Annals of Microbiology (2019) 69:151–160
https://doi.org/10.1007/s13213-018-1410-7

ORIGINAL ARTICLE

Genome sequence analysis of the Indian strain *Mannheimia haemolytica* serotype A2 from ovine pneumonic pasteurellosis

Swati Sahay[1,2] · Rajeswari Shome[1] · Jagadesan Sankarasubramanian[3] · Udayakumar S. Vishnu[3] · Awadhesh Prajapati[1] · Krithiga Natesan[1] · Bibek Ranjan Shome[1] · Habibur Rahman[4] · Jeyaprakash Rajendhran[3]

Received: 7 April 2018 / Accepted: 21 November 2018 / Published online: 13 December 2018
© Università degli studi di Milano 2018

Abstract

Mannheimia haemolytica is a leading causative agent of pasteurellosis in ruminants. Genome of *M. haemolytica* strains from different hosts has been sequenced worldwide to understand its pathogenesis. There are only few reports on the isolation of *M. haemolytica* in India with limited information on its molecular characteristics. The present study focuses on genome sequence analysis of a *M. haemolytica* strain isolated from pneumonic sheep. *Mannheimia haemolytica* A2 strain NIVEDI/MH/1 was isolated and identified by species and serotype-specific PCRs. Whole genome sequencing was performed using the Ion Torrent Personal Genome Machine. A comparative genomic analysis was performed to understand the virulence determinants of the Indian strain and its phylogenetic relationship with other global strains. Sequence data revealed a draft genome of 2,211,426 bp size with 41.3% GC content, assembled into 17 contigs, and contained 2379 genes. Five genomic islands identified in the genome showed high sequence identity with other respiratory pathogens of the Pasteurellaceae family. Phylogenetic analysis showed *M. haemolytica* A2 NIVEDI/MH/1 is very close to a *M. haemolytica* A2 strain from pneumonic calf. Further, the analysis revealed the presence of virulence, metal-, and multidrug resistance genes needed for pathogenesis and survival of the bacteria during infection. Also, we identified the presence of type I-C and type II-C of CRISPR-Cas arrays in the present sequenced genome. The study emphasizes the role of *M. haemolytica* in respiratory infections of ruminants in the Indian subcontinent and indicates the role of vertical and horizontal gene pools in pathogenicity and survivability of the bacteria.

Keywords *M. haemolytica* · India · Pneumonia · Genome · Pathogenic island · CRISPR-Cas array

Introduction

Mannheimia haemolytica is a Gram-negative, weakly hemolytic coccobacillus (Highlander 2001). It is an important etiological agent of pasteurellosis, also known as respiratory mannheimiosis (Mohamed and Abdelsalam 2008) and mastitis in ruminants (Omaleki et al. 2011, 2016). Based on the capsular antigens, *M. haemolytica* is classified into 12 serotypes (A1, A2, A5-A9, A12- A14, A16, and A17) (Klima et al. 2017). Of these, the serotypes A1 and A6, responsible for the bovine respiratory disease (BRD) in cattle, are extensively studied, whereas the serotypes A2 and A7, causing pneumonic pasteurellosis in small ruminants, are least characterized (Highlander 2001). Stress due to unfavorable environmental conditions, animal transportation, and prior viral and bacterial infections act as predisposing factors in triggering this opportunistic pathogen to cause disease outbreaks (Timsit et al. 2013). The molecular mechanisms involved in the conversion of this bacteria from commensal to pathogen are still unclear (Gioia et al. 2006).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13213-018-1410-7) contains supplementary material, which is available to authorized users.

✉ Jeyaprakash Rajendhran
jrajendhran@gmail.com

[1] Indian Council of Agricultural Research-National Institute of Veterinary Epidemiology and Disease informatics (ICAR-NIVEDI), Yelahanka, Bengaluru 560064, India

[2] Department of Microbiology, Centre for Research in Pure and Applied Sciences, Jain University, Bangalore 560011, India

[3] Department of Genetics, School of Biological Sciences, Madurai Kamaraj University, Madurai, Tamil Nadu 625021, India

[4] International Livestock Research Institute, Block-C, First Floor, NASC Complex, CG Centre, DPS Marg, Pusa, New Delhi 110012, India

The virulence factors determine the pathogenicity of an organism (Mohamed and Abdelsalam 2008). These virulence factors of *M. haemolytica* include ruminant-specific leukotoxin, anti-phagocytic capsule, lipopolysaccharide, iron-regulated outer membrane proteins, lipoproteins, sialoglycoprotease, neuraminidase, and two immunoglobulin proteases (Highlander 2001). In addition, the emergence of multi-antibiotic resistance to beta-lactams, tetracycline, streptomycin, sulfonamides, macrolides, and sulfamethazine in *M. haemolytica* has been reported (Hörmansdorfer and Bauer 1996).

Genomes of different serotypes of *M. haemolytica* were sequenced to identify novel targets for diagnostics and vaccine development (Klima et al. 2016), and to understand the phylogenetic relationship with other strains reported across the globe (Moustafa et al. 2015). Genome sequence analysis of *M. haemolytica* A1, strain ATCC BAA-410, from a lung of a calf that died of BRD identified the presence of virulence, natural competence, accessory, and transcription regulation genes and revealed widespread genes conserved in *M. haemolytica* as with other members of the Pasteurellaceae family. The study also explained the similarity of *M. haemolytica* virulence proteins with *Neisseria meningitidis*, suggesting a convergent evolution (Gioia et al. 2006). Lawrence et al. (2010b) reported the presence of phage regions and pseudogenes in the genome of *M. haemolytica* serotype A2, indicating pathogenic islands as a consequence of the horizontal gene transfer (HGT)). HGT is an important factor in the emergence of the pathogen by the evolution of the novel virulence genes and the spread of antimicrobial and metal resistance genes (Keen 2012). A previous study has substantiated the role of *Mannheimia* genomic islands (GIs) and HGT in the expression and regulation of virulence factors (Rao and Jayakumar 2017). Antibiotic and metal resistance genes were found to be associated with plasmids, chromosome (Watts et al. 1994), and integrative conjugative elements (ICEs) (Eidam et al. 2015). *M. haemolytica* genome also contains the CRISPR-Cas systems (clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins) (Klima et al. 2016), which act as an adaptive immune system for bacterial protection against invading genetic elements like phages and plasmids (Makarova et al. 2011).

There are fewer reports on the prevalence of *M. haemolytica* in India (Sharma et al. 2011; Dar et al. 2012; Kumar et al. 2015). However, no reports are available on the genome sequencing and the molecular characterization of *M. haemolytica* isolated from India. Hence, the present study was focused on the whole genome sequence analysis of *M. haemolytica* isolated from a pneumonic sheep to unravel the genes involved in the virulence and its phylogenetic relatedness to the global isolates.

Materials and methods

Sheep herd The sheep samples were collected from a private herd at Chikkajala village, Urban Bengaluru, Karnataka. The herd was maintained with a semi-intensive rearing method. The animals were not vaccinated and a sudden death of six sheep due to the respiratory infection was reported. The sample was collected from the sheep having the typical case of uncomplicated pneumonia with symptoms including dyspnea, pyrexia, and mucopurulent nasal discharge.

Isolation of *M. haemolytica* and extraction of genomic DNA

M. haemolytica was isolated from the nasopharynx of the sheep (*Ovis aries*) with a history of pneumonia. The isolate was tentatively identified as *M. haemolytica* based on conventional culture characteristics including colony morphology and standard biochemical tests (Marru et al. 2013), and confirmed by species and capsule-specific PCR assays (Alexander et al. 2008; Klima et al. 2017). The *M. haemolytica* A2 isolate was designated as strain NIVEDI/MH/1. The strain was initially cultured on tryptic soy agar (TSA) supplemented with 5–7% sheep blood at 37 °C for 24 h and subsequently subcultured on brain heart infusion agar (BHI) for genomic DNA extraction. DNA was extracted using the DNeasy kit as per the manufacturer's protocol (Qiagen, Hilden, Germany). The quantity and quality of the extracted DNA were determined by using NanoDrop2000 (Thermo Scientific, Waltham, USA) and by agarose gel (0.8%) electrophoresis, respectively.

Genome sequencing, assembly, and annotation

The whole genome of *M. haemolytica* strain A2 NIVEDI/MH/1 was sequenced using the Ion Torrent Personal Genome Machine (Life Technologies, Carlsbad, CA). The de novo assembly was performed using MIRA (Mimicking Intelligent Read Assembly) version 3.9.18 (Chevreux et al. 1999). For the genome annotation, the assembled sequence of *M. haemolytica* A2 NIVEDI/MH/1 was submitted to the NCBI Prokaryotic Genomes Annotation Pipeline (NCBI PGAP) (http://www.ncbi.nlm.nih.gov/ genome/ annotation prok/ process/). Further, annotation was also performed by using Rapid Annotations Subsystems Technology (RAST) (Aziz et al. 2008). The rRNAs and tRNAs genes were predicted using RNAmmer (Lagesen et al. 2007) and tRNAscan-SE 1.21 (Lowe and Eddy 1997), respectively. Protein-coding sequences were analyzed by the Cluster of Orthologous Group (COG) database (Tatusov et al. 2000) on WebMGA (Wu et al. 2011). Graphical view of the genome was generated using Artemis and DNAplotter. The total number of

contigs, their predicted coding sequences in forward and reverse direction, rRNA, and tRNA were plotted.

16S rRNA and genome-based phylogenetic analysis

The 16S rRNA gene from *M. haemolytica* A2 NIVEDI/MH/1 genome was identified using RNAmmer. The 16S rRNA genes of other genome sequences were retrieved from the NCBI RefSeq database (Supplementary Table S1). A phylogenetic tree based on the 16S rRNA genes was constructed using Molecular Evolutionary Genetics Analysis (MEGA) version 7 (Tamura et al. 2013) by the maximum likelihood method with the default bootstrap value as 500. The whole genome sequences of 62 *M. haemolytica*, 3 *Mannheimia varigena*, and one each of *Mannheimia granulomatis*, *Mannheimia massilioguelmaensis*, *Mannheimia succiniciproducens*, and *Mannheimia* sp. were retrieved from NCBI (Supplementary Table S1). These genome sequences were compared with the genome sequence of *M. haemolytica* A2 NIVEDI/MH/1. The Composition Vector (CV) method (Qi et al. 2004) was used to generate the distance matrix. The whole genome-based phylogenetic tree was constructed using the Neighbor-Joining (NJ) method and visualized using MEGA version 7.

Identification and prediction of virulent genes and genomic islands

The virulence factors were predicted using the Virulence Factor Database (VFDB) (Chen et al. 2016). The Genomic Islands (GIs) were envisaged using Island viewer 4, a computational tool integrated with four different genomic island prediction methods such as IslandPath-DIMOB, SIGI-HMM, Islander, and IslandPick. The IslandPath-DIMOB is based on several common characteristics of GIs such as abnormal sequence composition and presence of mobility genes (Hsiao et al. 2003). The SIGI-HMM is based on codon usage bias with the Hidden Markov Model approach (Waack et al. 2006). The Islander is based on mechanistic consequences of their typical site-specific integration into tRNA/tmRNA genes (Hudson et al. 2014), and the IslandPick is based on a comparative genomic approach (Langille et al. 2008). The Panseq server (Laing et al. 2010) was used to find the novel regions in the predicted GIs of *M. haemolytica* A2 NIVEDI/MH/1 by comparing with other existing genome sequences.

Prediction of the CRISPR-Cas system

The CRISPR-Cas system was predicted with the use of web tools CRISPRFinder (Grissa et al. 2007) and CRISPRone (Zhang and Ye 2017). CRISPRFinder, a web service tool, was used to detect CRISPR-Cas array in the sequence and define the direct repeats and spacers. CRISPRone was used for the annotation of the predicted CRISPR-Cas system to detect Cas genes and type of the CRISPR-Cas system present in the *M. haemolytica* A2 NIVEDI/MH/1 genome.

Results and discussion

Mannheimiosis is an emerging infectious disease of small ruminants in India. This study was focused on the isolation, identification, and whole genome sequencing of the *M. haemolytica* serotype A2 strain from a pneumonic sheep aimed to unravel its genetic background circulating in Indian terrain. The whole genome sequencing of *M. haemolytica* A2 NIVEDI/MH/1 using Ion Torrent PGM yielded a total of 1,556,783 reads with an average read length of 167 bp. The total sequence output was ~ 260 Mbp, which is approximately 189-fold coverage. The de novo assembly using MIRA yielded 17 contigs, and the largest contig was 586,935 bp long. The draft genome of *M. haemolytica* A2 NIVEDI/MH/1 was 2,211,426 bp with 41.32% GC content. Genome annotation by RAST predicted a total of 2379 protein-coding regions. Of these, 2023 genes were predicted with functions, and 356 genes were annotated as hypothetical proteins. RNAmmer and tRNAscan-SE 1.21 predicted a total of 17 rRNA and 58 tRNA genes, respectively (Fig. 1; Table 1). The COG category of the predicted genes is shown in Table 2.

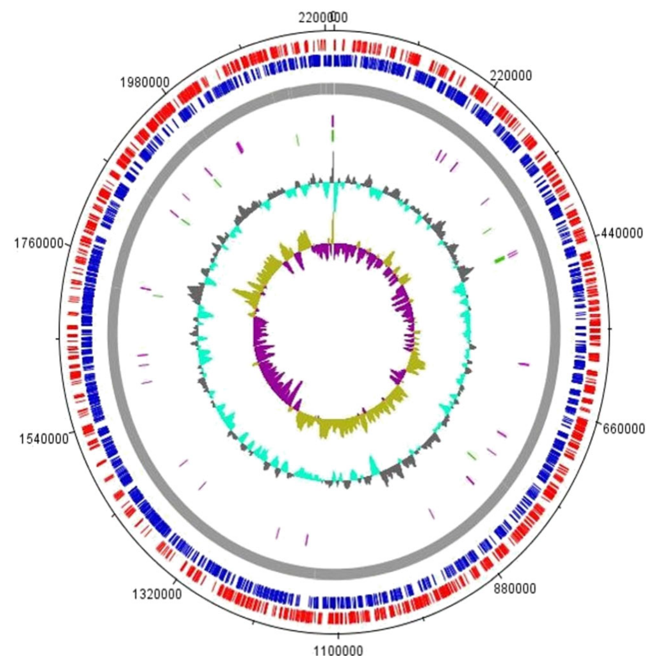


Fig. 1 Circular representation of *M. haemolytica* A2 NIVEDI/MH/1 genome. From the outer circle to the inner circle, black indicates the total base pair, red indicates forward strand, blue indicates reverse strand, gray indicates a total number of contigs, pink indicates tRNA, green indicates rRNA, gray and light blue indicate GCskew, and green and pink indicate GCplot

Table 1 Genome statistics of *M. haemolytica* A2 NIVEDI/MH/1

Attribute	Value	Percent[a]
Genome size (bp)	2,211,426	100
DNA G + C content (bp)	–	41.32
Total genes	2460	100
RNA genes	81	3.29
rRNA genes	19	0.77
tRNA genes	58	2.36
ncRNAs	4	0.16
Protein-coding genes	2379	96.71
Genes with function prediction (protein)	2023	82.23
Genes assigned to COGs	2152	87.46
CRISPR arrays	2	0.09

[a] These values are based on either the size of the genome in base pairs or the total number of genes in the annotated genome

The 16S rRNA gene sequence can be used for the species identification and phylogeny analysis (Clarridge 2004). However, the whole genome sequences are generally preferred to study the phylogenetic relatedness on the basis of geographical origin, serotypes, and host predilection (Moustafa et al. 2015). The 16S rRNA gene-based phylogenetic tree confirmed that the strain NIVEDI/MH/1 belongs to the species *M. haemolytica* as it was clustered with other strains of *M. haemolytica* in the same clade (Fig. 2). The whole genome-based phylogenetic tree constructed using 70 *Mannheimia* strains showed that the strain *M. haemolytica* A2 NIVEDI/MH/1 was the closest neighbor of *M. haemolytica* D35 serotype A2 (accession no. AUNK00000000), which was isolated from a pneumonic calf lung (Hauglund et al. 2015). Similarly, the *M. haemolytica* serotype A2 ovine strain from a pneumonic lung (accession no. ACZX00000000) (Lawrence et al. 2010a) was closely related to the *M. haemolytica* A2 NIVEDI/MH/1 strain and clustered together in the same clade. The phylogenetic tree based on both the 16S rRNA gene and whole genome sequence showed the similarity of *M. granulomatis* with the *M. haemolytica* A1/A6 PLK10 strain (Supplementary Fig. S1) and three *M. varigena* strains forming a separate clade but clustering with *Mannheimia* sp.

A total of 33 virulent genes were predicted by the VFDB database (Supplementary File S2). Of these, *lpxA* (ODQ38859), *lpxH* (ODQ37174), *lpxA* (ODQ38611), *lpxB* (ODQ35950), *rfaE* (ODQ38829), *wecA*(ODQ36883), *waaQ* (ODQ37174), and *lsgE* (ODQ37931) are needed for the biosynthesis of lipopolysaccharide (LPS). Previous studies have reported that all these genes are conserved among the members of the Pasteurellaceae family (Xu et al. 2008; Peng et al.

Table 2 Number of genes associated with general COG functional categories

COG class	Value	Percentage[a]	Description
A	1	0.04	RNA processing and modification
C	149	6.26	Energy production and conversion
D	25	1.05	Cell cycle control, cell division, chromosome partitioning
E	209	8.79	Amino acid transport and metabolism
F	76	3.19	Nucleotide transport and metabolism
G	201	8.45	Carbohydrate transport and metabolism
H	112	4.71	Coenzyme transport and metabolism
I	42	1.77	Lipid transport and metabolism
J	167	7.02	Translation, ribosomal structure, and biogenesis
K	113	4.75	Transcription
L	124	5.21	Replication, recombination, and repair
M	142	5.97	Cell wall/membrane/envelope biogenesis
N	5	0.21	Cell motility
O	102	4.29	Posttranslational modification, protein turnover, chaperones
P	136	5.72	Inorganic ion transport and metabolism
Q	21	0.88	Secondary metabolites biosynthesis, transport, and catabolism
R	222	9.33	General function prediction only
S	185	7.78	Function unknown
T	46	1.93	Signal transduction mechanisms
U	50	2.10	Intracellular trafficking, secretion, and vesicular transport
V	22	0.92	Defense mechanisms
W	2	0.08	Extracellular structures

[a] Percentage of annotated genes. The total is based on the total number of protein-coding genes in the genome

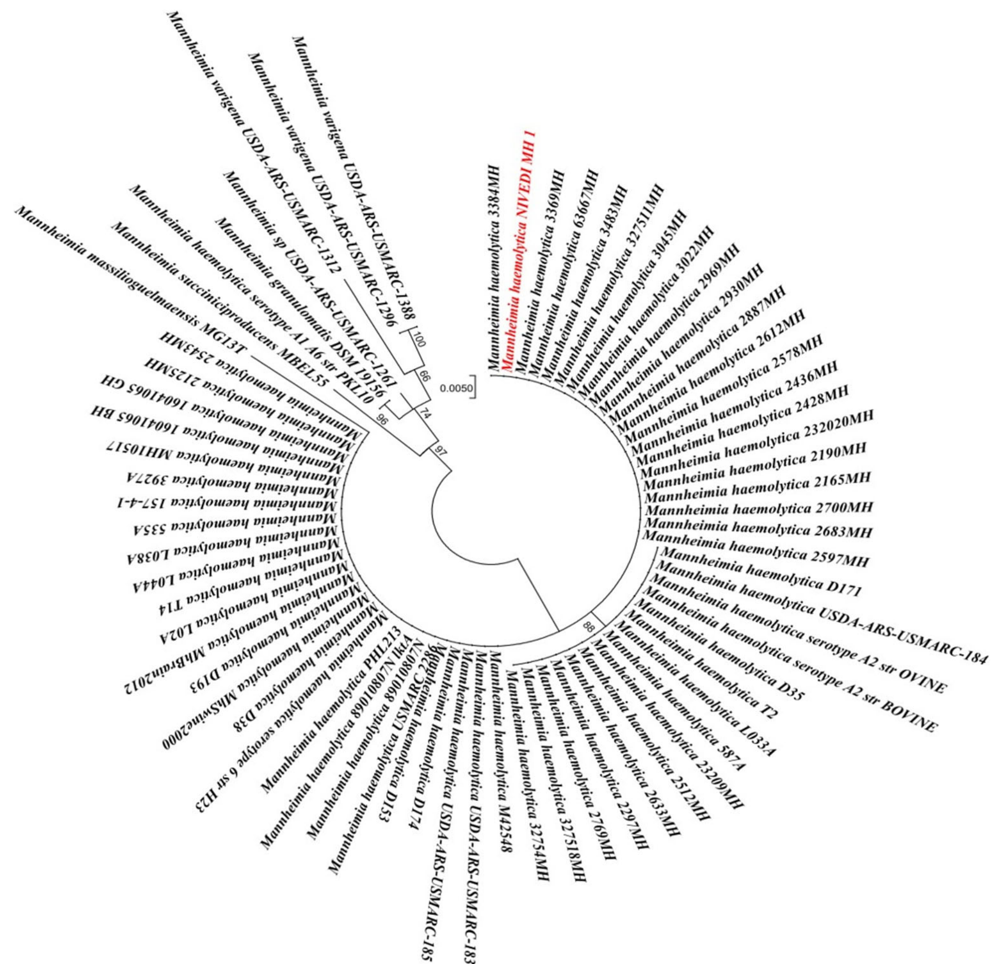


Fig. 2 16S rRNA gene-based phylogenetic analysis of *M. haemolytica* A2 NIVEDI/MH/1 and other *Mannheimia* sp. genome retrieved from the NCBI RefSeq database (access date: 2/05/17) using the Maximum likelihood method in the MEGA software version 7. The analysis revealed that *M. haemolytica* A2 NIVEDI/MH/1 isolate (in red) clustered to the same clade with other *M. haemolytica*

2016). LPS is an endotoxin responsible for the stimulation of IL-1beta, IL-8, and neutrophil influx leading to inflammation and damage of bovine pulmonary endothelial cells (Paulsen et al. 1989; Yoo et al. 1995; Lafleur et al. 1998). The gene *kdkA* (ODQ37905) is needed for the LPS glycoform generation (Peng et al. 2016). According to a previous study, LPS enhances the cytolytic activity of leukotoxin, (repeats in-*to*xin (RTX toxin)) of *M. haemolytica* (Lafleur et al. 1998). RAST annotation identified 28 genes are having virulence potential (Supplementary File S3). NCBI PGAP and RAST annotations identified the leukotoxin-activating lysine-acyltransferase *LktC* (ODQ39212) needed for the posttranslational acylation of the leukotoxin protein encoded by *LktA* in the *LktCABD* operon (Highlander et al. 1989).

The presence of *NanA* (ODQ39248), *NanK* (ODQ39247), and a hypothetical protein (ODQ37835) having similarity with sialidase (neuraminidase) are needed for the sialometabolism that helps bacteria to colonize, persist, and cause disease in mammalian species (Sakarya and Oncü 2003; Steenbergen et al. 2005; Severi et al. 2007) were identified. The gene for UDP-N-acetylglucosamine 2-epimerase (ODQ37932.1), a rate-limiting enzyme in sialometabolism (Klima et al. 2016), was identified in *M. haemolytica* A2 NIVEDI/MH/1. Also, adhesin (ODQ39272) and a type IV pilus biogenesis/stability protein PilW (ODQ38480) required for adhesion and motility as noticed in *Haemophilus influenzae*, *Pseudomonas aeruginosa*, and *Neisseria* sp. were identified (Bakaletz et al. 2005). The genome contained the TonB transport system, TonB-ExbB-ExbD (ODQ38608, ODQ38777, and ODQ38778) (Peng et al. 2016), ATP-binding iron ABC transporter (ODQ38724, ODQ38726, and ODQ38478), heme utilization proteins (HutZ and Hut W) (ODQ38471, ODQ38473), and iron-binding protein (sitA) (ODQ38479) needed for the iron acquisition. It also contained a ferric uptake regulator (Fur family) (ODQ38822), a transcriptional repressor required for regulation of iron acquisition genes. Other putative virulence factors are superoxide dismutase (ODQ38739), virulence protein (ODQ38476), antitoxin, component of a toxin-antitoxin (TA) (ODQ38475), lipoprotein localization factor LolB (ODQ37719), S-ribosyl homocysteine lyase, *Lux S* (ODQ38823), secretin, and ComE (ODQ38752) (Yu et al. 2016; Klima et al. 2016).

Five GIs were predicted in the genome of *M. haemolytica* A2 NIVEDI/MH/1 by the Island Viewer4, which codes for 100 genes. Among these, 67 were annotated as functional genes and 33 for hypothetical proteins (Fig. 3). Comparison

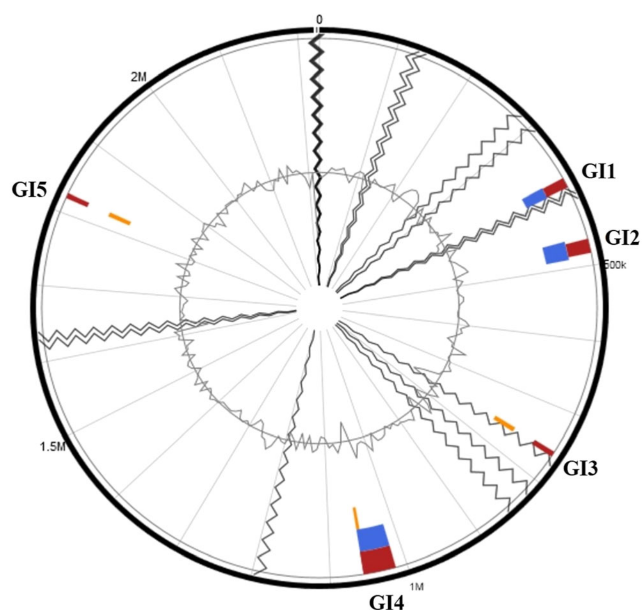


Fig. 3 Five GIs predicted in *M. haemolytica* A2 NIVEDI/MH/1 using IslandViewer. The black circle represents the total base pair, the inner circle represents the GC content, and the curved line indicates the total number of contigs. Orange, blue, and red represent the predicted GIs as per SIGI-HMM, IslandPath-DIMOB, and predicted by at least one method (unstated) within the island viewer tool, respectively

of *M. haemolytica* A2 NIVEDI/MH/1 with other *M. haemolytica* genome sequences using Panseq revealed that the GI3, GI4, and GI5 are unique to NIVEDI/MH/1. Among identified GIs, three regions showed higher similarity with *Bibersteinia trehalosi*, a bovine respiratory pathogen belonging to the Pasteurellaceae family (Blackall et al. 2007). The other two regions showed higher similarity with *Actinobacillus suis*, a swine pathogen (Michael J. Yaeger 1996) and *Ornithobacterium rhinotracheale*, a poultry respiratory pathogen (Thachil et al. 2009) (Table 3). Functional genes identified in the GIs include the plasmid stabilization system protein (ODQ39236), transposase (ODQ39232, ODQ38617), SprT family protein (ODQ38548), toxin MazF (ODQ38536), and various transcriptional regulators. It was found that GI2 has a gene for mercuric reductase (ODQ39227) needed for the mercury resistance by reduction of mercuric to elemental Hg (Freedman et al. 2012). Also, multi-copper polyphenol oxidoreductase (ODQ37742) (Guo et al. 2016), tellurite resistance methyltransferase *TehB* (ODQ38795), arsenate reductase (ODQ37201) (Jackson and Dugas 2003), and multidrug transporter (ODQ39218) genes were present in the GIs. Presence of these resistance genes enhances adaptability and helps the bacteria to survive in a changing hostile environment during infection. Thus, the horizontal transfer of genes among the members of the polymicrobial community might be responsible for the continuous exchange of the gene pool among the respiratory tract bacteria (Schroeder et al. 2017). HGT might lead to the development of virulence (Ho Sui et al. 2009), antibiotic resistance, and metal resistance (Hall 2010; Gilmore et al. 2013; Pagano et al. 2016).

The genome of *M. haemolytica* A2 NIVEDI/MH/1 also contains CRISPR-Cas array. The CRISPR-Cas is the adaptive immune system of the bacteria against invading phage and plasmid, which comprises of conserved direct repeat (DR) sequences, variable spacer sequences derived from the phage DNA, and Cas proteins (Koonin and Makarova 2009; Horvath and Barrangou 2010; Deveau et al. 2010). The CRISPR-Cas system provides immunity to bacteria against these external elements via the three-stage process of adaptation, expression, and interference (Karginov and Hannon 2010; Makarova et al. 2011). Besides the adaptive immunity against phages, the CRISPR-Cas system is also associated with gene expression and regulation of cellular processes like biofilm formation, spore formation, replicon maintenance and segregation, and DNA repair-recombination (Szczepankowska 2012). The CRISPR-Cas system is divided into three types and further into a few subtypes. Cas proteins adjacent to CRISPR sequences have the characteristic of the nucleases, helicase, polymerase, and various RNA-binding proteins (Jansen et al. 2002). Earlier, the presence of type I-C/Dvulg CRISPR-Cas system in serotypes A1, A2, and A6 of *M. haemolytica* were reported (Klima et al. 2016). With the use of CRISPRfinder, we identified the presence of two confirmed and two questionable CRISPR-Cas arrays in the genome of *M. haemolytica* NIVEDI/MH/1. Contig 4 and contig 7 of the genome contained type I-C and type II-C of the confirmed CRISPR-Cas array system, respectively. CRISPRone predicted the presence of various Cas proteins in CRISPR-Cas

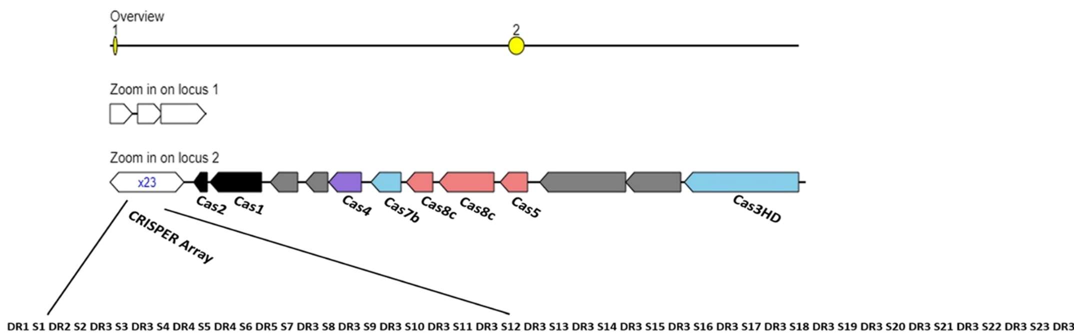
Table 3 Identification of genomic islands (GIs) in *M. haemolytica* A2 NIVEDI/MH/1

Accession ID	GI	Island region	GI (bp)	Locus start	Locus end	Identity (%)	HGT
MEHR01000006	GI1	4294-17134	12,841	BHC25_09075	BHC25_08985	93	*Actinobacillus suis*
MEHR01000001	GI2	283713-302447	18,735	BHC25_01545	BHC25_01465	93	*Bibersteinia trehalosi*
MEHR01000001	GI3	3720-10980	7261	BHC25_00065	BHC25_00025	88	*Ornithobacterium rhinotracheale*
MEHR01000003	GI4	167099-209087	41,989	BHC25_04710	BHC25_04925	93	*Bibersteinia trehalosi*
MEHR01000004	GI5	186571-193348	6778	BHC25_06380	BHC25_06425	81	*Bibersteinia trehalosi*

a ♦ **Seq 1: MEHR01000004.1_1**

⇒**Summary: seq-len = 586935 bp; # of CRISPR array = 1; # of cas gene = 11 (type I, IV)**

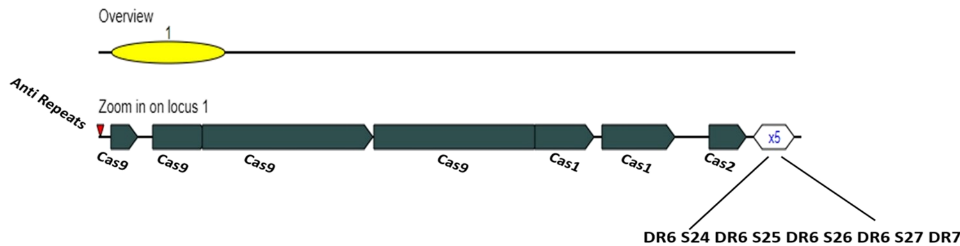
⇒Visulization of predicted CRISPR-Cas system(s)



♦ **Seq 1: MEHR01000007.1_1**

⇒**Summary: seq-len = 31096 bp; # of CRISPR array = 1; # of cas gene = 7 (type II)**

⇒Visulization of predicted CRISPR-Cas system(s)



b

Spacer	Sequence	Contig
S1	GTGCTATTGATTTAACTACTTCAGCATCTTGTAA	MEHR01000004.1
S2	ATTATCCAGCCGTTTGAGCAGGAGGTGAAGTGATG	MEHR01000004.1
S3	AGTTTAAAGAGACTTTAAAAATGGCAAAGGATTTA	MEHR01000004.1
S4	CGATGACTAAAACTAAACACAAACACCCACTTGCA	MEHR01000004.1
S5	ATCGTTACCTTTAATTAAAGGGTTAAACTTAGA	MEHR01000004.1
S6	TTTGACAAAATACACTCAACACCACAGCCATTGAT	MEHR01000004.1
S7	GCCTTAACCGAACGTGAGGGCAAAAGCGAAAAC	MEHR01000004.1
S8	CTATCCATAATCTGCTCTTTGCTCATCTCCAGCGA	MEHR01000004.1
S9	ATATCACAACCCCAAACGGCAAGAGAATTAAGA	MEHR01000004.1
S10	TCCGAATATCCCTCTTTTCGATGGCTTCTTGGAG	MEHR01000004.1
S11	ATTCAGGCTAATATGCTCCCACTTCCGCAAATCG	MEHR01000004.1
S12	TCCGCATTAAGACTATTGACTGCAATGTGGTTA	MEHR01000004.1
S13	TAATGGATTTTGTGTATTTGGGGTTGAGATAAGAA	MEHR01000004.1
S14	AGGGTCAGATGCTCTACCATTCGCAGCTTGTATGA	MEHR01000004.1
S15	ACCAAGCGATTAGTTCTTACGAGAACGGGTTTAGA	MEHR01000004.1
S16	CTGGCGGTTTTCTTTCAGGGCTTCACCTTTCGCAC	MEHR01000004.1
S17	GAGAGTACCGGCTTGCCGTTTGGCTCAGGGTGC	MEHR01000004.1
S18	ACAACACCCTCACCTTTTTCATTTACGCCATTAT	MEHR01000004.1
S19	AAAACAAGAGCGAAACAACTTGAAACACAAAAA	MEHR01000004.1
S20	TCTCGCTTAATGCGATCATAGTCTTTGACTAAACG	MEHR01000004.1
S21	TTTGATACTCCTTTAAGGGTGGTTGGTTGATAAAA	MEHR01000004.1
S22	GAAGGTTTGGTGCATTTAATGAAATTTAGATTGT	MEHR01000004.1
S23	TTGGGATATTGCGATTGGTATCAAAAATTGACGAA	MEHR01000004.1
S24	AGCTACAACCTATTTGATATGCGATTAGGCGTAAATAAA	MEHR01000007.1
S25	AGCTACAACTGCGGCTACCGTTTTAACAAAGAAACGGCG	MEHR01000007.1
S26	AGCTACAACTTATCAGACCAATTCGCATCACTCTATGCG	MEHR01000007.1
S27	AGCTACAACCTACTTGACCGCTTACCACCACGAAGATGA	MEHR01000007.1

Direct Repeats	Sequence	Contig
DR1	ATTTCAATATATATCGATGCGAAGATGGCTAC	MEHR01000004.1
DR2	GTTTCAATACACAACCACGCATAGGTGGCTGC	MEHR01000004.1
DR3	GTTTCAATTCACAGCCACGCGTAGGTGGCTGC	MEHR01000004.1
DR4	GTTTCAATTCACAGCCACGCGAAGGTGGCTGC	MEHR01000004.1
DR5	GTTTCAATTCACAGCCACGCAAAGGTGGCTGC	MEHR01000004.1
DR6	ATTGTAGCACTGCGAAATGAGAAAGGG	MEHR01000007.1
DR7	ATTGTAGCACTGCGAAATGAAAAAGGG	MEHR01000007.1

Fig. 4 Schematic representation of the CRISPR-Cas system in the *M. haemolytica* NIVEDI/MH/1 serotype A2. (**a**) CRISPR-Cas loci which include CRISPR array of direct repeats, a spacer sequence, and a series of Cas proteins. (**b**) Sequence of seven unique direct repeats (DR) and 27 spacers present in the CRISPR-Cas system present in two confirmed CRISPR array as shown in (**a**)

array. Type I-C CRISPR-Cas array of *M. haemolytica* NIVEDI/MH/1 consisted of Cas1, Cas2, Cas3HD, Cas4, Cas5, Cas7b, and Cas8c proteins and 23 spacers of phage DNA with 5 different direct repeats. Type II-C CRISPR-Cas array consisted of four spacers with direct repeats and Cas1, Cas2, and Cas9 proteins. Cas9 is the characteristic of the type II system and is a large protein needed for the crRNA synthesis and cleavage of the target DNA (Makarova et al. 2011). A schematic representation of the CRISPR loci along with spacers is shown in Fig. 4 (Supplementary File S4).

In summary, the whole genome of *M. haemolytica* serotype A2 isolated from a pneumonic sheep in India was sequenced. The phylogeny analysis showed that the *M. haemolytica* strain NIVEDI/MH/1 is closely related to other pathogenic *M. haemolytica* serotype A2 strains of bovine and ovine origin. The genome contained various virulence factor determinants, five GIs, and metal- and antibiotic resistance genes. Most of the GIs have shown sequence identity with other respiratory pathogens belonging primarily to the family Pasteurellaceae. Further, the strain contains type II-C of CRISPR-Cas array in its genome, in addition to the earlier reported type I-C CRISPR-Cas array. Overall, the presence of these elements substantiates the virulence potential of this Indian strain and emphasizes the role of HGT in the pathogenesis of an opportunistic respiratory pathogen. Further studies are needed on the various serotypes of *M. haemolytica* from different anatomical locations of the multiple hosts to understand the genotypic and pathogenic variants circulating in the Indian subcontinent.

Nucleotide sequence accession number

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number MEHR00000000 with version MEHR01000000.

Acknowledgments The authors acknowledge the UGC-CAS, UGC-CEGS, UGC-NRCBS, DBT-IPLS, and DST-PURSE programs of SBS, Madurai Kamaraj University, Madurai, Tamil Nadu, for providing the facilities for genome sequencing.

Funding This work was partially funded by the Department of Biotechnology, Govt. of India, New Delhi, through the twinning project entitled "Aetio-Pathology and molecular epidemiology of bacterial and viral diseases associated with the respiratory problems of yak in the North Eastern Regions of India" (Grant No: BT/391/NE/TBP/2012).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study was approved by the Institutional Ethics Committee. All applicable international, national, and institutional guidelines for the animal's care were followed during the sample collection.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Alexander TW, Cook SR, Yanke LJ et al (2008) A multiplex polymerase chain reaction assay for the identification of *Mannheimia haemolytica*, *Mannheimia glucosida*, and *Mannheimia ruminalis*. Vet Microbiol 130:165–175. https://doi.org/10.1016/j.vetmic.2008.01.001

Aziz RK, Bartels D, Best AA et al (2008) The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75

Bakaletz LO, Baker BD, Jurcisek JA et al (2005) Demonstration of type IV pilus expression and a twitching phenotype by *Haemophilus influenzae*. Infect Immun 73:1635–1643. https://doi.org/10.1128/IAI.73.3.1635-1643.2005

Blackall PJ, Bojesen AM, Christensen H, Bisgaard M (2007) Reclassification of [*Pasteurella*] *trehalosi* as *Bibersteinia trehalosi* gen. nov., comb. nov. Int J Syst Evol Microbiol 57:666–674. https://doi.org/10.1099/ijs.0.64521-0

Chen L, Zheng D, Liu B et al (2016) VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. Nucleic Acids Res 44:D694–D697. https://doi.org/10.1093/nar/gkv1239

Chevreux B, Wetter T, Suhai S (1999) Genome sequence assembly using trace signals and additional sequence information. in: computer science and biology: Proceedings of the German Conference on Bioinformatics (GCB) '99, GCB, Hannover, Germany. pp 45–56

Clarridge JE (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin Microbiol Rev 17:840–862. https://doi.org/10.1128/CMR.17.4.840-862.2004

Dar ML, Darzi MM, Mir MS et al (2012) Prevalence and pathological studies on ovine pneumonic pasteurellosis in Kashmir valley, India. Eurasian J Vet Sci 28:199–203

Deveau H, Garneau JE, Moineau S (2010) CRISPR/Cas system and its role in phage-bacteria interactions. Annu Rev Microbiol 64:475–493. https://doi.org/10.1146/annurev.micro.112408.134123

Eidam C, Poehlein A, Leimbach A et al (2015) Analysis and comparative genomics of ICEMh1, a novel integrative and conjugative element (ICE) of *Mannheimia haemolytica*. J Antimicrob Chemother 70:93–97. https://doi.org/10.1093/jac/dku361

Freedman Z, Zhu C, Barkay T (2012) Mercury resistance and mercuric reductase activities and expression among chemotrophic *Thermophilic Aquificae*. Appl Environ Microbiol 78:6568–6575. https://doi.org/10.1128/AEM.01060-12

Gilmore MS, Lebreton F, van Schaik W (2013) Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. Curr Opin Microbiol 16:10–16. https://doi.org/10.1016/j.mib.2013.01.006

Gioia J, Qin X, Jiang H et al (2006) The genome sequence of *Mannheimia haemolytica* A1: insights into virulence, natural competence, and *Pasteurellaceae* phylogeny. J Bacteriol 188:7257–7266. https://doi.org/10.1128/JB.00675-06

Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35:W52–W57. https://doi.org/10.1093/nar/gkm360

Guo X, Zhou S, Wang Y et al (2016) Characterization of a highly thermostable and organic solvent-tolerant copper-containing polyphenol oxidase with dye-decolorizing ability from Kurthia huakuii LAM0618T. PLoS One 11:e0164810. https://doi.org/10.1371/journal.pone.0164810

Hall RM (2010) *Salmonella* genomic islands and antibiotic resistance in *Salmonella enterica*. Future Microbiol 5:1525–1538. https://doi.org/10.2217/fmb.10.122

Hauglund MJ, Tatum FM, Bayles DO et al (2015) Genome sequences of *Mannheimia haemolytica* serotype A2 isolates D171 and D35, recovered from bovine pneumonia. Genome Announc 3:e00093–e00015. https://doi.org/10.1128/genomeA.00093-15

Highlander SK (2001) Molecular genetic analysis of virulence in *Mannheimia* (*Pasteurella*) *haemolytica*. Front Biosci a J virtual Libr 6:D1128–D1150

Highlander SK, Chidambaram M, Engler MJ, Weinstock GM (1989) DNA sequence of the *Pasteurella haemolytica* leukotoxin gene cluster. DNA (Mary Ann Liebert, Inc) 8:15–28

Ho Sui SJ, Fedynak A, Hsiao WWL et al (2009) The association of virulence factors with genomic islands. PLoS One 4:e8094. https://doi.org/10.1371/journal.pone.0008094

Hörmansdorfer S, Bauer J (1996) Resistance pattern of bovine *Pasteurella*. Berliner und Munchener Tierarztl Wochenschrift 109:168–171

Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. Science 327:167–170. https://doi.org/10.1126/science.1179555

Hsiao W, Wan I, Jones SJ, Brinkman FSL (2003) IslandPath: aiding detection of genomic islands in prokaryotes. Bioinforma (Oxford, England) 19:418–420

Hudson CM, Lau BY, Williams KP (2014) Islander: a database of precisely mapped genomic islands in tRNA and tmRNA genes. Nucleic Acids Res 43:48–53. https://doi.org/10.1093/nar/gku1072

Jackson CR, Dugas SL (2003) Phylogenetic analysis of bacterial and archaeal arsC gene sequences suggests an ancient, common origin for arsenate reductase. BMC Evol Biol 3:1–10. https://doi.org/10.1186/1471-2148-3-18

Jansen R, van EJDA, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 43:1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x

Karginov FV, Hannon GJ (2010) The CRISPR system: small RNA-guided defense in Bacteria and Archaea. Mol Cell 37:7–19. https://doi.org/10.1016/j.molcel.2009.12.033

Keen EC (2012) Paradigms of pathogenesis: targeting the mobile genetic elements of disease. Front Cell Infect Microbiol 2:161. https://doi.org/10.3389/fcimb.2012.00161

Klima CL, Cook SR, Zaheer R et al (2016) Comparative genomic analysis of *Mannheimia haemolytica* from bovine sources. PLoS One 11:1–23. https://doi.org/10.1371/journal.pone.0149520

Klima CL, Zaheer R, Briggs RE, McAllister TA (2017) A multiplex PCR assay for molecular capsular serotyping of *Mannheimia haemolytica* serotypes 1, 2, and 6. J Microbiol Methods 139:155–160. https://doi.org/10.1016/j.mimet.2017.05.010

Koonin EV, Makarova KS (2009) CRISPR-Cas: an adaptive immunity system in prokaryotes. F1000 Biol Reports 1:95. https://doi.org/10.3410/B1-95

Kumar J, Dixit SK, Kumar R (2015) Rapid detection of *Mannheimia haemolytica* in lung tissues of sheep and from bacterial culture. Vet World 8:1073–1077. https://doi.org/10.14202/vetworld.2015.1073-1077

Lafleur RL, Abrahamsen MS, Maheswaran SK (1998) The biphasic mRNA expression pattern of bovine interleukin-8 in *Pasteurella haemolytica* lipopolysaccharide-stimulated alveolar macrophages is primarily due to tumor necrosis factor alpha. Infect Immun 66:4087–4092

Lagesen K, Hallin P, Rødland EA et al (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160

Laing C, Buchanan C, Taboada EN et al (2010) Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. BMC Bioinforma 11:461. https://doi.org/10.1186/1471-2105-11-461

Langille MG, Hsiao WW, Brinkman FS (2008) Evaluation of genomic island predictors using a comparative genomics approach. BMC Bioinforma 9:329. https://doi.org/10.1186/1471-2105-9-329

Lawrence PK, Kittichotirat W, Bumgarner RE et al (2010a) Genome sequences of *Mannheimia haemolytica* serotype A2: ovine and bovine isolates. J Bacteriol 192:1167–1168. https://doi.org/10.1128/JB.01527-09

Lawrence PK, Kittichotirat W, McDermott JE, Bumgarner RE (2010b) A three-way comparative genomic analysis of *Mannheimia haemolytica* isolates. BMC Genomics 11. https://doi.org/10.1186/1471-2164-11-535

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Makarova KS, Haft DH, Barrangou R et al (2011) Evolution and classification of the CRISPR–Cas systems. Nat Rev Microbiol 9:467–477. https://doi.org/10.1038/nrmicro2577

Marru HD, Anijajo TT, Hassen AA (2013) A study on ovine pneumonic pasteurellosis: isolation and identification of *Pasteurellae* and their antibiogram susceptibility pattern in Haramaya District, Eastern Hararghe, Ethiopia. BMC Vet Res 9:239. https://doi.org/10.1186/1746-6148-9-239

Mohamed RA, Abdelsalam EB (2008) A review on pneumonic pasteurellosis (respiratory mannheimiosis) with emphasis on pathogenesis, virulence mechanisms and predisposing factors. Bulg J Vet Med 11:139–160

Moustafa AM, Seemann T, Gladman S et al (2015) Comparative genomic analysis of Asian haemorrhagic septicemia-associated strains of *Pasteurella multocida* identifies more than 90 hemorrhagic septicemia-specific genes. PLoS One 10:1–16. https://doi.org/10.1371/journal.pone.0130296

Omaleki L, Browning GF, Allen JL, Barber SR (2011) The role of *Mannheimia* species in ovine mastitis. Vet Microbiol 153:67–72. https://doi.org/10.1016/j.vetmic.2011.03.024

Omaleki L, Browning GF, Allen JL et al (2016) Molecular epidemiology of an outbreak of clinical mastitis in sheep caused by *Mannheimia haemolytica*. Vet Microbiol 191:82–87. https://doi.org/10.1016/j.vetmic.2016.06.005

Pagano M, Martins AF, Barth AL (2016) Mobile genetic elements related to carbapenem resistance in *Acinetobacter baumannii*. Brazilian J Microbiol 47:785–792. https://doi.org/10.1016/j.bjm.2016.06.005

Paulsen DB, Mosier DA, Clinkenbeard KD, Confer AW (1989) Direct effects of *Pasteurella haemolytica* lipopolysaccharide on bovine pulmonary endothelial cells in vitro. Am J Vet Res 50:1633–1637

Peng Z, Liang W, Liu W et al (2016) Genomic characterization of *Pasteurella multocida* HB01, a serotype A bovine isolate from China. Gene 581:85–93. https://doi.org/10.1016/j.gene.2016.01.041

Qi J, Luo H, Hao B (2004) CVTree: a phylogenetic tree reconstruction tool based on whole genomes. Nucleic Acids Res 32:W45–W47. https://doi.org/10.1093/nar/gkh362

Rao RT, Jayakumar K (2017) Genomic islands of *Mannheimia haemolytica*—in silico analysis. Curr Bioinforma 16:1–9. https://doi.org/10.2174/15748936106661507302034

Sakarya S, Oncü S (2003) Bacterial adhesins and the role of sialic acid in bacterial adhesion. Med Sci Monit Int Med J Exp Clin Res 9:RA76–RA82

Schroeder M, Brooks BD, Brooks AE (2017) The complex relationship between virulence and antibiotic resistance. Genes 8. https://doi.org/10.3390/genes8010039

Severi E, Hood DW, Thomas GH (2007) Sialic acid utilization by bacterial pathogens. Microbiology 153:2817–2822. https://doi.org/10.1099/mic.0.2007/009480-0

Sharma RK, Patil RD, Kishtwaria RS, Asrani RK (2011) An outbreak of pneumonic mannheimiosis in a livestock farm in sub-temperate region of India. Haryana Vet 50:89–91

Steenbergen SM, Lichtensteiger CA, Caughlan R et al (2005) Sialic acid metabolism and systemic pasteurellosis. Infect Immun 73:1284–1294. https://doi.org/10.1128/IAI.73.3.1284-1294.2005

Szczepankowska A (2012) Role of CRISPR/cas system in the development of bacteriophage resistance. Adv Virus Res:289–338

Tamura K, Stecher G, Peterson D et al (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197

Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36

Thachil AJ, Velayudhan BT, Shaw DP et al (2009) Pathogenesis of *Ornithobacterium rhinotracheale* in egg-laying hens with coexisting infectious bronchitis virus and *Escherichia coli* infections. J Appl Poult Res 18:780–788. https://doi.org/10.3382/japr.2009-00039

Timsit E, Christensen H, Bareille N et al (2013) Transmission dynamics of *Mannheimia haemolytica* in newly-received beef bulls at fattening operations. Vet Microbiol 161:295–304

Waack S, Keller O, Asper R et al (2006) Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinforma 7:142. https://doi.org/10.1186/1471-2105-7-142

Watts JL, Yancey RJ, Salmon SA, Case CA (1994) A 4-year survey of antimicrobial susceptibility trends for isolates from cattle with the bovine respiratory disease in North America. J Clin Microbiol 32:725–731

Wu S, Zhu Z, Fu L et al (2011) WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics 12:444. https://doi.org/10.1186/1471-2164-12-444

Xu Z, Zhou Y, Li L et al (2008) Genome biology of *Actinobacillus pleuropneumoniae* JL03, an isolate of serotype 3 prevalent in China. PLoS One 3:e1450. https://doi.org/10.1371/journal.pone.0001450

Yaeger MJ (1996) An outbreak of *Actinobacillus suis* septicemia in grow/finish pigs. J Vet Diagn Investig 8:381–383

Yoo HS, Maheswaran SK, Lin G et al (1995) Induction of inflammatory cytokines in bovine alveolar macrophages following stimulation with *Pasteurella haemolytica* lipopolysaccharide. Infect Immun 63:381–388

Yu C, Sizhu S, Luo Q et al (2016) Genome sequencing of a virulent avian *Pasteurella multocida* strain GX-Pm reveals the candidate genes involved in the pathogenesis. Res Vet Sci 105:23–27. https://doi.org/10.1016/j.rvsc.2016.01.013

Zhang Q, Ye Y (2017) Not all predicted CRISPR–Cas systems are equal: isolated cas genes and classes of CRISPR like elements. BMC Bioinforma 18:92. https://doi.org/10.1186/s12859-017-1512-4