Control Number:	2022-A-52-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Secret Ingredient: NGS to Uncover the Role of Microbes in Agricultural and Food Systems
Publishing Title:	Whole-genome sequencing-based subtyping and analyses reveal population structure and genetic diversity of *Salmonella* Typhimurium in wild birds
Abstract Body:	Within-host evolution of bacterial pathogens can lead to host-associated variants of the same species or serovar. Identification and characterization of closely related variants from diverse host species are crucial to public health and host-pathogen adaptation research. However, the work remained largely underexplored at a strain level until the advent of whole-genome sequencing (WGS). Here, we performed WGS-based subtyping and analyses of *Salmonella enterica* serovar Typhimurium (*n* = 787) from different wild birds across 18 countries over a 75-year period. We revealed seven avian host-associated *S.* Typhimurium variants/lineages. These lineages emerged globally over short timescales and presented genetic features (*e.g.*, lack of antimicrobial resistance, presence of lineage-specific virulence gene signatures) distinct from *S.* Typhimurium lineages circulating among humans and domestic animals. We further showed that, in terms of virulence, host adaptation of these variants was driven by genome degradation. Our results provide a snapshot of the population structure and genetic diversity of *S.* Typhimurium within avian hosts. We also show the importance of within-host evolution in shaping host specificity of bacterial pathogens. WGS-based subtyping and analyses are needed to unravel these emerging host-associated variants considering their close genetic relatedness at a strain level.

---PAGEBREAK---

Control Number:	2022-A-58-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Secret Ingredient: NGS to Uncover the Role of Microbes in Agricultural and Food Systems
Publishing Title:	Whole genome sequencing and analysis of *Vibrio parahaemolyticus* isolates from Alaskan oysters
Abstract Body:	*Vibrio parahaemolyticus* are pathogenic marine bacteria that are endemic to coastal waters throughout the world. The CDC estimates there are 45,000 cases of *V. parahaemolyticus* infections per year in the United States, with most foodborne infections from consumption of raw or undercooked seafood. Outbreaks and infections are typically restricted to warm water areas, but in 2004 an outbreak of illnesses on a cruise ship off the coast of Alaska was attributed to *V. parahaemolyticus* and consumption of Alaskan raw oysters. This study aimed to understand the prevalence of *V. parahaemolyticus* strains and sequence types (STs) of concern in Alaskan oysters. Oysters from several locations in the Gulf of Alaska were homogenized according to FDA methods and spread on TCBS and CHROMagar selective media to obtain *V. parahaemolyticus* isolates. Isolates were screened using real-time PCR for presence of *tlh*, *trh*, and *tdh* genes and RAPD-PCR done for coarse typing. Whole genome sequencing was performed on *tlh* positive isolates (n=54) using Illumina Miseq. On the GalaxyTrakr platform, sequencing reads were cleaned and trimmed (Trimmomatic/fastQC), assembled into contigs (SPADES/QUAST), and annotated (Prokka). Pangenome analysis with Roary found 3,445 core genes and 17,082 accessory genes across the *V. parahaemolyticus* isolates. MLST analysis using seven housekeeping genes (*dna*E, *gyr*B, *rec*A, *dtd*S, *pnt*A, *pyr*C, and *tna*A) identified seven STs: 12 (n=13), 28 (n=14), 43 (n=17), 121 (n=1), 152 (n=1), 153 (n=3), and 631 (n=1). The ST of three *V. parahaemolyticus* isolates was not able to be determined. Diversity among the isolates varied by location of isolation and included emerging STs of public health concern (ST43 and ST631). This study served as an opportunity to identify genetic profiles of native *V. parahaemolyticus* strains and assess potential emergence of strains in Alaskan waters. It is expected that with a changing climate, coastal waters in Alaska will get warmer in summer months and create a more habitable environment for *V. parahaemolyticus*.

---PAGEBREAK---

Control Number:	2022-A-65-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Secret Ingredient: NGS to Uncover the Role of Microbes in Agricultural and Food Systems
Publishing Title:	Large-Scale Genomic Analysis Identifies Antimicrobial Resistant and Potentially Pathogenic *Escherichia coli* from Dairy Calves at Commercial Settings
Abstract Body:	**Background:** Dairy calves carry a disproportionately large population of antimicrobial resistant (AMR) *Escherichia coli* compared with older animals. Currently, there is a paucity of information on the population structures, antimicrobial resistance gene (ARG) profiles, as well as virulence potential of AMR *E. coli* from dairy calves in commercial settings. Addressing these knowledge gaps will help strategize interventions and allow more informed antimicrobial stewardship programs to be developed. **Methods:** In total, the genomes of >1,000 non-redundant *E. coli* isolates (mostly AMR) previously isolated from feces collected from individual calves on 12 commercial dairy farms were sequenced (Illumina NextSeq 500, 2 x 150 bp reads). The genomes were evaluated with VirulenceFinder, ResFinder, MLST, and SerotypeFinder programs available at the Center for Genomic

Epidemiology webserver and other publicly available programs.

Results: The population of AMR *E. coli* in dairy calf feces was polyphyletic with at-least 150 known sequence types (STs) representing all major *E. coli* phylogroups, i.e., A, B1, B2, C, D, E, F, and G. The number of known *E. coli* STs in a single feces sample ranged between 1 and 8 (median 3). STs associated with extra-intestinal pathogenic *E. coli* (ExPEC), e.g., ST69 and ST117 were repeatedly detected. A diverse profile of ARGs and virulence factors were found in this group of isolates, including ARGs known to confer resistance to the drug classes Aminoglycoside, *β*-lactam, Fosfomycin, Macrolide, Phenicol, Sulfonamide, Tetracycline, and Trimethoprim. In this analysis, 31 Shiga-toxin gene (*stx*) harboring *E. coli* isolates (STEC) were identified. The STEC isolates belonged to at-least 15 STs and 11 serogroups including two of the "big six" serogroups, O103 and O111. All the O103 and O111 isolates in this study harbored the *stx* gene along with the intimin gene, *eae*, and the translocated intimin receptor coding gene, *tir*. Two of the STEC isolates harbored gene markers of an ExPEC (e.g., aerobactin and Afa/Dr-adhesin) suggesting that they may be hybrid ExPEC/STECs.

Conclusion: This study described the genomic attributes of potentially pathogenic AMR *E. coli* isolated from commercially raised dairy calves. Information presented herein will be useful for assessing public health risk and may help guide the development of preharvest prevention strategies of pathogenic and AMR *E. coli* in this important food animal reservoir.

---PAGEBREAK---

Control Number:	2022-A-105-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Secret Ingredient: NGS to Uncover the Role of Microbes in Agricultural and Food Systems
Publishing Title:	Development and optimisation of high throughput sequencing for genotypic characterisation of norovirus in shellfish
Abstract Body:	Norovirus is the leading cause of non-bacterial gastroenteritis in adults. Foodborne outbreaks of norovirus have the potential to spread internationally due to the complex global supply chain of contaminated foodstuffs such as fresh and frozen berries, leafy greens and shellfish. In this study, we have addressed both dry and wet lab methodologies for the high-throughput sequencing of norovirus in foodstuffs. For dry-lab, we used a simulated sequencing dataset to compare denoising-based pipelines (DADA2, Deblur and USEARCH-UNOISE3) and clustering-based pipelines (VSEARCH and FROGS) in an Illumina MiSeq (v3) run, as well as open source classifiers (RDP, BLASTn, IDTAXA, QIIME2 naive Bayes and sintax). Classifiers were trained using 3 different databases; a custom database, the noronet database and the HuCaT database. For the wet-lab work validation we generated a panel of spiked oysters, containing selected concentrations of norovirus and of varying genotypic compositions. The impact of DNase polymerase enzymes and reverse transcriptase enzymes were compared with regards to sequencing quality and accuracy. VSEARCH and USEARCH-UNOISE3 returned data more closely reflecting the expected composition, on all measures, ranking and phylogenetic. VSEARCH performed the best in terms of similarity to expected composition. However, DADA2 and Deblur performed poorly overall, while FROGS' performance varied across simulation. Classification was more strongly impacted by classifier rather than database; though disagreement increased with capsid variant designation. RDP classifier was determined to be a viable alternative to external classification. Wet lab studies demonstrated that DNA polymerase had a greater impact on quality of sequencing than reverse transcriptase enzyme. Classification accuracy was impacted by the starting concentration of norovirus RNA. We found that technical triplicates were necessary for accurate characterisation in lower concentration samples. Spurious genotypic characterisation results were observed in samples spiked with lower concentrations of norovirus, indicating issues with DNA polymerase performance and applicability of the method for genotypic characterisation in samples containing < 500 gc/g of norovirus RNA (RT-qPCR). Using the optimised method, we were able to successful replicate the expected composition as per Sanger sequencing in naturally contaminated shellfish. This study demonstrates the importance of the application of appropriate bioinformatic pipelines for accurate representation of sequencing data. Furthermore, we provide a standardised wet-lab method for genotypic characterisation of norovirus in shellfish that can be widely applied for both surveillance and outbreak investigations.

---PAGEBREAK---

Control Number:	2022-A-117-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Epidemiological Cues: NGS in Clinical and Public Health Microbiology
Publishing Title:	Comparative phylogenomic analyses reveal highly similar *Campylobacter jejuni* strains circulating between cattle and humans in Michigan
Abstract Body:	*Campylobacter jejuni* is the most common cause of gastroenteritis in the United States. More recently, there has been an increase in the number of *C. jejuni* outbreaks attributed to cattle products as well as an increased frequency of antibiotic resistant infections. Our prior study of *C. jejuni* strains from Michigan patients detected several cattle-associated multilocus sequence types (STs), such as ST-982, which predominated and was more frequently resistant to tetracycline. Tetracycline resistance frequencies were also higher in Michigan cases relative to national frequencies, suggesting that ST-982 and other cattle-associated tetracycline resistant strains

may contribute to regional variation. To address this hypothesis, we used whole-genome sequencing (WGS) to examine 286 *C. jejuni* strains, from human (n = 214) and cattle (n = 72) in Michigan and compared these strains by distance partition around medoids clustering (PAM) of the average nucleotide identity (ANI) using the Microbial Genomes Atlas (MiGA). The PAM medoid clustering tree, generated using the neighbor joining algorithm, identified cattle strains intermingled amongst human *C. jejuni* strains, in five unique STs or clonal complexes (CC). The cattle strains represented 15 unique STs. For example, ST-982 strains from patients (n=15) and cattle (n=8), were clustered closely together on the tree. High-quality, single nucleotide polymorphism (hqSNP) clustering analysis was then used to discriminate these related strains within ST and CC. Notably, these methods could distinguish highly related populations and identify clusters comprising related isolates from both humans and cattle. For instance, human and cattle derived ST-982 isolates differed by greater than 800 SNPs, whereas human ST-982 isolates differed by 11-36 SNPs, despite being from different years. These findings highlight the potential for sustained reservoirs of highly similar and distinct strains of *C. jejuni*. Such studies illustrate the usefulness of WGS analytical tools for better defining strains that may be more transmissible between species and demonstrate that both unique and diversifying pathogen populations are circulating within specific geographic locations.

---PAGEBREAK---

Control Number:	2022-A-123-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Epidemiological Cues: NGS in Clinical and Public Health Microbiology
Publishing Title:	Whole Genome Sequencing of *Listeria monocytogenes* from Zambian Meat Foods: a Focus on Hyper-virulent Clones

Abstract Body: Listeriosis is a severe global foodborne disease mainly linked to ready to eat (RTE) foods; however, so far few data are available from developing countries. Aim of this study was to characterize *Listeria monocytogenes* (*Lm*) strains detected from Zambian RTE foods, by whole genome sequencing (WGS). The sequences of 18 *Lm*, previously isolated in Zambian RTE meat, were obtained by Illumina platform. Subsequently, the BIGSdb-*Lm* database was used to acquire multilocus sequence typing (MLST) and core genome MLST (cgMLST), according to the Institut Pasteur's scheme of 1748 target loci (chewBBACA allele calling algorithm). BIGSdb-*Lm*, PHASTER and PlasmidFinder tools were used to identify virulence and resistance genes, intact prophages and plasmid replicons of hyper-virulent clones (n=9), respectively. Their cgMLST were also compared to available African genomic data. MLST identified 7 Clonal Complexes (CCs): CC1 (n=5), CC2 (n=4), CC9 (n=4), CC5 (n=2), CC121 (n=1), CC155 (n=1) and CC3 (n=1). CC1 and CC2, known to be hyper-virulent clones often linked to clinical cases, were predominant, carrying several virulence factors: a full-length *inlA* and *inlB, vip, virR* and *virS* genes and *Listeria* Pathogenicity Island 1 (LIPI-1). A complete LIPI-3, encoding for Listeriolysin S, was found in all CC1 strains, increasing *Lm* virulence potential. Stress and disinfectant resistance *bcrABC* cassette was found in 4 of 5 CC1 strains, while the remaining CC1 and one CC2 harbored *Tn6188_qac* for benzalkonium chloride tolerance. Various prophages were detected, while one plasmid, involved in stress response, was observed in one CC2 strain. Intrinsic antibiotic resistance genes for fosfomycin, quinolones, cationic peptides and sulfonamides were found, confirming phenotypic results. Clustering with other African genomic data highlighted a cluster (allelic distance of 3 alleles) with one CC1 strain isolated in 2016 from a meat retailer in South Africa (SA). In addition, allelic distances of 26 and 17 alleles were found for CC1 and CC2 strains isolated in 2016 and 2018 in SA meat retailers, respectively. These preliminary results, obtained also thanks to international scientific network "Enhancing Research For Africa Network" (ERFAN), show the presence of hyper-virulent CCs in Zambian RTE meat foods tested, posing a public health risk for consumers. Moreover, the presence of stress resistance factors could help these hyper-virulent clones to adapt, survive and persist over the years. WGS represents a powerful method useful to easily investigate *Lm* characteristics, monitor the worldwide diffusion and distribution of hyper-virulent clones and improve food safety. Further studies are crucial due to the global increasing request, trade and consumption of foreign foods.

---PAGEBREAK---

Control Number:	2022-A-125-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Epidemiological Cues: NGS in Clinical and Public Health Microbiology
Publishing Title:	Evolutionary dynamics of core and accessory genome of *Staphylococcus aureus* from multiple hosts

Abstract Body: *Staphylococcus aureus* is an opportunistic pathogen that causes a wide variety of infections in humans and animals. *S. aureus* exhibits a remarkable ability to switch and adapt to new hosts, yet the genetic basis of this trait remains poorly understood. We used whole genome sequencing to elucidate the structure and dynamics of the core and accessory genome of *S. aureus* isolates from human and animal hosts. In this study, we generated whole genome sequences of 437 *S. aureus* isolates recovered from human- and animal- clinical cases in New Hampshire, USA using Illumina HiSeq platform. Phylogenetic analysis revealed three major clades, each consisting of isolates from multiple host species: ST5 (n=77 genomes), ST8 (n=77 genomes) and ST30 (n=41 genomes). Phylogenetic reconstruction also allowed to identify host-specific lineages. We next sought to interrogate the genomes to

determine whether the accessory gene pool was homogeneously diverse or constrained by host. Human isolates carried on average more accessory genes (831 genes per genome) than animal isolates (p < 0.05). However, the number of unique genes per genome was statistically higher in animal isolates (mean = 9.74 genes) than in human (mean = 4.63 genes). *In silico* detection of antimicrobial resistance and virulence genes revealed that the diversity of those genes in human isolates was significantly higher than in animal isolates (p < 0.05). Plasmids analysis detected 62 different *rep* proteins from six families. Isolates from humans carried on average > two different *rep* families per genome, while animal isolates had on average only one *rep* family type per genome (p < 0.05). Core SNP distance ranged between 12,038 SNPs and 12,044 SNPs withing animal and human isolates, respectively. When compared the two groups, the core distance was 12,428 SNPs. The impacts of recombination events in the core genome were estimated in average of 29.74 events per genome in animal isolates and 22.28 in human isolates. Although recombination events seemed to have more impact in the evolution of the core genome of animal isolates, no statistical support was found. Lastly, principal component analysis of the core and accessory genome revealed low variation between human and animal isolates. However, further analysis splitting the isolates according to the STs was able to convey the high variation in the accessory genome. Our study highlights the importance of the accessory gene pool in shaping the evolutionary history of *S. aureus* and its ability to adapt to multiple host species. Our study has important implications for both animal and human health, including epidemiological tracking, disease control and treatment of resistant bacteria.

---PAGEBREAK---

Control Number: 2022-A-126-NGS

Session Title: Secret Ingredient

Topic 1: Epidemiological Cues: NGS in Clinical and Public Health Microbiology

Publishing Title: Leaky barriers to inter-species gene sharing between locally co-existing coagulase-negative *Staphylococcus*

Abstract Body: *Staphylococcus* species are generally grouped as coagulase-negative or coagulase positive. Coagulase negative *Staphylococcus* (CoNS) have been implicated as opportunistic pathogens in human and animal infections. Compared to the more well-known coagulase-positive *Staphylococcus aureus*, little is known of the contribution of genetic recombination in CoNS evolution. In this study, we aimed to elucidate the phylogenetic relationships and frequencies of genome-wide recombination in diverse CoNS species. Here, we generated whole genome sequences of 191 CoNS isolates from ten animal hosts using the Illumina NextSeq sequencing platform, which can be delineated into 15 different species. Genes mediating resistance to beta-lactams (*bla*I [n=50 genomes, 26.17%], *bla*Z [n=42 genomes, 21.98%] and *mecA* [n=18 genomes, 9.42%]) and fosfomycin (*fosB* [n=24 genomes, 12.56%]) were most common. We identified 11 types of the mobile SCC*mec*, which carried the *mecA* gene. At least one member of the 14 CoNS species harboured genetic determinants for metals resistance [arsenic, copper, cadmium and mercury]. The virulence genes *icaC* (n= 30 genomes, 15.70%), *esxA* (n= 25 genomes, 13.08%) and *hld* (n=14 genomes, 7.32%). We detected the presence of inter- and intra-species homologous recombination in core and shared accessory genes. Intra-species recombination was most frequent in *Staphylococcus simulans, Staphylococcus felis, Staphylococcus xylosus* and *Staphylococcus epidermidis*. Inter-species gene sharing was most frequent between the following species pairs: *Staphylococcus pseudoxylosus + Staphylococcus xylosus* and *Staphylococcus epidermidis + Staphylococcus hominis*. This study highlights the importance of CoNS as hubs for transfer and receipt of DNA, including resistance genes. Understanding the evolutionary drivers that shape the genetic diversity of CoNS is critical to effective surveillance, diagnostics, and control of CoNS infections.

---PAGEBREAK---

Control Number: 2022-A-160-NGS

Session Title: Secret Ingredient

Topic 1: Secret Ingredient: NGS to Uncover the Role of Microbes in Agricultural and Food Systems

Publishing Title: Toward a national microbial genomic surveillance infrastructure - a retrospective analysis of 4,891 Australian *Escherichia coli* genomes

Abstract Body: Linkages between human, animal and environmental microbial communities increasingly highlight the need for national and cross-sectoral microbial genomic surveillance systems for the identification of reservoirs and vectors of pathogenic and antimicrobial resistant (AMR) bacteria. *Escherichia coli* is a commensal of many vertebrate species, a common pathogen of both humans and livestock and a contaminant of food products and natural environments, making it an ideal bacterial species for the focus of a One Health genomic surveillance system. To guide the development of such a system we collated 4,891 Australian *E. coli* genomes of diverse host/source origin including multiple livestock species, companion animals, wild animals, natural environments, foodstuffs and healthy and diseased humans. Genotypic, pangenomic and phylogenetic analyses were used to study the relatedness of the strains, their carriage of virulence and AMR determinants and their association with mobile genetic elements, particularly those important in the dissemination of AMR and virulence. Among the 617 sequence types (STs) the top 20 were representative of the pandemic extraintestinal pathogenic *E. coli* (ExPEC) STs including STs 131, 95, 1193, 10, 73, 38 and 69. We also identified 84 novel sequence types. While these ExPEC STs were enriched in strains associated with sepsis and urinary tract infections in humans

we also observed their presence in many other sources including wastewater, livestock, wild animals and companion animals. Notably, core-genome multilocus sequence typing and reference-based phylogenetic analyses revealed extensive genomic similarity among strains isolated from different sources. This data is indicative of potential intersource transmission events and highlights the value of One Health genomic epidemiology. More generally, we present insights into the phylogenetic structure of Australian *E. coli*, highlight trends in their carriage of antibiotic resistance genes and mobile genetic elements and the association of these characteristics with source attribution. We also frame the utility of this data in informing the development of a prospective national genomic surveillance system for *E. coli* aiming to facilitate harmonised cross-sectoral data sharing and analysis to improve human, animal and environmental health outcomes.

---PAGEBREAK---

Control Number:	2022-A-190-NGS
Session Title:	**Secret Ingredient**
Topic 1:	Secret Ingredient: NGS to Uncover the Role of Microbes in Agricultural and Food Systems
Publishing Title:	Characterization of Frozen and Freeze-dried Commercial Raw Pet Food by Shotgun Metagenomics

Abstract Body:

An increasing number of reports have indicated that "raw" pet food can serve as a reservoir for zoonotic foodborne pathogens and antimicrobial resistance genes, and dogs fed raw diets have higher rates of shedding of resistant bacteria in their feces. In this study, shotgun metagenomic sequencing was employed to characterize the microbial communities of frozen (n=7) and freeze-dried (n=7) commercial "raw" pet foods. All samples were processed in triplicate. To complement culture-independent metagenomic profiling, selective enrichments for *Campylobacter*, *Listeria*, *Salmonella*, and *Enterococcus* were evaluated using shotgun metagenomic sequencing. Bacterial taxonomic classifications and antimicrobial resistance genes were analyzed using Kraken and AMRFinderPlus, respectively. Communities from culture-independent frozen samples were dominated by Firmicutes, Actinobacteria, Proteobacteria, and Bacteroidetes and shared a small core microbiome consisting of the genera *Carnobacterium*, *Brochothrix*, *Bacillus*, *Acinetobacter*, *Lactococcus*, and *Psychrobacter*. *Enterococcus* spp. were also detected from all frozen samples, with a relative abundance of ~0.02-1.1% and ~3-99% of in culture-independent and enriched samples respectively. *Salmonella*, *Campylobacter,* and *Listeria* were not detected at elevated levels (> 1.0%) in any of the frozen or freeze-dried products. Freeze-dried communities were distinct from those identified in frozen products and were dominated by the genera *Bacillus*, *Pediococcus*, and *Lactobacillus*. Products with poultry (duck or chicken) as the main protein source had a higher relative abundance of *E. coli* ($1.54 \pm 1.03\%$) in culture-independent samples than non-poultry-based products ($0.08 \pm 0.15\%$; $P < 0.01$). No other associations between primary protein source and specific foodborne pathogens were detected. Antimicrobial resistance genes conferring resistance to aminoglycosides, beta-lactams, folate pathway inhibitors, glycopeptides, macrolides, phenicols, quinolones, and tetracyclines were detected with considerable variation in antimicrobial gene profiles observed between product types (frozen vs freeze-dried) where higher detection rates and diversity were found in frozen samples. While shotgun metagenomics is a powerful tool for characterizing the microbial communities of diverse environments, the complex nature of pet foods matrices makes the detection of low-abundance pathogens challenging. Continued characterization of foodborne pathogens and resistant bacteria is needed to better understand the risks of these types of pet foods to the health of pets and their owners.

---PAGEBREAK---