AGRICULTURAL RESEARCH FOUNDATION
INTERIM REPORT
FUNDING CYCLE 2021 – 2023

TITLE: Method development for *Listeria monocytogenes* profiling to assist the food industry with controlling transient and persistent contamination reservoirs: an artisan cheese facility case study

RESEARCH LEADER: Jovana Kovacevic (PI), Stephanie Brown (co-PI), Joy Waite-Cusic (co-PI)

COOPERATORS: Lorraine McIntyre, British Columbia Centre for Disease Control

EXECUTIVE SUMMARY: Whole genome sequencing (WGS) is a tool often utilized during outbreak tracing; however, it is less often used to trace contamination events in food processing facilities. While the technology is available, associated costs and complexities with data processing, data storage and personnel training, are often limiting factors for many processors to use WGS in their operations. The goal of this project is to gather and use WGS data to identify a set of genetic elements that can be used for 'partial-genome' profiling of *L. monocytogenes* strains. The findings will be used towards the development of a simplified profiling method for *L. monocytogenes* isolates to help food processing facilities identify and mitigate contamination events from transient and persistent reservoirs. Answering whether contamination is from a transient source or potentially established reservoir in the production environment is critical for mitigation of risk, effective environmental monitoring programs and appropriate reaction to positive findings.

OBJECTIVES: The overall objective of this project is to identify a set of genetic elements that can be used for 'partial-genome' profiling of *L. monocytogenes* strains. This profiling method is intended to help food processing facilities identify and mitigate contamination events from transient and persistent reservoirs. This will be done through three specific objectives:

1. Whole genome sequencing of *L. monocytogenes* isolates (n=72) that have been recovered from an artisan cheese facility over six sampling events in 2007-2016 (facility closed);
2. Investigating the genetic relatedness of isolates recovered over time to identify putative transient and persistent strains; and
3. Identifying key genetic markers and their distribution within the sequenced *L. monocytogenes* population for inclusion in the 'partial-genome' profiling method for *L. monocytogenes* strains.

PROCEDURES:

Objective 1. Genomic DNA was extracted from 72 *L. monocytogenes* isolates using the Qiagen DNeasy Blood and Tissue kit (QIAGEN, Germantown, MD) according to the manufacturer's recommendations for Gram positive bacteria. DNA quality was measured and an A260/280 of 1.8-2.0 was confirmed using the DeNovix DS-11+ (DeNovix, New Castle, Delaware). Isolates that

did not meet the 260/280 quality range were reprecipitated following Qiagen protocol. DNA was sent to the Center for Quantitative Life Sciences (Oregon State University, Corvallis, OR) for library preparation, using the plexWell kit (seqWell, Beverly, MA), and sequencing. Paired-end (2 x 150 bp) sequencing was performed using the Illumina HiSeq platform (Illumina, San Diego, CA, United States). Raw sequence reads were quality checked with FastQC, followed by quality trimming with Trimmomatic (v 0.39). Trimmed sequence reads were *de novo* assembled using SPAdes optimized with unicycler (v 0.4.8) to construct contiguous sequences (contigs). Genomes were annotated using Prokka (v 1.12).

Objective 2. To assess diversity and relatedness of isolates, multilocus sequence typing (MLST), core genome MLST (cgMLST) and single nucleotide polymorphism analysis was carried out. Assembled sequences were run through the *Listeria* Pasteur database which compares sequence information of seven housekeeping genes (*acbZ, bglA, cat, dapE, dat, ldh,* and *lhkA)* to that of existing sequence type (ST) profiles within a database. ST and clonal complexes (CCs) were assigned based on the definition of Ragon et al. (2008) and those available in the database. Core genome MLST (cgMLST) was performed on the same platform using a scheme consisting of 1,748 conserved core genes (Moura et al., 2016). Following comparison of isolates to cgMLST profiles existing in the database, assemblies of all isolates were submitted to the database to receive a cgMLST type (CT). Single nucleotide polymorphisms (SNP) analysis was conducted with FastANI (v 1.1) to calculate pairwise average nucleotide identity (ANI) among isolates and confirm species-level grouping. BWA (v 0.7.17) was used to map raw reads to respective reference sequence within the group. These alignments were then annotated, sorted and duplicate reads identified using Picard tools (v 2.0.1). Graphtyper (v 2.6.2) was run on each dataset with the default parameters (Eggertsson et al., 2017). SNPs were filtered with vcffilter in vcflib (v 1.0.0), with calls annotated as “FAIL” or “heterozygous” filtered to “no-call”. The bitwise.dist function from the R package popper (v 2.9.2) was used to construct pairwise SNP distance tables from the fasta alignment, which were then used to construct and visualize a minimum spanning network. Geneious (v 2020.1.2) was used to further investigate mutations.

Objective 3. Genome assemblies (n=72) were screened for the presence or absence of virulence and stress tolerance genes using BLAST+, and the *Listeria* Pasteur database. Genes associated with sanitizer tolerance (*bcrABC, emrE, emrC, tetR, tnpABC, qacC, qacH*) were screened using BLASTN algorithm with a minimum nucleotide identify and alignment length coverage of 80% (Hurley et al., 2019). Prevalence of virulence factors including *inlA* and *inlB* genes, *Listeria* pathogenicity island 1 (LIPI-1), and stress survival islet 1 (SSI-1), were assessed using 80% minimum identity.

SIGNIFICANT ACCOMPLISHMENTS TO DATE:

To date, all 72 *L. monocytogenes* isolates have been whole genome sequenced (WGS) at the Center for Quantitative Life Sciences. Results detailing genetic markers and WGS analysis have been shared in a technical presentation at the International Association for Food Protection

Conference in Phoenix, AZ, 2021 (abstract below). Two manuscripts are currently in preparation, focusing on the genetic markers as well as genetic relatedness of the isolates.

Objectives 1 and 2.
All 72 isolates have been sequenced, and raw sequence reads processed into draft assembly files for subsequent analysis. The majority of isolates recovered belonged to ST11/CC11 (65/72), followed by ST7/CC7 (2/72), ST224 CC224 (1/72), and ST397/CC4 (1/72) (Figure 1). Notably, ST11 isolates were recovered over a 10 year period (2007-2016) across the facility on both food contact and non-food contact surfaces, as well as in cheese products. Three isolates belonged to hypervirulent clones (CC1, CC4).

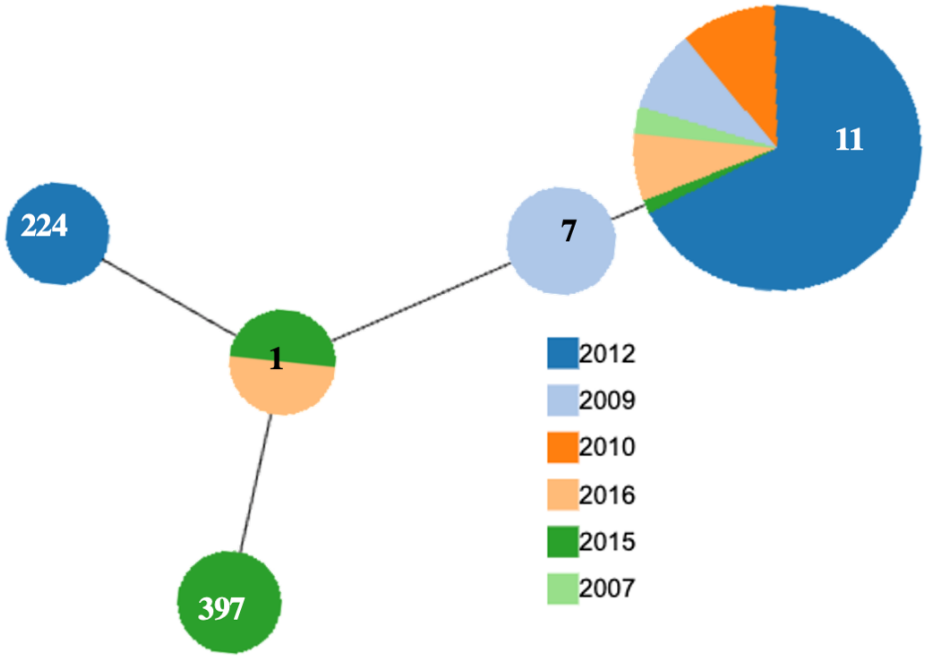

Figure 1 | Minimum spanning tree of *L. monocytogenes* sequence types (STs) recovered over 10 years. Numbers inside represent STs; node diameter is in positive correlation to number of isolates; node colors represent years when isolates were recovered**.**

Objective 3.
All isolates possessed *Listeria* pathogenicity island 1 (LIPI-1), which is present across the *Listeria* species. In the *actA* gene of LIPI-1, a 35 amino acid internal deletion was identified in 3/72 isolates. Premature stop codons (PMSCs) were not observed in either *inlA* or *inlB* genes across the isolate set. *Listeria* pathogenicity island-3 (LIPI-3) was seen in four isolates, recovered from cheeses and a raw milk pump, while one isolate possessed LIPI-4. None of the known genes associated with sanitizer tolerance were found within the tested isolate set. The five-gene stress survival islet 1 (SS1-1), known to assist *L. monocytogenes* with survival under acidic environments and high salt concentrations, was present in 6% (4/72) of isolates. *In progress: Presence, absence and mutations in additional virulence genes will be identified for 'partial-genome' screening method.*

Presented work:

International Association for Food Protection Annual Meeting, Phoenix, AZ. July 18-21, 2021.

Title: Genetic Diversity and Virulence of *Listeria monocytogenes* recovered from an Artisan Cheese Facility over a Decade
Technical talk presentation

Authors: Rebecca Bland, Stephanie Brown, Lorraine McIntyre, Sion Shyng and Jovana Kovacevic

Introduction: Whole genome sequencing (WGS) and isolate genotyping are utilized during outbreak tracing; however, they are less often used to trace contamination events in food processing facilities. Here we describe a retrospective use of WGS to better understand *Listeria monocytogenes* (*Lm*) contamination in an artisan cheese facility over a decade.

Purpose: Assess diversity and properties of *Lm* recovered from an artisan cheese facility over 10 years.

Methods: *Lm* (n=72) were recovered from 32 samples (13 cheese; 1 food contact; 18 non-food contact surfaces) collected during six sampling events in 2007-2016 (facility closed). Extracted DNA (Qiagen) was pair-end sequenced on Illumina HiSeq. Sequences were trimmed using Trimmomatic and *de novo* assembled with SPAdes. Multi-locus sequence types (STs), clonal complexes (CCs) and virulence profiles were assessed *in silico* using *Listeria* Pasteur database. Isolates were screened for antimicrobial tolerance-associated genes (*bcrABC, emrE*, *emrC*, *qacC*, *qacH, tetR*, *tnpABC*) with BLASTN (380% identity/alignment).

Results: Among the five STs and CCs recovered, the majority of isolates belonged to ST/CC11 (65/72), followed by ST/CC7 (3/72), ST/CC1 (2/72), ST/CC224 (1/72), and ST397/CC4 (1/72). Notably, ST/CC11 isolates were repeatedly recovered during 2007-2016, from drains and draining racks over multiple years, on mobile equipment (cart and table wheels, 2016), and in soft cheeses (2012). All isolates possessed full-length *inlA*, a known virulence factor. *Listeria* pathogenicity island-3 (LIPI-3) was seen in four isolates, recovered from cheeses and a raw milk pump. Three isolates belonged to hypervirulent clones (CC1, CC4).

Significance: Results showed systemic and widespread contamination with some *Lm* clones (ST/CC11; ST/CC1), recovered from multiple food

ADDITIONAL FUNDING RECEIVED DURING PROJECT TERM: BUILD Dairy.

FUTURE FUNDING POSSIBILITIES: USDA NIFA AFRI Foundational Program.

REFERENCES:

Eggertsson, H. P., Jonsson, H., Kristmundsdottir, S., Hjartarson, E., Kehr, B., Masson, G., Zink, F., Hjorleifsson, K. E., Jonasdottir, A., Jonasdottir, A., Jonsdottir, I., Gudbjartsson, D. F., Melsted, P.,

Stefansson, K., & Halldorsson, B. V. (2017). Graphtyper enables population-scale genotyping using pangenome graphs. *Nat Genet*, *49*(11), 1654-1660. https://doi.org/10.1038/ng.3964

Hurley, D., Luque-Sastre, L., Parker, C. T., Huynh, S., Eshwar, A. K., Nguyen, S. V., Andrews, N., Moura, A., Fox, E. M., Jordan, K., Lehner, A., Stephan, R., & Fanning, S. (2019). Whole-genome sequencing-based characterization of 100 *Listeria monocytogenes* isolates collected from food processing environments over a four-year period. *mSphere*, *4*(4). https://doi.org/10.1128/mSphere.00252-19

Moura, A., Criscuolo, A., Pouseele, H., Maury, M. M., Leclercq, A., Tarr, C., Bjorkman, J. T., Dallman, T., Reimer, A., Enouf, V., Larsonneur, E., Carleton, H., Bracq-Dieye, H., Katz, L. S., Jones, L., Touchon, M., Tourdjman, M., Walker, M., Stroika, S., Cantinelli, T., Chenal-Francisque, V., Kucerova, Z., Rocha, E. P., Nadon, C., Grant, K., Nielsen, E. M., Pot, B., Gerner-Smidt, P., Lecuit, M., & Brisse, S. (2016). Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol*, *2*, 16185. https://doi.org/10.1038/nmicrobiol.2016.185

Ragon, M., Wirth, T., Hollandt, F., Lavenir, R., Lecuit, M., Le Monnier, A., & Brisse, S. (2008). A new perspective on *Listeria monocytogenes* evolution. *PLoS Pathog*, *4*(9), e1000146. https://doi.org/10.1371/journal.ppat.1000146