MICROBIAL GENOMICS

RESEARCH ARTICLE

Omaleki *et al.*, *Microbial Genomics* 2022;8:000772

DOI 10.1099/mgen.0.000772

Phase variation in the glycosyltransferase genes of *Pasteurella multocida* associated with outbreaks of fowl cholera on free-range layer farms

Lida Omaleki[1,2,3,*], Patrick J. Blackall[1], Thom Cuddihy[4]†, Rhys T. White[2,3], Jodi M. Courtice[5], Conny Turni[1], Brian M. Forde[2,3,*],† and Scott A. Beatson[2,3,*]

Abstract

Fowl cholera caused by *Pasteurella multocida* has re-emerged in Australian poultry production since the increasing adoption of free-range production systems. Currently, autogenous killed whole-cell vaccines prepared from the isolates previously obtained from each farm are the main preventative measures used. In this study, we use whole-genome sequencing and phylogenomic analysis to investigate outbreak dynamics, as well as monitoring and comparing the variations in the lipopolysaccharide (LPS) outer core biosynthesis loci of the outbreak and vaccine strains. In total, 73 isolates from two different free-range layer farms were included. Our genomic analysis revealed that all investigated isolates within the two farms (layer A and layer B) carried LPS type L3, albeit with a high degree of genetic diversity between them. Additionally, the isolates belonged to five different sequence types (STs), with isolates belonging to ST9 and ST20 being the most prevalent. The isolates carried ST-specific mutations within their LPS type L3 outer core biosynthesis loci, including frameshift mutations in the outer core heptosyltransferase gene (*htpE*) (ST7 and ST274) or galactosyltransferase gene (*gatG*) (ST20). The ST9 isolates could be separated into three groups based on their LPS outer core biosynthesis loci sequences, with evidence for potential phase variation mechanisms identified. The potential phase variation mechanisms included a tandem repeat insertion in *natC* and a single base deletion in a homopolymer region of *gatG*. Importantly, our results demonstrated that two of the three ST9 groups shared identical rep-PCR (repetitive extragenic palindromic PCR) patterns, while carrying differences in their LPS outer core biosynthesis loci region. In addition, we found that ST9 isolates either with or without the *natC* tandem repeat insertion were both associated with a single outbreak, which would indicate the importance of screening more than one isolate within an outbreak. Our results strongly suggest the need for a metagenomics culture-independent approach, as well as a genetic typing scheme for LPS, to ensure an appropriate vaccine strain with a matching predicted LPS structure is used.

DATA SUMMARY

Genome sequence data generated in this study have been deposited with the National Center for Biotechnology Information (NCBI) under BioProject accession number PRJNA643588. Raw Illumina sequence read data have been deposited in the Sequence Read Archive (SRA) under the accession numbers SRR12130700 to SRR12130770. The programs used to analyse raw sequence reads, for polymorphism discovery, and whole-genome-sequencing based phylogenies are available as described in Methods.

Received 13 April 2021; Accepted 30 December 2021; Published 10 March 2022

Author affiliations: [1]Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St Lucia, QLD 4072, Australia; [2]Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of Queensland, St Lucia, QLD 4072, Australia; [3]Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, University of Queensland, St Lucia, QLD 4072, Australia; [4]QFAB Bioinformatics – Research Computing Centre, University of Queensland, St Lucia, QLD 4072, Australia; [5]Division of Research and Innovation, University of Southern Queensland, Toowoomba, QLD 4350, Australia.

***Correspondence:** Lida Omaleki, l.omaleki@uq.edu.au; Scott A. Beatson, s.beatson@uq.edu.au; Brian M. Forde, b.forde@uq.edu.au

Keywords: fowl cholera; lipopolysaccharides; *Pasteurella multocida*; phase variation; whole-genome sequencing.

Abbreviations: LPS, lipopolysaccharide; MLST, multilocus sequence typing; rep-PCR, repetitive extragenic palindromic PCR; SEQ, South-East Queensland; ST, sequence type; WGS, whole-genome sequencing.

†Present address: University of Queensland Centre for Clinical Research, Royal Brisbane and Women's Hospital Campus, Herston, QLD 4029, Australia.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Two supplementary tables are available with the online version of this article.

Impact Statement

The genomic and metadata analysis of two unique sets of *Pasteurella multocida* isolates, both from free-range egg production farms, provides apparent explanations, based on detected diversity in the lipopolysaccharide (LPS) outer core biosynthesis loci, for the repeated vaccination failures seen on both farms. Our data revealed a high degree of sequence diversity in the LPS outer core biosynthesis loci type L3, carried by *P. multocida* isolates across multiple years (22 years for one farm and 9 years for the other farm). Our identification of two different potential phase variation mechanisms in the glycosyltransferase genes of the sequence type (ST)9 isolates, possibly allowing evasion of vaccine protection, is a novel finding. In addition, the phylogenomic analysis demonstrated that the 2013 outbreak in farm A was as a result of two overlapping outbreaks caused by ST9 populations carrying two different alleles of *natC* (phase off or phase on). These findings demonstrate the benefit of genomic analysis in providing more detailed information on the possible structure of outer LPS, which is very important for the selection of strains for use in killed autogenous fowl cholera vaccines.

INTRODUCTION

Pasteurella multocida is a heterogeneous species from the family *Pasteurellaceae*, found on the mucosal surfaces of multiple hosts, that causes a range of markedly different diseases in endotherms, including haemorrhagic septicaemia in cattle and buffaloes, snuffles in rabbits, fowl cholera in poultry, and atrophic rhinitis in pigs [1]. *P. multocida* is also known as a zoonotic agent accounting for more than 50% of human wound infections following animal bites in the USA [2]. Fowl cholera has become a re-emerging problem in the free-range layer industry in Australia [3]. In a typical Australian free-range layer farm, hens have access to an outdoor range during the day and then stay in closed sheds at night to be safe from predators. The sheds are also equipped with special nesting boxes for the hens to lay their eggs with no feed or water outside to minimize biosecurity risks. In Australia, free-range eggs make up to 47% of the total retail egg sales currently (https://www.australianeggs.org.au/egg-industry).

Fowl cholera can be an acute disease with sudden death being the first observed manifestation [1]. Chronic infection can also follow an acute disease, with recovering birds becoming the reservoir of the bacteria. Mortality can be 20% and higher in naturally infected animals, with up to 100% in experimental conditions [1]. Mortalities of 55% have been reported in free-range chicken meat flocks through to processing [4].

Lipopolysaccharides (LPSs) are one of the major important immunogenic virulence factors of *P. multocida*. Based on their LPS antigens, *P. multocida* strains are classified into 16 somatic (Heddleston) serovars (1 to 16) [5, 6]. Vaccination with killed whole cells has been previously thought to provide protection only against isolates of the same Heddleston serovar as present in the vaccine [1]. Eight LPS genotypes (L1 to L8), based on the LPS outer core biosynthesis loci, have been recognized in the 16 Heddleston serovar reference strains of *P. multocida* [7]. LPS type L3 (represented by Heddleston serovars 3 and 4) has been recognized as the most prevalent LPS type associated with fowl cholera in Australia [8, 9]. The outer core biosynthesis loci of LPS type L3 is a 6175 bp region consisting of six glycosyltransferase genes required for the assembly of the outer core. These transferase genes are *htpE* (required for the addition of heptose, Hep IV, the first outer core sugar), *gctC* (required for the addition of the two outer core glucose molecules, Glc III and Glc IV, to the heptose), *gatF* (required for the addition of the first galactose, Gal I, to Glc IV), *gatG* (required for the addition of the second galactose molecules, Gal II, to Gal I), and *natB* and *natC* (required for the addition of the two *N*-acetylgalactosamines, GalNAcI and GalNAcII, respectively) [8]. Analysis of the LPS from the *P. multocida* subspecies *multocida* strain Pm70 (GenBank accession no. AE004439) revealed the capacity to produce a full LPS [8]. There has been no published record of phase variations in switching of the glycosyltransferase gene expression of LPS outer core biosynthesis loci in *P. multocida* [10]. The capacity of the LPS type L3 *P. multocida* isolates to produce one of the six possible LPS outer core structures has been previously recognized as a result of inactivating mutations, such as a frameshift or the introduction of a premature stop codon [8]. Phase variation is a mechanism of switching a gene between on and off in many host-adapted bacterial pathogens, such as non-typeable *Haemophilus influenzae*, *Pseudomonas aeruginosa*, *Glaesserella parasuis*, *Mannheimia haemolytica* and *Actinobacillus pleuropneumoniae* [11].

Recently, it has been shown that killed whole-cell vaccines give protection only against strains with identical or nearly identical LPS structures [12]. In Australia, multilocus sequence typing (MLST), Heddleston serotyping, and DNA genotyping by repetitive extragenic palindromic PCR (rep-PCR) fingerprinting have all been used for studies into the epidemiology of disease outbreaks, and also to guide autogenous vaccination programmes [9]. However, none of these techniques are discriminatory enough to provide the required insight into the LPS structure of the strains. Whole-genome sequencing (WGS) data are routinely used in human bacterial pathogen research to investigate strain relatedness and monitor the emergence of bacterial antimicrobial resistance and/or bacterial virulence factors [13]. Here, we use WGS to examine strain relatedness and LPS outer core biosynthesis loci of *P. multocida* isolates associated with fowl cholera outbreaks from two free-range layer farms over a period of either 22 years (farm A) or 9 years (farm B).

Table 1. Details of layer farm A fowl cholera outbreaks and the associated isolates of *P. multocida*

Year	Outbreak isolate	no. of sheds affected	no. of isolates	no. of chickens	ST	rep-PCR pattern	LPS L3 specification	Vaccine used at rearing
1994	PM204 to PM214	1	11	4	20	Not done	*gatG* 14 bp deletion	Unvaccinated flock
2002	PM979 to PM984	3	6	4	20	3	*gatG* 14 bp deletion	Unvaccinated flock
2011	PM1533 and PM1542*	2	2	2	9	2	*natC* 7 bp insertion	PM979 at 12 weeks
2013	PM1775 to PM1777	2	3	3	9	1	No missense mutation	Commercial tetra-valent at 10 weeks and PM979, PM1533 at 16 weeks
	PM1778†		1			2	*natC* 7 bp insertion	
2014	PM1947 to PM1949	2	3		9	2	*gatG* 1 bp insertion	PM979, PM1533
2015	PM2142	NA	1	1	9	1	No missense mutation	PM979, PM1542
2016	PM2232 to PM2240	NA	9	6	9	1	No missense mutation	PM979, PM1947‡

NA, Not available.
*PM1542 was isolated from hens of a different shed and different age to that of PM1533.
† PM1778 was isolated at the same time as PM1775 to PM1777. PM1778 also harboured a *gatF* E63K change compared to the 2011 strains.
‡PM979 and PM2240 were used as an autogenous bivalent vaccine after the 2016 outbreak.

METHODS

Farms and isolates

In total, 73 isolates of *P. multocida* associated with fowl cholera from two free-range layer farms (layer A and layer B) were included in this study. The selected farms have been using killed autogenous vaccines for prevention and control of fowl cholera. Each of the farms had several layer sheds housing the hens at night, with easy access to the open area during the day.

Isolates were collected during the disease investigations through submitting chicken (*Gallus gallus domesticus*) carcasses or swab samples by the farm veterinarian to front-line veterinary diagnostic laboratories. The isolates were then sent to our laboratory for typing and storing, with some isolates used in killed autogenous vaccines subsequently given during the rearing stage on farms from which those isolates were sourced. The vaccination routine included administration of two doses of a killed autogenous vaccine at rearing (usually around 10 and 19 weeks of age). However, due to financial constraints linked to vaccine costs and the cost of administration of the vaccine, it was reduced to once at rearing or even eliminated from time to time. The number of available isolates per outbreak was different across the years ranging from only one isolate to multiple isolates from the same chicken and also isolates from different chickens within an outbreak (Table S1, available with the online version of this article).

Layer farm A

In total, 36 isolates, obtained from outbreaks over a 22 year period (1994 to 2016), were included in this study (Table 1). This included 17 isolates from the first two outbreaks (1994 and 2002) also described by Zhang *et al.* [14], together with 19 isolates from the more recent outbreaks (Tables 1 and S1).

The farm was a multi-age layer operation located in coastal South-East Queensland (SEQ) that continuously had chickens with no down time nor any period of total flock depopulation. In 1994, the farm had five sheds and a total of 8000 layers, expanding to 15 sheds and 50 000 layers in 2002.

Layer farm B

The second free-range layer farm (farm B) was a large company with farms located at four different sites (B1 to B4), in SEQ close to the border with New South Wales, with each site having several free-range layer sheds. In total, 37 isolates from outbreaks of fowl cholera between 2010 and 2019 were included. This total included two of the three original vaccine strains PM842 (obtained from the previously owned operation at site B5) and PM844 (obtained from the farm at site B3) associated with outbreaks in 2001 (Table 2). The third vaccine strain (PM843) was an isolate from farm B3, which was not available for WGS at the time of the current project.

LPS typing and DNA fingerprinting

Genomic DNA of all the isolates was prepared using the PrepMan ultra sample preparation reagent (Life Technology; Applied Biosystems). The isolates were subjected to LPS multiplex-PCR [7], as well as rep-PCR [15].

Table 2. Details of the layer farm B fowl cholera outbreaks and the associated isolates of *P. multocida*

Farm	Year	Outbreak isolate	Age (weeks)	no. of isolates	ST	LPS type L3 specification	Vaccine used at rearing*
B5	2001	PM842	–	1	20	*gatG* 14 bp deletion	NA
B3	2001	PM844	–	1	8	No stop codon or frame shift	NA
B3	2010	PM1447	45	1	274	*htpE* insertion position 126	–
B3	2011	PM1541†	59	1	274	*htpE* insertion position 126	–
B3	2012	PM1582†	70	1	274	*htpE* insertion position 126	–
B3	2014	PM1845 to PM1853 and PM1855	–	10	8	No stop codon or frame shift	–
B2	2016	PM2248	62	1	20	*gatG* 14 bp deletion	**PM842**, PM843, PM1855
B4	2016	PM2290	50	1	7	*htpE* del position 695	**PM842**, PM843, PM1855
B2	2017	PM2356	50	1	20	*gatG* 14 bp deletion	**PM842**, PM843, PM1855
B1-shed1‡	2016	PM2269 and PM2305	50	2	9	*natC* 7 bp insertion	**PM842**, PM843, PM1855
B1-shed1	2017	PM2369	–	1	9	*natC* 7 bp insertion	**PM842**, PM843, PM1855
B1-shed1	2017 DEC	PM2383 to PM2391	33	9	9	*natC* 7 bp insertion	**PM842**, PM843, PM1855
B1-shed2	2018 AUG	PM2398 and PM2401	66	2	9	*natC* 7 bp insertion	**PM842**, PM843, PM1855 at rearing and then PM1855, PM2269, PM2290 at 43 weeks
B1-shed4	2018 SEP	PM2445	57	1	9	*natC* 7 bp insertion	PM1855, PM2269, PM2290
B1-shed3	2019	PM2470 and PM2585 to PM2587	66	4	9	*natC* 7 bp insertion	PM1855, PM2269, PM2290

NA, Not available.
*The vaccine strain in bold font has the 14 bp deletion in the *gatG* gene, so should have theoretically protected against the 2016 and 2017 outbreaks in B2. These birds were vaccinated twice at 9 and 13 weeks old. The vaccine strain in underlined font had the 7 bp insertion in the *natC* gene.
†PM1541 and PM1582 were isolated from the same flock at different ages.
‡Eight sheds in total were suffering from fowl cholera in farm B1 from late 2016 to August 2017.

WGS

DNA was extracted using a DNeasy UltraClean microbial kit (Qiagen), according to the manufacturer's instructions, from overnight cultures grown on 5% sheep blood agar. The DNA samples were then sequenced on an Illumina NextSeq 500 platform (150 bp paired ends) by the Australian Centre for Ecogenomics Sequencing Service at the University of Queensland (St Lucia, QLD 4072, Australia). Trimmomatic v0.36 [16] was used for filtering the raw reads, with the quality of the reads assessed before and after filtering using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). SPAdes v3.10.1 [17] was used for *de novo* assemblies. *In silico* MLST [18] profiling was performed using MLST v2.8 (https://github.com/tseemann/mlst) and the *P. multocida* RIRDC-MLST scheme [19]. BLASTN v2.2.31+ was used to perform *in silico* LPS typing by extraction of the LPS region from the draft assemblies as previously described [20]. Geneious Prime v2020.2.2 (https://www.geneious.com) was used for annotation of the region with implemented ClustalW v2.1 [21] for performing the alignments. The Nesoni v0.132 (https://github.com/Victorian-Bioinformatics-Consortium/nesoni) pipeline was used, implementing Bowtie 2 [22, 23], for mapping the trimmed paired-end reads to the chromosome of Pm70. BamView [24] implemented in Artemis v16.0.0 [25] was used to view the read pile-ups.

To investigate the relationship between the isolates from the two farms and other *P. multocida* strains, 40 complete genomes in total were obtained from GenBank using ncbi-genome-download (https://github.com/kblin/ncbi-genome-download). QUAST v5.0.2 [26] was used for quality assessment of the assemblies. FastTree implemented in Parsnp v1.2 [27] was used for core-genome alignment of the draft genome assemblies. Phylogenetic analysis was performed using RAxML v8.2.9 [28] with a general-time reversible nucleotide substitution model with a gamma correction for site variation for tree reconstruction (bootstrap 1000 with Lewis ascertainment correction). The core-genome assemblies of the *P. multocida* isolates within each farm, as well as those from the same MLST across the two investigated farms, were further investigated. To do this, Gubbins v2.1.0 [28] was used for removing the recombination sites in the core-genome SNP alignment, followed by masking the recombination output into the original core genome alignments using maskrc-svg (https://github.com/kwongj/maskrc-svg). This was then followed by phylogenetic analysis

using RAxML as described above. The generated phylogenetic trees were visualized using FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) and Phandango [29].

RESULTS

Layer farm A outbreak history and vaccination programmes

Since 1994, seven outbreaks of fowl cholera (1994, 2002, 2011, 2013, 2014, 2015 and 2016) have been reported on farm A. The 1994 outbreak occurred in an unvaccinated flock. Vaccination was introduced after this outbreak, but was discontinued in 2000 after a 6 year absence of disease. Following the 2002 outbreak, an autogenous killed fowl cholera vaccine based on isolate PM979 from the 2002 outbreak was used twice during rearing. However, due to financial constraints, the vaccination schedule was changed to once at 12 weeks of age and fowl cholera re-emerged on the farm in 2011. After the re-emergence of the disease in 2011, isolate PM1533 was included together with PM979 as a bivalent autogenous killed vaccine. This decision was based on the serotyping assay done at the time that identified PM979 as Heddleston serovar 4 and PM1533 as Heddleston serovar 3 (Table S1). In October 2013, at least two of the sheds, with 37- and 51-week-old birds, were affected by fowl cholera despite the routine vaccination (Table 1, Fig. 1). The chickens were vaccinated with a commercial tetra-valent (non-autogenous) vaccine at 10 weeks of age and the bi-valent killed autogenous vaccine (PM979 and PM1533) at 16 weeks of age. Comparison of the rep-PCR profiles of the four 2013 outbreak isolates with the autogenous vaccine strains indicated that three of the isolates (PM1775 to PM1777) produced a different, albeit closely related, fingerprint (rep-PCR type 1 in Fig. 2a) to that of vaccine strain PM1533 (rep-PCR type 2 in Fig. 2a). The fourth 2013 field isolate (PM1778) had the same rep-PCR profile as vaccine strain PM1533 (rep-PCR type 2). The rep-PCR fingerprint of vaccine strain PM979 was not seen in the field isolates. However, all four 2013 isolates were identified as Heddleston serovar 3 by gel immunoprecipitation assay (Table S1) and LPS type L3 via LPS multiplex-PCR (Table 1). Therefore, a decision was made not to make any changes in the vaccine. In October 2014, another outbreak occurred in two flocks that were 47 (isolates PM1947 and PM1948) and 54 (isolate PM1949) weeks of age. Older flocks also present on the farm did not show any signs of the disease. The rep-PCR fingerprinting showed that these 2014 isolates shared the same fingerprint (rep-PCR type 2) as PM1778 from the 2013 outbreak and PM1533 (the autogenous vaccine strain adopted after the 2011 outbreak). A decision was made to substitute PM1533 with PM1542, a 2011 isolate known to have the same rep-PCR fingerprint as PM1533, in the autogenous vaccine (Fig. 2b). Following the 2015 outbreak from which the laboratory received only one isolate (PM2142) (Table 1), the vaccine was changed again to contain PM979 and PM1947. However, the 2016 outbreaks following this change yielded isolates that all shared the same rep-PCR as the 2013 isolates, i.e. rep-PCR type 1 (Fig. 2c, d). This necessitated a change in the autogenous vaccine in 2017, removing PM1947 and substituting in PM979 and PM2240 (the latter being a 2016 outbreak isolate).

Layer farm B outbreak history and vaccination programmes

The original vaccine strains used across different sites of the company were obtained from outbreaks of fowl cholera in 2001, with PM842 obtained from a previously owned farm (B5), and PM843 and PM844 from farm B3. In 2001, the farm at site B3 was not a free-range farm and only contained breeders in sheds, while B5 was a free-range layer farm. There was a transfer of nest boxes from B3 to B5 and eggs from both farms were sent to the same hatchery. The two farms were 100 km apart, with farm B3 later populated with layers in sheds. Initially, the fowl cholera vaccine was administered twice during rearing (typically at 9 and 12 weeks of age) which, since 2003, became once between 8–13 weeks of age for cost reasons. In mid 2010, farm B3 became a free-range layer farm, with three production sheds and one rearing shed.

The first of the series of recent outbreaks on farm B3 was recognized in November 2010 with *P. multocida* isolated from the heart sac and peritoneal cavity (PM1447) (Table 2). The outbreak happened in a shed containing 17400 45-week-old hens. The total mortality from 16 to 76 weeks of age was 16% in this shed. The second confirmed outbreak happened in another shed on the B3 farm, in September 2011, with isolate PM1541 obtained from the subcutaneous tissue of a hen with a swollen face and wattle. The third confirmed outbreak was the reappearance of the disease in this second shed with birds at 70 weeks of age in January 2012, with again only one isolate sent for analysis (PM1582). This shed had a total mortality of 10% up to processing. This was followed by another outbreak in 2014 on farm B3 from which *P. multocida* isolates from 10 different birds were available.

The first outbreak on farm B1 happened in December 2016 in one shed (B1-shed1) (Table 2), followed by outbreaks in many other sheds involving chickens between 33 to 66 weeks of age. Isolates from four different sheds (B1-shed1 to B1-shed 4) were collected by the farm veterinarian and included in the current study (Table 2). The chickens were vaccinated once at 11–14 weeks of age with an autogenous vaccine containing PM842, PM843 and PM1855. The vaccination was then changed to twice at rearing for the flocks in 2017 with the same vaccine. The flocks then received a boost at week 43 with a new vaccine containing PM1855, PM2269 and PM2290 in 2018. From October 2018, all new flocks received the new vaccine; however, fowl cholera cases still occurred despite two vaccinations during rearing (B1-shed3 and -shed4). PM2269 was later replaced by PM2445 in the trivalent vaccine.

Fig. 1. Schematic diagram showing the LPS outer core biosynthesis loci type L3 subtypes of the isolates observed in farm A and the vaccine strains used during rearing. (a) Comparisons of the four major different LPS type L3 genotypes with that of *P. multocida* strain Pm70 known to produce a fully expressed LPS type L3 outer core [8]. The nucleotide sequence of LPS from Pm70 is set as the reference. The top grey bar next to each outbreak strain (coloured circles) depicts the nucleotide sequence and the lower one depicts the amino acid translation. Black vertical lines within the grey bars show areas of difference between the sequence of each outbreak strain and that of Pm70. The green arrows show insertion sites and the red arrows show the deletion sites. The proposed outer structure is depicted on the left and is adapted from Harper *et al.* (2015) [7], with the dashed line separating inner core from outer core LPS. The white circle represents heptose, the blue circle represents glucose, the yellow circle represents galactose and the yellow box represents GalNAc. (b) Comparison between the outbreak strains (coloured circles) and vaccine strains (coloured rectangles) since the 2011 outbreak in layer A. Each circle represents one bacterial sample as described in Table 1.

Outbreaks on farm B2 were recorded in September 2016 (isolate PM2248) and October 2017 (PM2356). The only recorded outbreak on farm B4 was in 2016 (PM2290). These two farms (B2 and B4) were also using the same vaccination routine as farm B1 at rearing.

Population level genomic analysis and *in silico* MLST and LPS typing

In this work, WGS was used to contextualize the isolates within the known population structure of *P. multocida* by core-genome SNP analysis of 39 publicly available complete genomes, with the genome of Pm70 as a reference (Table S2). Maximum-likelihood phylogenetic analysis of core-genome SNPs clustered isolates into well-supported clades, largely consistent with sequence type (ST), distributed throughout the *P. multocida* species tree (Fig. 3). Of the STs identified in this study, the ST7 and ST20 clades showed the closest phylogenetic relationship with ST20s, having a mean of 2710 (SD ±49.5) core-genome SNPs to ST7 while between 11489 and 41589 core-genome SNPs to the isolates from other STs. This was consistent with the similarity in their respective ST profiles (triple locus variants with different alleles for *gdh*, *pgi* and *pmi*). ST274 was the most divergent with a mean pair-wise SNP distance of 39731 to strain Pm70, forming an outgroup clade together with GenBank isolates that could not be

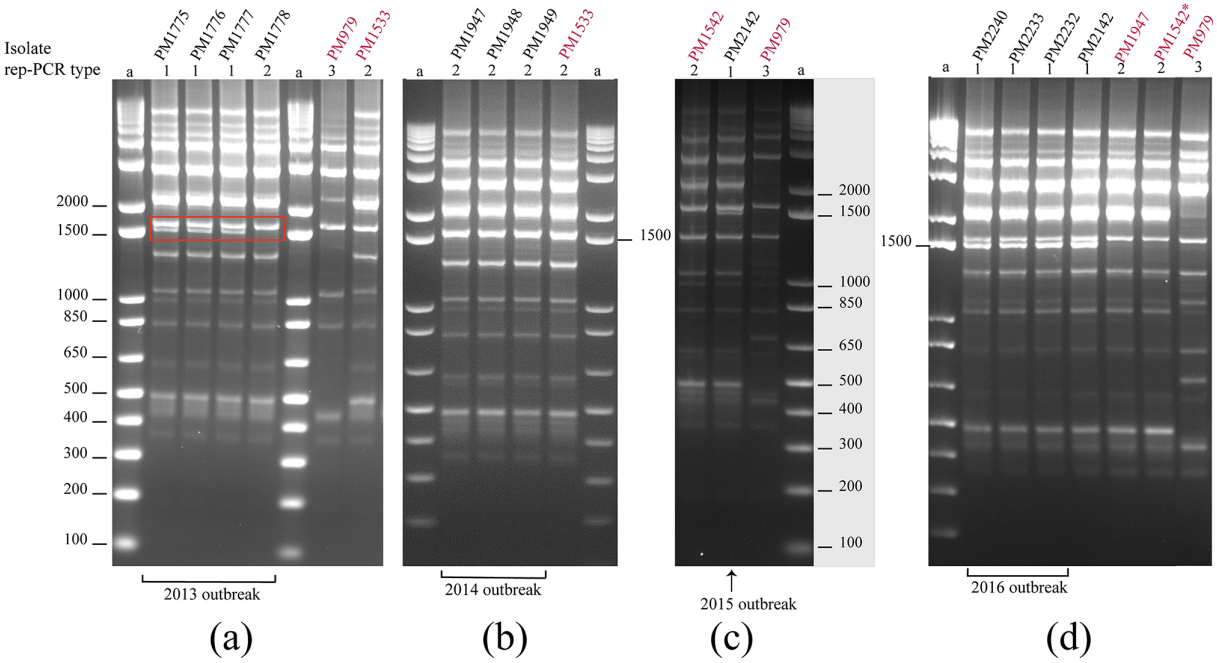

Fig. 2. Rep-PCR analysis of the outbreak *P. multocida* isolates (black font) and the vaccine strains (red font) used at the time on layer farm A. (a) A new rep-PCR type emerged in 2013 (rep-PCR type 1 – PM1775, PM1776, PM1778) with only one band difference with rep-PCR type 2 (PM1778) covered by the vaccine strain PM1533. The two rep-PCR types observed in the 2013 outbreak isolates showed only a one band difference (red box). (b) The 2014 isolates (PM1947, PM1948, PM1949) showed the same rep-PCR pattern (rep-PCR type 2) as the vaccine strain (PM1553). However, the genomic results showed that they had the capacity to produce a shorted LPS outer core compared to PM1533. (c, d) The rep-PCR pattern seen in 2013 (rep-PCR type 1) reappeared in the 2016 outbreak isolates (PM2240, PM2233, PM2232, PM2142). Lanes a, 1 kb plus DNA ladder (Invitrogen) – the annotated band sizes are in bp.

assigned to any known ST within the RIRDC scheme (Fig. 3). The ST9 clade included the Pm70 genome and isolates from both layer A and B farms clustered together (Fig. 3).

On layer farm A, the *in silico* MLST recognized that isolates from the earlier outbreaks (1994 and 2002) all belonged to ST20, while all isolates from the 2011 outbreak onwards belonged to ST9. However, *in silico* MLST typing identified five distinct STs (ST7, ST8, ST9, ST20 and ST274), between 2001 and 2019 in the layer B complex. While belonging to different STs, all the isolate genomes from the two investigated farms and 21 of the 39 complete genomes were predicted to encode LPS type L3.

Layer farm A genomic analysis

Phylogenomic analysis of ST20 isolates from farm A identified two well-supported clades, which corresponded to respective outbreaks in 1994 and 2002 (Fig. 4a). Isolates from these two outbreaks were highly similar at the core-genome level with a mean pair-wise non-recombinant SNP distance of 11.8 (SD ±10.6). The phylogeny also revealed that both outbreaks shared a common ancestor, rather than the 2002 outbreak isolates being directly descended from the 1994 outbreak. For the 2002 outbreak isolates to have directly descended from the 1994 isolates, the 2002 isolates would have to form a sub-tree of the 1994 isolates, which they clearly do not.

All isolates from the subsequent outbreaks (2011 to 2016) were ST9 and, thus, unrelated to the 1994 and 2002 outbreaks (Fig. 3). With a mean pair-wise non-recombinant SNP distance of 27.8 (SD ±13.9), the ST9 inter-host divergence was comparable to that seen in the ST20 isolates. However, in contrast to ST20, there was some phylogenetic evidence for the ST9-linked isolates for the sequential emergence of outbreaks from a common source. Two of the most basal ST9 taxa are isolates from the 2011 outbreak (PM1533 and PM1542), branching prior to a well-supported clade comprising all but one later isolate (PM1778). Within this large clade, sub-clade 2a contained all 2014 outbreak isolates, and sub-clade 2b contained 2013 (PM1775, PM1776, PM1777), 2015 (PM2142) and all 2016 outbreak isolates (Fig. 4b). However, the finding that 2013 isolates were not monophyletic (see, for example, PM1778 and PM1775), suggests that there were two overlapping outbreaks in 2013, involving closely related but genetically distinct ST9 populations, each potentially expressing different LPS structures (Fig. 4b). Collectively, these results point to an ongoing source of ST9 *P. multocida* established prior to 2011, which has seeded outbreaks in 2011 (PM1533, PM1542), 2013

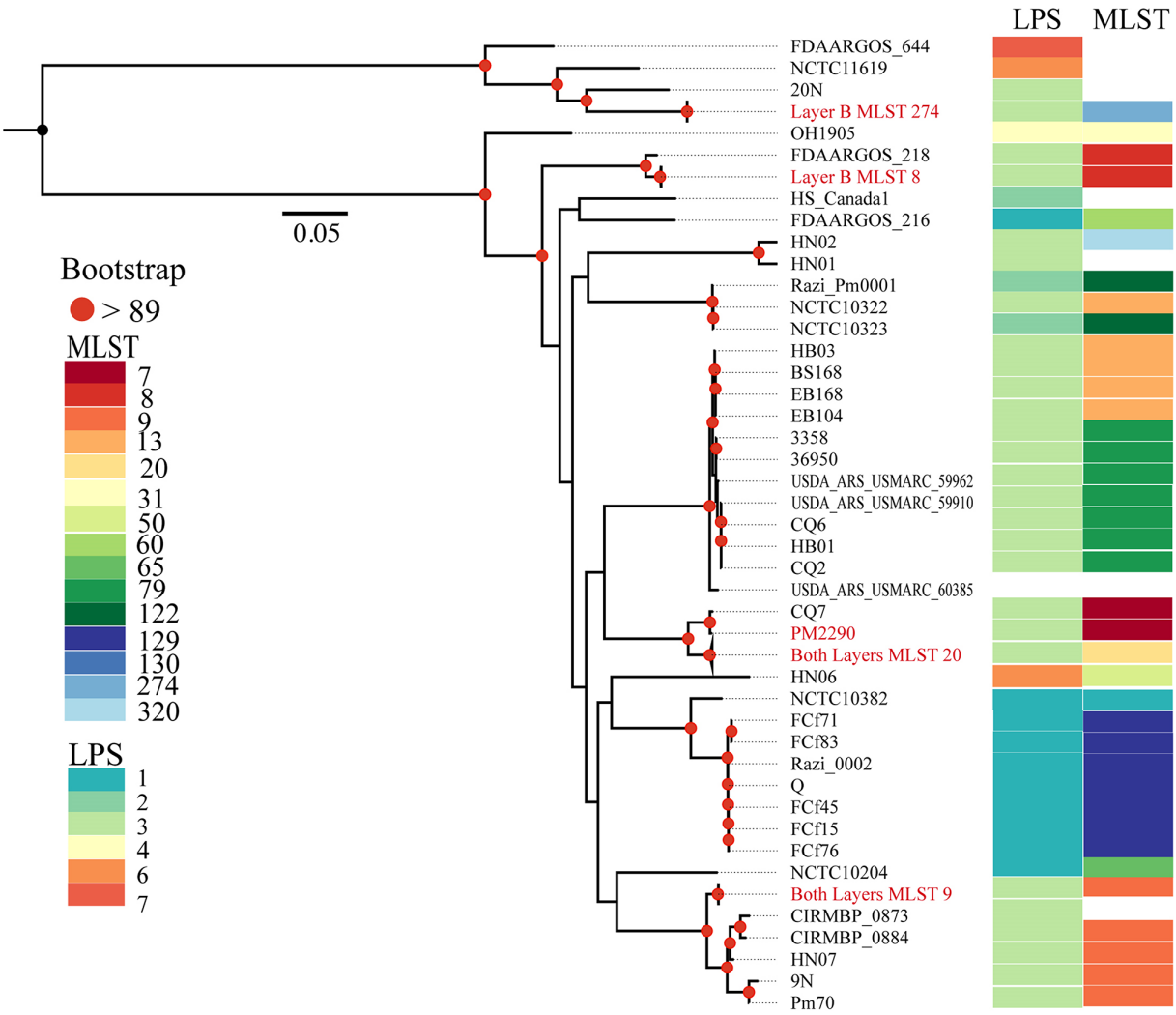

Fig. 3. Maximum-likelihood tree of the core-genome SNPs of *P. multocida* isolates from the two investigated free-range layer farms (red font) compared to those of the complete genomes obtained from GenBank. The tree is rooted at the midpoint, which corresponds to the actual root, by *Pasteurella canis* strain HL268 (GenBank accession no. CP083262), which has been omitted for visualization. In total, 84708 SNPs were identified from the 1470156 bp core-genome alignments using the genome of *P. multocida* strain Pm70 as the reference. Bar, number of substitutions per site.

(PM1775–PM1778) and 2014 (PM1947–PM1949), with the 2015 (PM2142) and 2016 outbreaks (PM2232–PM2240) directly linked to three of the isolates associated with the 2013 outbreaks (PM1775–PM1777) (Fig. 4b).

The LPS *in silico* typing revealed that, while all of the tested isolates within the two study farms carried LPS type L3, the nucleotide sequence of their LPS outer core biosynthesis loci was quite diverse and would be predicted to produce different LPS structures. When compared to the LPS type L3 locus of *P. multocida* strain Pm70, all ST20 isolates from the two farms had a 14 bp deletion within *gatG* (positions 491 to 505), and a 1 bp insertion within *natC* (313_314insA).

The ST9 isolates from farm A had three variant LPS outer core biosynthesis loci, each with a predicted different LPS structure, consistent with the ST9 phylogeny (Fig. 4b). The first ST9 variant is defined by a 7 bp sequence insertion (TTATTAT) in *natC* (position 730). This mutation is a single copy duplication of one half of an imperfect tandem repeat in Pm70 *natC*, producing a three-unit tandem repeat array (5′-TTAATAT-**TTATTAT**-TTATTAT-3′) in the most basal ST9 taxa (isolates PM1533 and PM1542 from 2011, and PM1778 from 2013). This 7 bp sequence is absent in all other ST9 isolates, suggesting that reversion to the Pm70 wild-type sequence in clade 2 isolates has occurred following deletion of a 7 bp repeat in *natC*. The 2014 outbreak isolates (clade 2a) also harboured a 1 bp deletion at position 722 of *gatG* (721_723delA) that was not present in clade 2b isolates (Fig. 4b). The observation that these mutations occur within a 7 bp tandem repeat array and a homopolymer tract, respectively, suggests that they are potential phase variation mechanisms. While the *gatG* deletion appears to have arisen only in the 2014 clade, the

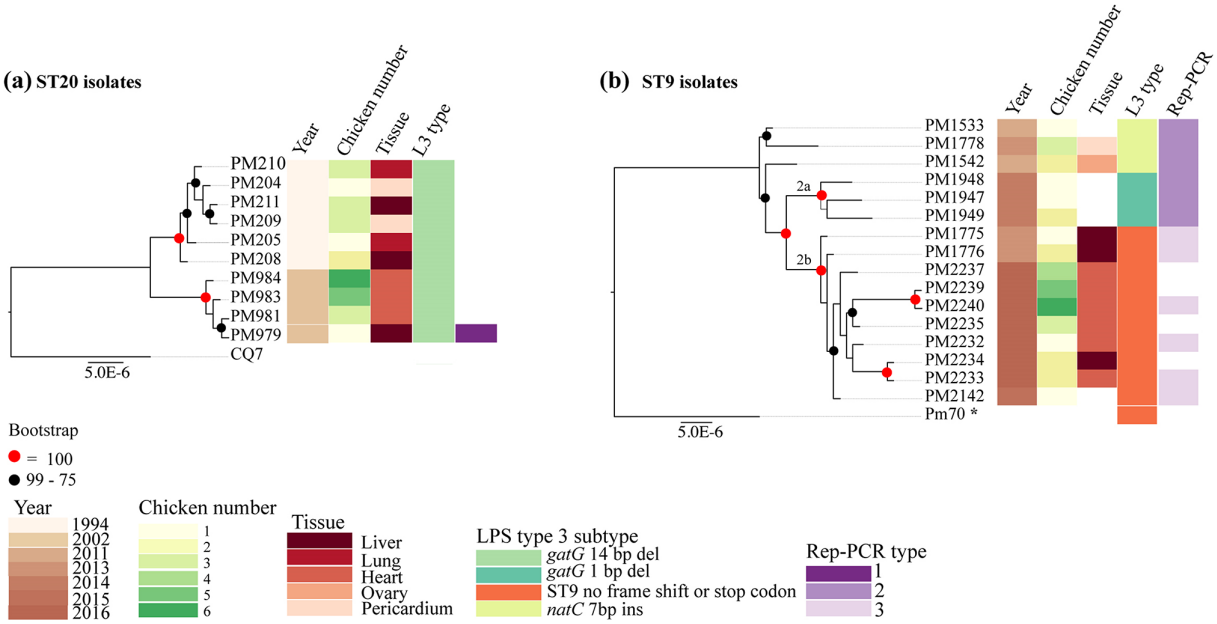

Fig. 4. Maximum-likelihood analysis of the *P. multocida* isolates from layer farm A with filtered recombination sites. (a) Maximum-likelihood analysis of the core-genome SNPs of the ST20 isolates. PM204 was identical to PM206, PM207, PM213 and PM214 at the core-genome level. PM979 was identical to PM980 and PM982 at the core-genome level. In total, 107 non-recombinant core-genome SNPs were identified using strain CQ7 as the outgroup. (b) Maximum-likelihood tree of the core-genome SNPs of the ST9 isolates. PM1776 and PM1777 were identical at the core-genome level. In total, 220 non-recombinant core-genome SNPs were identified using strain Pm70 as the outgroup. *, The LPS outer core biosynthesis loci of Pm70 harboured two SNPs with the outbreak isolates. Bar, number of substitutions per site.

confinement of the *natC* insertion to the most basal taxa suggests that this mutation was ancestral and reversion to wild-type (through deletion of a 7 bp repeat) occurred prior to divergence of the two major ST9 sub-clades (2a and 2b).

Layer farm B genomic analysis

Phylogenomic analysis of the isolates from the four sites of farm B (B1–B4) with fowl cholera outbreaks examined in this study clustered them into four distinct clades consistent with ST (Fig. 5a, b): (i) three ST274 isolates, collected from chickens at farm B3 during outbreaks of fowl cholera during three successive years – 2010 (PM1447), 2011 (PM1541) and 2012 (PM1582); (ii) ST8 isolates including the original vaccine strain PM844 obtained in 2001 from B3, and the 2014 outbreak isolates obtained from the same farm 13 years later (PM1845 to PM1855); (iii) a single ST7 isolate, PM2290 (the only isolate from farm B4) and three ST20 isolates (PM842, PM2248 and PM2356 – PM842 being one of the original vaccine strains, and PM2248 and PM2356 were from farm B2 during the 2016 and 2017 outbreaks, respectively); and (iv) 19 ST9 isolates from four different sheds of farm B1 from 2016 to 2019. This farm had recurring outbreaks of fowl cholera since 2016 despite routine vaccinations.

Comparisons between the LPS outer core biosynthesis loci revealed that isolates from each ST carried distinct genetic mutations in their LPS type L3 outer core biosynthesis loci (Fig. 5a). The ST274 and ST7 isolates carried an insertion at different positions of *htpE*, positions 695 and 126, respectively; each causing frameshifts predicted to result in early termination of HtpE translation. This would result in production of a fully truncated LPS with no outer core, as previously described in other *htpE* mutants [8]. All three ST20 isolates from layer B were identical in their LPS outer core biosynthesis loci to those from layer A. The ST8 isolates appeared to have 44 SNPs in their LPS outer core biosynthesis loci relative to Pm70; however, no frameshift mutations were identified in the region.

Although the ST9 isolates had an LPS sequence near identical to that of Pm70 in five of the glycosyltransferase genes, all carried the three-unit tandem repeat array identified in the *natC* region of ST9 isolates from layer farm A. In contrast to layer farm A, all the layer farm B ST9 isolates, except PM2269, also carried an amino acid change in their *gatG* gene, relative to Pm70 (C128Y).

ST9 isolates in farm B1 shared a mean of 14.7 (SD ±5.2) non-recombinant core-genome SNPs to each other, with isolates obtained from the same outbreak in shed 1 (April 2018) sharing between 4 and 22 non-recombinant SNPs. Further genomic investigation of ST9 isolates from shed 1 in farm B1 revealed phylogenetic relationships consistent with an ongoing source of *P. multocida*, with isolates from this particular shed being the source of outbreaks in the other three investigated sheds (Fig. 5c). Three of the four 2019 isolates from one flock in farm B1-shed 3 (PM2585, PM2586, PM2587) were most closely related to a 2016 isolate

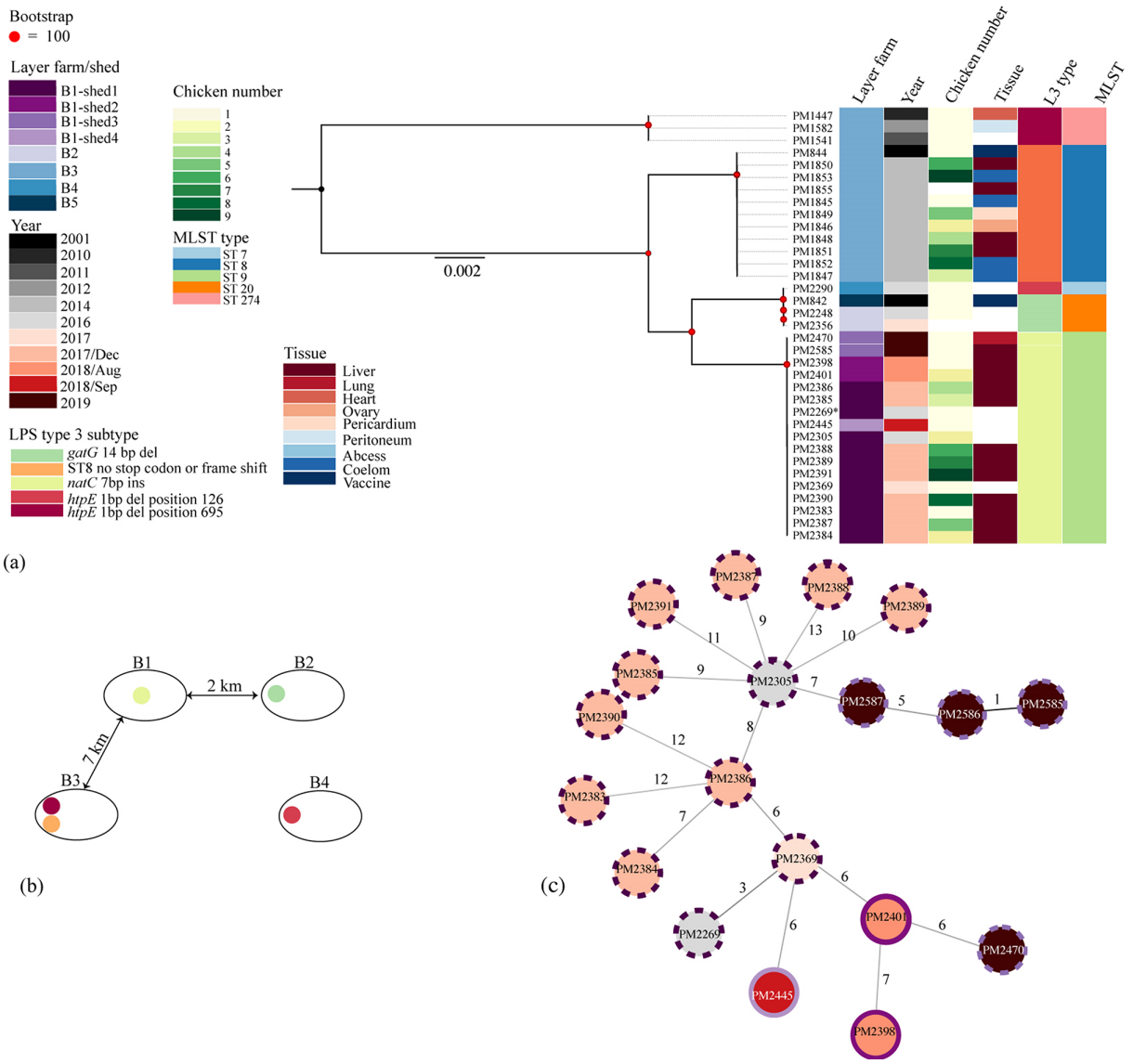

Fig. 5. Genomic relatedness of the isolates from the layer B complex. (a) Phylogenetic tree of the isolates obtained from all four farms across the complex between 2010 and 2019. In total, 62753 non-recombinant core-genome SNPs were identified using ST274 isolates as the outgroup. PM2585 was identical to PM2586 and PM2587 at the core-genome level. Bar, number of substitutions per site. (b) Schematic diagram representing the distance between the four farms. Coloured circles represent the major mutations in the LPS type 3 carried by the isolates obtained from each farm. (c) Minimum spanning tree of the core-genome SNPs of ST9 isolates obtained from B1 across the years and different sheds. Each circle represents an isolate as described in Table 2. The internal circle colour represents the year of isolation, the external circle colour represents the shed that the isolates came from. Shed 1 and shed 4 isolates are differentiated with dashed outer circles from the other two sheds (Table S1). * PM2269 also carried an amino acid change in the *gatG* relative to the other ST9 isolates from layer farm B.

from shed 1 (PM2305), suggesting direct transmission between the sheds. The fourth isolate, PM2470, was apparently indirectly transmitted from shed 1 via shed 2, being most closely related to PM2401.

Comparisons across layer farms A and B

Comparison of the tree topology of ST20 and ST9 isolates from both farms revealed ST-specific differences in the farm of origin. The ST20 isolates fell into two distinct clades, with the 1994 farm A isolates more closely related to isolates from farm B than to the 2002 farm A isolates (Fig. 6). The isolates obtained in 1994 from farm A had between 6 and 8 core-genome SNPs to PM842 (farm B5), and between 15 to 17 SNPs to isolate PM2248 (farm B5), compared with 26 to 29 core-genome SNPs with 2002 farm A isolates. Together, these results suggest outbreaks on farm A in 1994 and farm B in 2001 shared a common ancestor. In contrast to ST20, all ST9 isolates clustered into two well-supported clades according to their farm of origin (Fig. 7). This phylogenetic

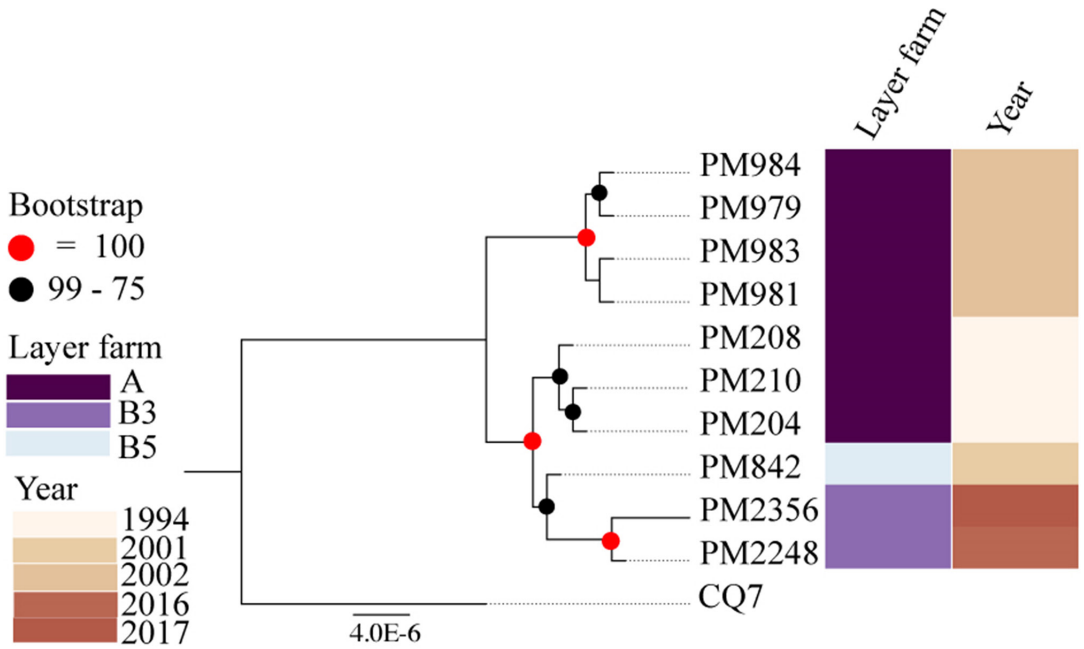

Fig. 6. Maximum-likelihood analysis of the ST20 isolates from layer farms A and B. PM210 was identical to PM209 and PM211 at the core-genome level. PM204 was identical to PM205, PM206, PM207, PM212, PM213 and PM214 at the core-genome level. PM979 was identical to PM980 and PM982 at the core-genome level. The tree was rooted with isolate CQ7 as the outgroup, with a total number of 111 non-recombinant SNPs identified. Bar, number of substitutions per site.

relationship supports the hypothesis that the three-unit tandem repeat array in *natC* was ancestral in ST9, with post-2013 outbreaks on farm A being dominated by a strain in which there is a reversion to wild-type (two-unit tandem repeat) (Fig. 7).

DISCUSSION

Fowl cholera is a major threat to the rapidly growing poultry industries in Australia [3] and other countries [1]. It has always been a problem in backyard poultry throughout the world, and, in developed countries, the recent shift to free-range/organic meat and egg production has resulted in a marked increase in the prevalence of fowl cholera in both chicken meat (broilers) and egg producing (layers) flocks [3]. The aim of this study was to provide some insight into the genetic diversity between outbreaks of fowl cholera as well as vaccination failure against this disease in two separate free-range layer farms.

Five different *P. multocida* MLST types (ST7, ST8, ST9, ST20 and ST274) were identified in association with fowl cholera on the layer B farms across 17 years, while only two (ST9 and ST20) were identified on layer farm A over a 22 year period. However, all 68 isolates from these two farms examined in this study harboured LPS outer core biosynthesis loci type L3, in line with previous observations [8, 9]. In addition, this study demonstrates that different genetic lineages of *P. multocida* have the capacity to carry this LPS type, supporting a role of horizontal gene transfer in the evolution of the LPS outer core biosynthesis loci, as suggested previously [4, 30].

The LPS type L3 outer core biosynthesis loci consist of six glycosyltransferase-encoding genes named *htpE*, *gctC*, *gatF*, *natB*, *gatG* and *natC* [8]. Major genetic mutations in *natC*, *gatG* and *htpE* and their associations with producing truncated LPS outer core have been reported [8]. In the current study, our genomic analysis identified new genetic disruptions in these three glycosyltransferase genes, as well as the previously reported deletion in the *gatG* gene of the ST20 isolates [4, 8]. No missense genetic mutations resulting in a frameshift or introduction of a premature stop codon were found in the other three glycosyltransferase genes *natB*, *gatF* and *gctC*, in agreement with previously published data [8].

Previously, using restriction endonuclease analysis (REA), the persistence of a single strain of *P. multocida* on farm A for over 10 years had been reported [14]. However, our genomic results disagreed with this finding and demonstrated that the isolates from farm A were not a single strain. Indeed, all the ST20 isolates from the 1994 outbreak on farm A were more closely related to those of the 2001 isolate from farm B5 and 2016 isolates from farm B2 than to isolates from the 2002 farm A outbreak. This finding demonstrates the far greater richness of understanding that is possible with genomic analysis as compared with techniques such as REA.

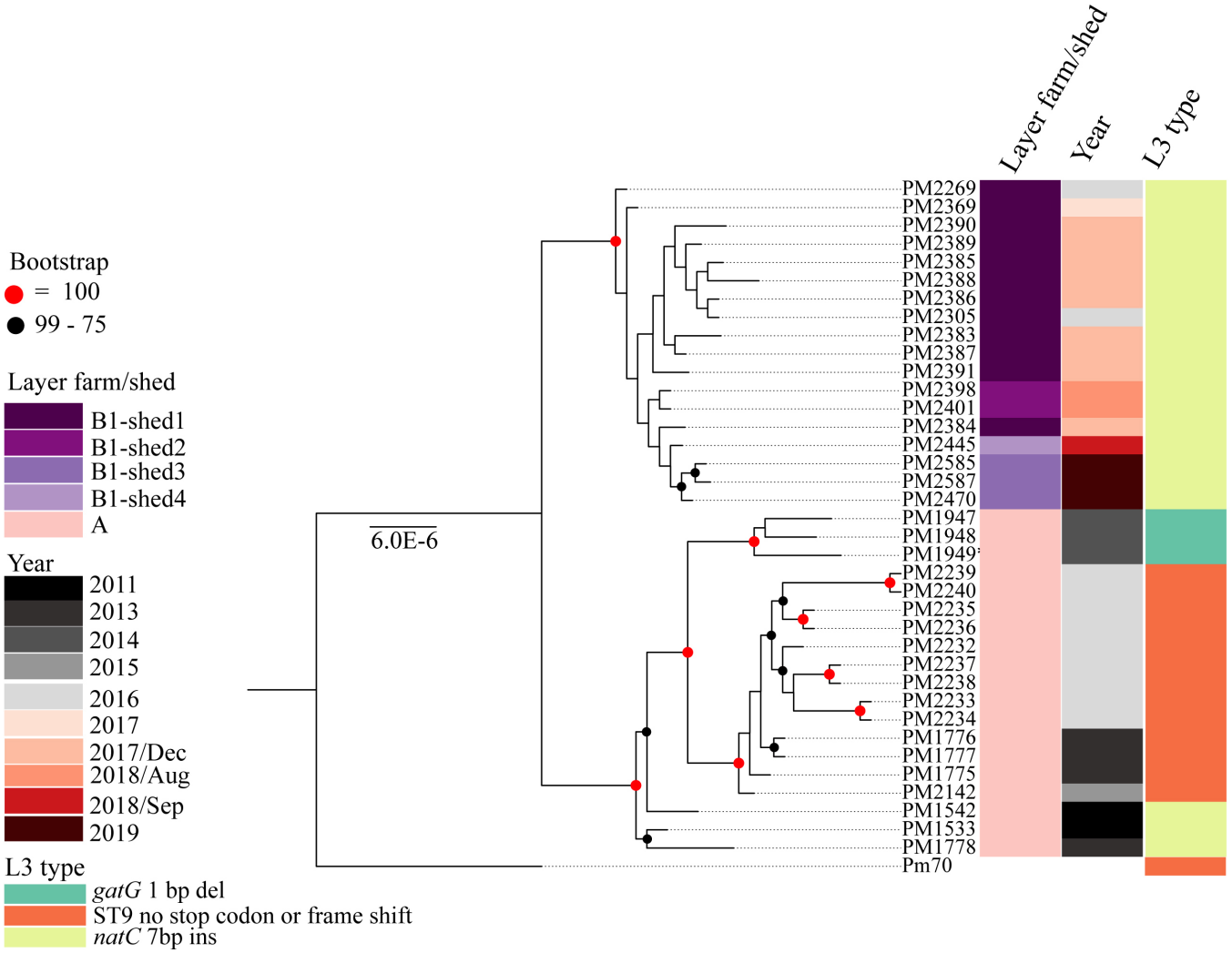

Fig. 7. Maximum-likelihood analysis of the ST9 isolates obtained from layer farms A and B. In total, 325 non-recombinant SNPs were identified. The tree was rooted with isolate Pm70 as the outgroup. Bar, number of substitutions per site.

The ST9 isolates on farm A, first detected in 2011 (strain PM1533), are only distantly linked to the ST20 isolates of the same farm (see Fig. 3) and, thus, are clearly a result of introduction of a new strain into the chicken population. The ST9 isolates were all obtained during the more recent outbreaks that emerged after almost a 10 year period where the farm was able to limit any impact of fowl cholera using a killed autogenous vaccine based on an ST20 isolate. However, diversity was observed in the LPS type L3 carried by the ST9 isolates on farm A with two putative novel phase variation mechanisms identified, none of which could be detected by any of the other currently in-use typing techniques. The first involved a tandem repeat insertion in the *natC* gene of the 2011 isolates (responsible for switching the *natC* gene off, predicting a failure to add the second *N*-acetylgalactosamine, GalNAcII, to the LPS outer core). The second mechanism was a 1 bp deletion in a homopolymer region of the *gatG* of the 2014 isolates (responsible for switching off the addition of the second galactose, GalII), while having the *natC* gene switched on. In 2013, while one ST9 isolate (PM1778) shared the tandem repeat insertion with the 2011 isolates, the other three isolates obtained from two other chickens (PM1775, PM1776 and PM1777) lacked the insertion in *natC* and carried LPS outer core biosynthesis loci that were identical to each other and those from isolates obtained in 2015 and 2016. These results suggest that there was a reversion to a wild-type functional *natC* gene, which would result in a full-length LPS structure. Identification of the co-existence of isolates with their *natC* gene switched on or off from a single outbreak in this farm is a novel and important finding. The introduction of PM1533, an isolate with the *natC* allele carrying the tandem repeat insertion (switched off), in the vaccine since 2011 was followed by identification of isolates carrying the switched on *natC* allele, as well as those carrying the single nucleotide putative phase variation in their *gatG* gene. This finding shows the capacity of *P. multocida* to switch a glycosyltransferase gene on or off

under pressure, such as that caused by a vaccination programme. This result shows the importance of investigating more than a single isolate per shed/flock to obtain the information on the possible diversity in the LPS outer core structure.

Phase variation has been identified in a number of LOS outer core biosynthesis loci of *H. influenzae* including galactosyltransferase, glucosyltransferase and *O*-acetylase [31, 32]. We have previously reported a virulence-related phase variation in the *latB* gene of *P. multocida* isolates carrying LPS type L4, associated with captive *Petaurus norfolcensis* (squirrel gliders) [20].

The phylogenetic analysis of ST9 isolates from the two farms confirmed that the tandem repeat insertion was acquired prior to the introduction of the organism into each farm, as it is split across the root of the tree and appeared in the most basal taxa of layer farm A isolates. On farm A, the organism reverted to wild-type by deleting a repeat, perhaps as a result of vaccination pressure.

Importantly, analysis of the sequence data indicated that for the 2011, 2013, 2014, 2015 and 2016 outbreaks, the LPS type L3 genotype of the field isolates was not represented in the vaccines used at the time. In particular, for the 2011, 2013, 2015 and 2016 outbreaks, at least one or more of the outbreak isolates were predicted to produce an LPS structure that was longer than any produced by the vaccine strains due to reversion to a wild-type *natC*. It now known that killed fowl cholera vaccines provide protection only against challenge from isolates with an identical LPS structure [12]. Hence, this difference in LPS structure between the vaccine and the field challenge may underpin the failure of the vaccination programme to provide protection. It is important to note that our genomic analysis suggests that the field isolates from the 2014 outbreak had the same LPS structure as the vaccine strain PM979, having a frameshift in their *gatG* gene, albeit at different positions. Hence, in theory, the outbreak should not have occurred as the vaccine was capable of providing protection. This suggests other factors such as stress, possible poor vaccination protocols of the vaccination crew, timing of the vaccination and/or waning immunity have a role to play in the occurrence of fowl cholera outbreaks in the face of a theoretically effective vaccination programme. These results highlight the importance of monitoring the predicted LPS structure as part of a *P. multocida* vaccination programme.

The layer farm B isolates could be divided into five different groups based on presence or absence of major genetic mutations in their LPS outer core biosynthesis loci. LPS type L3 carrying isolates of *P. multocida* belonging to ST7 and ST20 were previously identified in association with both pig and poultry hosts, with ST7 isolates also reported from fatal pasteurellosis in captive pinnipeds [33]. The isolates from these two closely related STs appear to carry different missense mutations in their glycosyltransferase genes. These differences were predicted to produce a different LPS outer structure with the only ST7 isolate from this farm having the capacity to produce a fully truncated LPS due to a newly identified missense mutation in *htpE*. In contrast, the ST20 isolates carried a missense mutation in *gatG*, and were predicted to produce an LPS that would end after the addition of the first galactose molecule, i.e. Hep.[Glc].Glc.Gal, as has been reported [4, 8]. We have previously reported a frameshift in the coding region of the *gatG* gene of ST7 *P. multocida* isolates obtained from diseased pinnipeds in SEQ [33]. These results indicate that ST7 *P. multocida* isolates have the capacity to produce different LPS structures within an LPS PCR type, a capacity also shared with the ST9 *P. multocida* isolates described in the current study.

The observed diversity between the isolates from different sites of the layer company B would make choosing a single autogenous vaccine to be used across all the farms of the complex very challenging. In addition to that, our data suggest that phase variations could contribute to the challenges faced in controlling fowl cholera caused by *P. multocida*. The presence of phase variable genes has been found to play a role in generating phenotypic variants within bacterial populations such as *Campylobacter jejuni* and *Streptococcus suis* [34, 35]. Similarly, our data suggested that phase variation could be contributing to LPS variation/diversity in *P. multocida* and provide the genetic mechanism for vaccine escape. No allelic difference was observed in the LPS type L3 outer core biosynthesis loci carried by the ST9 isolates of layer farm B1, unlike those of layer farm A. All of the investigated ST9 isolates from farm B1 had their *natC* gene switched off due to the three-unit tandem repeat array. The phylogenomic analysis confirms the observed putative phase variation mechanism of *natC* in the ST9 isolates is an ancestral characteristic and the isolates have been independently introduced to the two farms.

SNP detection is an essential step for investigating outbreak dynamics, as well as phenotype prediction [36]. The selection of the SNP difference cut-off that suggests isolates should be considered as the same outbreak remains a major challenge for WGS-based surveillances, as well as outbreak investigations in public health [37]. More information on the population diversity for each particular bacterial species, host and production system is required. For example, the within-outbreak population diversity observed in farm B1 during the current study was higher compared to our previous observation on fowl cholera outbreaks on a meat chicken farm [4]. This might reflect the effect of time on the population diversity of *P. multocida*, as layer chickens live much longer (typically over 50 weeks of age) than meat chickens (typically less than 50 days). The observation that overlapping outbreaks in 2013 were seeded by different ST9 subpopulations on farm A suggests that arbitrary SNP cut-offs should be interpreted in light of the overall phylogeny and not in isolation.

The genomic data also demonstrated the limited capacity of MLST and the DNA fingerprinting method, rep-PCR, in differentiating isolates carrying different genetic mutations in their LPS-encoding region, which would result in different LPS outer structure. We have previously reported the failure of rep-PCR in differentiating between two clades of ST20 *P. multocida* isolates that carried LPS types L3 or L1 [4].

The knowledge on the LPS outer core biosynthesis loci diversity gained in the current study and previous studies on poultry [4], pinnipeds [33] and squirrel gliders [20] has important implications in strain selection for killed autogenous vaccination for multiple host species. Importantly, the data could be used for application of metagenomic analyses and point of care diagnostics in the near future in investigations of diversity in the LPS outer core biosynthesis loci in disease outbreaks linked to *P. multocida*; as well as surveillance studies looking at asymptomatic carrier hosts. Genomic studies of two previously linked ST20 outbreaks in farm A revealed that each of the two were as a result of introduction of new strain. In addition, by using genomics, we reveal two newly identified potential phase variation mechanisms in the LPS outer core biosynthesis loci of LPS type L3 carried by ST9 *P. multocida*.

Despite the demonstrable benefits of WGS in this study, we acknowledge that our interpretation is limited by the selection of single isolates from each chicken in most cases. Previously, we have shown that the *P. multocida* core-genome diversity within chickens can exceed the between-chicken diversity, hampering outbreak investigation with WGS [4]. This limitation could be circumvented by collecting multiple samples for WGS from each chicken and/or supplementing the study with samples for shotgun metagenomic sequencing.

Funding information
This project was co-funded by AgriFutures Australia and Australian Eggs.

Acknowledgements
We thank the farms management and the veterinary consultants linked to the study farms for their assistance and expertise. This research was supported by QRIScloud and by use of the Nectar Research Cloud. The Nectar Research Cloud is a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy (NCRIS). All the sequencing data used in this study were generated within the study by L. Omaleki, P. J. Blackall, T. Cuddihy, R.T. White, J. M. Courtice, C. Turni, B. M. Forde and S. A. Beatson (NCBI BioProject number PRJNA643588).

Author contributions
The study was conceptualized by L.O., P.J.B., S.A.B., B.M.F, J.M.C., and C.T. The information on the outbreaks and vaccines was provided by J.M.C. Whole-genome analysis was performed by L.O. with assistance from S.A.B., T.C., R.T.W., and B.M.F. The analysis interpretations were performed by S.A.B., B.M.F., L.O., and P.J.B. The manuscript was drafted by L.O., and reviewed and edited by P.J.B., T.C., R.T.W., S.A.B., B.M.F., J.M.C., and C.T. All authors have read and approved the final version of the manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
This work was undertaken on stored bacterial cultures and no samples were collected from any birds during this study.

References

1. Blackall PJ, Hofacre CL. Fowl cholera. In: Swayne DE, Boulianne M, Logue CM, McDougald LR and Nair VL (eds). *Diseases of Poultry*. Hoboken, NJ: Wiley; 2020. pp. 831–845.
2. Wilson BA, Ho M. *Pasteurella multocida*: from zoonosis to cellular microbiology. *Clin Microbiol Rev* 2013;26:631–655.
3. Singh R, Blackall PJ, Remington B, Turni C. Studies on the presence and persistence of *Pasteurella multocida* serovars and genotypes in fowl cholera outbreaks. *Avian Pathol* 2013;42:581–585.
4. Omaleki L, Blackall PJ, Cuddihy T, Beatson SA, Forde BM, *et al*. Using genomics to understand inter- and intra- outbreak diversity of *Pasteurella multocida* isolates associated with fowl cholera in meat chickens. *Microb Genom* 2020;6:000346.
5. Brogden KA, Rhoades KR, Heddleston KL. A new serotype of *Pasteurella multocida* associated with fowl cholera. *Avian Dis* 1978;22:185–190.
6. Heddleston KL, Gallagher JE, Rebers PA. Fowl cholera: gel diffusion precipitin test for serotyping *Pasteurella multocida* from avian species. *Avian Dis* 1972;16:925–936.
7. Harper M, John M, Turni C, Edmunds M, St Michael F, *et al*. Development of a rapid multiplex PCR assay to genotype *Pasteurella multocida* strains by use of the lipopolysaccharide outer core biosynthesis locus. *J Clin Microbiol* 2015;53:477–485.
8. Harper M, St Michael F, John M, Vinogradov E, Steen JA, *et al*. *Pasteurella multocida* Heddleston serovar 3 and 4 strains share a common lipopolysaccharide biosynthesis locus but display both inter- and intrastrain lipopolysaccharide heterogeneity. *J Bacteriol* 2013;195:4854–4864.
9. Turni C, Singh R, Blackall PJ. Genotypic diversity of *Pasteurella multocida* isolates from pigs and poultry in Australia. *Aust Vet J* 2018;96:390–394.
10. Harper M, Boyce JD, Cox AD, St Michael F, Wilkie IW, *et al*. *Pasteurella multocida* expresses two lipopolysaccharide glycoforms simultaneously, but only a single form is required for virulence: identification of two acceptor-specific heptosyl I transferases. *Infect Immun* 2007;75:3885–3893.
11. Nahar N, Turni C, Tram G, Blackall PJ, Atack JM. *Actinobacillus pleuropneumoniae*: the molecular determinants of virulence and pathogenesis. *Adv Microb Physiol* 2021;78:179–216.
12. Harper M, John M, Edmunds M, Wright A, Ford M, *et al*. Protective efficacy afforded by live *Pasteurella multocida* vaccines in chickens is independent of lipopolysaccharide outer core structure. *Vaccine* 2016;34:1696–1703.
13. Walker MJ, Beatson SA. Outsmarting outbreaks. *Science* 2012;338:1161–1162.
14. Zhang P, Fegan N, Fraser I, Duffy P, Bowles RE, *et al*. Molecular epidemiology of two fowl cholera outbreaks on a free-range chicken layer farm. *J Vet Diagn Invest* 2004;16:458–460.
15. Gunawardana GA, Townsend KM, Frost AJ. Molecular characterisation of avian *Pasteurella multocida* isolates from Australia and Vietnam by REP-PCR and PFGE. *Vet Microbiol* 2000;72:97–109.
16. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
17. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
18. Jolley KA, Maiden MCJ. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 2010;11:595.
19. Subaaharan S, Blackall LL, Blackall PJ. Development of a multilocus sequence typing scheme for avian isolates of *Pasteurella multocida*. *Vet Microbiol* 2010;141:354–361.

20. Omaleki L, Beatson SA, Thomrongsuwannakij T, Blackall PJ, Buller NB, *et al*. Phase variation in *latB* associated with a fatal *Pasteurella multocida* outbreak in captive squirrel gliders. *Vet Microbiol* 2020;243:108612.
21. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, *et al*. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23:2947–2948.
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359.
23. Langmead B, Wilks C, Antonescu V, Charles R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* 2019;35:421–432.
24. Carver T, Harris SR, Otto TD, Berriman M, Parkhill J, *et al*. BamView: visualizing and interpretation of next-generation sequencing read alignments. *Brief Bioinform* 2013;14:203–212.
25. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 2012;28:464–469.
26. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
27. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 2014;15:524.
28. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
29. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, *et al*. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* 2018;34:292–293.
30. Harper M, Boyce JD. The myriad properties of *Pasteurella multocida* lipopolysaccharide. *Toxins* 2017;9:E254.
31. Fox KL, Atack JM, Srikhanta YN, Eckert A, Novotny LA, *et al*. Selection for phase variation of LOS biosynthetic genes frequently occurs in progression of non-typeable *Haemophilus influenzae* infection from the nasopharynx to the middle ear of human patients. *PLoS One* 2014;9:e90505.
32. Phillips ZN, Brizuela C, Jennison AV, Staples M, Grimwood K, *et al*. Analysis of invasive nontypeable *Haemophilus influenzae* isolates reveals selection for the expression state of particular phase-variable lipooligosaccharide biosynthetic genes. *Infect Immun* 2019;87:e00093-19.
33. Crawford RL, Blyde D, Blackall PJ, Forde BM, Beatson SA, *et al*. Novel insights into pasteurellosis in captive pinnipeds. *Vet Microbiol* 2019;231:232–237.
34. Lango-Scholey L, Aidley J, Woodacre A, Jones MA, Bayliss CD. High throughput method for analysis of repeat number for 28 phase variable loci of *Campylobacter jejuni* strain NCTC11168. *PLoS One* 2016;11:e0159634.
35. Atack JM, Weinert LA, Tucker AW, Husna AU, Wileman TM, *et al*. *Streptococcus suis* contains multiple phase-variable methyltransferases that show a discrete lineage distribution. *Nucleic Acids Res* 2018;46:11466–11476.
36. Bush SJ, Foster D, Eyre DW, Clark EL, De Maio N, *et al*. Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism-calling pipelines. *Gigascience* 2020;9:giaa007.
37. Payne M, Octavia S, Luu LDW, Sotomayor-Castillo C, Wang Q, *et al*. Enhancing genomics-based outbreak detection of endemic *Salmonella enterica* serovar Typhimurium using dynamic thresholds. *Microb Genom* 2021;7:000310.