ABSTRACT

Title of Thesis: DEVELOPMENT OF A MULTILOCUS SEQUENCE TYPING SCHEME FOR *AVIBACTIERUM PARAGALLINARUM*

Alyssa Meihua Harris, Master of Science, 2023

Thesis directed by: Assistant Professor, Mostafa Ghanem
Department of Veterinary Medicine

Avibacterium paragallinarum (*A. paragallinarum*), the causative agent of the respiratory disease Infectious Coryza (IC) in chickens, has seen a rising incidence in the United States. Current strain differentiation is inadequate for detailed epidemiological analysis. The objective of this study was to develop a Multilocus sequence typing (MLST) scheme for *A. paragallinarum* for outbreak investigations and to offer a better tool for strain differentiation. By evaluating whole genome sequences and clinical samples, we designed PCR amplicons for eighteen gene segments, selected six genes for their nucleotide diversity and discrimination potential. The MLST was used to differentiate seventy-five samples. Our MLST showed greater discriminatory power than existing HPG2-based methods, aligning closely with adhoc core genome MLST in 75 tested sample. Our newly developed MLST scheme enables more accurate strain differentiation, allowing for better understanding of *A. paragallinarum* epidemiology and population structure to help prevention and control efforts worldwide.

DEVELOPMENT OF A MULTILOCUS SEQUENCE TYPING SCHEME FOR *AVIBACTERIUM PARAGALLINARUM*

By

Alyssa Meihua Harris

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2023

Advisory Committee:
Dr. Mostafa Ghanem, Chair
Professor Nathaniel Tablante
Professor Debabrata Biswas
Dr. Sean Riley

Dedication

To my parents and sister. You all have been the biggest source of encouragement and love.
For the amazing people who have supported me throughout my education and more.
Thank you for making me see this adventure through to the end. On to the next one.

Acknowledgements

I am grateful to have been given the opportunity to be the first student to conduct research under the guidance of my advisor, Dr. Mostafa Ghanem, to whom I am indebted for his inspiration, support, and patience. He provided continuous encouragement throughout this project. Dr. Ghanem's knowledge and expertise in the poultry and epidemiology field has been an experience I will never forget and one I will be thankful for.

I would also like to thank Dr. Nathaniel Tablante, Dr. Sean Riley, and Dr. Debabrata Biwas for serving on my thesis committee and sharing their knowledge. Dr. Michele Williams, our previous lab director, was a great mentor and teacher for the time she spent teaching me many laboratory techniques and more. Dhiraj Chundru, my lab member, was an immense help with my research and my go-to person throughout our laboratory adventures. Fauzia Huq Nur, my other lab member, was so helpful on this project and someone I am so fortunate to have worked with.

The VMSC department, I would like to thank everyone for being so welcoming from the beginning of my journey here. Specifically, within the department I would like to thank several people. Dr. Mattias Corrotte for assistance with the Genetic Analyzer. Ms. Judy Knight for filling out the travel requests and the many notifications about package delivery. Mr. Tam Ong for the IT help and always spreading his positivity.

Avibacterium paragallinarum samples were kindly provided by outside sources. Dr. Mohamed El-Gazzar and Dr. Amro Hashish from Iowa State University School of Veterinary Medicine Diagnostics Laboratory, Ames, Iowa; Dr. Shankar Mondal from Salisbury Animal Health Laboratory, Salisbury, Maryland; Dr. Yan Zhang from Ohio Animal Disease Diagnostic Laboratory, Reynoldsburg, Ohio; CEVA Animal Health; and the services from Lasher Laboratory, Georgetown, Delaware. This thesis would not have been possible without the various

collaborations from these laboratories. Funding for this research was made possible through the University of Maryland.

Finally, I would like to thank my family again for their support and encouragement. My person, thank you for being you. To the other special people in my life, my life would not be the same without you all in it. I am so thankful to be surrounded by so many amazing and supportive people.

Table of Contents

Lists of Tables

List of Figures

Chapter 1: Introduction

1.1 Rationale and Objectives

The United States stands as a global leader in poultry production and ranks as the world's second-largest egg producer (*USDA ERS- Poultry & Eggs,* n.d.). With the rise in poultry production and consumption throughout the twenty-first century, the imperative to monitor and mitigate pathogens threatening poultry health and productivity has never been more critical. *Avibacterium paragallinarum* (*A. paragallinarum*), a gram-negative, non-motile bacterium, is the etiological agent of Infectious Coryza (IC), a respiratory disease that affects chickens globally. IC's increasing prevalence and pathogenicity in various U.S. regions highlight a need to re-evaluate pathogen behaviors, transmission dynamics, and the efficacy of current vaccines.

The current gap in knowledge about the transmission and introduction of this pathogen underpins the necessity for an advanced understanding of IC transmission dynamics and vaccine effectiveness. Traditional serotyping, the primary method for strain differentiation used in vaccine selection, is labor-intensive, expensive, and lacks availability in the U.S(Buter et al., 2023) (citation). While DNA fingerprinting has aided epidemiological studies, it requires the isolation of *A. paragallinarum*, which is often challenging due to the fastidious nature of *A. paragallinarum*. Single-locus-based sequence typing like HPG2, although employed for strain typing, does not yield sufficient or accurate epidemiological insights into the relatedness of *Avibacterium paragallinarum* isolates (Feberwee et al., 2019; Luna-Castrejón et al., 2020). The objective of this study is to develop a Multilocus sequence typing (MLST) scheme capable of distinguishing *Avibacterium paragallinarum* strains, thereby elucidating the epidemiology and population structure of *A. paragallinarum*. Furthermore, we intend to assess the performance of existing typing methods (HPG2) against the proposed MLST and adhoc core genome MLST, a whole genome-based method. The anticipated outcomes of this research are multifaceted and include:

1. Develop an accurate typing scheme (MLST) for *A. paragallinarum* strain differentiation.
2. Assess the performance of existing typing method (HPG2) against the proposed MLST.
3. Apply the new scheme to decipher the epidemiology of recent outbreaks in the Northeastern United States, determining if they are attributable to single or multiple strains.
4. Ascertain whether outbreak strains are shared across commercial and non-commercial flocks.
5. Elucidate the relatedness between local, regional, and international strains of *A. paragallinarum.*
6. Develop a scheme and share it online on PubMLST, a public database for molecular typing, to establish a worldwide connection for *A. paragallinarum* typing.

1.2 Literature Review

1.2.1 Definition and Synonyms

The poultry pathogen of interest in this study is *Avibacterium paragallinarum* (*A. paragallinarum*). *A. paragallinarum* is the causative agent of Infectious Coryza (IC). Belonging to the family, *Pasteurellaceae*, with a genome size of about 2.4 kilobase pairs (kbp) in length, this gram-negative bacterium primarily affects chickens and other wild and domestic birds. Infectious Coryza is commonly associated with inflammation of the nasal passages or nasal discharge (Beach & Schalm, 1936). Other symptoms of the disease are facial edema, conjunctivitis, anorexia, and diarrhea , all leading to decreased egg production and a significant economic impact on the poultry industry (Blackall, 1999).

1.2.2 Public Health & Economic Significance

Avibacterium paragallinarum is not a zoonotic disease and has never posed a threat to public health. On the other hand, *A. paragallinarum* causes great economic losses associated with IC in

Commented [AH1]: Is zoonotic since other animals can be carriers and loss of chickens is a public health concern

Commented [M2R1]: It is not zoonotic as it is not transmitted between animals and human but I has very limited effect on public health threat through threatening food security as it doesn't eliminate production to kill large number of birds so ignor that comments

Commented [M3R1]: It will not cause starvation

both layers and broilers. In 1934, the first case of Infectious Coryza was reported in California, United States (Page, 1962). Since then, the disease has been found to be isolated in both wild and domestic birds. *Avibacterium paragallinarum* affects broiler and layer chickens. IC can result in up to 10-40% egg drop production, due to decreased water and feed intake of infected birds. In broilers raised on antibiotic free farms, poor growth performance leads to economic losses. The substantial economic losses linked to IC result not only from impaired growth in young birds and a significant decrease in egg production among layers, but also from the extensive vaccination required for control. The consequences of the disease can be even more severe than the straightforward situation outlined above. For instance, an outbreak of the disease among older layer birds in California resulted in a total mortality rate of 48% and caused egg production to decline from 75% to 15.7% over three weeks.

Infectious coryza also has a significant impact on broiler chickens. In California, two instances of IC, one of which was complicated by the presence of *Mycoplasma synoviae*, resulted in increased condemnations primarily due to airsacculitis ranging from 8 % to 15%. Meanwhile, in Alabama, an outbreak of IC among broilers without any other disease agent involvement led to a condemnation rate of 69.8%, with all cases attributed to airsacculitis. When infectious coryza occurs in chicken flocks in developing countries, the economic consequences can be even more severe due to limited resources for vaccination and control measures (Blackall & Soriano-Vargas, 2020)

The effects of IC can be heightened when *infection* is also complicated by the presence of another avian disease like *Mycoplasma gallisepticum*, *Mycoplasma synoviae*, or *Ornithibacterium rhinotracheale* (Blackall, 1999). Once an infected chicken has been infected, it can be treated with antibiotics, but it remains a carrier of the bacteria for life. A chicken will be able to present clinical

signs of the disease again and reinfect chickens around it. This cycle poses a negative effect to farmers and a challenging situation in production systems that do not follow all-in-all-out programs like multiage layer facilities.

Together, these facts highlight the urgent need for a reliable and effective method of identifying and monitoring *A. paragallinarum* in poultry populations. Therefore, the development of a Multilocus sequence typing scheme for *A. paragallinarum* is crucial to accurately characterize and track the pathogen, ultimately aiding in the prevention and control of infectious coryza outbreaks.

1.2.3 History & Current Impact

Avibacterium paragallinarum was first isolated in the United States almost a century ago (Blackall & Soriano-Vargas, 2020) (citation). California has had sporadic outbreaks that negatively impacted the poultry industry. It became prevalent around the world and is only becoming more prevalent. Since 2019, there were multiple outbreaks in Pennsylvania, Alabama, and California where multiple isolates were collected for epidemiological analyses (Byukusenge et al., 2020). Now, *A. paragallinarum* continue to spread and being reported in the Mid-Atlantic and Midwestern regions (Ohio and Iowa) of the United States as well. This new shift in geographical range may be a result of pathogen evolution and adaptation to colder climate conditions or because of climate change leading to warmer temperatures in these new areas.

1.2.4 Growth Requirement

Avibacterium paragallinarum is a microaerophilic pathogen that requires decreased levels of O_2 to grow. The bacteria can grow in an atmosphere of 5% carbon dioxide and has a minimal and maximal temperature of growth of 25°C and 45°C, respectively, with an optimal range of 34°C–42°C. *Avibacterium paragallinarum* can either be nicotinamide adenine dinucleotide (NAD, V factor) independent or dependent. For this study, the *A. paragallinarum* strains were provided with

Nicotinamide adenine dinucleotide (NAD). *A. paragallinarum* is a fastidious and slow growing organism. The bacteria can be cultured on chocolate agar: red blood cells (RBC) lysed by heating to release NAD. Once the bacteria are cultured on chocolate agar for 48 hours, it can be transferred to grow in broth. Another alternative approach to grow *A. paragallinarum* is to grow it on blood agars with a nurse culture of *Staphylococcus aureus* (Page, 1962). The preferred broth used is brain heart infusion (BHI) along with a 5% serum supplement for optimum growth. The serum used in this study was HeIsoVitaleX- NAD: L-cystine, L-glutamine, L-cysteine, adenine, guanine, ferric nitrate (iron source), vitamin B1, and 4-amino benzoic acid (PABA, a folic acid/folate precursor). *Avibacterium paragallinarum* colonies appear as circular, transparent, and smooth dewdrops on culture plates and microscopically as gram-negative coccobacilli (Blackall & Soriano-Vargas, 2020). After the media is prepped for both plate agars and broth, the media can be stored at 37°C (shaking if in liquid) for 2 days in an elevated carbon dioxide atmosphere (5-10%). *A. paragallinarum* can be detected when the broth becomes turbid.

1.2.5 Classification

Avibacterium paragallinarum belongs to the family *Pasteurellaceae* which includes other poultry pathogens(Blackall & Soriano-Vargas, 2020). Previously known as *Haemophilus paragallinarum*, the new genus *Avibacterium* was proposed for four species (Blackall et al., 2005). After a comparative analysis of the 16S rRNA gene sequence, *Haemophilus paragallinarum* and three other strains (*Pasteurella gallinarum*, *Pasteurella avium*, and *Pasteurella volantium*) had over a 95% similarity so these species were added to the genus *Avibacterium* (Blackall et al., 2005). Any undiscovered organism that does not match the genotypic or phenotypic methods will also be identified as *Avibacterium sp. A*. The genus *Avibacterium* is currently classified into five species, which are *Avibacterium avium*, *Avibacterium endocarditidis*, *Avibacterium gallinarum*,

Avibacterium paragallinarum, and *Avibacterium volantium*. However, recent studies suggest that there may only be two or three species within the genus (Bisgaard et al., 2012). The current understanding of *A. paragallinarum* isolates is that both V-factor dependent and V-factor independent exist. *A. paragallinarum* is commonly acknowledged as a primary pathogen, while the other members of the genus are considered either opportunistic pathogens (*A. endocarditidis*, and *A. gallinarum*) or not known to cause disease (*A. avium*, and *A. volantium*) (Blackall & Soriano-Vargas, 2020).

1.2.6 Diagnosis, Identification & Molecular Detection

Infectious coryza is often diagnosed based on history, clinical signs, and gross lesion during necropsy. Identifying and isolating the bacteria can be difficult due to its fastidious growth needs and coexistence with other bacteria in the respiratory tract. Previous antimicrobial treatment during chronic infection and delays in sample processing have been shown to hinder the successful retrieval of *A. paragallinarum* using traditional bacteriological methods(Dufour-Zavala, 2008). For definitive laboratory diagnosis of *A. paragallinarum* for IC, there are several molecular assays available. A conventional polymerase chain reaction (PCR) test was developed by Chen et. al 1996 and has been widely utilized for the detection of *A. paragallinarum* in clinical specimens (Chen et al., 1996). A 5′ Taq nuclease assay targeting the HPG-2 gene region of *A. paragallinarum* was established in 2008 (Corney et al., 2008) and its primers were updated in 2019 (Clothier et al., 2019). More recently, a highly sensitive and specific probe-based real-time PCR for the detection of *A. paragallinarum* in clinical samples has been developed (Kuchipudi et al., 2021). Furthermore, a recent lateral flow test has been introduced for rapid pathogen detection (Morales Ruiz et al., 2018); however, this test lacks specificity and is unable to distinguish *A.*

paragallinarum from related commensal species like *A. avium, A. endocarditis, A. gallinarum,* and *A. volantium*

1.2.7 Strain Classification

Through the utilization of a plate agglutination assay with whole cells and chicken antisera, Page was the first to develop a serotyping scheme for *A. paragallinarum*. The Page scheme recognizes three serovars; A, B, and C (Page, 1962). The Kume scheme was another serologic classification that was later established based on hemagglutination inhibition (HI) test (Blackall, Eaves, et al., 1990). This newly developed scheme agreed with Page serovars A, B, and C but recognizes nine additional serovars; A-1, A-2, A-3, A-4, B-1, C-1, C-2, C-3, and C-4 (Blackall, Eaves, et al., 1990). The usage of the HI is more preferred than the Page serovar for isolates. Kume's scheme has not been replicated due to the technicality of the system. Strain classification is important for understanding the epidemiology of *A. paragallinarum* for effective control measures and prevention. The two serotyping schemes can differentiate the strains into the three or nine serovars which is critically needed for vaccine selection as the cross-protection between Page serotypes is limited and often they are recognized as three different immuovars A, B and C (Blackall & Soriano-Vargas, 2020). Despite that, it seems that neither approach has been widely used or adopted due to technical challenges, cost, and time considerations (Feberwee et al., 2019).

Nucleic acid-based techniques for molecular serotyping and molecular strain typing have also been developed. DNA fingerprinting using restriction endonuclease analysis was successful in both typing and providing epidemiological insights for *A. paragallinarum* (Blackall, Morrow, et al., 1990). Ribotyping has also been beneficial, particularly in understanding the relationships among *A. paragallinarum* isolates in different regions such as South Africa (Miflin et al., 1995) and China (Mifflin et al., 1997). Several recent studies have investigated the validity of using the whole or

parts of *HMTp210* gene, that encode for the outer membrane hemagglutinin (HA) for molecular prediction of serotype (Buter et al., 2023; Tan et al., 2020). The combination of region 1 of HMTp210 and the hypervariable region (HVR) of the same gene HMTp210 sequences resulted in 14 clusters strongly correlated with Page serovars and the nine known Kume serovars, providing a potential new molecular serotyping approach for *A. paragallinarum* (Buter et al., 2023). However, these serotype classifications do not correlate with the geographical origin of the samples and their overall phylogenetic relatedness. Thus, there is an urgent need for a molecular typing scheme for epidemiological investigation of *A. paragallinarum*. For epidemiological typing purposes, the available schemes include HPG2 gene region and ERIC-PCR, another DNA fingerprinting method, has been used for typing *A. paragallinarum*(Feberwee et al., 2019). Using enterobacterial repetitive intergenic consensus (ERIC-PCR) as a technique for DNA fingerprinting has been useful to identify different strains based on the different banding patterns of the genomic DNA that is generated by gel electrophoresis. This method has low reproducibility, time consuming and limited to pure cultures (Luna-Castrejón et al., 2020). Recently, strain typing *of A. paragallinarum* has been performed by using the HPG2 region and ERIC-PCR for epidemiological investigation of Dutch outbreaks (Feberwee et al., 2019). HPG2 region also, has been used for characterization of few U.S. strains in California (Gallardo et al., 2020; Luna-Castrejón et al., 2020). HPG2 region showed agreement with ERIC-PCR results for the limited samples set used in the Dutch study(Feberwee et al., 2019). However, there are typing controversies with epidemiological information regarding the correlation of its results and the geographical origin of several samples that have been reported for using HPG2 region included in our study. Furthermore, these approaches were developed nearly three decades ago, before even a complete genome for *A. paragalinarum* became available. Thus, it is time to update our typing

approaches and leverage recent advances in next-generation sequencing to develop MLST and cgMLST schemes.

Multi-locus sequence typing (MLST) represents a rapid, practical, transferable, and comprehensive method for sequence typing. Introduced in 1998 by Maiden et al., MLST was suggested as an enhanced approach to sequence typing for distinguishing between Neisseria meningitidis strains during the investigation of clinical outbreaks. MLST involves amplification, sequencing, and comparison of internal segments (500 to 700 base pairs) from five to seven housekeeping genes. Regarding poultry pathogens, MLST has been successfully developed for multiple MLST assays have been successfully developed such as *Mycoplasma synoviae* (El-Gazzar et al., 2017), *Mycoplasma gallisepticum* (Ghanem & El-Gazzar, 2019), *Mycoplasma iowae* (Ghanem & El-Gazzar, 2016), and *Ornithobacterium rhinotracheale* (Thieme et al., 2016). MLST is a robust and reliable method for typing bacterial strains and investigating their population structure.

1.2.8 Pathobiology & Epizootiology

Avibacterium paragallinarum is prevalent wherever chickens are raised. IC has become a problem to the poultry industry worldwide. Chicken is the natural host for *A. paragallinarum.* At any age, a chicken is susceptible to IC and clinical signs can begin to develop 24-48 hours post infection. Studies have shown that infected juvenile chickens have less severe symptoms than mature birds (Blackall & Soriano-Vargas, 2020). However, in mature birds, the disease can affect the bird more severely and cause mortalities (Blackall & Soriano-Vargas, 2020).

1.2.9 Pathology

Avibacterium paragallinarum creates inflammation of the mucous membrane of the sinus. Clinical signs of infection include edema on and around the face, nasal discharge, and anorexia that can

last weeks (Guo, Liu, Chen, et al., 2022). The severity of the disease can vary by age, health and living conditions, co-infection, and breed. Co-infection with other bacterial infections or viruses can lead to mortalities in layer flocks and increase the severity (Morales-Erasto et al., 2016). Morbidity and mortality are enhanced with the presence of *A. paragallinarum* and other bacteria like *Mycoplasma gallisepticum, Ornithobacterium rhinotracheale,* and *Mycoplasma synoviae.* Studies have shown that the pathogen is able to invade the host and replicate in the blood. Affected areas were primarily in the upper respiratory tract (Blackall & Soriano-Vargas, 2020).

1.2.10 Transmission & Carrier

Once a chicken becomes infected with *A. paragallinarum*, the bird remains a carrier for life. After being treated with antibiotics to reduce clinical signs of infection, the birds subsequently become a reservoir for the disease. The primary mode of transmission is horizontally by nasal discharge, secretions, and shared environments between the infected and noninfected birds. Direct contact of birds can spread the pathogen through nasal discharge, respiratory secretions, or ocular exudates. Chickens raised on farms often share space and equipment that can also be a source of transmission. A study in California found that as little as 11 colony-forming units (CFU) was needed for a chicken to become susceptible to *A. paragallinarum* (Page, 1962). Droplets produced by infected birds are also capable of facilitating its transmission through the air. In addition, poultry farms that have different age groups together, infected birds typically spread the disease to the subsequent age group within a few weeks(Blackall & Soriano-Vargas, 2020).

1.2.11 Incubation Period and Clinical Signs

Commented [AH4]: cit

Clinical signs of *A. paragallinarum* develop around 24-48 hours post infection or exposure. If a chicken does not have a concurrent infection, IC will affect the bird for 2-3 weeks. The initial clinical sign of

A. paragallinarum infection is usually facial edema or any other inflammation affecting the upper respiratory tract (Blackall & Soriano-Vargas, 2020). Arthritis and septicemia have also been clinical signs in layers and broilers, respectively. Birds will also experience a decrease in food and water consumption because of the clinical symptoms. Due to decrease in appetite, broilers will not reach optimum growth and layers will have a 10-40% drop in egg production.

1.2.12 Biosecurity & Management Procedures

The best method to be free of IC is to be proactive in having biosecurity practices and vaccinate the birds. To prevent the spread of disease, producers should buy their chickens from reputable breeders that are free of any avian diseases. Infected birds should be isolated from old stock as well as all infected birds. If birds are contained in an area where infection has occurred, the proper vaccine that matches the outbreak strain should be given for immunity. When managing infected birds, proper personal protective equipment (PPE) should be worn and changed out to prevent spreading it to healthy birds. Any equipment or materials used for infected birds should be disinfected after each use and should not be used for 2-3 weeks before use in healthy flocks. The birds' environment and drinking water should also be properly maintained daily.

1.2.13 Vaccines

Inactivated vaccines (bacterins) are common for this *A. paragallinarum* and are commercially available worldwide. The vaccine dosage varies, but on average the first dose is recommended to be given as early as 5 weeks, and the second dosage given four weeks after that (Blackall & Soriano-Vargas, 2020). Chickens should be given annual boosters for prevention. The first vaccine developed was a bivalent vaccine that was developed using Page serovars. At the time, serovar B was not believed to be distinct from A or C, hence a bivalent vaccine could provide cross protection against all strains (Blackall, 1999). Now, there are several vaccines available for protection.

Studies have shown that emulsification increases production of the number of antibodies than aluminum hydroxide (Gong et al., 2014).

CEVA Animal Health and MSD Animal Health each have an oil emulsified vaccines to protect against *A. paragallinarum* (Soriano-Vargas, n.d.). Both vaccines are administered intramuscularly, and two dosages are recommended. A study done in China showed that antibody titers were higher in double vaccinated chickens than single vaccinated (Guo, Liu, Xu, et al., 2022). The three serovars are not cross protective of each other (Blackall & Soriano-Vargas, 2020). When selecting a vaccine for protection, it is important to know what serovar is present in the target population.

1.2.14 Treatment

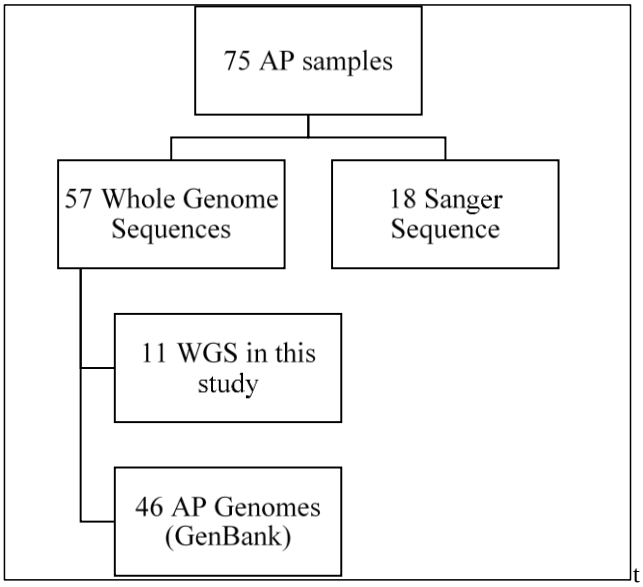
Treatment of IC involves the use of antibiotics such as erythromycin or oxytetracycline. Both antibiotics have shown efficacy in alleviating cases. Treating the disease early is important to prevent further severity. Administration of the antibiotics can be done through the drinking water. The proper dosage and guidelines prescribed by a licensed veterinarian must always be followed for proper control (2,8). Sulfonamides like trimethoprim-sulfamethoxazole, have also been proven to be effective. Although there are antibiotics and sulfonamides available for mitigating IC, relapse is common after treatment. With the rise of antimicrobial resistance and the popularity of antibiotic free production, it has become increasingly difficult to manage and control IC outbreaks in the field.

Chapter 2: Materials and Methods

2.1 *Avibacterium paragallinarum* samples

75 *A. paragallinarum* samples were used in the study. Two distinct sample types include clinical samples (PCR-positive bacterial isolates or tissue swabs) and whole genome sequence (WGSs) isolates obtained from National Center for Biotechnology Information (NCBI) in March 2023. The seventy-five samples included 45 WGSs from NCBI (Figure 1). The remaining 30 samples were clinical samples collected from cases submitted to the following state diagnostic laboratories: Salisbury Animal Health Laboratory (SAHL), Salisbury, Maryland, Iowa State Diagnostic Laboratory, Ames, Iowa, Lasher Laboratory, Georgetown, Delaware, Ohio Animal Diseases Diagnostic Laboratory (ADDL), Reynoldsburg, Ohio, and CEVA Animal Health. These samples were initially tested positive for *A. paragallinarum* according to the protocols used by these laboratories. Received samples were retested for *A. paragallinarum* by conventional PCR test according to Chen et al., 1998 (Chen et al., 1998). Thirty-two samples were isolated in the USA and twenty-two originated from China. The origin of the remaining samples is as followed: four samples from Germany, three samples from Mexico, two samples each from Peru, Japan, England, Russia, and France, and one sample each from South Africa and Gabon. United States samples were isolated from six different states. These states were Iowa (8 samples), Maryland (7 samples), Pennsylvania (7 samples), Delaware (5 samples), Ohio (4 samples), and Colorado (1 sample) (Appendix A). The oldest sample is the reference strain NCTC11296 (AP49_58) that was collected in 1958 and the most recent samples was collected in 2022 from Iowa. All seven samples from Maryland and two samples from Delaware (AP61_DEL_17 and AP62_DEL_17) are from non-commercial backyard poultry operations. The remaining samples are either from commercial or unknown poultry operation. The serotype for most of the samples was unknown except for a few samples from GenBank that included serotypes A1, A2, B1, C1, C2, and C3. In this study, all

samples were given a sample ID that starts with AP (*A. paragallinarum*) and the number of the sample in our study, followed by the state or country of origin if known, and two-digit year of isolation. The original sample name is listed under other aliases in Appendix A.



t

Figure 1. Flowchart the seventy-five types of *A. paragallinarum* used in this sample.

2.2 Selection of the gene candidates with possible MLST loci

To identify candidate gene segments to use as MLST loci, we have developed an adhoc core genome MLST (adhoc cgMLST) scheme for the 13 *A. paragallinarum* WGS's that were available from GenBank as of June 2021 using Ridom SeqSphere+ (Version7.0.4.). The complete genome of AP01_PE_FARPER_15 was chosen as the reference genome and the rest of the genomes were used as query genomes. According to that analysis, 1170 genes were defined as core genes, shared between all query samples and the reference sample, and 935 targets were used as accessory genes. Core genes were ranked according to the number of alleles which indicates a higher degree of DNA variation. Genes that had the highest allele numbers were selected for further evaluation.

These candidate genes were selected based on the following criteria: present in all samples typed by cgMLST, nucleotide diversity, and predicted gene function. Then, a multiple sequence alignment of each gene was visually inspected to select genes in which primers can be located to amplify the target segment (location of the single-nucleotide polymorphisms (SNPs)). Eighteen candidate genes with suitable loci were identified based on that analysis (Table 1).

Table 1. Eighteen genes with candidate loci selected based on core genome MLST analysis of 13 A. paragallinarum genomes and their locus tag, forward and reverse primers, and amplicon size.

Number	Locus Tag	Predicted Protein Function	Primer Sequence (5' to 3')	Amplicon Size (bp)
1	EIA51_00135	potassium transporter	F: CGGCTATAAACCCACCCTCA	417
			R: CTTGCCATTTCTGCYACGCT	
2	EIA51_00965	transcriptional regulator	F: TGATGGCGAACAAATCAGYGA	455
			R: GCTCAACTTGCTCTGGTGTAA	
3	EIA51_01020	glutamate synthase small subunit	F: CGAACGCCAACAAGAAATCG	490
			R: TGCCGCACCTTCTTCTTTTG	
4	EIA51_02965	filamentous hemagglutinin N-terminal domain-containing protein	F: GYKCTTTCGGCATTGGAAGT	535
			R: CCGTCCGCAATGATGAAGTC	
5	EIA51_03445	putative lipid II flippase FtsW	F: ACGGCTATGTGTTTTGGCTC	570
			R: CAAAGTCGGTATGAGCTTCGG	
6	EIA51_04255	pyruvate kinase	F: TTTCTCTCACGGAACCCCTG	568
			R: AGCACGTTCAACTTTAGCGA	
7	EIA51_06250	chromosome partition protein MukF	F: GAACATCATTGGCGGCGTAA	732
			R: GGCGAGCGGATAGTTTTCC	
8	EIA51_06635	methylenetetrahydrofolate reductase	F: GATTGATGCCACGCCTGAAG	568
			R: GCTGCGGTTTAGGGTGTAGA	
9	EIA51_06735	FAD-binding oxidoreductase	F: CACCGATCTGCCTTCCATTG	694
			R: TAAACGCAACCATTCACGCA	
10	EIA51_07360	DNA topoisomerase (ATP-hydrolyzing) subunit A	F: AAAACGAACGTATCACCGCC	705
			R: CTACTTGCTCATCTTCGGCG	
11	EIA51_07390	pyrroline-5-carboxylate reductase	F: CTTTAGCGGCCCACAATCC	555
			R: TGAAACAACCAATCCCGAGC	
12	EIA51_07505	bifunctional biotin--[acetyl-CoA-carboxylase] synthetase/biotin operon repressor	F: CATTGCCCCTTATCAAGCGG	668
			R: CCAGTAACAAATGCCCCTCG	
13	EIA51_08000	inference by VFDB, Virulence factor: Toxin	F: AGTTGATGTTGTTGCGGTTG	687
			R: ACTTGCTCATTTGCCTGACG	
14	EIA51_08215	exodeoxyribonuclease VII large subunit	F: CGTATCGGGGCATTGGTATC	574
			R: AATAACGGGAATGGCGGAGT	
15	EIA51_08385	30S ribosomal protein S1	F: AGATCCTTGGGTTGCGATTG	638
			R: AACCTTCAACTCCGCCATCT	
16	EIA51_08435	peptide chain release factor 3	F: AGGGGCATCGAACGAATTTG	635
			R: CAAATTGCAGCACACCAACC	
17	EIA51_10630	DUF853 domain-containing protein	F: AGGGCTTGCTCAATTTGGTG	742
			R: TAGGCGATTTCCACTGGTGT	
18	EIA51_10680		F: GTTGAACTGACGGGCAATCA	781

translocation/assembly module TamB	R: CGCTTTGGGTTTGTGCTTTG

2.3 Genomic DNA extraction of *A. paragallinarum* samples

Clinical samples were received as either aliquots (200 µl to 2 mL) of brain heart infusion (BHI) broth of *A. paragallinarum* PCR-positive cases, cryogenic storage vials with 1 mL frozen isolates, or as live culture on plates. Genomic DNA was extracted using the QIAamp DNA mini-Kit (Qiagen, Valencia, CA) from bacterial culture and swabs samples. The manufacturer's protocol for purification from tissues and isolation of genomic DNA from bacterial suspension cultures was followed. Extracted DNA samples were reassessed for *A. paragallinarum* according to Chen et al., 1998. Extracted DNA was kept in the freezer at - 20℃ until further use.

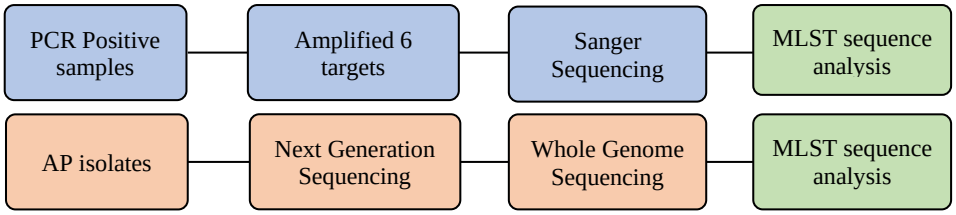


Figure 2. Flowchart of the sample processing workflow for PCR positive samples and bacterial isolates used in this study.

2.4 PCR primer design and amplification conditions

Following identification of the candidate genes, primers were designed to amplify the 18 candidate loci from clinical samples using Primer 3 (6). For amplification of the 18-target segment, 36 primers (18 sets of forward and reverse primers) were initially used. All primers were designed to have an optimum annealing temperature of ~60℃. Each locus selected for sequence typing was shorter in length than the original PCR amplicon size. The PCR reaction was performed in a total volume of 25 µL of reaction mixture, as follows: 12.25 µL of GoTaq Colorless Hot Start, 1.25 µL of 10 mM primer mix, 10.50 µL of nuclease-free water, and 1.0 µL of DNA template. All reactions

were performed by the MiniAmp™ Thermal Cycler (Thermo Fisher Scientific Inc., USA). The cycling program was as follows; a hot start at 95℃ for 2 min, 35 cycles of denaturing at 94℃ for 30 sec, annealing at 60℃ for 30 sec, extension at 72℃ for 90 sec, and a final extension at 72℃ for 5 min. Correct amplicon sizes were confirmed by visualization on 0.5% agarose gel stained with 5 µL ethidium bromide (10 mg/mL) in a 0.5 x Tris Glacial Acetate EDTA (TAE) buffer after electrophoresis and an ultraviolet trans illumination using.

2.5 Specificity and sensitivity of the PCR primers

To evaluate the specificity of the designed primers for the final selected loci, the primers were queried using primer BLAST tool at NCBI against standard nr/nt database (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). To test the sensitivity of the primers for selected loci and the HPG2 primers, DNA was extracted from AP10_DE_20 and measured using Qubit (Invitrogen™, USA). Then, a ten-fold serial dilution of the DNA extracted was prepared from a starting concentration of 37 ng/ µL which was equivalent to 1.4×10^7 genome copies per µl (10^7-10^3). The genome copy number was calculated using DNA copy number dilution calculator available at https://www.technologynetworks.com/tn/tools/copynumbercalculator. The lowest DNA concentrations yielding visible products during agarose gel electrophoresis represented the detection limit for the PCR primers of each locus.

2.6 PCR purification and sequencing of PCR amplicons

PCR products showing appropriate band sizes were purified from the PCR mix with the QIAquick PCR Purification Kit (Qiagen, Valencia, CA) following manufacturer's recommendations. Purification was used to maximize the sequence quality by removing unnecessary residuals of the PCR reaction that may hinder the sequencing reaction. DNA concentration and purity of purified PCR product was determined by NanoDrop 1000 Spectrophotometer V3.5 (*NanoDrop*

Technologies, Inc., USA) and Qubit.4 (Invitrogen™, USA). The sequenced amplicons were trimmed to avoid poor sequence quality associated with the ends of Sanger-sequenced amplicons.

2.7 Targeted sequencing

For samples where WGS was not available, the PCR amplicon of MLST loci and HPG2 were sequenced using Applied Biosystems 3130/3130xl Genetic Analyzers at the core facility of the Department of Veterinary Medicine, University of Maryland. For amplicon preparation, cycle sequencing was done using the BigDye Terminator v3.1 Kit according to manufacturer's protocol followed by cleaning of BigDye Sequencing Reactions using HighPrep DTR (MagBio Genomics, Gaithersburg, Maryland) following manufacturer's protocol.

2.8 Whole genome sequencing

For received samples where pure isolates were available, WGSs were generated. The extracted DNA was used for sequencing library preparation with an Illumina DNA preparation-tagmentation kit with IDT for Illumina DNA/RNA unique dual (UD) indexes, set A (Illumina, USA). The sequencing was performed using the Illumina MiSeq system (v3 reagent kit, with 2 × 300-bp paired-end reads) at the Molecular Epidemiology Laboratory at University of Maryland.

2.9 WGS assembly and annotation

Raw reads from Illumina sequencing were preprocessed and assembled through the standard pipeline designed by the BV-BRC: Bacterial and Viral Bioinformatics Resource Center (https://www.bv-brc.org/).

2.10 MLST Sequence analysis and sequence type assignment

For WGS, MLST targets were queried, and allelic profiles and sequence types were identified using processing WGS option within the SeqSphere (Version 7.0.4). For clinical samples with Sanger sequencing data, target sequence assembly and quality check were performed using the

same software. For each of the six loci, every new sequence was assigned a unique allele number in the order they arose, and any unique combination of the six allele numbers was assigned a unique sequence type number in the order they occurred. Clonal complexes were identified as a collection of samples that share at least 4 similar alleles from the ancestral sequence type of the cluster. Analysis of MLST data was performed using Ridom SeqSphere+ Version7.0.4. The analysis included percent of G+C content, number of alleles, polymorphic sites, percent of DNA variability, and the Hunter-Gaston Discriminatory index (DI). These values were either calculated manually or automatically using SeqSphere (Version7.0.4).

2.11 HPG2 analysis

The reference sequence of the HPG2 hypervariable region segment (444 nucleotides) was queried against 57 whole genome sequences. The HPG2 sequence from the clinical samples was generated by direct amplification and Sanger sequencing. The final phylogenetic tree was generated for 68 *A. paragallinarum* samples as a HPG2 gene could not be identified in seven WGS samples.

2.12 cgMLST analysis for WGS samples

An adhoc cgMLST target definer analysis and filtering procedure, results and list of core, accessory and discarded genes are listed in Appendix B. Fifty-seven WGS samples were typed using the adhoc-cgMLST scheme. Each genome was screened for 1170 genes using SeqSphere (Version7.0.4). A phylogenetic tree was generated using the neighbor joining method.

Chapter 3: Results

3.1 Selected gene candidates for MLST loci

18 candidate loci were successfully amplified from 3 clinical samples (Figure 3). All amplicons were sequenced, and the sequence was used for further evaluation and selection of the final loci. Different combinations of seven loci of the eighteen candidates were assessed to type 42 samples, and six loci with the highest discriminatory power were selected to formulate the final scheme. Table 2 shows their primer sequence, amplicon size, and the final segment size of each locus.

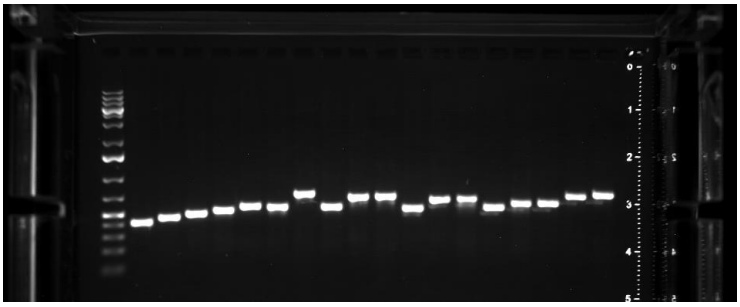


Figure 3. Agarose gel image of PCR amplicons of eighteen candidate loci that were successfully amplified with single bands using DNA from AP09_DEL_20. Starting from lane 2, lane 2 to 20 were loaded with 1000 bp DNA ladder, EIA51_00135, EIA51_00965, EIA51_01020, EIA51_02965, EIA51_03445, EIA51_04255, EIA51_06250, EIA51_06635, EIA51_06735, EIA51_07360, EIA51_07390, EIA51_07505, EIA51_08000, EIA51_08215, EIA51_08385, EIA51_08435, EIA51_10630, EIA51_10680.

Table 2 The final six loci selected for the A. paragallinarum MLST scheme, their forward and reverse primers, amplicons size, final segment size, start and end point from start codon.

Gene	Locus Tag	Primer Sequence (5' to 3')	Amplicon Size (bp)	Final Segment		
				Size (bp)	Start	End
metF	EIA51_RS06635	F: GATTGATGCCACGCCTGAAG R: GCTGCGGTTTAGGGTGTAGA	568	495	315	810
prfC	EIA51_RS08435	F: AGGGGCATCGAACGAATTTG R: CAAATTGCAGCACACCAACC	635	549	759	1308
rpsA	EIA51_RS08385	F: AGATCCTTGGGTTGCGATTG R: AACCTTCAACTCCGCCATCT	638	534	852	1386
rseA	EIA51_RS00965	F: TGATGGCGAACAAATCAGYGA R: GCTCAACTTGCTCTGGTGTAA	455	408	51	459
ufp1	EIA51_RS10630	F: AGGGCTTGCTCAATTTGGTG R: TAGGCGATTTCCACTGGTGT	742	669	441	1110
xseA	EIA51_RS08215	F: CGTATCGGGGCATTGGTATC R: AATAACGGGAATGGCGGAGT	574	513	162	675

3.2 Allelic profiles and sequence types

The analysis involved amplifying and sequencing six specific loci (*metF, prfC, rpsA, rseA, ufp1,* and *xseA*) in a total of 75 samples. These loci yielded 31 distinct sequence types as shown in Table 3. Among the six housekeeping genes examined, *prfC* and *rseA* displayed the highest number of alleles with 24 each. Following them were *metF* with 19 alleles, *ufp1* with 18 alleles, *xseA* with 17 alleles, and *rpsA* with 16 alleles. Based on the Hunter- Gaston discriminatory index (DI), the DI for the individual loci ranged from 0.822 for *ufp1* and 0.843 for *rseA*. The concatenated sequences were 0.832 with a 95% confidence interval between [0.823-0.934] (Table 4).

Table 3. The 31 unique sequence types, their allelic profiles, and their frequency percentage out of the 75 *A. paragallinarum* samples used.

Sequence Types (ST)	Profile						Frequency %
	metF	*prfC*	*rpsA*	*rseA*	*ufp1*	*xseA*	
ST-1	1	1	1	1	1	1	1.3
ST-2	1	1	1	1	2	1	18.7
ST-3	1	1	2	2	2	2	2.7
ST-4	1	1	11	1	2	10	1.3
ST-5	1	16	1	1	2	1	1.3
ST-6	1	18	1	1	2	1	1.3
ST-7	2	2	3	3	3	3	29.3
ST-8	2	2	3	15	3	3	1.3
ST-9	3	3	4	4	4	4	1.3
ST-10	4	4	5	5	5	5	4.0
ST-11	5	5	6	6	6	6	4.0
ST-12	5	17	6	20	6	6	1.3
ST-13	5	20	6	6	6	6	1.3
ST-14	6	6	7	7	7	7	1.3
ST-15	7	7	8	8	8	8	4.0
ST-16	7	7	12	13	11	11	1.3
ST-17	7	8	8	8	4	9	1.3
ST-18	8	6	7	10	9	7	4.0
ST-19	8	6	7	11	9	7	1.3
ST-20	9	1	1	22	2	1	1.3
ST-21	10	10	10	12	10	5	1.3
ST-22	11	11	8	8	4	2	2.7
ST-23	12	2	3	3	3	3	1.3
ST-24	13	12	10	16	12	12	1.3
ST-25	13	21	14	23	18	17	1.3
ST-26	14	14	14	18	14	14	1.3
ST-27	15	15	10	19	15	15	1.3
ST-28	16	12	10	16	16	12	1.3
ST-29	17	13	14	17	13	13	1.3
ST-30	18	19	14	21	17	16	1.3
ST-31	19	22	15	24	14	15	1.3

Table 4. The final selected 6 loci, concatenated sequence and their segment size, percent of G+C content, number of alleles, percent of DNA variability, polymorphic sites, and the Hunt-Gaston Diversity Index

	metF	*prfC*	*rpsA*	*rseA*	*ufp1*	*xseA*	Concatenated sequence
Segment size	495	549	534	408	669	513	3168
% G+C content	45%	50%	36%	43%	47%	49%	45%
Number of alleles	19	24	16	24	18	17	118
Polymorphic sites	27	79	21	20	55	47	249
%DNA variability	5.5%	14.4%	3.9%	4.9%	8.2%	9.2%	7.6%
Hunter-Gaston DI	0.829	0.837	0.827	0.843	0.822	0.832	0.832
DI Confidence interval (95% CI)	[0.772-0.887]	[0.777-0.890]	[0.772-0.882]	[0.786-0.900]	[0.762-0.881]	[0.774-0.890]	[0.823 - 0.934]

3.3 Sensitivity and Specificity of the PCR primers

The detection limit for the PCR primers of each locus was to 1.4×10^3 genome copies per µL, indicating higher sensitivity for amplifying *A. paragallinarum* DNA from clinical sample. Additionally, the BLAST results confirmed that all primer pairs were highly specific to *A. paragallinarum* with no cross-amplification observed from other species, ensuring accurate and reliable detection of the target pathogen.

3.4. MLST data analysis

Figure 4 shows the final 6 loci and their number of alleles, polymorphic sites, and percent of DNA variability amongst the 75 *A. paragallinarum* samples. The figure also displays values of HPG2 for comparison. Of the six housekeeping genes, *prfC* and *ufp1* had the highest percentage of DNA variability (14.4% and 9.3% respectively), followed by *xseA* (9.2%). The average nucleotide diversity for the concatenated sequence of the six MLST loci is 7.6%, which is almost double the DNA variability rate of the HPG2 region (3.4%) among typed samples. There was no consistent correlation between the number of alleles and percent of DNA variability and the number of polymorphic sites for each MLST locus. A wide range of GC content among MLST loci was observed ranging from 36% for *rpsA* to 50% for *prfC*. This variation in GC content may indicate

that the selected MLST loci are dispersed across the genome of *Avibacterium paragallinarum*, suggesting diverse evolutionary origins and functions of MLST targets. This finding is consistent with previous studies that have shown MLST loci to be in different regions of the genome in various bacterial species. Furthermore, the range of GC content observed among the MLST loci highlights the importance of considering the genomic context when selecting MSLT targets and designing primers for MLST analysis (Table 4).

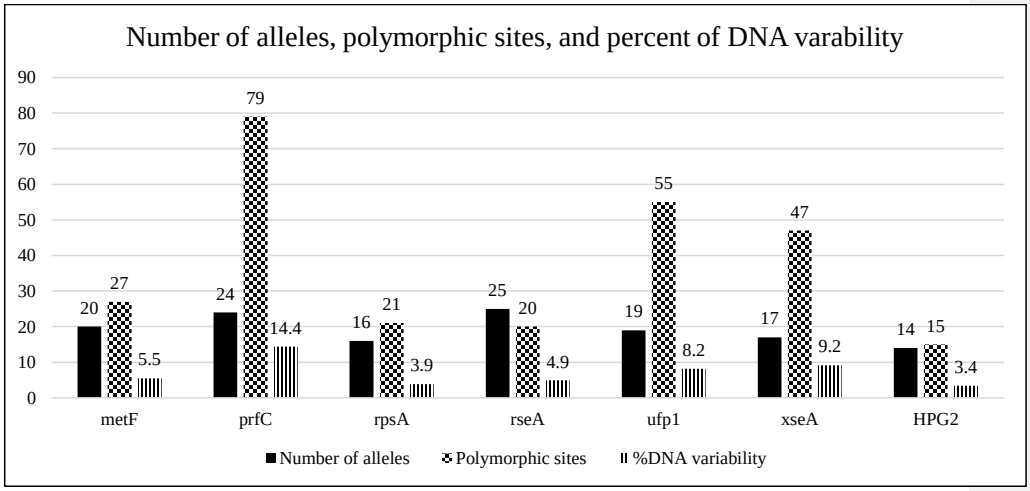


Figure 4. The final selected 6 loci and their number of alleles, Polymorphic sites, and percent of DNA variability in 75 AP samples. In addition, values of HPG2 are displayed.

3.5 MLST phylogenetic analysis of phylogenetic analysis of *A. paragallinarum*

This study aimed to investigate the population structure and diversity of *A. paragallinarum* using a collection of 75 samples. The results obtained from MLST analysis were utilized to create a minimum spanning tree, as depicted in Figure 5. Each circle on the tree represents a distinct sequence type identified through MLST analysis, with the numbers along the connecting lines indicating the number of alleles that differ between each sequence type. Moreover, the color-coding on the spanning tree represents different clonal clusters observed within this sample

collection. By employing our newly developed MLST scheme, we classified these 75 samples into 31 unique sequence types and identified eight clonal clusters or complexes. In figure 7, the ST of each sample is displayed to the right of the sample in the dendrogram. Additionally, Figure 6 displays the same minimum spanning tree where samples are color-coded based on their geographical origin for further insights into distribution patterns.

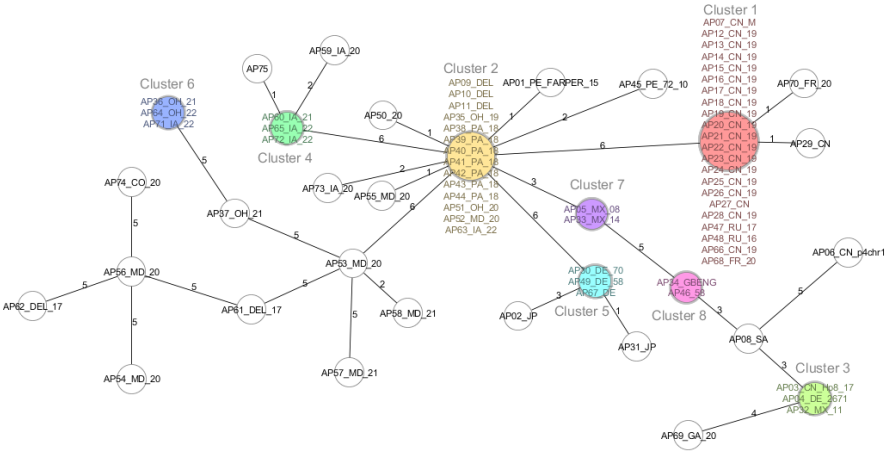


Figure 5. A minimum spanning tree for the 75 samples and 8 clusters typed by the six loci MLST. Each unique sequence type is represented by one circle. Eight clusters were identified by having no more than 2 different alleles from the ancestral sequence type. The number of differing alleles between different sequence types are displayed on the connecting lines. The tree and the distance matrix underlying the network were generated using SeqSphere+ software. Note the sample name includes an initial arbitrary number followed by the state or country from which the sample was collected followed by the 2-digit number representing the year of isolation. Samples were colored coded based matching cluster number.

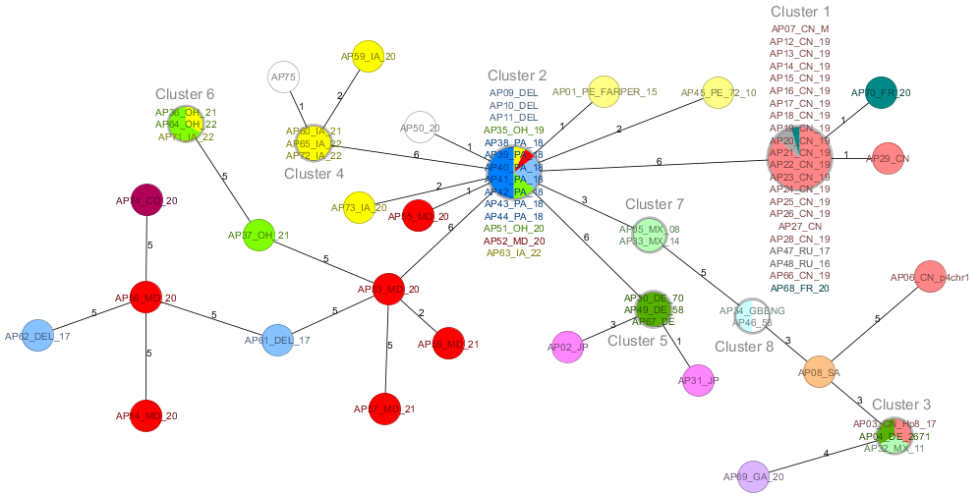


Figure 6. Minimum spanning tree analysis based on the allelic profiles of six MLST gene targets from 75 whole genome and target sequenced AP samples using the SeqSphere^{+} software. The numbers on the connecting lines illustrate the numbers of target genes with differing alleles. Samples are color-coded according to the geographic origin as presented in the legend. The distance matrix underlying the network was built from all pairwise allelic profile comparisons of Six MLST targets, using the pairwise, no missing values option in SeqSphere^{+} software. Note the sample name includes an initial arbitrary number followed by the state from which the sample was collected followed by the 2-digit number representing the year of isolation.

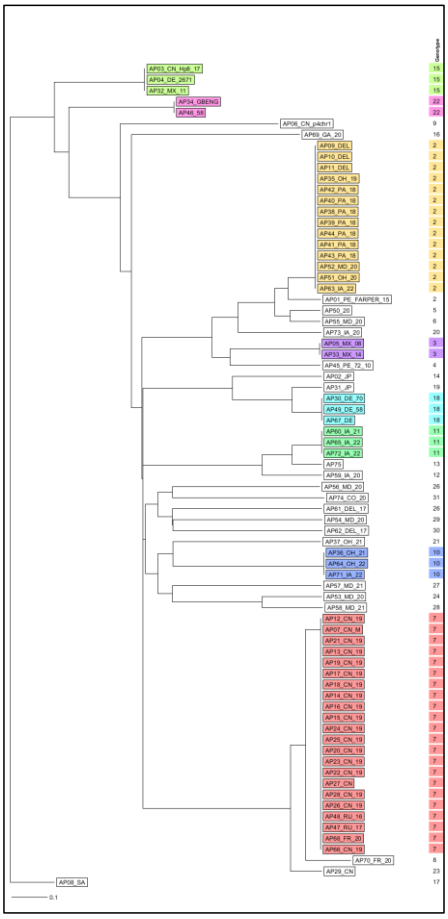


Figure 7. A six-locus MLST dendrogram diplaying 75 samples including sanger sequenced clinical samples. The distance matrix was built from all pairwise allelic profile comparisons of seven MLST loci, using the pairwise no missing values option in SeqSphere+software. The dendrogram is rooted by AP08_SA generated using neighbor joining method in SeqSphere+software. Samples are color coded according to their cluster. ST of each sample are displayed to the right of the sample.

3.6 *A. paragallinarum* cgMLST analysis for WGS samples

To assess the accuracy of the newly developed MSLT scheme in predicting *A. paragallinarum* phylogeny, we compared it to a widely accepted gold standard method like adhoc cgMLST that utilizes whole genome sequence information (Figure 7). The MLST study involved analyzing a total of 75 samples, out of which 57 WGS samples only were used for adhoc cgMLST evaluation purposes. Out of these, eleven samples were sequenced on the Illumina MiSeq platform specifically for this research. The remaining 46 *A. paragallinarum* genomes were obtained from NCBI GenBank. By analyzing the genetic information from these samples using the data set containing 1170 core genes, it was possible to determine whether our genomes represent a diverse set of samples. Interestingly, there were 47 different cgMLST type observed in total across all 57 samples examined (Figure 8). Additionally, an analysis of the discriminatory index (DI) of our cgMLST yielded a DI value of 0.988 with a confidence interval ranging between [0.976-1.0], indicating the high discriminatory power of cgMLST.

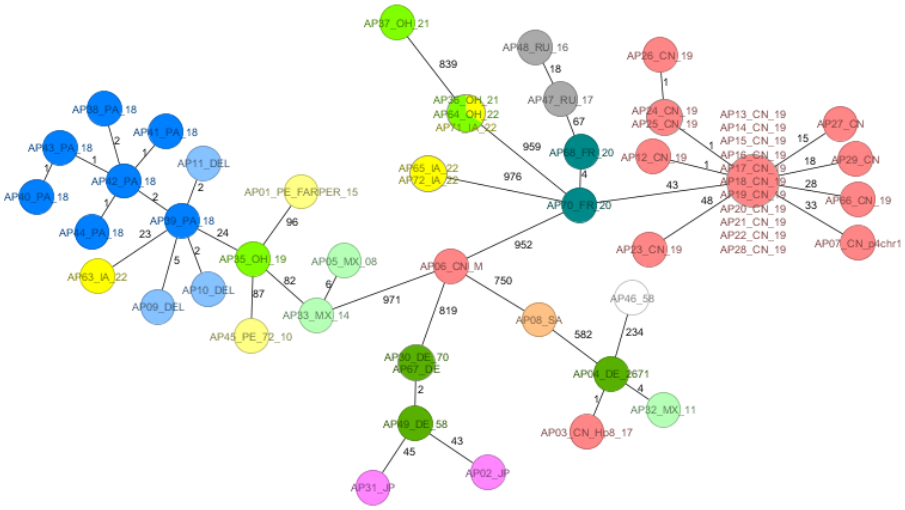


Figure 8. Minimum spanning tree analysis based on the allelic profiles of 1170 core gene targets from 57 whole genome AP samples using the SeqSphere^{+} software.

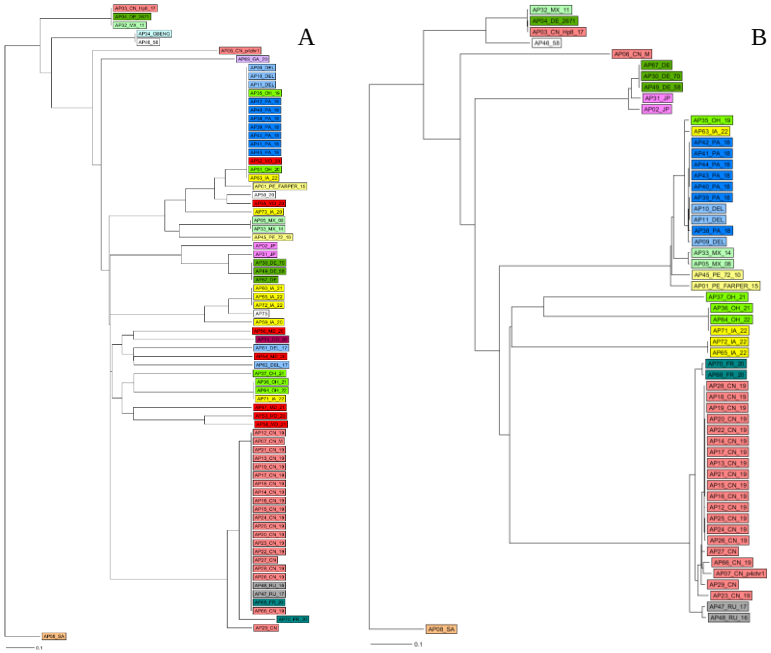


Figure 9 A.) MLST dendrogram displaying 75 samples including whole genome and target sequenced samples. The sequence type (ST) and cluster number for each sample is displayed on the columns labeled genotype and cluster. The distance matrix was built from all pairwise allelic profile comparisons of six MLST loci in SeqSphere+ software and B) cgMLST dendrogram displaying only 57 whole genome sequenced *A. paragallinarum* samples. The distance matrix was built from all pairwise allelic profile comparisons of 1170 cgMLST targets, using the pairwise ignoring missing values option in SeqSphere+ software. Using this option, genes with at least one missing value are not completely removed from the comparison but are ignored only during a pairwise comparison in case of a missing value. Dendrogram are rooted by AP08_SA sample.

3.7 HPG2 region analysis of *A. paragallinarum*

To compare the newly developed MLST scheme with the existing sequence targets for typing *A. paragallinarum*, a comprehensive analysis was conducted on 68 samples. The partial HPG2 region served as the basis for this typing, resulting in these samples being categorized into 12 distinct sequence types (ST) (Figure 9), with a calculated diversity index of 0.788 and falling within a confidence interval of [0.724-0.851]. Despite that, it became clear that while the HPG2

region analysis revealed limited genetic variation among all 68 samples, it showed lower discriminatory power compared to MLST in general and often contradicted both the MLST typing (Figure 10) as well as epidemiological information regarding such as geographical origin associated with each sample. For instance, several international samples, including AP05_MX_08, AA33_MX_14, and AP01_PE_FARPER_15, were found to share the same HPG2 sequence type as several US samples such as AP40_PA_18, AP52_MD_20, AP35_OH_19, and AP63_IA_22. Moreover, a United States sample named AP54_MD_20 showed a remarkably close genetic relationship with the Chinese samples analyzed in this study compared to other US samples. These findings indicate that the MLST scheme offers a more accurate and reliable method for typing *A. paragallinarum* when compared to analyzing only the HPG2 hypervariable region.

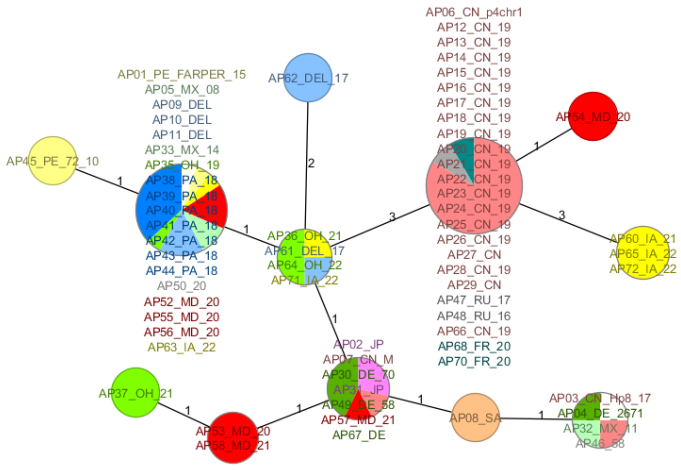


Figure 10. Spanning tree generated for 68 samples based on HPG2 sequence comparison. Distance based on nucleotide columns from 15 SNV positions with no missing values. Note the limited discriminatory power of HPG2 where samples are typed in to only 12 diverse types.

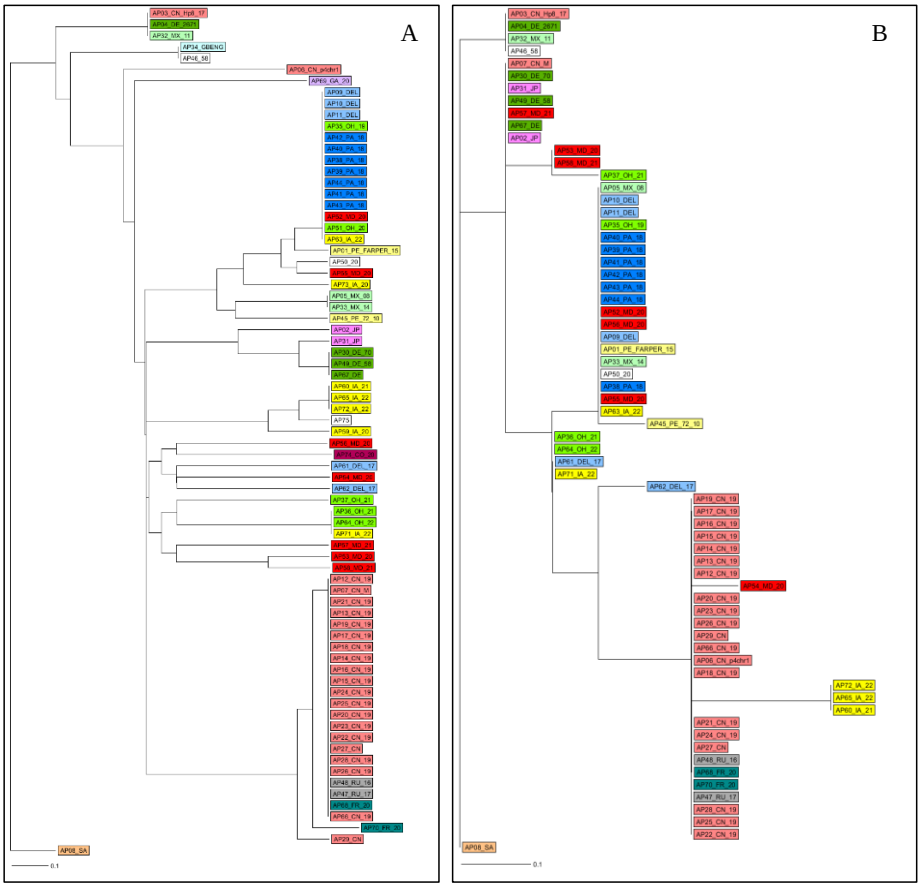


Figure 11 A.) MLST dendrogram displaying 75 samples including whole genome and target sequenced samples. The sequence type (ST) and cluster number for each sample is displayed on the columns labeled genotype and cluster. The distance matrix was built from all pairwise allelic profile comparisons of six MLST loci in SeqSphere+ software and B) Shows a dendrogram generated for 68 samples based on HPG2 sequence comparison. Distance based on nucleotide columns from 15 SNV positions with no missing values. The dendrogram is rooted by AP08_SA sample. The dendrogram was generated using neighbor joining method in SeqSphere+ software.

Note the limited discriminatory power of HPG2 where samples are typed in to only 12 different types.

3.8 Discrimination Index (DI) of cgMLST, MLST, and current HPG2

The Discriminatory Index (Hunter Gaston) is the probability a genotyping method will classify different strains into different types (Hunter & Gaston, 1988). The DI values range from [0-1], indicating that strains in a population are different when the DI is closer to 1. For the *A. paragallinarum* MLST, the DI was 0.879 and higher than the HPG2 DI of 0.788. This significance reveals the newly developed MLST can strain differentiate more accurately when we consider that MLST results agreed with the epidemiological information of different strains in contrast to HPG2 results. Despite the superiority of MLST over HPG2 region, cgMLST showed higher discriminatory power than MLST as shown in Table 5. The cgMLST results were in agreement with the epidemiological information and confirmed the MLST results however one key difference to note that MSLT results can be generated based on the sequence information of 6 gene segments only while cgMLST is based on 1170 full genes which require the generation of WGS data from pure isolates that is challenging to obtain from clinical cases. This fact highlights that MLST is still a valuable tool to type PCR positive samples even the era of next generation sequencing when obtaining WGS is not attainable.

Table 5. The number of WGS and target-sequenced A. paragallinarum samples, targets used for genotyping and the discriminatory index for the newly developed MLST, HPG2, and the cgMLST.

	# Samples	*# Samples Skipped*	*# Different Types*	*Discriminatory Index*	*Confidence Interval (95% CI)*
HPG2	68	0	12	0.788	[0.724-0.851]
MLST	75	0	31	0.879	[0.823-0.934]
cgMLST	57	0	47	0.988	[0.976-1.0]

Chapter 4: Discussion

Recently, there has been an increased prevalence of Infectious Coryza (IC) in commercial and noncommercial chickens including layers and broilers in the Northeastern United States. Current strain differentiation methods of *A. paragallinarum* such as traditional and molecular serotyping have limited value and discriminatory power to investigate the epidemiology of *A. paragallinarum* outbreaks and population structure. Therefore, a feasible and accurate strain differentiation method was needed to investigate the epidemiology of IC and to prevent future outbreaks. We hypothesized that Multilocus Sequence Typing (MLST) could provide better discriminatory power and a more accurate method for typing *A. paragallinarum* directly from clinical samples without isolation. MLST is a relatively new sequence typing approach that has been used successfully for typing both animal and human pathogens. It uses the sequence information of several internal fragments of up to seven conserved housekeeping genes (HKG) to differentiate between bacterial strains. We have studied all *A. paragallinarum* whole genome sequences available at GenBank and identified six gene fragments to be used as MLST targets based on nucleotide diversity and discriminatory power. The phylogenetic relationship between all *A. paragallinarum* genomes and clinical samples were compared using MLST, Adhoc cgMLST, and HPG2 when sequences were available. The MLST appeared to have higher discriminatory power compared to HPG2. Moreover, a high degree of agreement was found between our newly developed MLST and adhoc cgMLST results using the currently tested samples. In total, 75 samples, mixed of WGS and clinical samples, were evaluated for this study where the results showed strain differentiation between *A. paragallinarum* strains from clinical samples.

Generally, MLST showed a diverse population structure of the global and national *A. paragallinarum* population. The sample collection exhibits a diverse range of geographic origins and sources where many sequence types are represented with only one sample, however

there are instances where certain samples show less variation when considering allelic distance. Moreover, large clusters of clonal samples were revealed among US and Chinese samples. This may be due to the sampling bias related to the increased sampling of strains related to multiple commercial outbreaks. It was observed that all samples from Cluster 1 originated from China and shared identical STs, except for one sample which appeared to have come from France according to its attached metadata. cgMLST revealed that this French strain (AP70_FR_20) was indeed different from the Chinese samples with 43 different alleles and similar observation was noted for the Russian samples (AP47_RU_17 and AP48 RU_16). Moreover, the superior discriminatory power of cgMLST has differentiated the twenty Chinese samples within MLST cluster 1 into 9 different cgMLST types with 15 samples having no more than one allele different from each other's. Similarly, the majority of *A. paragallinarum* samples from the Mid-Atlantic and Midwestern states (PA, MD, DE, OH, and IA) and had most recent outbreak been grouped together into a single MLST sequence type under Cluster 2. This suggests that these strains might be closely related to each other. However, according to cgMLST, cluster 2 appears to be more diverse than revealed by MLST. Despite that there was unambiguous evidence of clonality between several commercial samples from Pennsylvania and Delaware. This may suggest that a single sequence type or strain may be responsible for the outbreak in the Northeastern region of the United States. Furthermore, it may indicate a biosecurity breach within that region.

On the other hand, very diverse sequence types were discovered for many noncommercial poultry samples collected from submission to the Salisbury Animal Health Diagnostic Laboratory Maryland in 2020 and 2021 (AP53_MD_20, AP54_MD_20, AP55_ MD_20, AP56_MD_20, AP57_MD_21, and AP58_MD_21) and auction market samples from Delaware (AP62_DEL_17 and AP61_DEL_17). This finding is expected among these noncommercial poultry as non-

commercial poultry flocks are most likely privately owned and do not have the same connections and transmission routes as commercial farms. The diversity of sequence types in the noncommercial farms could confirm that there is no or weak chance of transmission amongst these flocks. This also, potentiates the explanation of the clonality of sequence types seen among commercial poultry farms where they may share multiple links that may contribute to the transmission of *A. paragallinarum* like vaccine crews, feed mill trucks, and other maintenance services.

MLST successfully grouped samples according to their geographical origin up to the state level as we can see samples from Ohio and Iowa grouped into different clusters that are distantly related from other Northeastern U.S. samples. In addition, MLST was able to differentiate between samples from the same state as in the case of Maryland samples where they are collected mainly from noncommercial flocks. It is worth noting that MLST results also suggest that there are multiple sequence types circulating among commercial operation within different states. However, more samples are needed to confirm this observation.

Multiple examples within this data set have proved the accuracy of MLST compared to cgMLST especially when it comes to identifying similar or related sequence types. For instance, MLST was able to identify cluster 3 samples (AP03_CN_HP8_17, AP04_DE_2671, and AP32_MX_11) as a single sequence type using MLST despite their WGS were submitted to GenBank from three different countries and years apart. This finding was confirmed by cgMLST as well. The explanation for this controversy is that this sample is a reference strain that was sequenced and submitted from different countries and despite that MLST was able to correctly relate them.

Another example where MLST can help in understanding the epidemiological situation over time is clearly displayed Cluster 4 samples. Cluster 4 includes samples AP60_IA_21, AP65_IA_22,

and AP72_IA_22 which originated from the same state and were collected in two successive years. This may indicate the persistence and transmission of the same sequence type over time. This also may be consistent with the carrier status of *A. paragallinarum* infected birds that facilitates the survival and dissemination of *A. paragallinarum* in older infected birds as in multiage layer facilities.

In this study, MLST helped better evaluate the validity of HPG2 region for phylogenetic analysis of *A. paragallinarum*. The HPG2 showed lower discriminatory power compared to MLST in general and often contradicted both the MLST typing (Figure 6A) as well as epidemiological information regarding geographical origin associated with each sample. For instance, several international samples, including AP05_MX_08, Ap33_MX_14, and AP01_PE_FARPER_15, were found to share the same HPG2 sequence type as several US samples such as AP40_PA_18, AP52_MD_20, AP35_OH_19, and AP63_IA_22. Moreover, A United States sample named AP54_MD_20 showed a remarkably close genetic relationship with the Chinese samples analyzed in this study compared to other US samples. These findings indicate that the MLST scheme offers a more accurate and reliable method for typing *A. paragallinarum* when compared to analyzing only the HPG2 region.

While cgMLST has superior discriminatory power than MLST and reveals a more accurate relationship between samples, it is use is limited by the need for WGS and isolation of *A. paragallinarum* in pure isolates that is difficult to obtain as *A. paragallinarum* is a fastidious organism. In addition, there is no standard cgMLST scheme for *A. paragallinarum* typing. Therefore, MLST is very relevant and needed to allow typing of PCR positive *A. paragallinarum* clinical samples directly without isolation.

In this study, it was not possible to evaluate the correlation between MLST and traditional or molecular serotyping as we do not have access to antisera to perform serological testing and access to all strains used in the study and. However, comparing the reported serotypes for some of the public samples from GenBank, we found that there is no correlation between MLST types and serotypes. This is consistent with the nature of MLST as MLST analysis is used to reveal the phylogenetic relatedness between different samples based on housekeeping genes and not surface antigens like HMTp*210* which has been shown to be not useful for epidemiological typing or phylogenetic relatedness analysis (Tan et al. 2021). In that paper, samples from different countries were classified as the same serotype.

Chapter 5: Summary and Conclusion

The concluding chapter of this thesis encapsulates the significant advancements made in the development of Multilocus sequence typing (MLST) as a benchmark method for typing *A. paragallinarum.* This method stands out for not requiring the isolation of pure cultures and offers a more nuanced discriminatory power as compared to existing DNA or serum-based schemes, aligning well with geographical and temporal data. The utility of MLST extends to outbreak investigations, enhancing our grasp of *A. paragallinarum* molecular epidemiology and bolstering prevention and control measures.

This research aimed to develop an epidemiological tool that would assist in understanding the population structure of *A. paragallinarum* across various levels. The thorough analysis of diverse samples has enabled a detailed evaluation of this population structure as represented within our diverse but limited data set. The objectives of this study have been met with distinction; the acquired data facilitates the differentiation of *A. paragallinarum* strains and elucidates the connections between local, regional, and international strains.

The MLST analysis has contributed some insights into the epidemiology of *A. paragallinarum* strains involved in recent outbreaks in the Mid-Atlantic and Midwestern United States and has allowed to the differentiation of strains in both commercial and non-commercial chickens. However, more samples and better analysis methods are needed to enhance that understanding and verify our findings. With 75 samples successfully sequenced, I identified 31 distinct strains using MLST, contributing to a more robust understanding of the epidemiology of IC, which is critical for enhancing outbreak management and control strategies. The newly developed scheme could be a promising tool for strain differentiation of *A. paragallinarum* for new samples. It enables a better understanding of *A. paragallinarum* epidemiology and population structure. In addition, it

provides non-ambiguous, portable, and universal typing for AP that would facilitate prevention and control efforts worldwide.

Looking forward, the intent is to integrate this MLST scheme into the PubMLST database (https://pubmlst.org/), fostering the submission and exchange of new sequence types and isolate data nationally and across the world. This will expand our knowledge of *A. paragallinarum* epidemiology on both a local and global scale. Moreover, the establishment of a standardized core genome MLST (cgMLST) is a much-needed next step. This step would leverage the affordability of sequencing technologies and the extensive data obtainable from whole genome sequencing for applications beyond phylogenetic analysis, such as tracking antimicrobial resistance and pinpointing virulence factors. Until these advancements come to fruition, MLST is predicted to serve as a potent and invaluable interim tool, meeting the immediate demands of the field.

Appendices

Appendix A. The 75 *A. paragallinarum* samples used in this study and their related information as indicated in the header column. Please note that the table includes 75 rows with the same sample ID and table header across multiple pages.

Sample ID	Sequence Type	*metF*	*prfC*	*rpsA*	*rseA*	*ufp1*	*xseA*	Group	Serovar
AP01_PE_FARPER_15	1	1	1	1	1	1	1	PE- Peru	C2
AP02_JP	14	6	6	7	7	7	7	JP- Japan	A1
AP03_CN_Hp8_17	15	7	7	8	8	8	8	CN- China	A1
AP04_DE_2671	15	7	7	8	8	8	8	DE- Germany	B1
AP05_MX_8	3	1	1	2	2	2	2	MX- Mexico	C1
AP06_CN_p4chr1	9	3	3	4	4	4	4	CN- China	Unknown
AP07_CN_M	7	2	2	3	3	3	3	CN- China	
AP08_SA	17	7	8	8	8	4	9	SA- South Africa	C3
AP09_DEL_20	2	1	1	1	1	2	1	DEL- Delaware, USA	?
AP10_DEL_20	2	1	1	1	1	2	1	DEL- Delaware, USA	?
AP11_DEL_20	2	1	1	1	1	2	1	DEL- Delaware, USA	?
AP12_CN_19	7	2	2	3	3	3	3	CN- China	?
AP13_CN_19	7	2	2	3	3	3	3	CN- China	?
AP14_CN_19	7	2	2	3	3	3	3	CN- China	?
AP15_CN_19	7	2	2	3	3	3	3	CN- China	?
AP16_CN_19	7	2	2	3	3	3	3	CN- China	?
AP17_CN_19	7	2	2	3	3	3	3	CN- China	?
AP18_CN_19	7	2	2	3	3	3	3	CN- China	?
AP19_CN_19	7	2	2	3	3	3	3	CN- China	?
AP20_CN_19	7	2	2	3	3	3	3	CN- China	?
AP21_CN_19	7	2	2	3	3	3	3	CN- China	?
AP22_CN_19	7	2	2	3	3	3	3	CN- China	?
AP23_CN_19	7	2	2	3	3	3	3	CN- China	?
AP24_CN_19	7	2	2	3	3	3	3	CN- China	?
AP25_CN_19	7	2	2	3	3	3	3	CN- China	?
AP26_CN_19	7	2	2	3	3	3	3	CN- China	?
AP27_CN	7	2	2	3	3	3	3	CN- China	?
AP28_CN_19	7	2	2	3	3	3	3	CN- China	?
AP29_CN	23	12	2	3	3	3	3	CN- China	?
AP30_DE_70	18	8	6	7	10	9	7	DE- Germany	?
AP31_JP	19	8	6	7	11	9	7	JP- Japan	?
AP32_MX_11	15	7	7	8	8	8	8	MX- Mexico	?
AP33_MX_14	3	1	1	2	2	2	2	MX- Mexico	?
AP34_GBENG	22	11	11	8	8	4	2	GBENG- England	?
AP35_OH_19	2	1	1	1	1	2	1	OH- Ohio, USA	?
AP36_OH_21	10	4	4	5	5	5	5	OH- Ohio, USA	?

AP37_OH_21	21	10	10	10	12	10	5	OH- Ohio, USA	?
AP38_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP39_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP40_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP41_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP42_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP43_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP44_PA_18	2	1	1	1	1	2	1	PA- Pennsylvania, USA	?
AP45_PE_72_10	4	1	1	11	1	2	10	PE- Peru	?
AP46_58	22	11	11	8	8	4	2	No Group Assigned	?
AP47_RU_17	7	2	2	3	3	3	3	RU- Russia	?
AP48_RU_16	7	2	2	3	3	3	3	RU- Russia	?
AP49_58	18	8	6	7	10	9	7	DE- Germany	?
AP50_20	5	1	16	1	1	2	1	No Group Assigned	?
AP51_OH_20	2	1	1	1	1	2	1	OH- Ohio, USA	?
AP52_MD_20	2	1	1	1	1	2	1	MD- Maryland, USA	?
AP53_MD_20	24	13	12	10	16	12	12	MD- Maryland, USA	?
AP54_MD_20	29	17	13	14	17	13	13	MD- Maryland, USA	?
AP55_MD_20	6	1	18	1	1	2	1	MD- Maryland, USA	?
AP56_MD_20	26	14	14	14	18	14	14	MD- Maryland, USA	?
AP57_MD_21	27	15	15	10	19	15	15	MD- Maryland, USA	?
AP58_MD_21	28	16	12	10	16	16	12	MD- Maryland, USA	?
AP59_IA_20	12	5	17	6	20	6	6	IA- Iowa, USA	?
AP60_IA_21	11	5	5	6	6	6	6	IA- Iowa, USA	?
AP61_DEL_17	25	13	21	14	23	18	17	DEL- Delaware, USA	?
AP62_DEL_17	30	18	19	14	21	17	16	DEL- Delaware, USA	?
AP63_IA_22	2	1	1	1	1	2	1	IA- Iowa, USA	?
AP64_IA_22	10	4	4	5	5	5	5	OH- Ohio, USA	?
AP65_IA_22	11	5	5	6	6	6	6	IA- Iowa, USA	?
AP66_ZJC	7	2	2	3	3	3	3	CN- China	?
AP67_CCUGA2	18	8	6	7	10	9	7	DE- Germany	A2
AP68_FR_20	7	2	2	3	3	3	3	FR- France	?
AP69_GA_20	16	7	7	12	13	11	11	GA- Gabon	?
AP70_FR_20	8	2	2	3	15	3	3	FR- France	?
AP71_IA_22	10	4	4	5	5	5	5	IA- Iowa, USA	?
AP72_IA_22	11	5	5	6	6	6	6	IA- Iowa, USA	?
AP73_IA	20	9	1	1	22	2	1	IA- Iowa, USA	?
AP74_CO_20	31	19	22	15	24	14	15	CO- Colorado, USA	?
AP75	13	5	20	6	6	6	6	No Group Assigned	?

Sample ID	Sample Source	Sample Type	Sequencing Approach	Country of Isolation
AP01_PE_FARPER_15	GenBank	WGS	WGS	Peru
AP02_JP	GenBank	WGS	WGS	Japan
AP03_CN_Hp8_17	GenBank	WGS	WGS	China
AP04_DE_2671	GenBank	WGS	WGS	Germany
AP05_MX_8	GenBank	WGS	WGS	Mexico
AP06_CN_p4chr1	GenBank	WGS	WGS	China
AP07_CN_M	GenBank	WGS	WGS	China
AP08_SA	GenBank	WGS	WGS	South Africa
AP09_DEL_20	This Study	Clinical Sample	WGS/ Sanger Sequencing	United States
AP10_DEL_20	This Study	Clinical Sample	WGS/ Sanger Sequencing	United States
AP11_DEL_20	This Study	Clinical Sample	WGS/ Sanger Sequencing	United States
AP12_CN_19	GenBank	WGS	WGS	China
AP13_CN_19	GenBank	WGS	WGS	China
AP14_CN_19	GenBank	WGS	WGS	China
AP15_CN_19	GenBank	WGS	WGS	China
AP16_CN_19	GenBank	WGS	WGS	China
AP17_CN_19	GenBank	WGS	WGS	China
AP18_CN_19	GenBank	WGS	WGS	China
AP19_CN_19	GenBank	WGS	WGS	China
AP20_CN_19	GenBank	WGS	WGS	China
AP21_CN_19	GenBank	WGS	WGS	China
AP22_CN_19	GenBank	WGS	WGS	China
AP23_CN_19	GenBank	WGS	WGS	China
AP24_CN_19	GenBank	WGS	WGS	China
AP25_CN_19	GenBank	WGS	WGS	China
AP26_CN_19	GenBank	WGS	WGS	China
AP27_CN	GenBank	WGS	WGS	China
AP28_CN_19	GenBank	WGS	WGS	China
AP29_CN	GenBank	WGS	WGS	China
AP30_DE_70	GenBank	WGS	WGS	Germany
AP31_JP	GenBank	WGS	WGS	Japan
AP32_MX_11	GenBank	WGS	WGS	Mexico
AP33_MX_14	GenBank	WGS	WGS	Mexico
AP34_GBENG	GenBank	WGS	WGS	England
AP35_OH_19	This Study	Clinical Sample	WGS	United States
AP36_OH_21	This Study	Clinical Sample	WGS	United States
AP37_OH_21	This Study	Clinical Sample	WGS	United States
AP38_PA_18	GenBank	WGS	WGS	United States
AP39_PA_18	GenBank	WGS	WGS	United States
AP40_PA_18	GenBank	WGS	WGS	United States

AP41_PA_18	GenBank	WGS	WGS	United States
AP42_PA_18	GenBank	WGS	WGS	United States
AP43_PA_18	GenBank	WGS	WGS	United States
AP44_PA_18	GenBank	WGS	WGS	United States
AP45_PE_72_10	GenBank	WGS	WGS	Peru
AP46_58	GenBank	WGS	WGS	?
AP47_RU_17	GenBank	WGS	WGS	Russia
AP48_RU_16	GenBank	WGS	WGS	Russia
AP49_58	GenBank	WGS	WGS	?
AP50_20	This Study	Clinical Sample	Sanger Sequencing	?
AP51_OH_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP52_MD_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP53_MD_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP54_MD_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP55_MD_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP56_MD_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP57_MD_21	This Study	Clinical Sample	Sanger Sequencing	United States
AP58_MD_21	This Study	Clinical Sample	Sanger Sequencing	United States
AP59_IA_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP60_IA_21	This Study	Clinical Sample	Sanger Sequencing	United States
AP61_DEL_17	This Study	Clinical Sample	Sanger Sequencing	United States
AP62_DEL_17	This Study	Clinical Sample	Sanger Sequencing	United States
AP63_IA_22	This Study	Clinical Sample		United States
AP64_IA_22	This Study	Clinical Sample		United States
AP65_IA_22	This Study	Clinical Sample		United States
AP66_ZJC	GenBank	WGS	WGS	China
AP67_CCUGA2	This Study	WGS	WGS	Germany
AP68_FR_20	This Study	WGS reads	WGS	France
AP69_GA_20	This Study	WGS reads	WGS	Gabon
AP70_FR_20	This Study	WGS reads	WGS	France
AP71_IA_22	GenBank	WGS	WGS	United States
AP72_IA_22	This Study	WGS	WGS	United States
AP73_IA	This Study	Clinical Sample	Sanger Sequencing	United States
AP74_CO_20	This Study	Clinical Sample	Sanger Sequencing	United States
AP75	This Study	Clinical Sample	Sanger Sequencing	United States

Sample ID	State of Isolation	County of Isolation	Year of Isolation	Bird Type	Production Type
AP01_PE_FARPER_15	?	?	2015	Chicken	?
AP02_JP	?	?	?	Chicken	?
AP03_CN_Hp8_17	?	?	2017	Chicken	?
AP04_DE_2671	?	?	?	Chicken	?
AP05_MX_8	?	?	?	Chicken	?
AP06_CN_p4chr1	?	?	?	Chicken	?
AP07_CN_M	?	?	?	Chicken	?
AP08_SA	?	?	?	Chicken	?
AP09_DEL_20	Delaware	?	2020	Chicken	?
AP10_DEL_20	Delaware	?	2020	Chicken	?
AP11_DEL_20	Delaware	?	2020	Chicken	?
AP12_CN_19	?	?	2019	Chicken	?
AP13_CN_19	?	?	2019	Chicken	?
AP14_CN_19	?	?	2019	Chicken	?
AP15_CN_19	?	?	2019	Chicken	?
AP16_CN_19	?	?	2019	Chicken	?
AP17_CN_19	?	?	2019	Chicken	?
AP18_CN_19	?	?	2019	Chicken	?
AP19_CN_19	?	?	2019	Chicken	?
AP20_CN_19	?	?	2019	Chicken	?
AP21_CN_19	?	?	2019	Chicken	?
AP22_CN_19	?	?	2019	Chicken	?
AP23_CN_19	?	?	2019	Chicken	?
AP24_CN_19	?	?	2019	Chicken	?
AP25_CN_19	?	?	2019	Chicken	?
AP26_CN_19	?	?	2019	Chicken	?
AP27_CN	?	?	?	Chicken	?
AP28_CN_19	?	?	2019	Chicken	?
AP29_CN	?	?	?	Chicken	?
AP30_DE_70	?	?	1970	Chicken	?
AP31_JP	?	?	?	Chicken	?
AP32_MX_11	?	?	2011	Chicken	?
AP33_MX_14	?	?	2014	Chicken	?
AP34_GBENG	?	?	?	Chicken	?
AP35_OH_19	Ohio	?	2019	Chicken	?
AP36_OH_21	Ohio	?	2021	Chicken	?
AP37_OH_21	Ohio	?	2021	Chicken	?
AP38_PA_18	Pennsylvania	?	2018	Chicken	?
AP39_PA_18	Pennsylvania	?	2018	Chicken	?
AP40_PA_18	Pennsylvania	?	2018	Chicken	?

AP41_PA_18	Pennsylvania	?	2018	Chicken	?
AP42_PA_18	Pennsylvania	?	2018	Chicken	?
AP43_PA_18	Pennsylvania	?	2018	Chicken	?
AP44_PA_18	Pennsylvania	?	2018	Chicken	?
AP45_PE_72_10	?	?	2010	Chicken	?
AP46_58	?	?	1958	Chicken	?
AP47_RU_17	?	?	2017	Chicken	?
AP48_RU_16	?	?	2016	Chicken	?
AP49_58	?	?	1958	Chicken	?
AP50_20	?	?	2020	Chicken	?
AP51_OH_20	Ohio	?	2020	Chicken	?
AP52_MD_20	Maryland	Wicomico	2020	Chicken	Non-commercial
AP53_MD_20	Maryland	Wicomico	2020	Chicken	Non-commercial
AP54_MD_20	Maryland	Wicomico	2020	Chicken	Non-commercial
AP55_MD_20	Maryland	Wicomico	2020	Chicken	Non-commercial
AP56_MD_20	Maryland	Wicomico	2020	Chicken	Non-commercial
AP57_MD_21	Maryland	Wicomico	2021	Chicken	Non-commercial
AP58_MD_21	Maryland	Wicomico	2021	Chicken	Non-commercial
AP59_IA_20	Iowa	?	2020	Chicken	?
AP60_IA_21	Iowa	?	2021	Chicken	?
AP61_DEL_17	Delaware	Wyoming	2017	Chicken	Non-commercial
AP62_DEL_17	Delaware	Wyoming	2017	Chicken	Non-commercial
AP63_IA_22	Iowa	?	2022	Chicken	?
AP64_IA_22	Iowa	?	2022	Chicken	?
AP65_IA_22	Iowa	?	2022	Chicken	?
AP66_ZJC	?	?	?	Chicken	?
AP67_CCUGA2	?	?	?	Chicken	?
AP68_FR_20	?	?	2020	Chicken	?
AP69_GA_20	?	?	2020	Chicken	?
AP70_FR_20	?	?	2020	Chicken	?
AP71_IA_22	Iowa	?	?	Chicken	?
AP72_IA_22	Iowa	?	?	Chicken	?
AP73_IA	Iowa	?	?	Chicken	?
AP74_CO_20	Colorado	?	2020	Chicken	?
AP75	?	?	?	Chicken	?

Sample ID	Genome Accession(s)	Organism/Name	Status
AP01_PE_FARPER_15	NZ_CP034110.1	Avibacterium paragallinarum	Complete Genome
AP02_JP	RQXP01	Avibacterium paragallinarum	Contig
AP03_CN_Hp8_17	VXDA01	Avibacterium paragallinarum	Scaffold
AP04_DE_2671	RQXQ01	Avibacterium paragallinarum	Contig
AP05_MX_8	NZ_CP050316.1	Avibacterium paragallinarum	Complete Genome
AP06_CN_p4chr1	NZ_CP081939.1	Avibacterium paragallinarum	Complete Genome
AP07_CN_M	CP086713.1	Avibacterium paragallinarum	Complete Genome
AP08_SA	RQXS01	Avibacterium paragallinarum	Contig
AP09_DEL_20	?	Avibacterium paragallinarum	?
AP10_DEL_20	?	Avibacterium paragallinarum	?
AP11_DEL_20	?	Avibacterium paragallinarum	?
AP12_CN_19	JACEVM01	Avibacterium paragallinarum	Scaffold
AP13_CN_19	JACEVN01	Avibacterium paragallinarum	Scaffold
AP14_CN_19	JACEVO01	Avibacterium paragallinarum	Scaffold
AP15_CN_19	JACEVP01	Avibacterium paragallinarum	Scaffold
AP16_CN_19	JACEVQ01	Avibacterium paragallinarum	Scaffold
AP17_CN_19	JACEVR01	Avibacterium paragallinarum	Scaffold
AP18_CN_19	JACEVS01	Avibacterium paragallinarum	Scaffold
AP19_CN_19	JACEVT01	Avibacterium paragallinarum	Scaffold
AP20_CN_19	JACEVU01	Avibacterium paragallinarum	Scaffold
AP21_CN_19	JACEVV01	Avibacterium paragallinarum	Scaffold
AP22_CN_19	JACEVW01	Avibacterium paragallinarum	Scaffold
AP23_CN_19	JACEVX01	Avibacterium paragallinarum	Scaffold
AP24_CN_19	JACEVY01	Avibacterium paragallinarum	Scaffold
AP25_CN_19	JACEVZ01	Avibacterium paragallinarum	Scaffold
AP26_CN_19	JACEWA01	Avibacterium paragallinarum	Scaffold
AP27_CN	NZ_CP113786.1	Avibacterium paragallinarum	Chromosome
AP28_CN_19	JACEWC01	Avibacterium paragallinarum	Scaffold
AP29_CN	NZ_CP113955.1	Avibacterium paragallinarum	Chromosome
AP30_DE_70	PQVK01	Avibacterium paragallinarum	Contig
AP31_JP	AOGF01	Avibacterium paragallinarum	Contig
AP32_MX_11	LAEN01	Avibacterium paragallinarum	Contig
AP33_MX_14	NZ_CP058307.1	Avibacterium paragallinarum	Complete Genome
AP34_GBENG	UFSW01	Avibacterium paragallinarum	Contig
AP35_OH_19	?	Avibacterium paragallinarum	?
AP36_OH_21	?	Avibacterium paragallinarum	?
AP37_OH_21	?	Avibacterium paragallinarum	?
AP38_PA_18	NZ_CP051642.1	Avibacterium paragallinarum	Complete Genome
AP39_PA_18	NZ_CP051641.1	Avibacterium paragallinarum	Complete Genome
AP40_PA_18	NZ_CP051640.1	Avibacterium paragallinarum	Complete Genome

AP41_PA_18	NZ_CP051639.1	Avibacterium paragallinarum	Complete Genome
AP42_PA_18	NZ_CP051638.1	Avibacterium paragallinarum	Complete Genome
AP43_PA_18	NZ_CP051637.1	Avibacterium paragallinarum	Complete Genome
AP44_PA_18	NZ_CP051636.1	Avibacterium paragallinarum	Complete Genome
AP45_PE_72_10	AFFP02	Avibacterium paragallinarum	Contig
AP46_58	?	Avibacterium paragallinarum	?
AP47_RU_17	QJPJ01	Avibacterium paragallinarum	Contig
AP48_RU_16	QJPI01	Avibacterium paragallinarum	Contig
AP49_58	UGHK01	Avibacterium paragallinarum	Contig
AP50_20	?	Avibacterium paragallinarum	?
AP51_OH_20	?	Avibacterium paragallinarum	?
AP52_MD_20	?	Avibacterium paragallinarum	?
AP53_MD_20	?	Avibacterium paragallinarum	?
AP54_MD_20	?	Avibacterium paragallinarum	?
AP55_MD_20	?	Avibacterium paragallinarum	?
AP56_MD_20	?	Avibacterium paragallinarum	?
AP57_MD_21	?	Avibacterium paragallinarum	?
AP58_MD_21	?	Avibacterium paragallinarum	?
AP59_IA_20	?	Avibacterium paragallinarum	?
AP60_IA_21	?	Avibacterium paragallinarum	?
AP61_DEL_17	?	Avibacterium paragallinarum	
AP62_DEL_17	?	Avibacterium paragallinarum	
AP63_IA_22	?	Avibacterium paragallinarum	
AP64_IA_22	?	Avibacterium paragallinarum	
AP65_IA_22	?	Avibacterium paragallinarum	
AP66_ZJC	NZ_CP095161.1	Avibacterium paragallinarum	Complete Genome
AP67_CCUGA2	?	Avibacterium paragallinarum	
AP68_FR_20	?	Avibacterium paragallinarum	
AP69_GA_20	?	Avibacterium paragallinarum	
AP70_FR_20	?	Avibacterium paragallinarum	
AP71_IA_22	CP104914.1, CP104915.1,CP104916.1	Avibacterium paragallinarum	Complete Genome
AP72_IA_22	CP104917.1	Avibacterium paragallinarum	Complete Genome
AP73_IA	?	Avibacterium paragallinarum	
AP74_CO_20	?	Avibacterium paragallinarum	
AP75	?	Avibacterium paragallinarum	

Sample ID	Genus	Species	Size (Mb)	Scaffolds	Genes	Release Date	Chromosomes/ RefSeq
AP01_PE_FARPER_15	Avibacterium	paragallinarum	2.43	1	2334	43171	NZ_CP034110.1
AP02_JP	Avibacterium	paragallinarum	2.66	82	2781	19-02-2019	?
AP03_CN_Hp8_17	Avibacterium	paragallinarum	2.36	182	2546	23-09-2019	?
AP04_DE_2671	Avibacterium	paragallinarum	2.34	135	2372	19-02-2019	?
AP05_MX_8	Avibacterium	paragallinarum	2.52	1	2469	29-03-2020	NZ_CP050316.1
AP06_CN_p4chr1	Avibacterium	paragallinarum	2.77	1	2818	31-08-2021	NZ_CP081939.1
AP07_CN_M	Avibacterium	paragallinarum	2.77	1		20-11-2021	?
AP08_SA	Avibacterium	paragallinarum	2.39	139	2534	19-02-2019	?
AP09_DEL_20	Avibacterium	paragallinarum	?	?	?	?	?
AP10_DEL_20	Avibacterium	paragallinarum	?	?	?	?	?
AP11_DEL_20	Avibacterium	paragallinarum	?	?	?	?	?
AP12_CN_19	Avibacterium	paragallinarum	2.59	120	2578	44146	?
AP13_CN_19	Avibacterium	paragallinarum	2.59	123	2572	44146	?
AP14_CN_19	Avibacterium	paragallinarum	2.59	124	2577	44146	?
AP15_CN_19	Avibacterium	paragallinarum	2.59	122	2569	44146	?
AP16_CN_19	Avibacterium	paragallinarum	2.59	119	2569	44146	?
AP17_CN_19	Avibacterium	paragallinarum	2.59	119	2571	44146	?
AP18_CN_19	Avibacterium	paragallinarum	2.59	126	2575	44146	?
AP19_CN_19	Avibacterium	paragallinarum	2.59	124	2575	44146	?
AP20_CN_19	Avibacterium	paragallinarum	2.59	128	2575	44146	?
AP21_CN_19	Avibacterium	paragallinarum	2.59	125	2571	44146	?
AP22_CN_19	Avibacterium	paragallinarum	2.59	127	2573	44146	?
AP23_CN_19	Avibacterium	paragallinarum	2.56	65	2507	44146	?
AP24_CN_19	Avibacterium	paragallinarum	2.59	124	2572	44146	?
AP25_CN_19	Avibacterium	paragallinarum	2.59	125	2586	44146	?
AP26_CN_19	Avibacterium	paragallinarum	2.59	128	2573	44146	?
AP27_CN	Avibacterium	paragallinarum	2.61	1	2602	44877	NZ_CP113786.1
AP28_CN_19	Avibacterium	paragallinarum	2.59	122	2572	44146	?
AP29_CN	Avibacterium	paragallinarum	2.59	1	2592	44907	NZ_CP113955.1
AP30_DE_70	Avibacterium	paragallinarum	2.46	335	2598	43283	?
AP31_JP	Avibacterium	paragallinarum	2.67	135	2857	22-03-2013	?
AP32_MX_11	Avibacterium	paragallinarum	2.41	154	2543	42159	?
AP33_MX_14	Avibacterium	paragallinarum	2.53	1	2497	30-06-2020	NZ_CP058307.1
AP34_GBENG	Avibacterium	paragallinarum	2.79	4	2984	43259	?
AP35_OH_19	Avibacterium	paragallinarum	?	?	?	?	?
AP36_OH_21	Avibacterium	paragallinarum	?	?	?	?	?
AP37_OH_21	Avibacterium	paragallinarum	?	?	?	?	?
AP38_PA_18	Avibacterium	paragallinarum	2.42	1	2333	27-04-2020	NZ_CP051642.1
AP39_PA_18	Avibacterium	paragallinarum	2.42	1	2335	27-04-2020	NZ_CP051641.1
AP40_PA_18	Avibacterium	paragallinarum	2.42	1	2335	27-04-2020	NZ_CP051640.1

AP41_PA_18	Avibacterium	paragallinarum	2.42	1	2340	27-04-2020	NZ_CP051639.1
AP42_PA_18	Avibacterium	paragallinarum	2.42	1	2345	27-04-2020	NZ_CP051638.1
AP43_PA_18	Avibacterium	paragallinarum	2.42	1	2340	27-04-2020	NZ_CP051637.1
AP44_PA_18	Avibacterium	paragallinarum	2.42	1	2336	27-04-2020	NZ_CP051636.1
AP45_PE_72_10	Avibacterium	paragallinarum	2.45	75	2422	28-07-2011	?
AP46_58	Avibacterium	paragallinarum	?	?	?	?	?
AP47_RU_17	Avibacterium	paragallinarum	2.56	67	2513	43379	?
AP48_RU_16	Avibacterium	paragallinarum	2.55	66	2521	43379	?
AP49_58	Avibacterium	paragallinarum	2.86	3	3110	30-07-2018	?
AP50_20	Avibacterium	paragallinarum	?	?	?	?	?
AP51_OH_20	Avibacterium	paragallinarum	?	?	?	?	?
AP52_MD_20	Avibacterium	paragallinarum	?	?	?	?	?
AP53_MD_20	Avibacterium	paragallinarum	?	?	?	?	?
AP54_MD_20	Avibacterium	paragallinarum	?	?	?	?	?
AP55_MD_20	Avibacterium	paragallinarum	?	?	?	?	?
AP56_MD_20	Avibacterium	paragallinarum	?	?	?	?	?
AP57_MD_21	Avibacterium	paragallinarum	?	?	?	?	?
AP58_MD_21	Avibacterium	paragallinarum	?	?	?	?	?
AP59_IA_20	Avibacterium	paragallinarum	?	?	?	?	?
AP60_IA_21	Avibacterium	paragallinarum	?	?	?	?	?
AP61_DEL_17	Avibacterium	paragallinarum	?	?	?	?	?
AP62_DEL_17	Avibacterium	paragallinarum	?	?	?	?	?
AP63_IA_22	Avibacterium	paragallinarum	?	?	?	?	?
AP64_IA_22	Avibacterium	paragallinarum	?	?	?	?	?
AP65_IA_22	Avibacterium	paragallinarum	?	?	?	?	?
AP66_ZJC	Avibacterium	paragallinarum	2.68	1	2665	13-04-2022	NZ_CP095161.1
AP67_CCUGA2	Avibacterium	paragallinarum	?	?	?	?	?
AP68_FR_20	Avibacterium	paragallinarum	?	?	?	?	?
AP69_GA_20	Avibacterium	paragallinarum	?	?	?	?	?
AP70_FR_20	Avibacterium	paragallinarum	?	?	?	?	?
AP71_IA_22	Avibacterium	paragallinarum	2.41	3	2310	28-09-2022	?
AP72_IA_22	Avibacterium	paragallinarum	2.53	1	2475	28-09-2022	?
AP73_IA	Avibacterium	paragallinarum	?	?	?	?	?
AP74_CO_20	Avibacterium	paragallinarum	?	?	?	?	?
AP75	Avibacterium	paragallinarum	?	?	?	?	?

Sample ID	Chromosomes/ INSDC	Plasmids/ RefSeq	Plasmids/INSDC	WGS	Strain/aliases
AP01_PE_FARPER_15	CP034110.1	?	?	?	FARPER-174
AP02_JP	?	?	?	RQXP01	AVPG 221
AP03_CN_Hp8_17	?	?	?	VXDA01	Hp8
AP04_DE_2671	?	?	?	RQXQ01	2671

AP05_MX_8	CP050316.1	?	?	?	ESV-135
AP06_CN_p4chr1	CP081939.1	?	?	?	p4chr1
AP07_CN_M	CP086713.1	?	?	?	M
AP08_SA	?	?	?	RQXS01	SA-3
AP09_DEL_20	?	?	?	?	MEL-AP1
AP10_DEL_20	?	?	?	?	MEL-AP2
AP11_DEL_20	?	?	?	?	MEL-AP3
AP12_CN_19	?	?	?	JACEVM01	AP1-1N-1
AP13_CN_19	?	?	?	JACEVN01	AP1-2N-1
AP14_CN_19	?	?	?	JACEVO01	AP1-3N-1
AP15_CN_19	?	?	?	JACEVP01	AP1-3S-1
AP16_CN_19	?	?	?	JACEVQ01	AP12-4N-1
AP17_CN_19	?	?	?	JACEVR01	AP12-5S-1
AP18_CN_19	?	?	?	JACEVS01	APX1-1N-2
AP19_CN_19	?	?	?	JACEVT01	APX1-1S-1
AP20_CN_19	?	?	?	JACEVU01	APX1-2N-1
AP21_CN_19	?	?	?	JACEVV01	APX2-2N-2
AP22_CN_19	?	?	?	JACEVW01	APX3-2S-2
AP23_CN_19	?	?	?	JACEVX01	Y2S-2
AP24_CN_19	?	?	?	JACEVY01	Z1N-1-2
AP25_CN_19	?	?	?	JACEVZ01	Z1N-2-1
AP26_CN_19	?	?	?	JACEWA01	Z1N-2-2
AP27_CN	CP113786.1	?	?	?	C-AP1
AP28_CN_19	?	?	?	JACEWC01	Z2S-1-2
AP29_CN	CP113955.1	?	?	?	C-AP4
AP30_DE_70	?	?	?	PQVK01	221
AP31_JP	?	?	?	AOGF01	CL
AP32_MX_11	?	?	?	LAEN01	AVPG2015
AP33_MX_14	CP058307.1	?	?		NCTC10926
AP34_GBENG	?	?	?	UFSW01	B1927814
AP35_OH_19	?	?	?	?	AG21-0333
AP36_OH_21	?	?	?	?	AG21-0332
AP37_OH_21	?	?	?	?	ADL-AP01
AP38_PA_18	CP051642.1	?	?	?	ADL-AP02
AP39_PA_18	CP051641.1	?	?	?	ADL-AP07
AP40_PA_18	CP051640.1	?	?	?	ADL-AP10
AP41_PA_18	CP051639.1	?	?	?	ADL-AP15
AP42_PA_18	CP051638.1	?	?	?	ADL-AP16
AP43_PA_18	CP051637.1	?	?	?	ADL-AP17
AP44_PA_18	CP051636.1	?	?	?	72
AP45_PE_72_10	?	?	?	AFFP02	

AP46_58	?	?	?		SCPM-O-B-8406
AP47_RU_17	?	?	?	QJPJ01	SCPM-O-B-8407
AP48_RU_16	?	?	?	QJPI01	NCTC11296
AP49_58	?	?	?	UGHK01	2020054022 AP-1 Isolate
AP50_20	?	?	?	?	2020036834-1
AP51_OH_20	?	?	?	?	20-s2774
AP52_MD_20	?	?	?	?	20-s4486
AP53_MD_20	?	?	?	?	20-s2469-5
AP54_MD_20	?	?	?	?	20-f4058-2
AP55_MD_20	?	?	?	?	20-f4953-4
AP56_MD_20	?	?	?	?	20-f4519-1
AP57_MD_21	?	?	?	?	21-f170-5
AP58_MD_21	?	?	?	?	2020036656-2
AP59_IA_20	?	?	?	?	2021005974 AP-2 Isolate
AP60_IA_21	?	?	?	?	17-9123-NC-L
AP61_DEL_17	?	?	?	?	2021109589
AP62_DEL_17	?	?	?	?	
AP63_IA_22	?	?	?	?	AP-2 5974
AP64_IA_22	?	?	?	?	ZJ-C
AP65_IA_22	?	?	?	?	CCUG_12835_ NCTC11296
AP66_ZJC	CP095161.1				LS20-3748-1-1_2020
AP67_CCUGA2	?	?	?	?	LS20-3748-3-1_2020
AP68_FR_20	?	?	?	?	LS20-3748-5-1_2020
AP69_GA_20	?	?	?	?	AG21-0333
AP70_FR_20	?	?	?	?	AP-2
AP71_IA_22	CP104914.1	?	CP104915.1, CP104916.1	?	2020036676-1
AP72_IA_22	CP104917.1	?	?	?	2020095597 Trach-2
AP73_IA	?	?	?	?	IC Pro-2 Flock birds S-12 1
AP74_CO_20			?		
AP75	?	?	?	?	

Appendix B. adhoc cgMLST Target Definer Results and list of core, accessory, and discarded genes.

Date: Apr 22, 2021 3:42 PM
cgMLST Target Definer version: 1.5 (win)
Ridom SeqSphere+ version: 7.2.6
User: Mostafa Ghanem, University of Maryland
Server: 10.27.63.4 (SeqSphere+ Server on VDPAM-006219)

Resulting Targets:
1170 targets were defined for cgMLST (1097373 bases)
935 targets were used as Accessory targets (885978 bases)
148 targets were discarded.

SETTINGS

Seed Genome:
 * NZ_CP034110.1 (06-JAN-2021), 2425949 bases, 2253 genes with CDS (Avibacterium paragallinarum strain FARPER-174 chromosome, complete genome)

Penetration Query Genomes (41):
 * Fasta file 221.fasta, 2668385 bases, 135 contigs
 * Fasta file 2671.fasta, 2337807 bases, 135 contigs
 * Fasta file 72.fasta, 2453490 bases, 75 contigs
 * Fasta file ADL-AP01.fasta, 2415542 bases
 * Fasta file ADL-AP02.fasta, 2416187 bases
 * Fasta file ADL-AP07.fasta, 2415993 bases
 * Fasta file ADL-AP10.fasta, 2415552 bases
 * Fasta file ADL-AP15.fasta, 2415950 bases
 * Fasta file ADL-AP16.fasta, 2415855 bases
 * Fasta file ADL-AP17.fasta, 2415699 bases
 * Fasta file AP1-1N-1.fasta, 2592909 bases, 120 contigs
 * Fasta file AP1-2N-1.fasta, 2593701 bases, 123 contigs
 * Fasta file AP1-3N-1.fasta, 2593713 bases, 124 contigs
 * Fasta file AP1-3S-1.fasta, 2593266 bases, 122 contigs
 * Fasta file AP12-4N-1.fasta, 2592871 bases, 119 contigs
 * Fasta file AP12-5S-1.fasta, 2593384 bases, 119 contigs
 * Fasta file APX1-1N-2.fasta, 2592203 bases, 126 contigs
 * Fasta file APX1-1S-1.fasta, 2593393 bases, 124 contigs
 * Fasta file APX1-2N-1.fasta, 2592270 bases, 128 contigs
 * Fasta file APX2-2N-2.fasta, 2594295 bases, 125 contigs
 * Fasta file APX3-2S-2.fasta, 2592768 bases, 127 contigs
 * Fasta file AVPG 221.fasta, 2655376 bases, 82 contigs
 * Fasta file AVPG2015.fasta, 2532404 bases
 * Fasta file CCUG 12835.fasta, 2460178 bases, 335 contigs
 * Fasta file CL.fasta, 2410835 bases, 154 contigs
 * Fasta file ESV-135.fasta, 2521134 bases

* Fasta file FARPER-174.fasta, 2425949 bases
* Fasta file GCF_015355095.1_ASM1535509v1_genomic.fna, 2593266 bases, 122 contigs
* Fasta file GCF_015355095.1_ASM1535509v1_genomic.fna.gz, 2593266 bases, 122 contigs
* Fasta file Hp8.fasta, 2357467 bases, 182 contigs
* Fasta file Modesto.fasta, 2499017 bases, 138 contigs
* Fasta file NCTC10926.fasta, 2790137 bases, 4 contigs
* Fasta file NCTC11296.fasta, 2860434 bases, 3 contigs
* Fasta file SA-3.fasta, 2390514 bases, 139 contigs
* Fasta file SCPM-O-B-8406.fasta, 2556353 bases, 67 contigs
* Fasta file SCPM-O-B-8407.fasta, 2554710 bases, 66 contigs
* Fasta file Y2S-2.fasta, 2563134 bases, 65 contigs
* Fasta file Z1N-1-2.fasta, 2592887 bases, 124 contigs
* Fasta file Z1N-2-1.fasta, 2593077 bases, 125 contigs
* Fasta file Z1N-2-2.fasta, 2591829 bases, 128 contigs
* Fasta file Z2S-1-2.fasta, 2591797 bases, 122 contigs

Seed Genome Filters:
* Minimum Length Filter (requires >=50 bases)
* Start Codon Filter (requires start codon at beginning of the gene)
* Stop Codon Filter (requires single stop codon at end of gene)
* Homologous Gene Filter (requires no multiple copies of gene with BLAST overlap>=100bp, identity>=90.0%)
* Gene Overlap Filter (requires no overlap with other genes >4 bases)

Query Genome BLAST Search:
* Requires BLAST hit with overlap=100%, identity>=90.0% in every query genome
* BLAST options: Word size=11, Mismatch penalty=-1, Match reward=1, Gap open costs=5, Gap extension costs=2

Penetration Query Genomes Filters:
* Stop Codon Percentage Filter (requires single stop codon at end of gene in >80% penetration query genomes)
BLAST version 2.2.12
Citation: Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-410.
DETAILS
Progress:

* cgMLST target search start at Apr 22, 2021 3:38 PM
* 2253 targets found in seed genome NZ_CP034110.1 (06-JAN-2021), 2425949 bases, 2253 genes (Avibacterium paragallinarum strain FARPER-174 chromosome, complete genome)
* 1994 targets after filtering seed genome NZ_CP034110.1 (06-JAN-2021), 2425949 bases, 2253 genes (Avibacterium paragallinarum strain FARPER-174 chromosome, complete genome)
* 1519 targets after blasting against Fasta file 221.fasta, 2668385 bases, 135 contigs
* 1395 targets after blasting against Fasta file 2671.fasta, 2337807 bases, 135 contigs
* 1382 targets after blasting against Fasta file 72.fasta, 2453490 bases, 75 contigs

* 1377 targets after blasting against Fasta file ADL-AP01.fasta, 2415542 bases
* 1377 targets after blasting against Fasta file ADL-AP02.fasta, 2416187 bases
* 1377 targets after blasting against Fasta file ADL-AP07.fasta, 2415993 bases
* 1377 targets after blasting against Fasta file ADL-AP10.fasta, 2415552 bases
* 1377 targets after blasting against Fasta file ADL-AP15.fasta, 2415950 bases
* 1377 targets after blasting against Fasta file ADL-AP16.fasta, 2415855 bases
* 1377 targets after blasting against Fasta file ADL-AP17.fasta, 2415699 bases
* 1300 targets after blasting against Fasta file AP1-1N-1.fasta, 2592909 bases, 120 contigs
* 1300 targets after blasting against Fasta file AP1-2N-1.fasta, 2593701 bases, 123 contigs
* 1300 targets after blasting against Fasta file AP1-3N-1.fasta, 2593713 bases, 124 contigs
* 1300 targets after blasting against Fasta file AP1-3S-1.fasta, 2593266 bases, 122 contigs
* 1300 targets after blasting against Fasta file AP12-4N-1.fasta, 2592871 bases, 119 contigs
* 1300 targets after blasting against Fasta file AP12-5S-1.fasta, 2593384 bases, 119 contigs
* 1300 targets after blasting against Fasta file APX1-1N-2.fasta, 2592203 bases, 126 contigs
* 1300 targets after blasting against Fasta file APX1-1S-1.fasta, 2593393 bases, 124 contigs
* 1300 targets after blasting against Fasta file APX1-2N-1.fasta, 2592270 bases, 128 contigs
* 1300 targets after blasting against Fasta file APX2-2N-2.fasta, 2594295 bases, 125 contigs
* 1300 targets after blasting against Fasta file APX3-2S-2.fasta, 2592768 bases, 127 contigs
* 1300 targets after blasting against Fasta file AVPG 221.fasta, 2655376 bases, 82 contigs
* 1299 targets after blasting against Fasta file AVPG2015.fasta, 2532404 bases
* 1294 targets after blasting against Fasta file CCUG 12835.fasta, 2460178 bases, 335 contigs
* 1244 targets after blasting against Fasta file CL.fasta, 2410835 bases, 154 contigs
* 1244 targets after blasting against Fasta file ESV-135.fasta, 2521134 bases
* 1244 targets after blasting against Fasta file FARPER-174.fasta, 2425949 bases
* 1244 targets after blasting against Fasta file GCF_015355095.1_ASM1535509v1_genomic.fna, 2593266 bases, 122 contigs
* 1244 targets after blasting against Fasta file GCF_015355095.1_ASM1535509v1_genomic.fna.gz, 2593266 bases, 122 contigs
* 1243 targets after blasting against Fasta file Hp8.fasta, 2357467 bases, 182 contigs
* 1220 targets after blasting against Fasta file Modesto.fasta, 2499017 bases, 138 contigs
* 1198 targets after blasting against Fasta file NCTC10926.fasta, 2790137 bases, 4 contigs
* 1198 targets after blasting against Fasta file NCTC11296.fasta, 2860434 bases, 3 contigs
* 1181 targets after blasting against Fasta file SA-3.fasta, 2390514 bases, 139 contigs
* 1176 targets after blasting against Fasta file SCPM-O-B-8406.fasta, 2556353 bases, 67 contigs
* 1176 targets after blasting against Fasta file SCPM-O-B-8407.fasta, 2554710 bases, 66 contigs
* 1176 targets after blasting against Fasta file Y2S-2.fasta, 2563134 bases, 65 contigs
* 1176 targets after blasting against Fasta file Z1N-1-2.fasta, 2592887 bases, 124 contigs
* 1176 targets after blasting against Fasta file Z1N-2-1.fasta, 2593077 bases, 125 contigs
* 1176 targets after blasting against Fasta file Z1N-2-2.fasta, 2591829 bases, 128 contigs
* 1176 targets after blasting against Fasta file Z2S-1-2.fasta, 2591797 bases, 122 contigs
* cgMLST target search ended at Apr 22, 2021 3:42 PM

cgMLST Genome Coverage:
* 45.2% of Seed genome NZ_CP034110.1 (06-JAN-2021), 2425949 bases, 2253 genes with CDS (Avibacterium paragallinarum strain FARPER-174 chromosome, complete genome) bases covered by cgMLST targets

* 41.1% of Query genome Fasta file 221.fasta bases covered by cgMLST targets
* 46.9% of Query genome Fasta file 2671.fasta bases covered by cgMLST targets
* 44.7% of Query genome Fasta file 72.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP01.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP02.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP07.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP10.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP15.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP16.fasta bases covered by cgMLST targets
* 45.4% of Query genome Fasta file ADL-AP17.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file AP1-1N-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file AP1-2N-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file AP1-3N-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file AP1-3S-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file AP12-4N-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file AP12-5S-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file APX1-1N-2.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file APX1-1S-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file APX1-2N-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file APX2-2N-2.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file APX3-2S-2.fasta bases covered by cgMLST targets
* 41.3% of Query genome Fasta file AVPG 221.fasta bases covered by cgMLST targets
* 43.3% of Query genome Fasta file AVPG2015.fasta bases covered by cgMLST targets
* 44.6% of Query genome Fasta file CCUG 12835.fasta bases covered by cgMLST targets
* 45.5% of Query genome Fasta file CL.fasta bases covered by cgMLST targets
* 43.5% of Query genome Fasta file ESV-135.fasta bases covered by cgMLST targets
* 45.2% of Query genome Fasta file FARPER-174.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file GCF_015355095.1_ASM1535509v1_genomic.fna bases covered by cgMLST targets
* 42.3% of Query genome Fasta file GCF_015355095.1_ASM1535509v1_genomic.fna.gz bases covered by cgMLST targets
* 46.5% of Query genome Fasta file Hp8.fasta bases covered by cgMLST targets
* 43.9% of Query genome Fasta file Modesto.fasta bases covered by cgMLST targets
* 39.3% of Query genome Fasta file NCTC10926.fasta bases covered by cgMLST targets
* 38.4% of Query genome Fasta file NCTC11296.fasta bases covered by cgMLST targets
* 45.9% of Query genome Fasta file SA-3.fasta bases covered by cgMLST targets
* 42.9% of Query genome Fasta file SCPM-O-B-8406.fasta bases covered by cgMLST targets
* 43.0% of Query genome Fasta file SCPM-O-B-8407.fasta bases covered by cgMLST targets
* 42.8% of Query genome Fasta file Y2S-2.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file Z1N-1-2.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file Z1N-2-1.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file Z1N-2-2.fasta bases covered by cgMLST targets
* 42.3% of Query genome Fasta file Z2S-1-2.fasta bases covered by cgMLST targets

TARGET DETAILS

Seed Genome Start Codon Filter: filtered out 10 targets (discarded)

EIA51_RS01240, EIA51_RS01645, EIA51_RS02115, EIA51_RS05705, EIA51_RS08010, EIA51_RS08660, EIA51_RS08790, EIA51_RS09900, EIA51_RS10130, EIA51_RS10890

Seed Genome Stop Codon Filter: filtered out 79 targets (discarded)
EIA51_RS00145, EIA51_RS00350, EIA51_RS00420, EIA51_RS00800, EIA51_RS01040, EIA51_RS01385, EIA51_RS01415, EIA51_RS01600, EIA51_RS01880, EIA51_RS01950, EIA51_RS02085, EIA51_RS02375, EIA51_RS02730, EIA51_RS02845, EIA51_RS02850, EIA51_RS03015, EIA51_RS03025, EIA51_RS03040, EIA51_RS03085, EIA51_RS03355, EIA51_RS03420, EIA51_RS03730, EIA51_RS03870, EIA51_RS03910, EIA51_RS04120, EIA51_RS04285, EIA51_RS04345, EIA51_RS04430, EIA51_RS04815, EIA51_RS04965, EIA51_RS04990, EIA51_RS05090, EIA51_RS05135, EIA51_RS05280, EIA51_RS05420, EIA51_RS05505, EIA51_RS05540, EIA51_RS05570, EIA51_RS05740, EIA51_RS05860, EIA51_RS05875, EIA51_RS05990, EIA51_RS06070, EIA51_RS06100, EIA51_RS06105, EIA51_RS06140, EIA51_RS06215, EIA51_RS06240, EIA51_RS06255, EIA51_RS06305, EIA51_RS06380, EIA51_RS06605, EIA51_RS06715, EIA51_RS06825, EIA51_RS06860, EIA51_RS06895, EIA51_RS07120, EIA51_RS07315, EIA51_RS07485, EIA51_RS08025, EIA51_RS08090, EIA51_RS09310, EIA51_RS09510, EIA51_RS09570, EIA51_RS09680, EIA51_RS09710, EIA51_RS09765, EIA51_RS09770, EIA51_RS10135, EIA51_RS10390, EIA51_RS10505, EIA51_RS10525, EIA51_RS10595, EIA51_RS10685, EIA51_RS11225, EIA51_RS11400, EIA51_RS11760, EIA51_RS11815, EIA51_RS11820

Seed Genome Homologous Gene Filter: filtered out 59 targets (discarded)
EIA51_RS00020, EIA51_RS00060, EIA51_RS00065, EIA51_RS00770, EIA51_RS00900, EIA51_RS01190, EIA51_RS01500, EIA51_RS01605, EIA51_RS02055, EIA51_RS02230, EIA51_RS02235, EIA51_RS02270, EIA51_RS02320, EIA51_RS02325, EIA51_RS02330, EIA51_RS02335, EIA51_RS02340, EIA51_RS02400, EIA51_RS02420, EIA51_RS02430, EIA51_RS02450, EIA51_RS02535, EIA51_RS02675, EIA51_RS02700, EIA51_RS03150, EIA51_RS04730, EIA51_RS07580, EIA51_RS07680, EIA51_RS07685, EIA51_RS07715, EIA51_RS07740, EIA51_RS07805, EIA51_RS07840, EIA51_RS07845, EIA51_RS07850, EIA51_RS07855, EIA51_RS07860, EIA51_RS07875, EIA51_RS07905, EIA51_RS07910, EIA51_RS07960, EIA51_RS08855, EIA51_RS08945, EIA51_RS08970, EIA51_RS08990, EIA51_RS09005, EIA51_RS10075, EIA51_RS10080, EIA51_RS10500, EIA51_RS10870, EIA51_RS11255, EIA51_RS11260, EIA51_RS11580, EIA51_RS11630, EIA51_RS11770, EIA51_RS11775, EIA51_RS11795, EIA51_RS11805, EIA51_RS11810

Seed Genome Gene Overlap Filter: filtered out 111 targets (moved to Accessory)
EIA51_RS00075, EIA51_RS00250, EIA51_RS00515, EIA51_RS00540, EIA51_RS00755, EIA51_RS00790, EIA51_RS00825, EIA51_RS01000, EIA51_RS01065, EIA51_RS01070, EIA51_RS01145, EIA51_RS01255, EIA51_RS01360, EIA51_RS01615, EIA51_RS01625, EIA51_RS01635, EIA51_RS01655, EIA51_RS01825, EIA51_RS01890, EIA51_RS02105, EIA51_RS02125, EIA51_RS02135, EIA51_RS02210, EIA51_RS02220, EIA51_RS02390, EIA51_RS02550, EIA51_RS02585, EIA51_RS02650, EIA51_RS02680, EIA51_RS02980, EIA51_RS03140, EIA51_RS03280, EIA51_RS03435, EIA51_RS03450, EIA51_RS03635, EIA51_RS03845, EIA51_RS03890, EIA51_RS03995, EIA51_RS04140, EIA51_RS04300, EIA51_RS04375, EIA51_RS04445, EIA51_RS04665, EIA51_RS04750, EIA51_RS04775,

EIA51_RS05040, EIA51_RS05250, EIA51_RS05385, EIA51_RS05400, EIA51_RS05455, EIA51_RS05490, EIA51_RS05555, EIA51_RS05810, EIA51_RS05945, EIA51_RS06120, EIA51_RS06160, EIA51_RS06185, EIA51_RS06325, EIA51_RS06420, EIA51_RS06480, EIA51_RS06655, EIA51_RS06785, EIA51_RS06950, EIA51_RS06970, EIA51_RS07070, EIA51_RS07130, EIA51_RS07215, EIA51_RS07275, EIA51_RS07325, EIA51_RS07745, EIA51_RS07770, EIA51_RS07795, EIA51_RS07945, EIA51_RS08100, EIA51_RS08115, EIA51_RS08130, EIA51_RS08190, EIA51_RS08225, EIA51_RS08280, EIA51_RS08375, EIA51_RS08810, EIA51_RS08820, EIA51_RS08860, EIA51_RS08950, EIA51_RS09010, EIA51_RS09150, EIA51_RS09265, EIA51_RS09275, EIA51_RS09395, EIA51_RS09580, EIA51_RS09640, EIA51_RS09845, EIA51_RS10060, EIA51_RS10605, EIA51_RS10720, EIA51_RS10945, EIA51_RS11040, EIA51_RS11055, EIA51_RS11215, EIA51_RS11320, EIA51_RS11430, EIA51_RS11450, EIA51_RS11505, EIA51_RS11515, EIA51_RS11570, EIA51_RS11600, EIA51_RS11615, EIA51_RS11720, EIA51_RS11745, EIA51_RS11755, EIA51_RS11765

Query Genome BLAST Search: filtered out 818 targets (moved to Accessory)
EIA51_RS00010, EIA51_RS00015, EIA51_RS00025, EIA51_RS00030, EIA51_RS00035, EIA51_RS00040, EIA51_RS00045, EIA51_RS00050, EIA51_RS00055, EIA51_RS00070, EIA51_RS00105, EIA51_RS00110, EIA51_RS00185, EIA51_RS00190, EIA51_RS00195, EIA51_RS00200, EIA51_RS00205, EIA51_RS00215, EIA51_RS00260, EIA51_RS00320, EIA51_RS00325, EIA51_RS00335, EIA51_RS00340, EIA51_RS00345, EIA51_RS00355, EIA51_RS00360, EIA51_RS00365, EIA51_RS00370, EIA51_RS00375, EIA51_RS00380, EIA51_RS00385, EIA51_RS00390, EIA51_RS00395, EIA51_RS00400, EIA51_RS00405, EIA51_RS00410, EIA51_RS00415, EIA51_RS00425, EIA51_RS00430, EIA51_RS00465, EIA51_RS00495, EIA51_RS00505, EIA51_RS00520, EIA51_RS00535, EIA51_RS00565, EIA51_RS00670, EIA51_RS00675, EIA51_RS00690, EIA51_RS00705, EIA51_RS00710, EIA51_RS00715, EIA51_RS00720, EIA51_RS00725, EIA51_RS00730, EIA51_RS00735, EIA51_RS00740, EIA51_RS00745, EIA51_RS00750, EIA51_RS00760, EIA51_RS00765, EIA51_RS00780, EIA51_RS00785, EIA51_RS00795, EIA51_RS00840, EIA51_RS00850, EIA51_RS00860, EIA51_RS00905, EIA51_RS00930, EIA51_RS00935, EIA51_RS00940, EIA51_RS00945, EIA51_RS00990, EIA51_RS00995, EIA51_RS01005, EIA51_RS01025, EIA51_RS01055, EIA51_RS01060, EIA51_RS01075, EIA51_RS01080, EIA51_RS01090, EIA51_RS01125, EIA51_RS01130, EIA51_RS01135, EIA51_RS01205, EIA51_RS01225, EIA51_RS01230, EIA51_RS01285, EIA51_RS01290, EIA51_RS01295, EIA51_RS01345, EIA51_RS01355, EIA51_RS01370, EIA51_RS01390, EIA51_RS01405, EIA51_RS01435, EIA51_RS01505, EIA51_RS01510, EIA51_RS01515, EIA51_RS01520, EIA51_RS01525, EIA51_RS01530, EIA51_RS01535, EIA51_RS01540, EIA51_RS01545, EIA51_RS01550, EIA51_RS01555, EIA51_RS01560, EIA51_RS01565, EIA51_RS01570, EIA51_RS01575, EIA51_RS01580, EIA51_RS01585, EIA51_RS01590, EIA51_RS01595, EIA51_RS01610, EIA51_RS01620, EIA51_RS01630, EIA51_RS01640, EIA51_RS01685, EIA51_RS01700, EIA51_RS01750, EIA51_RS01765, EIA51_RS01770, EIA51_RS01790, EIA51_RS01820, EIA51_RS01850, EIA51_RS01870, EIA51_RS01905, EIA51_RS01935, EIA51_RS01945, EIA51_RS01995, EIA51_RS02000, EIA51_RS02015, EIA51_RS02060, EIA51_RS02065, EIA51_RS02070, EIA51_RS02075, EIA51_RS02080, EIA51_RS02090, EIA51_RS02100, EIA51_RS02110, EIA51_RS02120, EIA51_RS02130, EIA51_RS02140, EIA51_RS02145, EIA51_RS02150, EIA51_RS02155, EIA51_RS02160, EIA51_RS02165, EIA51_RS02170,

EIA51_RS02175, EIA51_RS02180, EIA51_RS02185, EIA51_RS02190, EIA51_RS02195, EIA51_RS02200, EIA51_RS02205, EIA51_RS02215, EIA51_RS02225, EIA51_RS02240, EIA51_RS02245, EIA51_RS02250, EIA51_RS02255, EIA51_RS02260, EIA51_RS02265, EIA51_RS02275, EIA51_RS02280, EIA51_RS02285, EIA51_RS02290, EIA51_RS02295, EIA51_RS02305, EIA51_RS02310, EIA51_RS02315, EIA51_RS02345, EIA51_RS02350, EIA51_RS02355, EIA51_RS02360, EIA51_RS02365, EIA51_RS02370, EIA51_RS02380, EIA51_RS02395, EIA51_RS02405, EIA51_RS02410, EIA51_RS02415, EIA51_RS02425, EIA51_RS02435, EIA51_RS02440, EIA51_RS02445, EIA51_RS02460, EIA51_RS02505, EIA51_RS02510, EIA51_RS02515, EIA51_RS02540, EIA51_RS02545, EIA51_RS02555, EIA51_RS02560, EIA51_RS02565, EIA51_RS02575, EIA51_RS02620, EIA51_RS02630, EIA51_RS02635, EIA51_RS02640, EIA51_RS02645, EIA51_RS02655, EIA51_RS02660, EIA51_RS02665, EIA51_RS02670, EIA51_RS02685, EIA51_RS02690, EIA51_RS02695, EIA51_RS02720, EIA51_RS02725, EIA51_RS02860, EIA51_RS02870, EIA51_RS02875, EIA51_RS02950, EIA51_RS03020, EIA51_RS03030, EIA51_RS03035, EIA51_RS03045, EIA51_RS03050, EIA51_RS03055, EIA51_RS03060, EIA51_RS03065, EIA51_RS03070, EIA51_RS03075, EIA51_RS03080, EIA51_RS03155, EIA51_RS03170, EIA51_RS03230, EIA51_RS03250, EIA51_RS03260, EIA51_RS03310, EIA51_RS03315, EIA51_RS03320, EIA51_RS03325, EIA51_RS03335, EIA51_RS03340, EIA51_RS03345, EIA51_RS03350, EIA51_RS03360, EIA51_RS03365, EIA51_RS03370, EIA51_RS03390, EIA51_RS03500, EIA51_RS03505, EIA51_RS03510, EIA51_RS03545, EIA51_RS03555, EIA51_RS03560, EIA51_RS03565, EIA51_RS03570, EIA51_RS03575, EIA51_RS03585, EIA51_RS03645, EIA51_RS03660, EIA51_RS03670, EIA51_RS03680, EIA51_RS03705, EIA51_RS03710, EIA51_RS03715, EIA51_RS03735, EIA51_RS03795, EIA51_RS03800, EIA51_RS03805, EIA51_RS03825, EIA51_RS03830, EIA51_RS03840, EIA51_RS03850, EIA51_RS03855, EIA51_RS03860, EIA51_RS03865, EIA51_RS03960, EIA51_RS03975, EIA51_RS03980, EIA51_RS04020, EIA51_RS04025, EIA51_RS04030, EIA51_RS04060, EIA51_RS04160, EIA51_RS04190, EIA51_RS04240, EIA51_RS04275, EIA51_RS04280, EIA51_RS04290, EIA51_RS04295, EIA51_RS04305, EIA51_RS04310, EIA51_RS04315, EIA51_RS04320, EIA51_RS04325, EIA51_RS04330, EIA51_RS04335, EIA51_RS04340, EIA51_RS04385, EIA51_RS04420, EIA51_RS04435, EIA51_RS04440, EIA51_RS04450, EIA51_RS04465, EIA51_RS04470, EIA51_RS04475, EIA51_RS04480, EIA51_RS04485, EIA51_RS04490, EIA51_RS04495, EIA51_RS04500, EIA51_RS04505, EIA51_RS04510, EIA51_RS04515, EIA51_RS04520, EIA51_RS04525, EIA51_RS04530, EIA51_RS04575, EIA51_RS04580, EIA51_RS04585, EIA51_RS04590, EIA51_RS04605, EIA51_RS04610, EIA51_RS04615, EIA51_RS04620, EIA51_RS04625, EIA51_RS04645, EIA51_RS04650, EIA51_RS04655, EIA51_RS04660, EIA51_RS04670, EIA51_RS04675, EIA51_RS04685, EIA51_RS04710, EIA51_RS04715, EIA51_RS04725, EIA51_RS04735, EIA51_RS04755, EIA51_RS04780, EIA51_RS04790, EIA51_RS04795, EIA51_RS04800, EIA51_RS04820, EIA51_RS04835, EIA51_RS04840, EIA51_RS04845, EIA51_RS04880, EIA51_RS04885, EIA51_RS04890, EIA51_RS04895, EIA51_RS04910, EIA51_RS04915, EIA51_RS05000, EIA51_RS05015, EIA51_RS05020, EIA51_RS05025, EIA51_RS05050, EIA51_RS05055, EIA51_RS05080, EIA51_RS05105, EIA51_RS05120, EIA51_RS05200, EIA51_RS05205, EIA51_RS05260, EIA51_RS05265, EIA51_RS05270, EIA51_RS05315, EIA51_RS05335, EIA51_RS05390, EIA51_RS05395, EIA51_RS05405, EIA51_RS05410, EIA51_RS05415, EIA51_RS05425, EIA51_RS05440, EIA51_RS05475, EIA51_RS05480, EIA51_RS05495, EIA51_RS05500, EIA51_RS05510, EIA51_RS05515, EIA51_RS05520, EIA51_RS05525, EIA51_RS05530,

EIA51_RS05535, EIA51_RS05560, EIA51_RS05565, EIA51_RS05600, EIA51_RS05605, EIA51_RS05645, EIA51_RS05650, EIA51_RS05670, EIA51_RS05675, EIA51_RS05710, EIA51_RS05725, EIA51_RS05730, EIA51_RS05735, EIA51_RS05745, EIA51_RS05780, EIA51_RS05785, EIA51_RS05790, EIA51_RS05795, EIA51_RS05800, EIA51_RS05805, EIA51_RS05815, EIA51_RS05835, EIA51_RS05905, EIA51_RS05930, EIA51_RS05970, EIA51_RS05980, EIA51_RS05985, EIA51_RS06045, EIA51_RS06055, EIA51_RS06060, EIA51_RS06065, EIA51_RS06085, EIA51_RS06090, EIA51_RS06095, EIA51_RS06110, EIA51_RS06115, EIA51_RS06130, EIA51_RS06135, EIA51_RS06145, EIA51_RS06150, EIA51_RS06155, EIA51_RS06170, EIA51_RS06175, EIA51_RS06220, EIA51_RS06245, EIA51_RS06260, EIA51_RS06265, EIA51_RS06275, EIA51_RS06285, EIA51_RS06290, EIA51_RS06310, EIA51_RS06320, EIA51_RS06330, EIA51_RS06360, EIA51_RS06390, EIA51_RS06395, EIA51_RS06400, EIA51_RS06430, EIA51_RS06485, EIA51_RS06510, EIA51_RS06525, EIA51_RS06600, EIA51_RS06615, EIA51_RS06620, EIA51_RS06630, EIA51_RS06660, EIA51_RS06675, EIA51_RS06680, EIA51_RS06775, EIA51_RS06790, EIA51_RS06795, EIA51_RS06830, EIA51_RS06835, EIA51_RS06845, EIA51_RS06865, EIA51_RS06880, EIA51_RS06900, EIA51_RS06935, EIA51_RS06955, EIA51_RS06960, EIA51_RS06990, EIA51_RS07005, EIA51_RS07010, EIA51_RS07015, EIA51_RS07020, EIA51_RS07025, EIA51_RS07030, EIA51_RS07035, EIA51_RS07040, EIA51_RS07045, EIA51_RS07050, EIA51_RS07055, EIA51_RS07065, EIA51_RS07075, EIA51_RS07150, EIA51_RS07160, EIA51_RS07175, EIA51_RS07185, EIA51_RS07220, EIA51_RS07230, EIA51_RS07235, EIA51_RS07240, EIA51_RS07245, EIA51_RS07250, EIA51_RS07255, EIA51_RS07260, EIA51_RS07265, EIA51_RS07270, EIA51_RS07280, EIA51_RS07285, EIA51_RS07290, EIA51_RS07385, EIA51_RS07395, EIA51_RS07400, EIA51_RS07410, EIA51_RS07415, EIA51_RS07425, EIA51_RS07430, EIA51_RS07435, EIA51_RS07440, EIA51_RS07450, EIA51_RS07470, EIA51_RS07475, EIA51_RS07480, EIA51_RS07490, EIA51_RS07495, EIA51_RS07500, EIA51_RS07515, EIA51_RS07520, EIA51_RS07525, EIA51_RS07535, EIA51_RS07540, EIA51_RS07545, EIA51_RS07550, EIA51_RS07555, EIA51_RS07560, EIA51_RS07565, EIA51_RS07570, EIA51_RS07575, EIA51_RS07585, EIA51_RS07590, EIA51_RS07595, EIA51_RS07600, EIA51_RS07605, EIA51_RS07610, EIA51_RS07615, EIA51_RS07620, EIA51_RS07625, EIA51_RS07630, EIA51_RS07635, EIA51_RS07640, EIA51_RS07645, EIA51_RS07650, EIA51_RS07655, EIA51_RS07660, EIA51_RS07665, EIA51_RS07670, EIA51_RS07675, EIA51_RS07690, EIA51_RS07695, EIA51_RS07700, EIA51_RS07720, EIA51_RS07725, EIA51_RS07730, EIA51_RS07735, EIA51_RS07750, EIA51_RS07755, EIA51_RS07760, EIA51_RS07765, EIA51_RS07775, EIA51_RS07780, EIA51_RS07785, EIA51_RS07790, EIA51_RS07800, EIA51_RS07810, EIA51_RS07815, EIA51_RS07820, EIA51_RS07825, EIA51_RS07830, EIA51_RS07835, EIA51_RS07865, EIA51_RS07870, EIA51_RS07880, EIA51_RS07890, EIA51_RS07895, EIA51_RS07900, EIA51_RS07915, EIA51_RS07920, EIA51_RS07925, EIA51_RS07930, EIA51_RS07935, EIA51_RS07940, EIA51_RS07950, EIA51_RS07955, EIA51_RS07965, EIA51_RS07970, EIA51_RS07975, EIA51_RS08015, EIA51_RS08030, EIA51_RS08035, EIA51_RS08040, EIA51_RS08045, EIA51_RS08050, EIA51_RS08055, EIA51_RS08080, EIA51_RS08120, EIA51_RS08125, EIA51_RS08135, EIA51_RS08145, EIA51_RS08150, EIA51_RS08160, EIA51_RS08175, EIA51_RS08180, EIA51_RS08220, EIA51_RS08245, EIA51_RS08265, EIA51_RS08270, EIA51_RS08275, EIA51_RS08310, EIA51_RS08315, EIA51_RS08325, EIA51_RS08335, EIA51_RS08395, EIA51_RS08400, EIA51_RS08405, EIA51_RS08420, EIA51_RS08425, EIA51_RS08480, EIA51_RS08510, EIA51_RS08580,

EIA51_RS08635, EIA51_RS08640, EIA51_RS08745, EIA51_RS08785, EIA51_RS08795, EIA51_RS08800, EIA51_RS08805, EIA51_RS08815, EIA51_RS08825, EIA51_RS08830, EIA51_RS08835, EIA51_RS08840, EIA51_RS08845, EIA51_RS08850, EIA51_RS08890, EIA51_RS08895, EIA51_RS08900, EIA51_RS08905, EIA51_RS08910, EIA51_RS08915, EIA51_RS08920, EIA51_RS08935, EIA51_RS08940, EIA51_RS08955, EIA51_RS08960, EIA51_RS08965, EIA51_RS08975, EIA51_RS08985, EIA51_RS08995, EIA51_RS09000, EIA51_RS09015, EIA51_RS09030, EIA51_RS09110, EIA51_RS09170, EIA51_RS09175, EIA51_RS09205, EIA51_RS09215, EIA51_RS09220, EIA51_RS09235, EIA51_RS09240, EIA51_RS09250, EIA51_RS09255, EIA51_RS09260, EIA51_RS09280, EIA51_RS09305, EIA51_RS09315, EIA51_RS09320, EIA51_RS09370, EIA51_RS09375, EIA51_RS09380, EIA51_RS09385, EIA51_RS09390, EIA51_RS09400, EIA51_RS09405, EIA51_RS09410, EIA51_RS09440, EIA51_RS09470, EIA51_RS09500, EIA51_RS09515, EIA51_RS09540, EIA51_RS09545, EIA51_RS09555, EIA51_RS09635, EIA51_RS09645, EIA51_RS09650, EIA51_RS09655, EIA51_RS09660, EIA51_RS09665, EIA51_RS09685, EIA51_RS09690, EIA51_RS09700, EIA51_RS09705, EIA51_RS09715, EIA51_RS09725, EIA51_RS09745, EIA51_RS09775, EIA51_RS09785, EIA51_RS09795, EIA51_RS09815, EIA51_RS09910, EIA51_RS09915, EIA51_RS09920, EIA51_RS09925, EIA51_RS09940, EIA51_RS09945, EIA51_RS09950, EIA51_RS09955, EIA51_RS10015, EIA51_RS10095, EIA51_RS10100, EIA51_RS10105, EIA51_RS10110, EIA51_RS10115, EIA51_RS10185, EIA51_RS10205, EIA51_RS10225, EIA51_RS10230, EIA51_RS10235, EIA51_RS10250, EIA51_RS10265, EIA51_RS10280, EIA51_RS10285, EIA51_RS10305, EIA51_RS10365, EIA51_RS10370, EIA51_RS10375, EIA51_RS10395, EIA51_RS10415, EIA51_RS10425, EIA51_RS10470, EIA51_RS10475, EIA51_RS10480, EIA51_RS10485, EIA51_RS10490, EIA51_RS10495, EIA51_RS10510, EIA51_RS10535, EIA51_RS10550, EIA51_RS10555, EIA51_RS10575, EIA51_RS10600, EIA51_RS10620, EIA51_RS10635, EIA51_RS10650, EIA51_RS10655, EIA51_RS10660, EIA51_RS10715, EIA51_RS10745, EIA51_RS10810, EIA51_RS10845, EIA51_RS10850, EIA51_RS10895, EIA51_RS10900, EIA51_RS10905, EIA51_RS10955, EIA51_RS10960, EIA51_RS10965, EIA51_RS10980, EIA51_RS10990, EIA51_RS11105, EIA51_RS11140, EIA51_RS11145, EIA51_RS11150, EIA51_RS11200, EIA51_RS11300, EIA51_RS11305, EIA51_RS11310, EIA51_RS11315, EIA51_RS11325, EIA51_RS11330, EIA51_RS11335, EIA51_RS11340, EIA51_RS11345, EIA51_RS11350, EIA51_RS11355, EIA51_RS11360, EIA51_RS11365, EIA51_RS11370, EIA51_RS11375, EIA51_RS11380, EIA51_RS11385, EIA51_RS11390, EIA51_RS11395, EIA51_RS11405, EIA51_RS11410, EIA51_RS11415, EIA51_RS11420, EIA51_RS11425, EIA51_RS11435, EIA51_RS11440, EIA51_RS11445, EIA51_RS11455, EIA51_RS11460, EIA51_RS11465, EIA51_RS11470, EIA51_RS11475, EIA51_RS11480, EIA51_RS11485, EIA51_RS11490, EIA51_RS11495, EIA51_RS11500, EIA51_RS11510, EIA51_RS11520, EIA51_RS11525, EIA51_RS11530, EIA51_RS11535, EIA51_RS11540, EIA51_RS11550, EIA51_RS11555, EIA51_RS11560, EIA51_RS11565, EIA51_RS11585, EIA51_RS11590, EIA51_RS11595, EIA51_RS11605, EIA51_RS11610, EIA51_RS11620, EIA51_RS11625, EIA51_RS11660, EIA51_RS11665, EIA51_RS11680, EIA51_RS11685, EIA51_RS11700, EIA51_RS11705, EIA51_RS11725, EIA51_RS11740, EIA51_RS11780, EIA51_RS11785

Penetration Query Genomes Stop Codon Percentage Filter: filtered out 6 targets (moved to Accessory)
EIA51_RS01235, EIA51_RS03675, EIA51_RS07370, EIA51_RS08195, EIA51_RS11100,

EIA51_RS11750
Messages for NZ_CP034110.1 (06-JAN-2021), 2425949 bases, 2253 genes with CDS
(Avibacterium paragallinarum strain FARPER-174 chromosome, complete genome):
 * Skipped EIA51_RS00005 because location is not contiguous.

Bibliography

Beach, J. R., & Schalm, O. W. (1936). Studies of the Clinical Manifestations and Transmissibility of Infectious Coryza of Chickens. *Poultry Science*, *15*(6), 466–472. https://doi.org/10.3382/ps.0150466

Bisgaard, M., Nørskov-Lauritsen, N., de Wit, S. J., Hess, C., & Christensen, H. (2012). Multilocus sequence phylogenetic analysis of Avibacterium. *Microbiology (Reading, England)*, *158*(Pt 4), 993–1004. https://doi.org/10.1099/mic.0.054429-0

Blackall, P. J. (1999). Infectious Coryza: Overview of the Disease and New Diagnostic Options. *Clinical Microbiology Reviews*, *12*(4), 627–632. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC88928/

Blackall, P. J., Christensen, H., Beckenham, T., Blackall, L. L., & Bisgaard, M. (2005). Reclassification of Pasteurella gallinarum, [Haemophilus] paragallinarum, Pasteurella avium and Pasteurella volantium as Avibacterium gallinarum gen. Nov., comb. Nov., Avibacterium paragallinarum comb. Nov., Avibacterium avium comb. Nov. And Avibacterium volantium comb. Nov. *International Journal of Systematic and Evolutionary Microbiology*, *55*(Pt 1), 353–362. https://doi.org/10.1099/ijs.0.63357-0

Blackall, P. J., Eaves, L. E., & Rogers, D. G. (1990). Proposal of a new serovar and altered nomenclature for Haemophilus paragallinarum in the Kume hemagglutinin scheme. *Journal of Clinical Microbiology*, *28*(6), 1185–1187. https://doi.org/10.1128/jcm.28.6.1185-1187.1990

Blackall, P. J., Morrow, C. J., McInnes, A., Eaves, L. E., & Rogers, D. G. (1990). Epidemiologic studies on infectious coryza outbreaks in northern New South Wales, Australia, using

serotyping, biotyping, and chromosomal DNA restriction endonuclease analysis. *Avian Diseases*, *34*(2), 267–276.

Blackall, P. J., & Soriano-Vargas, E. (2020). Infectious Coryza and Related Bacterial Infections. In D. E. Swayne, M. Boulianne, C. M. Logue, L. R. McDougald, V. Nair, D. L. Suarez, S. Wit, T. Grimes, D. Johnson, M. Kromm, T. Y. Prajitno, I. Rubinoff, & G. Zavala (Eds.), *Diseases of Poultry* (1st ed., pp. 890–906). Wiley. https://doi.org/10.1002/9781119371199.ch20

Blackall, P. J., & Soriano-Vargas, E. (2020). Infectious Coryza and Related Bacterial Infections. In *Diseases of Poultry* (pp. 890–906). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119371199.ch20

Buter, R., Feberwee, A., de Wit, S., Heuvelink, A., da Silva, A., Gallardo, R., Soriano Vargas, E., Swanepoel, S., Jung, A., Tödte, M., & Dijkman, R. (2023). Molecular characterization of the HMTp210 gene of Avibacterium paragallinarum and the proposition of a new genotyping method as alternative for classical serotyping. *Avian Pathology*, *52*(5), 362–376. https://doi.org/10.1080/03079457.2023.2239178

Byukusenge, M., Nissly, R. H., Li, L., Pierre, T., Mathews, T., Wallner-Pendleton, E., Dunn, P., Barnhart, D., Loughrey, S., Davison, S., Kelly, D. J., Tewari, D., Jayarao, B. M., & Kuchipudi, S. V. (2020). Complete Genome Sequences of Seven Avibacterium paragallinarum Isolates from Poultry Farms in Pennsylvania, USA. *Microbiology Resource Announcements*, *9*(27), e00654-20. https://doi.org/10.1128/MRA.00654-20

Chen, X., Miflin, J. K., Zhang, P., & Blackall, P. J. (1996). Development and Application of DNA Probes and PCR Tests for Haemophilus paragallinarum. *Avian Diseases*, *40*(2), 398. https://doi.org/10.2307/1592238

Clothier, K. A., Stoute, S., Torain, A., & Crossley, B. (2019). Validation of a real-time PCR assay for high-throughput detection of Avibacterium paragallinarum in chicken respiratory sites. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *31*(5), 714–718. https://doi.org/10.1177/1040638719866484

Corney, B. G., Diallo, I. S., Wright, L., Hewitson, G., De Jong, A., Tolosa, X., Burrell, P., Duffy, P., Rodwell, B., Boyle, D. B., & Blackall, P. J. (2008). Rapid and sensitive detection of Avibacterium paragallinarum in the presence of other bacteria using a 5' Taq nuclease assay: A new tool for diagnosing infectious coryza. *Avian Pathology: Journal of the W.V.P.A*, *37*(6), 599–604. https://doi.org/10.1080/03079450802449139

Dufour-Zavala, L. (2008). A Laboratory Manual for the Isolation, Identification and Characterization of Avian Pathogens. In *American Association of Avian Pathologists*.

El-Gazzar, M., Ghanem, M., McDonald, K., Ferguson-Noel, N., Raviv, Z., & Slemons, R. D. (2017). *Development of Multilocus Sequence Typing (MLST) for Mycoplasma synoviae.*

Feberwee, A., Dijkman, R., Buter, R., Soriano-Vargas, E., Morales-Erasto, V., Heuvelink, A., Fabri, T., Bouwstra, R., & de Wit, S. (2019). Identification and characterization of Dutch *Avibacterium paragallinarum* isolates and the implications for diagnostics. *Avian Pathology*, *48*(6), 549–556. https://doi.org/10.1080/03079457.2019.1641178

Gallardo, R. A., Da Silva, A. P., Egaña-Labrin, S., Stoute, S., Kern, C., Zhou, H., Cutler, G., & Corsiglia, C. (2020). Infectious Coryza: Persistence, Genotyping, and Vaccine Testing. *Avian Diseases*, *64*(2), 157–165. https://doi.org/10.1637/0005-2086-64.2.157

Ghanem, M., & El-Gazzar, M. (2016). Development of multilocus sequence typing (MLST) assay for Mycoplasma iowae. *Veterinary Microbiology*, *195*, 2–8. https://doi.org/10.1016/j.vetmic.2016.08.013

Ghanem, M., & El-Gazzar, M. (2019). Development of a Multilocus Sequence Typing Assay for Mycoplasma gallisepticum. *Avian Diseases*, *63*(4), 693–702. https://doi.org/10.1637/aviandiseases-D-19-00072

Gong, Y., Zhang, P., Wang, H., Zhu, W., Sun, H., He, Y., Shao, Q., & Blackall, P. J. (2014). Safety and efficacy studies on trivalent inactivated vaccines against infectious coryza. *Veterinary Immunology and Immunopathology*, *158*(1–2), 3–7. https://doi.org/10.1016/j.vetimm.2013.01.015

Guo, M., Liu, D., Chen, X., Wu, Y., & Zhang, X. (2022). Pathogenicity and innate response to Avibacterium paragallinarum in chickens. *Poultry Science*, *101*(1), 101523. https://doi.org/10.1016/j.psj.2021.101523

Guo, M., Liu, D., Xu, H., Zhang, H., Jin, Y., Tan, H., Wu, Y., & Zhang, X. (2022). The Protective Efficacy of an Inactivated Vaccine against Avibacterium paragallinarum Field Isolates. *Veterinary Sciences*, *9*(9), 458. https://doi.org/10.3390/vetsci9090458

Hunter, P. R., & Gaston, M. A. (1988). Numerical index of the discriminatory ability of typing systems: An application of Simpson's index of diversity. *Journal of Clinical Microbiology*, *26*(11), 2465–2466.

Kuchipudi, S. V., Yon, M., Surendran Nair, M., Byukusenge, M., Barry, R. M., Nissly, R. H., Williams, J., Pierre, T., Mathews, T., Walner-Pendleton, E., Dunn, P., Barnhart, D., Loughrey, S., Davison, S., Kelly, D. J., Tewari, D., & Jayarao, B. M. (2021). A Highly Sensitive and Specific Probe-Based Real-Time PCR for the Detection of Avibacterium

paragallinarum in Clinical Samples From Poultry. *Frontiers in Veterinary Science*, *8*. https://www.frontiersin.org/articles/10.3389/fvets.2021.609126

Luna-Castrejón, L. P., Buter, R., Pantoja-Nuñez, G. I., Acuña-Yanes, M., Ceballos-Valenzuela, K., Talavera-Rojas, M., Salgado-Miranda, C., Heuvelink, A., de Wit, S., Soriano-Vargas, E., & Feberwee, A. (2020). Identification, HPG2 Sequence Analysis, and Antimicrobial Susceptibility of Avibacterium paragallinarum Isolates Obtained from Outbreaks of Infectious Coryza in Commercial Layers in Sonora State, Mexico. *Avian Diseases*, *65*(1). https://doi.org/10.1637/aviandiseases-D-20-00103

McAuliffe, L., Gosney, F., Hlusek, M., de Garnica, M. L., Spergser, J., Kargl, M., Rosengarten, R., Ayling, R. D., Nicholas, R. A. J., & Ellis, R. J. (2011). Multilocus sequence typing of Mycoplasma agalactiae. *Journal of Medical Microbiology*, *60*(6), 803–811. http://jmm.microbiologyresearch.org/content/journal/jmm/10.1099/jmm.0.028159-0

Miflin, J. K., Chen, X., & Blackall, P. J. (1997). Molecular characterization of isolates of Haemophilus paragallinarum from China by ribotyping. *Avian Pathology: Journal of the W.V.P.A*, *26*(1), 119–127. https://doi.org/10.1080/03079459708419199

Miflin, J. K., Horner, R. F., Blackall, P. J., Chen, X., Bishop, G. C., Morrow, C. J., Yamaguchi, T., & Iritani, Y. (1995). Phenotypic and molecular characterization of V-factor (NAD)-independent haemophilus paragallinarum. *Avian Diseases*, *39*(2), 304–308.

Morales Ruiz, S., Bendezu, J., Choque Guevara, R., Montesinos, R., Requena, D., Choque Moreau, L., Montalván Ávalos, Á., & Fernández-Díaz, M. (2018). Development of a lateral flow test for the rapid detection of Avibacterium paragallinarum in chickens suspected of having infectious coryza. *BMC Veterinary Research*, *14*(1), 411. https://doi.org/10.1186/s12917-018-1729-0

Morales-Erasto, V., Falconi-Agapito, F., Luna-Galaz, G. A., Saravia, L. E., Montalvan-Avalos, A., Soriano-Vargas E, E., & Fernández-Díaz, M. (2016). Coinfection of Avibacterium paragallinarum and Ornithobacterium rhinotracheale in Chickens from Peru. *Avian Diseases*, *60*(1), 75–78. https://doi.org/10.1637/11265-082015-ResNote.1

Page, L. A. (1962). Haemophilus Infections in Chickens. III. Factors in Intraflock Transmission of Infectious Coryza and Its Chemical and Antibiotic Therapeusis. *Avian Diseases*, *6*(2), 211. https://doi.org/10.2307/1587862

Soriano-Vargas, E. (n.d.). *Infectious Coryza in Chickens*. Merck Manual Veterinary Manual.

Tan, D.-H., Ou, S.-C., Shien, J.-H., Huang, S.-W., Hsieh, M.-K., & Chang, P.-C. (2020). Serotypes and Hemagglutinin Gene Sequences of Avibacterium paragallinarum Isolated in Taiwan. *Avian Diseases*, *64*(2), 197–202. https://doi.org/10.1637/0005-2086-64.2.197

Thieme, S., Mühldorfer, K., Lüschow, D., & Hafez, H. M. (2016). Molecular Characterization of the Recently Emerged Poultry Pathogen Ornithobacterium rhinotracheale by Multilocus Sequence Typing. *PLOS ONE*, *11*(2), e0148158. https://doi.org/10.1371/journal.pone.0148158