

**Validation:** Emily Smith.

**Visualization:** Emily Smith.

**Writing – original draft:** Emily Smith.

**Writing – review & editing:** Emily Smith, Elizabeth Miller, Jeannette Munoz Aguayo, Cristian Flores Figueroa, Jill Nezworski, Marissa Studniski, Ben Wileman, Timothy Johnson.

## References

1. Harper M, Boyce JD, Adler B. *Pasteurella multocida* pathogenesis: 125 years after Pasteur. FEMS microbiology letters. 2006; 265(1):1–10. <https://doi.org/10.1111/j.1574-6968.2006.00442.x> PMID: [17107417](https://pubmed.ncbi.nlm.nih.gov/17107417/)
2. Wilson BA, Ho M. *Pasteurella multocida*: from zoonosis to cellular microbiology. Clinical microbiology reviews. 2013; 26(3):631–55. <https://doi.org/10.1128/CMR.00024-13> PMID: [23824375](https://pubmed.ncbi.nlm.nih.gov/23824375/)
3. Adler AC, Cestero C, Brown RB. Septic shock from *Pasteurella multocida* following a cat bite: case report and review of literature. Connecticut medicine. 2011; 75(10):603. PMID: [22216675](https://pubmed.ncbi.nlm.nih.gov/22216675/)
4. Heydemann J, Heydemann JS, Antony S. Acute infection of a total knee arthroplasty caused by *Pasteurella multocida*: a case report and a comprehensive review of the literature in the last 10 years. International Journal of Infectious Diseases. 2010; 14:e242–e5. <https://doi.org/10.1016/j.ijid.2009.09.007> PMID: [20116315](https://pubmed.ncbi.nlm.nih.gov/20116315/)
5. Jarvis WR, Banko S, Snyder E, Baltimore RS. *Pasteurella multocida* osteomyelitis following dog bites. American Journal of Diseases of Children. 1981; 135(7):625–7. <https://doi.org/10.1001/archpedi.1981.02130310031011> PMID: [7246490](https://pubmed.ncbi.nlm.nih.gov/7246490/)
6. Weber DJ, Wolfson JS, Swartz MN, Hooper DC. *Pasteurella multocida* infections. Report of 34 cases and review of the literature. Medicine. 1984; 63(3):133–54. PMID: [6371440](https://pubmed.ncbi.nlm.nih.gov/6371440/)
7. Carpenter T, Snipes K, Kasten R, Hird D, Hirsh D. Molecular epidemiology of *Pasteurella multocida* in turkeys. American journal of veterinary research. 1991; 52(8):1345–9. PMID: [1718191](https://pubmed.ncbi.nlm.nih.gov/1718191/)
8. McFadden A, Christensen H, Fairley R, Hill F, Gill J, Keeling S, et al. Outbreaks of pleuritis and peritonitis in calves associated with *Pasteurella multocida* capsular type B strain. New Zealand Veterinary Journal. 2011; 59(1):40–5. <https://doi.org/10.1080/00480169.2011.547168> PMID: [21328156](https://pubmed.ncbi.nlm.nih.gov/21328156/)
9. Marruchella G, Di Francesco CE, Trachtman AR, Mosca F, Sebastiani C, Di Provvio A, et al. Severe outbreak of pasteurellosis in sows: a case description. Large Animal Review. 2019; 25(3):111–3.
10. Sudaryatna PE, Nakamura K, Mekata H, Sekiguchi S, Kubo M, Kobayashi I, et al. Bovine respiratory syncytial virus infection enhances *Pasteurella multocida* adherence on respiratory epithelial cells. Veterinary microbiology. 2018; 220:33–8. <https://doi.org/10.1016/j.vetmic.2018.04.031> PMID: [29885798](https://pubmed.ncbi.nlm.nih.gov/29885798/)
11. Ross RF. *Pasteurella multocida* and its role in porcine pneumonia. Animal Health Research Reviews. 2006; 7(1–2):13. <https://doi.org/10.1017/S1466252307001211> PMID: [17389051](https://pubmed.ncbi.nlm.nih.gov/17389051/)
12. Dahl C, Permin A, Christensen J, Bisgaard M, Muhairwa A, Petersen KD, et al. The effect of concurrent infections with *Pasteurella multocida* and *Ascaridia galli* on free range chickens. Veterinary microbiology. 2002; 86(4):313–24. [https://doi.org/10.1016/s0378-1135\(02\)00015-9](https://doi.org/10.1016/s0378-1135(02)00015-9) PMID: [11955781](https://pubmed.ncbi.nlm.nih.gov/11955781/)
13. Perelman B, Hadash D, Meroz M, Gur-Lavie A, Abramson M, Samberg Y. Vaccination of young turkeys against fowl cholera. Avian pathology. 1990; 19(1):131–7. <https://doi.org/10.1080/03079459008418662> PMID: [18679920](https://pubmed.ncbi.nlm.nih.gov/18679920/)
14. Mbuthia PG, Njagi LW, Nyaga PN, Beboru L, Minga U, Kamundia J, et al. *Pasteurella multocida* in scavenging family chickens and ducks: carrier status, age susceptibility and transmission between species. Avian Pathology. 2008; 37(1):51–7. <https://doi.org/10.1080/03079450701784891> PMID: [18202950](https://pubmed.ncbi.nlm.nih.gov/18202950/)
15. Carter G. A haemagglutination test for the identification of serological types. Am J Vet Res. 1955; 16:481–4. PMID: [13238757](https://pubmed.ncbi.nlm.nih.gov/13238757/)
16. Carter G. A new serological type of *Pasteurella multocida* from Central Africa. Vet Rec. 1961; 73(1,052).

17. Boyce JD, Chung JY, Adler B. *Pasteurella multocida* capsule: composition, function and genetics. *Journal of biotechnology*. 2000; 83(1–2):153–60. [https://doi.org/10.1016/s0168-1656\(00\)00309-6](https://doi.org/10.1016/s0168-1656(00)00309-6) PMID: [11000471](#)
18. Chung JY, Wilkie I, Boyce JD, Townsend KM, Frost AJ, Ghoddusi M, et al. Role of capsule in the pathogenesis of fowl cholera caused by *Pasteurella multocida* serogroup A. *Infection and immunity*. 2001; 69(4):2487–92. <https://doi.org/10.1128/IAI.69.4.2487-2492.2001> PMID: [11254611](#)
19. Boyce JD, Adler B. The capsule is a virulence determinant in the pathogenesis of *Pasteurella multocida* M1404 (B: 2). *Infection and immunity*. 2000; 68(6):3463–8. <https://doi.org/10.1128/iai.68.6.3463-3468.2000> PMID: [10816499](#)
20. Jacques M, Kobisch M, Belanger M, Dugal F. Virulence of capsulated and noncapsulated isolates of *Pasteurella multocida* and their adherence to porcine respiratory tract cells and mucus. *Infection and immunity*. 1993; 61(11):4785–92. <https://doi.org/10.1128/IAI.61.11.4785-4792.1993> PMID: [8406879](#)
21. Heddleston K, Gallagher J, Rebers P. Fowl cholera: gel diffusion precipitin test for serotyping *Pasteurella multocida* from avian species. *Avian diseases*. 1972:925–36. PMID: [4628021](#)
22. Rimler R, Rebers P, Phillips M. Lipopolysaccharides of the Heddleston serotypes of *Pasteurella multocida*. *American journal of veterinary research*. 1984; 45(4):759–63. PMID: [6731991](#)
23. Harper M, John M, Turni C, Edmunds M, Michael FS, Adler B, et al. Development of a rapid multiplex PCR assay to genotype *Pasteurella multocida* strains by use of the lipopolysaccharide outer core biosynthesis locus. *Journal of clinical microbiology*. 2015; 53(2):477–85. <https://doi.org/10.1128/JCM.02824-14> PMID: [25428149](#)
24. Turni C, Singh R, Blackall P. Genotypic diversity of *Pasteurella multocida* isolates from pigs and poultry in Australia. *Australian veterinary journal*. 2018; 96(10):390–4. <https://doi.org/10.1111/avj.12748> PMID: [30255575](#)
25. Lu Y, Pakes S, Stefanu C. Capsular and somatic serotypes of *Pasteurella multocida* isolates recovered from healthy and diseased rabbits in Texas. *Journal of clinical microbiology*. 1983; 18(2):292–5. <https://doi.org/10.1128/JCM.18.2.292-295.1983> PMID: [6619283](#)
26. Turni C, Dayao D, Aduriz G, Cortabarría N, Tejero C, Ibabe JC, et al. A *Pasteurella multocida* strain affecting nulliparous heifers and calves in different ways. *Veterinary microbiology*. 2016; 195:17–21. <https://doi.org/10.1016/j.vetmic.2016.08.022> PMID: [27771064](#)
27. Heddleston KL, Gallagher JE, Rebers PA. Fowl Cholera: Gel Diffusion Precipitin Test for Serotyping *Pasteurella multocida* from Avian Species. *Avian Diseases*. 1972; 16(4):925–36. <https://doi.org/10.2307/1588773> PMID: [4628021](#)
28. Gerardo SH, Citron DM, Claros MC, Fernandez HT, Goldstein EJ. *Pasteurella multocida* subsp. *multocida* and *P. multocida* subsp. *septica* differentiation by PCR fingerprinting and  $\alpha$ -glucosidase activity. *Journal of clinical microbiology*. 2001; 39(7):2558–64. <https://doi.org/10.1128/JCM.39.7.2558-2564.2001> PMID: [11427568](#)
29. Biberstein EL, Jang SS, Kass PH, Hirsh DC. Distribution of indole-producing urease-negative pasteurellas in animals. *Journal of Veterinary Diagnostic Investigation*. 1991; 3(4):319–23. <https://doi.org/10.1177/104063879100300408> PMID: [1760464](#)
30. Holst E, Rollf J, Larsson L, Nielsen J. Characterization and distribution of *Pasteurella* species recovered from infected humans. *Journal of Clinical Microbiology*. 1992; 30(11):2984–7. <https://doi.org/10.1128/JCM.30.11.2984-2987.1992> PMID: [1452670](#)
31. Kuhnert P, Boerlin P, Emler S, Krawinkler M, Frey J. Phylogenetic analysis of *Pasteurella multocida* subspecies and molecular identification of feline *P. multocida* subsp. *septica* by 16S rRNA gene sequencing. *International journal of medical microbiology*. 2000; 290(7):599–604. [https://doi.org/10.1016/S1438-4221\(00\)80008-9](https://doi.org/10.1016/S1438-4221(00)80008-9) PMID: [11200541](#)
32. Jolley KA, Maiden MC. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC bioinformatics*. 2010; 11(1):1–11. <https://doi.org/10.1186/1471-2105-11-595> PMID: [21143983](#)
33. Subaaharan S, Blackall L, Blackall P. Development of a multi-locus sequence typing scheme for avian isolates of *Pasteurella multocida*. *Veterinary microbiology*. 2010; 141(3–4):354–61. <https://doi.org/10.1016/j.vetmic.2010.01.017> PMID: [20129742](#)
34. *Pasteurella multocida* Multi-host MLST database. Available from: [http://pubmlst.org/pmultocida\\_multihost/](http://pubmlst.org/pmultocida_multihost/).
35. *Pasteurella multocida* RIRDC MLST database. Available from: <https://pubmlst.org/organisms/pasteurella-multocida>.
36. García-Alvarez A, Vela AI, San Martín E, Chaves F, Fernández-Garayzábal JF, Lucas D, et al. Characterization of *Pasteurella multocida* associated with ovine pneumonia using multi-locus sequence typing (MLST) and virulence-associated gene profile analysis and comparison with porcine isolates.

- Veterinary microbiology. 2017; 204:180–7. <https://doi.org/10.1016/j.vetmic.2017.04.015> PMID: 28532799
37. Köndgen S, Leider M, Lankester F, Bethe A, Lübke-Becker A, Leendertz FH, et al. *Pasteurella multocida* involved in respiratory disease of wild chimpanzees. PLoS one. 2011; 6(9). <https://doi.org/10.1371/journal.pone.0024236> PMID: 21931664
  38. Hotchkiss EJ, Hodgson JC, Lainson FA, Zadoks RN. Multilocus sequence typing of a global collection of *Pasteurella multocida* isolates from cattle and other host species demonstrates niche association. BMC microbiology. 2011; 11(1):115. <https://doi.org/10.1186/1471-2180-11-115> PMID: 21612618
  39. Bisgaard M, Petersen A, Christensen H. Multilocus sequence analysis of *Pasteurella multocida* demonstrates a type species under development. Microbiology. 2013; 159(3):580–90. <https://doi.org/10.1099/mic.0.063461-0> PMID: 23329677
  40. Bertelsen MF, Bojesen AM, Bisgaard M, Petersen A, Christensen H. *Pasteurella multocida* carriage in red-necked wallabies (*Macropus rufogriseus*). Journal of Zoo and Wildlife Medicine. 2012; 43(4):726–9. <https://doi.org/10.1638/2011-0051R.1> PMID: 23272337
  41. Moustafa AM, Bennett MD, Edwards J, Azim K, Msaik MA, Choudhary MI, et al. Molecular typing of haemorrhagic septicaemia-associated *Pasteurella multocida* isolates from Pakistan and Thailand using multilocus sequence typing and pulsed-field gel electrophoresis. Research in veterinary science. 2013; 95(3):986–90. <https://doi.org/10.1016/j.rvsc.2013.07.003> PMID: 23916592
  42. Zhu W, Fan Z, Qiu R, Chen L, Wei H, Hu B, et al. Characterization of *Pasteurella multocida* isolates from rabbits in China. Veterinary Microbiology. 2020:108649. <https://doi.org/10.1016/j.vetmic.2020.108649> PMID: 32402342
  43. Peng Z, Wang H, Liang W, Chen Y, Tang X, Chen H, et al. A capsule/lipopolysaccharide/MLST genotype D/L6/ST11 of *Pasteurella multocida* is likely to be strongly associated with swine respiratory disease in China. Archives of microbiology. 2018; 200(1):107–18. <https://doi.org/10.1007/s00203-017-1421-y> PMID: 28825122
  44. Oh Y-H, Moon D-C, Lee YJ, Hyun B-H, Lim S-K. Genetic and phenotypic characterization of tetracycline-resistant *Pasteurella multocida* isolated from pigs. Veterinary microbiology. 2019; 233:159–63. <https://doi.org/10.1016/j.vetmic.2019.05.001> PMID: 31176403
  45. Massacci FR, Magistrali CF, Cucco L, Curcio L, Bano L, Mangili P, et al. Characterization of *Pasteurella multocida* involved in rabbit infections. Veterinary microbiology. 2018; 213:66–72. <https://doi.org/10.1016/j.vetmic.2017.11.023> PMID: 29292006
  46. Kehrenberg C, Schwarz S. Plasmid-borne florfenicol resistance in *Pasteurella multocida*. Journal of antimicrobial chemotherapy. 2005; 55(5):773–5. <https://doi.org/10.1093/jac/dki102> PMID: 15814600
  47. San Millan A, Escudero JA, Gutierrez B, Hidalgo L, Garcia N, Llagostera M, et al. Multiresistance in *Pasteurella multocida* is mediated by coexistence of small plasmids. Antimicrobial agents and chemotherapy. 2009; 53(8):3399–404. <https://doi.org/10.1128/AAC.01522-08> PMID: 19528282
  48. Rosenau A, Labigne A, Escande F, Courcoux P, Philippon A. Plasmid-mediated ROB-1 beta-lactamase in *Pasteurella multocida* from a human specimen. Antimicrobial agents and chemotherapy. 1991; 35(11):2419–22. <https://doi.org/10.1128/aac.35.11.2419> PMID: 1804018
  49. Pastorello I, Paccani SR, Rosini R, Mattera R, Navarro MF, Urosev D, et al. EsiB, a novel pathogenic *Escherichia coli* secretory immunoglobulin A-binding protein impairing neutrophil activation. MBio. 2013; 4(4):e00206–13. <https://doi.org/10.1128/mBio.00206-13> PMID: 23882011
  50. Webb JS, Nikolakakis KC, Willett JL, Aoki SK, Hayes CS, Low DA. Delivery of CdiA nuclease toxins into target cells during contact-dependent growth inhibition. PLoS one. 2013; 8(2). <https://doi.org/10.1371/journal.pone.0057609> PMID: 23469034
  51. Bera A, Herbert S, Jakob A, Vollmer W, Götz F. Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of *Staphylococcus aureus*. Molecular microbiology. 2005; 55(3):778–87. <https://doi.org/10.1111/j.1365-2958.2004.04446.x> PMID: 15661003
  52. Lacour S, Doublet P, Obadia B, Cozzone AJ, Grangeasse C. A novel role for protein-tyrosine kinase Etk from *Escherichia coli* K-12 related to polymyxin resistance. Research in Microbiology. 2006; 157(7):637–41. <https://doi.org/10.1016/j.resmic.2006.01.003> PMID: 16814990
  53. Botros S, Mitchell D, Van Ommen C. Deletion of the *Escherichia coli* K30 group I capsule biosynthesis genes *wza*, *wzb* and *wzc* confers capsule-independent resistance to macrolide antibiotics. J Exp Microbiol Immunol. 2015; 19.
  54. McENTEE K. Genetic analysis of the *Escherichia coli* K-12 srl region. J Bacteriol. 1977; 132(3):904–11. <https://doi.org/10.1128/JB.132.3.904-911.1977> PMID: 336611

55. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, et al. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Research*. 2016; 45(D1):D543–D50. <https://doi.org/10.1093/nar/gkw1003> PMID: 27899573
56. Carattoli A, Zankari E, García-Fernández A, Larsen MV, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial agents and chemotherapy*. 2014; 58(7):3895–903. <https://doi.org/10.1128/AAC.02412-14> PMID: 24777092
57. Liu M, Li X, Xie Y, Bi D, Sun J, Li J, et al. ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic acids research*. 2019; 47(D1):D660–D5. <https://doi.org/10.1093/nar/gky1123> PMID: 30407568
58. Petersen A, Bisgaard M, Townsend K, Christensen H. MLST typing of *Pasteurella multocida* associated with haemorrhagic septicaemia and development of a real-time PCR specific for haemorrhagic septicaemia associated isolates. *Veterinary microbiology*. 2014; 170(3–4):335–41. <https://doi.org/10.1016/j.vetmic.2014.02.022> PMID: 24636905
59. Mir R, Thomas P, Viswas K, Gupta S, Verma J, Singh A, et al. Multilocus sequence typing of Indian isolates of *Pasteurella multocida*. *Indian Journal of Comparative Microbiology, Immunology and Infectious Diseases*. 2011; 32(1and2):30–5.
60. Wilkie IW, Harper M, Boyce JD, Adler B. *Pasteurella multocida*: diseases and pathogenesis. *Pasteurella multocida*: Springer; 2012. p. 1–22.
61. Kubatzky KF. *Pasteurella multocida* and immune cells. *Pasteurella multocida*: Springer; 2012. p. 53–72.
62. Peng Z, Wang X, Zhou R, Chen H, Wilson BA, Wu B. *Pasteurella multocida*: genotypes and genomics. *Microbiology and Molecular Biology Reviews*. 2019; 83(4):e00014–19. <https://doi.org/10.1128/MMBR.00014-19> PMID: 31484691
63. Tang X, Zhao Z, Hu J, Wu B, Cai X, He Q, et al. Isolation, antimicrobial resistance, and virulence genes of *Pasteurella multocida* strains from swine in China. *Journal of clinical microbiology*. 2009; 47(4):951–8. <https://doi.org/10.1128/JCM.02029-08> PMID: 19158260
64. Furian TQ, Borges KA, Laviniki V, Rocha SLdS, Almeida CNd, Nascimento VPd, et al. Virulence genes and antimicrobial resistance of *Pasteurella multocida* isolated from poultry and swine. *Brazilian journal of microbiology*. 2016; 47(1):210–6. <https://doi.org/10.1016/j.bjm.2015.11.014> PMID: 26887247
65. Amonsin A, Wellehan JF, Li L-L, Laber J, Kapur V. DNA fingerprinting of *Pasteurella multocida* recovered from avian sources. *Journal of Clinical Microbiology*. 2002; 40(8):3025–31. <https://doi.org/10.1128/jcm.40.8.3025-3031.2002> PMID: 12149370
66. Sarangi LN, Thomas P, Gupta S, Priyadarshini A, Kumar S, Nagaleekar VK, et al. Virulence gene profiling and antibiotic resistance pattern of Indian isolates of *Pasteurella multocida* of small ruminant origin. *Comparative Immunology, Microbiology and Infectious Diseases*. 2015; 38:33–9. <https://doi.org/10.1016/j.cimid.2014.11.003> PMID: 25499527
67. Verma S, Sharma M, Katoch S, Verma L, Kumar S, Dogra V, et al. Profiling of virulence associated genes of *Pasteurella multocida* isolated from cattle. *Veterinary research communications*. 2013; 37(1):83–9. <https://doi.org/10.1007/s11259-012-9539-5> PMID: 23007877
68. Johnson TJ, Abrahante JE, Hunter SS, Hauglund M, Tatum FM, Maheswaran SK, et al. Comparative genome analysis of an avirulent and two virulent strains of avian *Pasteurella multocida* reveals candidate genes involved in fitness and pathogenicity. *BMC microbiology*. 2013; 13(1):106. <https://doi.org/10.1186/1471-2180-13-106> PMID: 23672515
69. Aski HS, Tabatabaei M. Occurrence of virulence-associated genes in *Pasteurella multocida* isolates obtained from different hosts. *Microbial pathogenesis*. 2016; 96:52–7. <https://doi.org/10.1016/j.micpath.2016.04.008> PMID: 27057674
70. Bosch M, Garrido ME, Llagostera M, de Rozas AMP, Badiola I, Barbé J. Characterization of the *Pasteurella multocida* *hgbA* gene encoding a hemoglobin-binding protein. *Infection and immunity*. 2002; 70(11):5955–64. <https://doi.org/10.1128/iai.70.11.5955-5964.2002> PMID: 12379670
71. Abdulrahman RF. Molecular evolution of outer membrane proteins and characterization of temperate bacteriophages of *Pasteurella multocida* strains from different host species: University of Glasgow; 2016.
72. Lee J, Kim YB, Kwon M. Outer membrane protein H for protective immunity against *Pasteurella multocida*. *The Journal of Microbiology*. 2007; 45(2):179–84. PMID: 17483806
73. Luo Y, Zeng Q, Glisson JR, Jackwood MW, Cheng I-HN, Wang C. Sequence analysis of *Pasteurella multocida* major outer membrane protein (OmpH) and application of synthetic peptides in vaccination of chickens against homologous strain challenge. *Vaccine*. 1999; 17(7–8):821–31. [https://doi.org/10.1016/S0264-410X\(98\)00266-7](https://doi.org/10.1016/S0264-410X(98)00266-7) PMID: 10067687

74. Gong Q, Qu N, Niu M, Qin C, Cheng M, Sun X, et al. Immune responses and protective efficacy of a novel DNA vaccine encoding outer membrane protein of avian *Pasteurella multocida*. *Veterinary immunology and immunopathology*. 2013; 152(3–4):317–24. <https://doi.org/10.1016/j.vetimm.2013.01.001> PMID: 23340446
75. Wu J-R, Shien J-H, Shieh HK, Chen C-F, Chang P-C. Protective immunity conferred by recombinant *Pasteurella multocida* lipoprotein E (PipE). *Vaccine*. 2007; 25(21):4140–8. <https://doi.org/10.1016/j.vaccine.2007.03.005> PMID: 17449151
76. Tabatabai LB, Zehr ES. Identification of five outer membrane-associated proteins among cross-protective factor proteins of *Pasteurella multocida*. *Infection and immunity*. 2004; 72(2):1195–8. <https://doi.org/10.1128/iai.72.2.1195-1198.2004> PMID: 14742575
77. Gunawardana GA, Townsend KM, Frost AJ. Molecular characterisation of avian *Pasteurella multocida* isolates from Australia and Vietnam by REP-PCR and PFGE. *Veterinary Microbiology*. 2000; 72(1–2):97–109. [https://doi.org/10.1016/s0378-1135\(99\)00191-1](https://doi.org/10.1016/s0378-1135(99)00191-1) PMID: 10699507
78. Saxena M, Singh V, Kumar A, Chaudhuri P, Singh VP, Shivachandra S, et al. REP-PCR analysis of *Pasteurella multocida* isolates from wild and domestic animals in India. *Veterinary research communications*. 2006; 30(8):851–61. <https://doi.org/10.1007/s11259-006-3321-5> PMID: 17139535
79. Versalovic J, Koeuth T, Lupski R. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic acids research*. 1991; 19(24):6823–31. <https://doi.org/10.1093/nar/19.24.6823> PMID: 1762913
80. Hooton SP, Timms AR, Cummings NJ, Moreton J, Wilson R, Connerton IF. The complete plasmid sequences of *Salmonella enterica* serovar Typhimurium U288. *Plasmid*. 2014; 76:32–9. <https://doi.org/10.1016/j.plasmid.2014.08.002> PMID: 25175817
81. Henry T, Couillault C, Rockenfeller P, Boucrot E, Dumont A, Schroeder N, et al. The *Salmonella* effector protein PipB2 is a linker for kinesin-1. *Proceedings of the National Academy of Sciences*. 2006; 103(36):13497–502. <https://doi.org/10.1073/pnas.0605443103> PMID: 16938850
82. Singer RS, Porter L. Estimates of On-Farm Antimicrobial Usage in Broiler Chicken and Turkey Production in the United States, 2013–2017.
83. Team STD. SRA-Tools. Available from: <http://ncbi.github.io/sra-tools/>.
84. Kai B. NCBI Genome Downloading Scripts. Available from: <https://github.com/kblin/ncbi-genome-download>.
85. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
86. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*. 2012; 19(5):455–77. <https://doi.org/10.1089/cmb.2012.0021> PMID: 22506599
87. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013; 29(8):1072–5. <https://doi.org/10.1093/bioinformatics/btt086> PMID: 23422339
88. Seemann T. MLST. Available from: <https://github.com/tseemann/mlst>.
89. Townsend KM, Boyce JD, Chung JY, Frost AJ, Adler B. Genetic Organization of *Pasteurella multocida* cap Loci and Development of a Multiplex Capsular PCR Typing System. *Journal of Clinical Microbiology*. 2001; 39(3):924–9. <https://doi.org/10.1128/JCM.39.3.924-929.2001> PMID: 11230405
90. Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. GenBank. *Nucleic acids research*. 2019; 47(D1):D94–D9. <https://doi.org/10.1093/nar/gky989> PMID: 30365038
91. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC bioinformatics*. 2009; 10(1):421. <https://doi.org/10.1186/1471-2105-10-421> PMID: 20003500
92. Seemann T. Rapid haploid variant calling and core genome alignment. Snippy. Available from: <https://github.com/tseemann/snippy>.
93. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic acids research*. 2015; 43(3):e15–e. <https://doi.org/10.1093/nar/gku1196> PMID: 25414349
94. Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*. 2017; 14(6):587. <https://doi.org/10.1038/nmeth.4285> PMID: 28481363
95. Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*. 2015; 32(1):268–74. <https://doi.org/10.1093/molbev/msu300> PMID: 25371430

96. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial genomics*. 2016; 2(4). <https://doi.org/10.1099/mgen.0.000056> PMID: [28348851](https://pubmed.ncbi.nlm.nih.gov/28348851/)
97. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution*. 2018; 35(2):518–22. <https://doi.org/10.1093/molbev/msx281> PMID: [29077904](https://pubmed.ncbi.nlm.nih.gov/29077904/)
98. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research*. 2019; 47(W1):W256–W9. <https://doi.org/10.1093/nar/gkz239> PMID: [30931475](https://pubmed.ncbi.nlm.nih.gov/30931475/)
99. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014; 30(14):2068–9. <https://doi.org/10.1093/bioinformatics/btu153> PMID: [24642063](https://pubmed.ncbi.nlm.nih.gov/24642063/)
100. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015; 31(22):3691–3. <https://doi.org/10.1093/bioinformatics/btv421> PMID: [26198102](https://pubmed.ncbi.nlm.nih.gov/26198102/)
101. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome biology*. 2016; 17(1):238. <https://doi.org/10.1186/s13059-016-1108-8> PMID: [27887642](https://pubmed.ncbi.nlm.nih.gov/27887642/)
102. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995; 57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
103. Seemann T. Abricate. Available from: <https://github.com/tseemann/abricate>.
104. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrobial agents and chemotherapy*. 2019; 63(11):e00483–19. <https://doi.org/10.1128/AAC.00483-19> PMID: [31427293](https://pubmed.ncbi.nlm.nih.gov/31427293/)
105. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome research*. 2004; 14(7):1394–403. <https://doi.org/10.1101/gr.2289704> PMID: [15231754](https://pubmed.ncbi.nlm.nih.gov/15231754/)
106. Michael GB, Kadlec K, Sweeney MT, Brzuszkiewicz E, Liesegang H, Daniel R, et al. ICE Pmu1, an integrative conjugative element (ICE) of *Pasteurella multocida*: analysis of the regions that comprise 12 antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*. 2012; 67(1):84–90.
107. Team R. RStudio: integrated development for R. RStudio, Inc, Boston, MA URL <http://www.rstudio.com>. 2015; 42:14.