OPTIMIZING MAIN SPECTRUM PROFILES FOR USE IN REAL-TIME MALDI-TOF MASS SPECTROMETRY IDENTIFICATION OF *LEPTOSPIRA* SEROVARS

BY

JENNIFER E. NICKOLYN-MARTIN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in VMS-Pathobiology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Master's Committee:

Professor Carol Maddox, Advisor
Associate Professor Gee Lau
Assistant Professor Matthew Allender

Clinical Associate Professor Marilyn O'Hara Ruiz (deceased)

ABSTRACT

Leptospirosis, a zoonotic disease with a growing number of reported cases worldwide and expansion into new geographic areas, is an increasingly important public and veterinary health concern. Leptospirosis epidemiology is changing for both human and non-human animals. Over 250 pathogenic serovars of *Leptospira*, the causative agents of leptospirosis, have been identified. Given the proclivity for certain associations between these serovars and reservoir hosts, identification of clinical isolates to the serovar-level is key to understanding the epidemiology of this disease. However, an economical and rapid serovar typing method that can be incorporated into routine diagnostic testing is still lacking. The companion animal most often diagnosed with leptospirosis is the dog. Routine epi-surveillance of canine leptospirosis could benefit both canine and human health. Given the widespread use of Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) by diagnostic laboratories for microorganism identification to the species-level, MALDI-TOF MS was explored as a potential method for serovar detection and identification in leptospirosis-positive canine urine samples.

Commercial MALDI platforms currently do not include *Leptospira* serovar mass spectral profiles in their reference databases. Therefore, a custom MALDI *Leptospira* Main Spectrum Profile (MSP) database, representing seven selected serovars of *Leptospira interrogans,* was created for this project. An initial database consisted of MSPs constructed according to the manufacturer's recommended guidelines. MSP specificity was tested using serial dilutions of serovar culture and culture spiked canine urine. Specificity of these original MSPs was found to be insufficient for accurate and consistent serovar identification. Various MALDI sample preparation and deposition methods were tested to determine if any produced higher quality sample spectra that would improve serovar MSP specificity. It was found that the standard ethanol/formic acid protein extraction sample preparation and dried droplet with matrix overlay deposition methods produced the best spectra for MSP creation. Several serovar raw spectra and peak analyses were performed using Bruker's FlexControl, Compass Explorer, and ClinProTools software to identify potentially-characteristic serovar peak patterns and outlier raw spectra. The data obtained were used in creating several additional MSP types for each serovar, with each type using a different combination of creation parameters. It was found that MSPs designed using creation parameters that differed from those recommended by the manufacturer resulted in the greatest MSP specificity. The best-performing MSP for each serovar was chosen to create a custom *Leptospira* MSP database. Database specificity was tested in two blind-coded trials using the Bruker MALDI Biotyper Realtime Classification software. In the first trial, which used serovar culture samples for testing, 111 of 112 serovar sample spots returned a correct first match. Specificity ranged from 99 to 100%, while sensitivity ranged from 81 to 100%. In the second trial, which used serovar-spiked canine urine samples, 105 of 112 serovar sample spots returned a correct first match. Specificity for this trial ranged from 97 to 100%, while sensitivity ranged from 75 to 100%. This work demonstrated that, by optimizing custom MSP creation parameters,

MALDI-TOF MS can be used to identify *Leptospira* isolates at the serovar-level within the real-time classification workflow.

The second part of this project explored MALDI's sensitivity for *Leptospira* serovar detection and identification. The goal was to learn whether MALDI-TOS MS can detect *Leptospira* at the concentrations typically seen in leptospirosis-positive canine urine specimens submitted to the University of Illinois at Urbana-Champaign's Veterinary Diagnostic Laboratory (VDL). All qPCR-canine-leptospirosis-positive cases diagnosed by the VDL over a 2.5-year period were analyzed. The average cycle threshold (C_T) value for these cases was 35.18. A standard curve was used to estimate a corresponding *Leptospira* concentration of 1.18×10^3 organisms/mL. Sensitivity trials were performed using serial two-fold dilutions of serovar cultures and serovar-spiked canine urine specimens tested against the custom *Leptospira* MSP database. The highest dilutions which returned accurate MALDI identifications to the serovar- and genus-levels were noted. *Leptospira* concentrations and C_T values corresponding to these dilutions were determined. The lowest average concentrations that returned accurate MALDI identifications to the serovar- and genus-levels 3.55×10^8 organisms/mL and 3.33×10^8 organisms/mL, respectively. No significant difference was found between the lowest concentrations that returned accurate identifications for serovar culture versus culture-spiked canine specimen dilutions. The difference between the average concentration of leptospirosis-positive specimens received by the VDL and the lowest average concentrations identified by the MALDI revealed that the MALDI's sensitivity for *Leptospira* is too low to be used as a routine leptospirosis diagnostic or epi-surveillance tool. Various concentration methods, including centrifugation and filtration, were tested to determine whether they could sufficiently concentrate *Leptospira* samples for MALDI direct detection. None proved successful.

In summary, MALDI's potential use for routine leptospirosis diagnostics and surveillance remains possible. Different matrices, development of selective enrichment cultures, novel sample preparation and concentration methods, modifications to MALDI software parameters, or a combination thereof, may yet offer a path forward. The work presented here proposes a technique by which custom *Leptospira* MSPs may be created for use in real-time serovar identification. To the author's knowledge, this is the first study to report MALDI-TOF MS *Leptospira* serovar identification within the real-time, rather than offline, classification workflow. Though further testing of additional serovars and clinical isolates is needed, this technique may prove applicable to a variety of serovars other than those used in this study. If so, this technique would offer a method for diagnostic laboratories to create custom *Leptospira* MSPs for identification and surveillance of the predominant circulating serovars in their respective geographic regions.

ACKNOWLEDGEMENTS

I first want to express my sincere gratitude to my advisor, Dr. Carol Maddox. Thank you for all of your feedback, motivation and support. I next want to thank my committee members, Dr. Gee Lau and Dr. Matthew Allender, for their valuable advice, insight and encouragement. I also want to convey my appreciation for the time Dr. Marilyn O'Hara Ruiz, who had also been on my committee, devoted to my project prior to her passing. I further want to acknowledge and thank my advisor, committee members, Dr. Dongwan Yoo, Karen Nichols and the Pathobiology Department for their flexibility and accommodations that allowed me to complete my thesis while managing a chronic illness.

I also gratefully acknowledge the Companion Animal Research Fund that supported this project.

I next want to thank the University of Illinois at Urbana-Champaign's Veterinary Diagnostic Laboratory (VDL) for sharing laboratory space and providing the canine urine specimens and *Leptospira* cultures used in this study. I also want to recognize the following current and prior members of the VDL: Marwa Adawe, Larissa Bokan, Vanessa DeShambo, Therese Eggett, Kristin Ellis, Namjung Jung, Saraswathi Lanka, Christina Nowicki, Miranda O'Dell, Shannon Pierce, Samantha Soukup, and Amy Stevenson. Thank you for training me in various diagnostic techniques, sharing your technical expertise, and offering advice when I faced challenges. Your help was invaluable.

To my coach and mentor, Cara Sachs, thank you for your steadfast encouragement and help in navigating the challenging landscape of graduate studies with a chronic illness. To my parents, sister, and husband's family, a heartfelt thank you for your continuous support and encouragement. To Jiji, Caspian, and Natsu, thank you for your company during the many hours of writing, and for always brightening my mood when I got discouraged. To Sophie and Hera, thank you for providing a sense of calm and encouragement during this journey.

Most importantly, I want to thank my wonderful husband for his love, patience, and belief in me. You have been my pillar of strength throughout my master's studies and thesis work. This wouldn't have been possible without you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviation	Meaning
2D	two-dimensional
ACN	acetonitrile
BDAL	Bruker Daltonics database
BTS	bacterial test standard (Bruker)
CAAT	cross-absorption agglutination test
CPT	ClinProTools
C_T	cycle threshold
CV	cross validation
DT	direct transfer
EtOH	ethanol
FA	formic acid
GA	Genetic Algorithm (CPT model type)
HCCA	α-cyano-4-hydroxycinnamic acid (MALDI matrix solution)
LOD	limit of detection
MALDI-TOF MS	matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MAT	microscopic agglutination test
MBT	MALDI Biotyper
MS	mass spectrometry
MSP	main spectrum profile
m/z	mass-to-charge ratio
OMP	outer membrane protein
P80	Polysorbate 80
PADs	Anderson-Darling p-values (in CPT)
PES	polyethylsulfone membrane
PSR	Peak statistic report (in CPT)
PTTA	Welch's T-test or ANOVA p-value (in CPT)

PVDF	polyvinylidene difluoride
PWKWs	Wilcoxon or Kruskal-Wallis test p-values (in CPT)
QC	Quick Classifier (CPT model)
RC	recognition capability
RTC	Realtime Classification (software)
SC	starting concentration
S/N	signal-to-noise ratio
TE	tube extraction
TLR2	Toll-like receptor 2
TLR4	Toll-like receptor 4
TOF	time-of-flight
UHPLC	ultrahigh-performance liquid chromatography
VDL	veterinary diagnostic laboratory
WKW	Wilcoxon or Kruskal-Wallis test (in CPT)
X Original **(A Original, B Original, C Original, G Original, H Original, I Original, P Original)**	References the original MSP created for each serovar. This MSP was created using one set of raw spectra, a maximum desired peak number of 70, and a lower boundary of 3000 Da for peak detection. The "X" here is a placeholder for the first letter of the respective serovar. A = Autumnalis, B = Bratislava, C = Canicola, G = Grippotyphosa, H = Hardjo, I = Icterohaemorrhagiae, and Pomona = Pomona hence forth.
X 70 **(A 70, B 70, C 70, G 70, H 70, I 70, P 70)**	References the second MSP created for each serovar. This MSP was created using three sets of raw spectra, a maximum desired peak number of 70, and a lower boundary of 3000 Da for peak detection.
X 100 **(A 100, B 100, C 100, G 100, H 100, I 100, P 100)**	References the third MSP created for each serovar. This MSP was created using three sets of raw spectra, a maximum desired peak number of 100, and a lower boundary of 3000 Da for peak detection.
X 125 **(A 125, B 125, C 125, G 125, H 125, I 125, P 125)**	References the fourth MSP created for each serovar. This MSP was created using three sets of raw spectra, a maximum desired peak number of 125, and a lower boundary of 3000 Da for peak detection.
X 2000 M100 **(B 2000 M100, G 2000 M 100)**	References an additional type of MSP created for serovars Bratislava and Grippotyphosa. This MSP was created using three sets of raw spectra, a maximum desired peak number of 100, and a lower boundary of 2000 Da for peak detection.

X 2D **(B 2D, G 2D)**	References an additional type of MSP created for serovars Bratislava and Grippotyphosa. This MSP was created using three sets of raw spectra, a maximum desired peak number of 125, a lower boundary of 3000 Da for peak detection, and the removal of selected raw spectra identified via 2D peak analysis in CPT.
X Revision_3000_70 **(B Revision_3000_70, G Revision_3000_70)**	References an additional type of MSP created for serovars Bratislava and Grippotyphosa. This MSP was created using three sets of raw spectra, a maximum desired peak number of 70, a lower boundary of 3000 Da for peak detection, and the removal of selected raw spectra identified via 2D peak analysis, PCA, and external validation of classification models in CPT. The spectra omitted in this MSP differed from those omitted in the X 2D MSP.
X Revision_3000_100 **(B Revision_3000_100, G Revision_3000_100)**	References an additional type of MSP created for serovars Bratislava and Grippotyphosa. This MSP was created using three sets of raw spectra, a maximum desired peak number of 100, a lower boundary of 3000 Da for peak detection, and the removal of selected raw spectra identified via 2D peak analysis, PCA, and external validation of classification models in CPT. The spectra omitted in this MSP differed from those omitted in the X 2D MSP.
X Revision_2000_70 **(B Revision_2000_70, G Revision_2000_70)**	References an additional type of MSP created for serovars Bratislava and Grippotyphosa. This MSP was created using three sets of raw spectra, a maximum desired peak number of 70, a lower boundary of 2000 Da for peak detection, and the removal of selected raw spectra identified via 2D peak analysis, PCA, and external validation of classification models in CPT. The spectra omitted in this MSP differed from those omitted in the X 2D MSP.
X Revision_2000_100 **(B Revision_2000_100, G Revision_2000_100)**	References an additional type of MSP created for serovars Bratislava and Grippotyphosa. This MSP was created using three sets of raw spectra, a maximum desired peak number of 100, a lower boundary of 2000 Da for peak detection, and the removal of selected raw spectra identified via 2D peak analysis, PCA, and external validation of classification models in CPT. The spectra omitted in this MSP differed from those omitted in the X 2D MSP.
%T	percent transmittance

Chapter 1: INTRODUCTION

1.1 THE CHANGING EPIDEMIOLOGY OF LEPTOSPIROSIS

Leptospirosis, caused by pathogenic serovars of spirochete bacteria in the genus *Leptospira*, is one of the leading zoonotic causes of morbidity and mortality worldwide (Lau et al. 2010). It has been estimated to cause one million cases and approximately 60,000 human deaths each year (Costa et al. 2015). Since the disease often presents with non-specific symptoms and can be difficult to diagnose, it is thought to be under-reported (Guerra 2013). While recognized for over a century, it is now considered a re-emerging disease due to a changing epidemiology. These changes include an increase in the number of reported cases worldwide, an expansion into urban areas and more temperate regions, and the appearance of more severe forms of the disease (Bharti et al. 2003, WHO 2003, Meites et al. 2004, Vijayachari et al. 2008, Lau et al. 2010, Chen et al. 2011, Guerra 2013, Pijnacker et al. 2016).

Dogs are the most common companion animal diagnosed with this disease (Bowles 2015, Lunn 2019). They are known reservoir hosts for *Leptospira interrogans* serovar Canicola (Levett 2001, Ellis 2015), and often become accidental hosts for different serovars carried by other domestic or wild animals. Leptospirosis in canines can vary from subclinical illness to severe disease and even death (Adler and Faine 2006, van de Maele et al. 2008, Lunn 2019). Chronically infected dogs can also act as carriers, shedding leptospires via urine into the environment, where the bacteria may stay viable and infectious for days to weeks, depending on the environmental conditions (Batista et al. 2004, Barragan et al. 2017).

As seen with human leptospirosis, the incidence of canine leptospirosis has been on the rise. Higher case numbers have been reported in North America (Rentko et al. 1992, Sykes 2001, Prescott et al. 2002, Ward et al. 2002), parts of Europe (Majetic et al. 2014, Major et al. 2014, Pijnacker et al. 2016), and the Asia Pacific region (Victoriano et al. 2009). Case numbers may also be on the rise in other regions where leptospirosis is known to cause substantial human disease but for which incidence data is lacking (e.g. Africa) (Allan et al. 2015, Roqueplo et al. 2019). Additionally, the number of severe cases, such as those presenting with leptospirosis pulmonary hemorrhagic syndrome, has increased (Gendron et al. 2014, Schuller et al. 2015). Over the last four decades, there has also been a gradual shift in the predominant circulating serovars responsible for the majority of canine clinical cases (Rentko et al. 1992, Sykes 2001, Prescott et al. 2002, Ward et al. 2004, Sykes et al. 2011). Seroprevalence shifts in pathogenic serovars can directly impact the efficacy of canine leptospirosis vaccines. This is because the current canine vaccines used in particular geographic regions are designed to protect against serovars thought to be predominant in those particular regions and are generally serovar/serogroup specific. With over 250 pathogenic serovars currently recognized (Cerqueira and Picardeau 2009, Adler and de la Peña Moctezuma 2010) and the wide variety of potential *Leptospira* maintenance hosts, detection of seroprevalence shifts is vital for disease management. The extent of zoonotic transmission of

leptospirosis from dogs to humans is still unknown. However, urinary excretion of leptospires by asymptomatic dogs has been reported (Rojas et al. 2010, Miotto et al. 2018). Consequently, infected dogs, both with and without obvious clinical signs, could serve as a source of infection for humans and other animals.

The increase in clinical leptospirosis cases and expansion of cases into non-endemic areas are thought to be related to changes in population growth and distribution, land use, and climate (Lau et al. 2010). The expansion into previously wild areas cleared for new construction or agriculture increases the potential for interactions between humans, domestic animals, and wildlife. This results in greater opportunities for leptospirosis transmission between wildlife reservoirs and incidental hosts (Ward et al. 2002, Guerra 2013). Once considered a disease found primarily in rural settings, the number of leptospirosis cases in urban areas has continued to rise. Urban slums found in the larger cities of some developing countries have been hit particularly hard, with regular outbreaks during seasonal periods of heavy rainfall (Barcellos and Sabroza 2001, LaRocque et al. 2005, Minter et al. 2018). Poor housing conditions, crowded living quarters and lack of adequate sanitation in these areas increase the likelihood of exposure to reservoir hosts and areas that have been contaminated by host urine (Romero et al. 2003). It is important to note that this disease is not just a problem for resource-poor countries in tropical regions. Incidence has increased in developed countries and in more temperate regions (Bharti et al. 2003, Andersen-Ranberg et al. 2016). It is predicted that approximately 66% of the world's human population will reside in urban areas by 2050 (United Nations 2015). The number of leptospirosis cases in these areas is expected to rise accordingly (Haake and Levett 2015).

An increasingly globalized population is another contributing factor (Bandara et al. 2014, Day et al. 2016). The rise in international travel and migration has helped facilitate the transfer of *Leptospira* serovars from endemic regions into new, non-endemic regions (Hoffmeister et al. 2010, Leshem et al. 2010, Goris et al. 2013). Additionally, the large-scale movement of livestock and other animals between countries has increased opportunities for the introduction of novel serovars and different host populations into new regions. Since most mammalian species can act as carriers of pathogenic *Leptospira* (Levett 2001, Bharti et al. 2003, WHO 2003), serovars introduced to a new region are likely to find suitable carriers.

Climate change is the third contributing factor. The shift toward a higher mean global temperature has affected precipitation, temperature and humidity patterns (McMichael 2013). The warming climate has been linked to an increase in the frequency and severity of extreme weather events (Field et al. 2012). Outbreaks of leptospirosis often correspond with extreme weather events associated with heavy rainfall, such as hurricanes and flooding (Barcellos and Sabroza 2001, Maskey et al. 2006, Pellizzer et al. 2006). Therefore, an increase in such weather events facilitates an increase in leptospirosis outbreaks. Global temperatures are expected to rise between approximately 1–6 °C by 2100 (McMichael 2013). Wide-scale warmer temperatures can extend the range and time over which leptospires shed into the environment

can remain viable and infectious (Levett 2001, Lau et al. 2010, Chen et al. 2011). Between climate change and the effects of increasing population growth, the global burden of leptospirosis is predicted to escalate (Lau et al. 2010, Chen et al. 2011, Guerra 2013, Haake and Levett 2015).

The predicted increase in leptospirosis cases and potential impact on both public and veterinary health, calls for a surveillance strategy that incorporates serovar typing of clinical cases. Identifying the infecting serovar(s) in clinical cases allows for identification (ID) of the predominant, pathogenic, region-specific, circulating serovars. This data is crucial for the vaccine design, the early detection of outbreaks, identification of infection reservoirs, and the application of outbreak control measures (Guerra 2013).

However, no rapid and efficient tool is currently available for routine serovar typing. The gold standard for serological diagnosis is the Microscopic Agglutination Test (MAT). In this assay, a patient's serum is tested against a select panel of serovar antigens. Test results are reported as the level of titer in the patient's serum against these antigens. Results have often been interpreted as indicating a current or previous infection with a specific serovar. However, the test can only reveal whether a patient has been exposed to a serovar in the same serogroup as the serovar(s) for which the test shows a positive titer. Studies that have examined the MAT as a tool for predicting the infecting serovar found that it accurately suggested the infecting serovar in fewer than 50% of cases (Levett 2003, Smythe et al. 2009). Molecular testing for leptospirosis is most often performed via the polymerase chain reaction (PCR) assay or one of its variants (e.g. real-time PCR) on patient urine samples. PCR assays test for the presence of *Leptospira* DNA (Adler and de la Peña Moctezuma 2010, Rojas et al. 2010). Though these assays offer high sensitivity and early diagnosis prior to seroconversion (Picardeau et al. 2014), the primers used in currently available commercial assays target conserved regions of *Leptospira* DNA and so cannot predict the infecting serovar (Picardeau et al. 2014, Reagan and Sykes 2019). A few methods of post-amplification analysis have been employed to differentiate between select *Leptospira* species or species clusters. However, these techniques were found to be impractical for routine, high-throughput testing (Merien et al. 2005, Perez and Goarant 2010). The lack of serovar identification in positive PCR tests should not have a significant negative impact on treatment decisions for individual patients. Nonetheless, these test results cannot be used for tracking the distribution of *Leptospira* serovars and reservoir hosts, detecting seroprevalence shifts in real-time, or informing vaccine design. Other test that may be used for leptospirosis diagnostics, such as enzyme-linked immunosorbent assays including several point-of-care variants, again offer no serovar identification. Serovar typing is primarily done via whole genome sequencing, which is primarily performed in reference laboratories and cannot be practically applied to routine testing.

Over the last few decades, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF-MS) has become an increasingly popular tool in clinical diagnostic laboratories. Compared to many of the conventional, biochemical tests used in diagnostic microbiology, MALDI-TOF MS offers a more cost-effective, rapid, and high-throughput method of microbial identification

(Seng et al. 2009, Croxatto et al. 2012, Heaton and Patel 2017). As such, it is now commonly used in routine diagnostics to identify pathogenic bacteria, fungi, and yeast with a high degree of accuracy to the genus and species levels (De Carolis et al. 2014b, Heaton and Patel 2017).

Like other types of mass spectrometry (MS), MALDI-TOF produces a characteristic mass spectrum for a sample based on a subset of the sample's ionized molecules. This spectrum represents specific components of the sample and can be used for sample identification. Unlike other mass spectrometric forms, MALDI-TOF employs a matrix solution that is mixed with sample material and aids in the ionization process. The matrix is most often a weak organic acid which can act a proton donor and protects sample molecules from fragmentation during ionization. This allows for the ionization of large, intact biomolecules (Karas et al. 1985, De Carolis et al. 2014b). Generated ions travel through a time-of-flight mass analyzer at a velocity that is inversely proportional to their mass-to-charge (*m/z)* ratios. The ions strike an ion detector at the end of the TOF tube which records the number of ions at each *m/z* and corresponding times-of-flight (Croxatto et al. 2012). This data is then used to create a unique spectrum for the sample. In the analysis of microorganisms, the majority of ions recorded represent the conserved and highly abundant microbial proteins in the sample organism. As such, the resulting sample spectrum serves as a mass spectral "fingerprint" unique to that organism (Carbonnelle et al. 2011). Identification of the organism then proceeds either via comparison of the sample spectrum with reference spectra in the MALDI database, or by comparison of sample spectrum peak masses with those in a reference proteome database (Singhal et al. 2015).

Given the advantages MALDI offers over many conventional tests, and the widespread adoption of MALDI-TOF as a routine diagnostic tool, there has been a growing interest in its potential for pathogen detection directly from patient specimens (Ferreira et al. 2010, Rodriguez-Sanchez et al. 2014). For example, urine samples are one of the main specimen types submitted for microbiological testing. Identification of a urine sample pathogen usually takes 18 to 48 hours due to the required culture and isolation of the organism prior to most tests (Li et al. 2019). Testing a patient's urine sample directly would remove the culture and isolation steps, reducing the amount of time between sample submission and diagnosis. This would allow physicians to prescribe targeted and appropriate antibiotic therapy more quickly, which should translate into faster resolution of patients' infections. Several studies have demonstrated MALDI's ability to accurately identify organisms directly from urine (Kim et al. 2015, Rosselló et al. 2015). Since urine is one of the two most common patient sample types submitted for canine leptospirosis testing, the ability to directly test these specimens via MALDI could result in faster and improved patient care. MALDI pathogen subtyping is another growing area of interest. A growing body of work reports successful identification at a sub-species level, including aerobic Gram-negative bacteria (Seibold et al. 2010, Dieckmann and Malorny 2011), Gram-positive bacteria (Lartigue et al. 2009, Josten et al. 2013), anaerobic Gram-negative bacteria (Mencacci et al. 2013), and *Mycobacterium* spp. (Fangous et al. 2014). Considering the many advantages MALDI offers for routine diagnostics, and its

potential for serovar-level typing directly from patient urine samples, MALDI-TOF MS was explored for canine leptospirosis serovar typing.

The purpose of this thesis is to determine whether whole-cell MALDI-TOF MS can be used as a rapid and cost-effective complementary tool for routine diagnostics and epidemiological surveillance of canine leptospirosis.

Main Hypothesis: Whole-cell MALDI-TOF MS has the specificity and sensitivity to detect and identify *Leptospira* in urine samples of leptospirosis-positive canines.

1.2 AIMS AND OBJECTIVES

Specific Aim 1: Investigate the ability of MALDI-TOF MS to detect and identify selected serovars of *Leptospira interrogans* (Chapter 3).

Working Hypothesis 1: MALDI-TOF MS can be used to identify *Leptospira* isolates to the serovar-level in real-time.

Objective 1: Create a custom Main Spectral Profile reference library for seven selected *Leptospira* serovars.

Objective 2: Test the custom library's specificity in blind-coded trials.

Specific Aim 2: Evaluate whether the presence of urine matrix in *Leptospira* test samples affects serovar identification specificity (Chapter 3).

Working Hypothesis 2: MALDI identification specificity for serovar-spiked urine samples will be the same as that for culture-only samples.

Objective 3: Test MSP library specificity on both *Leptospira* serovar culture samples and serovar-spiked urine samples.

Specific Aim 3: Determine whether MALDI-TOS MS can detect *Leptospira*-positive samples at the concentrations typically seen in positive canine samples submitted to the University of Illinois at Urbana-Champaign's Veterinary Diagnostic Laboratory (Chapter 4).

Working Hypothesis 3: Whole-cell MALDI-TOF MS has the sensitivity to detect *Leptospira* serovars in the urine of clinically affected canines.

Objective 4: Create a standard curve to estimate the concentration of *Leptospira* in samples of unknown concentration.

Objective 5: Identify the average C_T value and corresponding *Leptospira* concentration for qPCR-positive canine leptospirosis cases from a 2.5-year period.

Objective 6: Evaluate MALDI-TOF MS sensitivity for *Leptospira* by testing the custom *Leptospira* MSP library using serial two-fold dilutions of serovar cultures and serovar-spiked canine urine samples.

1.3 THESIS OVERVIEW

Background information on leptospirosis and MALDI-TOF MS is presented in Chapter 2. Experimental work done for this project is described in Chapters 3 and 4. Chapter 3 describes the creation of the *Leptospira* Main Spectral Profile (MSP) database, raw mass spectra and peak analyses used to improve MSP identification accuracy, and blind trials to test the specificity of the MSP database. Chapter 4 discusses MALDI-TOF MS sensitivity trials used to learn whether MALDI can detect *Leptospira* at concentrations typically seen in leptospirosis-positive canine urine samples received by the University of Illinois at Urbana-Champaign's (UIUC) Veterinary Diagnostic Laboratory (VDL) for testing. This thesis concludes with Chapter 5, which reviews the findings of this study, presents reflections on the project, and offers suggestions for future work.

Chapter 2: LITERATURE REVIEW

2.1 LEPTOSPIROSIS – THE DISEASE

Leptospirosis, caused by pathogenic *Leptospira* spirochetes (Xu et al. 2016), is a globally pervasive zoonotic disease that can cause significant morbidity and mortality in both humans and other animals (Bharti et al. 2003, Langston and Heuter 2003, Costa et al. 2015). Its widespread distribution is due in part to the large number of hosts which can serve as reservoirs for infection (Levett 2001, Ko et al. 2009). *Leptospira* have been reported to occur in a wide variety of mammals (Bharti et al. 2003, Sessions and Greene 2004, Mgode et al. 2015). Antibodies against various pathogenic serovars have also been found in reptiles (Lindtner-Knific et al. 2013, Rodrigues et al. 2016, Pérez-Flores et al. 2017) and amphibians (Everard et al. 1988), though it is unclear whether non-mammal species can serve as a reservoir of infection for humans.

Leptospirosis occurs seasonally. In temperate regions, where ambient temperature is the limiting factor in the survival of *Leptospira* outside a host animal, the majority of reported cases occur in late summer and throughout the fall (Goldstein 2010). Generally, few cases are diagnosed in the winter and spring (Harkin and Gartrell 1996). In tropical and sub-tropical regions, where higher temperatures and humidity levels allow *Leptospira* to remain viable outside of a host for longer periods of time, peak incidence usually occurs in the rainy season (Levett 2001). Periods of heavy rainfall cause increased runoff and contribute to flooding, which can spread *Leptospira* contaminated water and soils over large areas. This, in turn, increases opportunities for transmission to naïve hosts (Lau et al. 2010, Muñoz-Zanzi et al. 2014).

Infection with pathogenic *Leptospira* typically begins with the bacteria's entry into a host through skin cuts or abrasions (Levett 2001, Sessions and Greene 2004), exposed mucous membranes or conjunctiva, or via aerosol droplets containing the bacteria. Since the bacteria are dependent on the availability of fresh water in order to remain viable in the environment (Adler and de la Peña Moctezuma 2010), a host must typically come in contact with *Leptospira*-contaminated urine, water, or moist soil to become infected. Pathogenic strains cannot reproduce outside of a host (Mohammed et al. 2011). However, once inside a host, in the presence of optimum growth conditions, they reproduce quickly. If *Leptospira* survive host entry and the initial host immune response, they swiftly move into the bloodstream and lymphatic system and are quickly disseminated throughout the body. This leptospiremic phase generally lasts seven days from the date of infection (Adler and Faine 2006). When the concentration of leptospires in the body reaches a critical threshold, symptoms emerge. With the production of circulating antibodies, leptospires are opsonized, removed by phagocytosis, and the bacteremic phase ends. Damaged tissues may recover, as can occur with the liver and kidneys, though permanent tissue damage may be a complication (Adler and Faine 2006).

In humans, the incubation period for leptospirosis is generally one to two weeks, though it can range from two to 30 days (Guerra 2009). It is thought that most infections are subclinical. However, infection can cause severe disease and even death. Clinical cases are usually biphasic and generally present in either an anicteric or icteric form. Approximately 85-90% of patients develop the anicteric form, which is the milder of the two clinical syndromes and often biphasic (Guerra 2009). In the first (septicemic) phase, which lasts approximately a week, symptoms are often nonspecific and include fever, chills, aseptic meningitis, nausea and vomiting, myalgia, conjunctival suffusion, and headache (Levett 2001, Guerra 2009, Haake and Levett 2015). Patients may also develop a rash. In this phase, leptospires can be isolated from cerebrospinal fluid, blood, and tissues. This phase is followed by a period of 1- to 3-days in which fever subsides and symptoms improve. The second (immune) phase, characterized by leptospiurea, then begins. Anti-leptospiral IgM antibodies appear in the patient's serum. *Leptospira* are eliminated from all areas of the body except for the glomeruli, and perhaps the eyes and brain where the organisms may remain for weeks or months. At this point, some of the earlier clinical symptoms may reappear, and the patient may develop complications such as iridocyclitis and peripheral neuropathy. Overall, the anicteric form of the disease usually resolves within a month (Farr 1995). The icteric form, also known as Weil's syndrome, occurs in around 5%–10% of patients and is more severe with a higher mortality rate. The leptospiremic and immune phases in this form are not as clearly delimited (Farr 1995). Some of the symptoms seen with this form include high fever, jaundice, renal dysfunction, disseminated intravascular coagulation, neurological changes, and hepatic necrosis (Farr 1995, Maroun et al. 2011, Schuller et al. 2015). An additional manifestation, severe pulmonary hemorrhagic syndrome, is being reported with increasing frequency. This syndrome is associated with a high mortality rate. Patients who develop this form experience hypoxemia, pulmonary hemorrhage, alveolar edema and bilateral lung consolidation (Truong and Coburn 2012). While the symptoms described above for each form and phase of the disease have been reported as those most common, a wider range of symptoms have also been reported and mimic those seen in other febrile illnesses, such as influenza, meningitis, and hepatitis (WHO 2003, Musso and La Scola 2013). This can delay treatment and lead to more serious outcomes, including death.

Nonhuman animals with leptospirosis also experience a broad spectrum of symptoms (Ellis 2015). Subclinical illness can occur, most often in an animal infected by a serovar for which it serves as a maintenance host (Lunn 2019). In dogs, the disease can range from subclinical illness to death. The incubation period ranges from 5 to 15 days (van de Maele et al. 2008) and presentation of clinical disease takes one of several forms. In the peracute form, death can occur quickly, often with very few symptoms. Acute (fulminant) leptospirosis presents as the sudden onset of fever with other symptoms including vomiting, anorexia, diarrhea, myalgia, tachypnea, conjunctivitis, aseptic meningitis, dyspnea and potentially, shock (Adler and Faine 2006, van de Maele et al. 2008, Ellis 2015). The deterioration of the patient in this form may occur so quickly that renal and hepatic damage, and subsequent symptoms, do

not have time to manifest. In the subacute form, symptoms can include those seen in the acute form, but at a level that does not result in the same rapid decline. Additional symptoms common to this form include those associated with the progressive decline of renal and liver function, which include vasculitis, chronic hepatitis with icterus, polyuria and polydipsia, and weight loss. Several studies have reported that the severe pulmonary hemorrhagic syndrome seen in humans is also now recognized in dogs (Gendron et al. 2014, Schuller et al. 2015). As in humans, this syndrome leads to a higher rate of mortality in canines compared to other forms of leptospirosis. The mechanism by which this syndrome occurs is still being elucidated but is thought to result from a patient's exaggerated immune response to bacterial toxins or other bacterial components rather than direct infection of the lungs by leptospires. Dogs may also develop chronic infection, which may include symptoms such as chronic interstitial nephritis, uveitis, hepatic fibrosis, hepatic encephalopathy, and ascites (André-Fontaine and Hernandez 2008, Ellis 2015). This type of infection can also be asymptomatic. Chronically infected dogs may act as carriers, periodically shedding leptospires via their urine for extended periods of time and serving as a source of infection for both their owners and other animals (Feigin et al. 1973, Goldstein 2010, Ellis 2015). While antibiotic treatment is available, success of treatment is correlated with early diagnosis and aggressive treatment of the disease (Sykes 2011).

The cycle of leptospirosis infection in nature is perpetuated by chronic infection of the proximal renal tubules in animal maintenance (reservoir) hosts (Haake and Levett 2015). Maintenance hosts are those in which infection with a particular host-adapted serovar is endemic within the species and which typically experience only subacute or mild illness (Blackmore and Hathaway 1979) and limited antibody response (Zuerner 2015). Colonization of the kidneys in maintenance hosts persists because the *Leptospira* are protected from the host immune response in the tubular epithelial cells of the kidneys. Leptospires are then shed via urine either intermittently or continuously, for months or throughout their lifetime (van de Maele et al. 2008, Guerra 2009). The concentration of excreted bacteria can range up to 10^8 leptospires per mL of urine (Adler and Faine 2006). While the majority of maintenance hosts will only develop a mild illness, some carriers may develop chronic renal disease over time (Guerra 2009). Primary reservoir hosts for serovar Autumnalis are mice, and possibly rats (Levett 2001, Adler and Faine 2006). Serovar Bratislava is maintained in horses, pigs, and sheep and may be maintained in skunks, opossums, and mice (Ward et al. 2004, Hensley 2016). Serovar Canicola is maintained by dogs (Levett 2001, Bharti et al. 2003, Ellis 2015). Reservoir hosts for serovar Grippotyphosa include raccoons, voles, skunks, and opossums (Sykes 2014). Serovar Hardjo is maintained cattle and sheep (Ellis 2015). Rats are a known primary reservoir for serovar Icterohaemorrhagiae (Levett 2001, Adler and Faine 2006). Potential maintenance hosts for serovar Pomona include cattle, pigs, horses, opossums, skunks and sea lions (Sykes 2014, Lunn 2019). Reservoir hosts and the serovars they maintain may differ based on geographic region (Hartskeerl and Terpstra 1996, Levett 2001).

Incidental (accidental) hosts for a particular serovar are those which, when infected with that serovar,

experience acute and often severe disease, and potentially even death. Infection is usually acquired via indirect contact with a maintenance host. Animals can be incidental hosts of some serovars, yet maintenance hosts of others (Levett 2001).

2.2 *LEPTOSPIRA* - THE ORGANISM

The causative agents of leptospirosis are pathogenic, Gram-negative spirochetes within the family Leptospiraceae and the genus *Leptospira* (Faine et al. 1999). The name "*Leptospira*" is derived from the Greek leptos, meaning thin and speira, meaning coiled (Chen 2002). Suitably named, these slender, helical organisms with pointed ends have an average diameter of approximately 0.1–0.3 μm, and a length ranging from 6-20 μm (Haake and Zückert 2015, Zuerner 2015). However, recently isolated pathogenic *Leptospira* are commonly shorter and more compact than serially-passaged laboratory strains (Ellis et al. 1983). Leptospires have two endoflagella, each arising from opposite ends of the cell and wrapping around the bacterium in a right-handed helical conformation within the periplasmic space. The rotation of these flagella, and the spiral shape of the cell, allow for *Leptospira*'s characteristic corkscrew movement and the distinctive hook typically seen at one or both ends of the cell (Malmstrom et al. 2009, Evangelista and Coburn 2010). The *Leptospira* genome consists of two circular chromosomes: one approximately 4.3 Mbp and the second, approximately 350 kbp (Nascimento et al. 2004).

Cell membrane architecture is similar to that of many other Gram-negative bacteria. It consists of an inner, cytoplasmic membrane and an outer cell membrane separated by a periplasmic space, which holds a thin layer of peptidoglycan. The inner membrane consists of a phospholipid bilayer, which is closely affiliated with the peptidoglycan (Cullen et al. 2004). The outer membrane (OM) has three layers. The inner two layers consist of a phospholipid bilayer. The third, outermost layer is comprised of lipopolysaccharide (LPS), the major component of this layer, and a variety of structural and functional proteins all anchored in the outer membrane. While most Gram-negative bacteria have LPS in their outer membrane, *Leptospira* are unique among spirochetes in this sense because only leptospires contain LPS; *Treponema* and *Borrelia* do not (Haake and Matsunaga 2010).

Leptospira LPS is similar in composition to that found in many other Gram-negative bacteria (Mohammed et al. 2011). It is made up of three covalently-linked components or regions: the O-antigen polysaccharide, the core (R) polysaccharide, and lipid A (Patra et al. 2015). O-antigen is the bacteria's main antigen and the outermost portion of the LPS molecule that projects away from the bacterium (Cullen et al. 2004). The polymorphic composition of the sugars in this region, both between species and among strains within a species, result in epitopes that contribute to immunological specificity. The O-polysaccharide is attached to the core (R) polysaccharide, which is composed of a short sugar chain. This core is then attached to lipid-A, which secures LPS to the outer membrane (Nahori et al. 2005). LPS has several functions. It serves as a physical barrier to block the entry of most harmful substances into the bacterial cell. It also contributes to bacterial integrity by stabilizing the cell membrane (Bruslind 2017).

Probably the most well-known function of LPS is that of an endotoxin, which activates the host immune response via the alternative complement pathway. Lipid A is the toxic component of LPS, and its composition is highly conserved among most Gram-negative bacteria (Nahori et al. 2005).

However, the LPS of *Leptospira* is unique. It exhibits lower endotoxicity than that of other Gram-negative bacteria. This toxicity difference is thought to be related to some atypical features of the *Leptospira* lipid A moiety (Que-Gewirth et al. 2004), which include fatty acids that differ in length from those generally found in lipid A, an atypical backbone structure, and a single methylated phosphate in place of the more common double unmethylated phosphates. Methylated phosphates are uncommon in biology and until now, have not been found in lipid A (Adler and de la Peña Moctezuma 2010, Haake and Zückert 2015). Most intriguing is the way in which *Leptospira* LPS has been found to trigger the innate immune response in humans. The LPS of Gram-negative bacteria is generally recognized by the Toll-like receptor 4 (TLR4) (Chow et al. 1999), whereas TLR2 typically recognizes components of Gram-positive bacteria, such as peptidoglycan and lipoproteins (Takeuchi et al. 1999). While *Leptospira* is Gram-negative, its LPS is recognized by TLR2, not TLR4 (Chow et al. 1999, Werts et al. 2001).

In addition to LPS, lipoproteins make up a large portion of the proteins found in the outer membrane of pathogenic *Leptospira*. The most abundant of these is LipL32, which is the second most recognized *Leptospira* antigen after LPS (Hauk et al. 2009). It is known to be a hemolysis-associated protein (*Hap-1*), and is used by leptospires to bind fibronectin, collagen and laminin (Haake et al. 2000, Ko et al. 2009). LipL32 induces B and T cell immune responses in mice (Lin et al. 2010). It has also been found to bind to TLR2 during the innate immune response (Hsu et al. 2010). Other lipoproteins present in the OM include LipL21, LipL41, LipL46, LipL53 (Cullen et al. 2003, Matsunaga et al. 2006, Evangelista and Cobern 2010, Haake and Matsunaga 2010, Luo et al. 2010).

Another major group of outer membrane proteins (OMPs) are the leptospiral immunoglobulin-like proteins (Lig), which include LigA, LigB and LigC. These proteins are expressed only during infection (Palaniappan et al. 2005). Like LipL32, Lig proteins are found only in pathogenic *Leptospira* (Matsunaga et al. 2003, Ko et al. 2009). These proteins interact with several complement regulators (Meri et al. 2005, Barbosa et al. 2009) and appear to help leptospires evade the host immune response. Additionally, they are involved in adhesion and invasion of host cells (Murray 2015). A portion of LigA has shown promise as a potential vaccine candidate (Silva et al. 2007, Faisal et al. 2008).

Several other OMPs of interest include Loa22, a lipoprotein known to be necessary for *in vivo* infection. It is also currently being explored for its potential as a vaccine target (Haake and Matsunaga 2010). A transmembrane protein, OmpL1, has also been well-studied. This protein acts as a porin and is known to express epitopes during infection which induce a response in CD4+ T cells, including Th1 cytokine response (Lin et al. 2011).

The variations in *Leptospira* LPS are the basis for serological classification and identification of serovars (Levett 2001, Bharti et al. 2003) which are the basic taxonomic unit of *Leptospira* (Levett 2001,

Ahmed et al. 2006). Prior to 1989, *Leptospira* were divided into two groups. Pathogenic leptospires were grouped together in species *Leptospira interrogans,* while non-pathogenic strains, usually found in the environment and rarely isolated from hosts (Dikken and Kmety 1978), were placed in species *Leptospira biflexa* (Levett 2001). Members in *Leptospira interrogans* were further divided into those that are pathogenic versus intermediately pathogenic based on molecular and biological features (Fouts et al. 2016). Both species were additionally classified into serovars, which were historically determined based on the cross-absorption agglutination test (CAAT). This test uses a homologous antigen to determine differences between serovars based on the epitopes expressed by the O side-chains of *Leptospira* LPS. However, the CAAT is considered complicated and time-consuming. Other serological typing methods that have since been employed include factor analysis (Dikken and Kmety 1978) and the use of monoclonal antibodies (Adler et al. 1989). Currently over 250 pathogenic and more than 60 nonpathogenic serovars have been identified. To help organize the large number of serovars, they have been arranged into serogroups based on the similarity of their surface antigens. Serovars that cross-agglutinate are placed in the same serogroup (Ahmed et al. 2006, Ko et al. 2009, Evangelista and Cobern 2010). Twenty-four serogroups are currently recognized (André-Fontaine 2006, Levett 2015a). While serogroups are not considered a formal taxon, their use in *Leptospira* classification has proven valuable in serological diagnosis and epidemiological studies (Fouts et al. 2016).

More recently, a genotypic classification system has been introduced that uses DNA hybridization analysis to identify genomospecies (Brenner et al. 1999, Levett 2001). Currently, 21 genomospecies are recognized (Bharti et al. 2003, Adler and de la Peña Moctezuma 2010, Smythe et al. 2013). Of these species, seven are responsible for the majority of clinical leptospirosis cases: *L. interrogans, L. borgpetersenii, L. santarosai, L. noguchii, L. weilli, L. kirschneri* and *L. alexander* (Ahmed et al. 2006). Unfortunately, the two classification systems are independent of each other. Both pathogenic and nonpathogenic serovars can occur within the same species and serovars belonging to the same serogroup may be found in multiple genomospecies (Robinson et al. 1982, Brenner et al. 1999). For example, serovars within serogroup Pomona may be found in the genomospecies *L. borgpetersenii, L. noguchii,* and *L. interrogans* (Brenner et al. 1999, Feresu et al. 1999). The same genes for the production of LPS have been found in strains belonging to different genomospecies. This is thought to be due to horizontal gene transfer between the species (Brenner et al. 1999, Feresu et al. 1999, Haake et al. 2004). For this reason, neither an isolate's serogroup nor serovar can predict the species to which it belongs (Levett 2015a).

Leptospira are aerobic organisms. Optimum growth in vitro for pathogenic species occurs at temperatures ranging from 28°C-30°C, but they can also grow at 37°C. Saprophytic leptospires have the same temperature range for optimum growth in vitro as pathogenic species, but will also grow at temperatures as low as 11°C-13°C. All species favor a pH range of 7.2-7.6 (Cameron 2015). Nutritional requirements for growth of pathogenic strains include particular vitamins and nutritional supplements,

sources of nitrogen, and long chain unsaturated fatty acids, which are an essential source of carbon. These fatty acids are also toxic. Therefore, it is necessary to add detoxicating agents, such as sorbitol-complexed fatty acids (Tweens) or albumin, to culture medium (Faine et al. 1999). EMJH culture medium, which is the Johnson and Harris modification (Johnson and Harris 1967) of the Ellinghausen McCullough medium (Ellinghausen and McCullough 1965), is the most frequently used liquid medium for *Leptospira* culture. EMJH is based on polysorbate (Tween), bovine serum albumin, and oleic acid (Cameron 2015). Liquid media is preferred since growth of leptospires on solid media is generally slow, difficult, and unobtainable for some fastidious strains (Cameron 2015). Laboratory strains of pathogenic *Leptospira* have a doubling interval of approximately 6-8 hours, whereas a new culture of freshly collected leptospires has a generation time of roughly 14-18 hours, with initial growth delayed for several days to weeks. Maximum growth for pathogenic leptospires is reached in approximately 4-7 days. Saprophytic strains exhibit faster growth, with a generation time of 3.5-4.5 hours and maximum growth reached in 2-3 days (Adler and de la Peña Moctezuma 2010).

Until recently, a solid agar medium selective for *Leptospira* was not available. The fastidious nature of these organisms makes culture on standard solid agar media difficult, requiring between 10 days to six weeks for a colony to appear (Cameron 2015), if growth is successful. This long incubation period makes dehydration and contamination of cultures a common problem. Due to the time and technical know-how needed to successfully grow *Leptospira* using this method, culture on solid media has been reserved mainly for research (Bharti et al. 2003, Ahmad 2005). This may change with the recent development of Leptospira Vanaporn Wuthiekanun (LVW) agar (Wuthiekanun et al. 2013). This solid medium, selective for *Leptospira*, was designed to facilitate faster *Leptospira* growth. The number of studies that have reported use of this medium is still small. However, with an approximate growth time of seven days after inoculation, use of solid media culture may become more common (Wuthiekanun et al. 2013).

2.3 EPIDEMIOLOGY

In 1886, Dr. Adolf Weil described several cases of a particular form of an icteric fever (Weil 1886) that would later come to be termed Weil's disease. Leptospires would not be discovered as the infectious agent until around 1917 (Levett 2001). Since then, we have learned much about various forms of leptospirosis, considered to be the most widespread zoonotic disease (Lau et al. 2010). However, almost a century later, the accurate diagnosis of leptospirosis and classification of specific *Leptospira* remain difficult, as does determining the global burden of this disease. Though long recognized as a global zoonosis with the potential to cause severe disease, it is now identified as a re-emerging disease of global public health importance (WHO 2003, Meites et al 2004, Vijayachari et al. 2008) and was reinstated by the Centers for Disease Control and Prevention (CDC) as a nationally notifiable disease in 2013 (CDC 2020). This new characterization is due to a changing epidemiology which includes an

upsurge in the number of reported cases in both humans and non-human animals worldwide (Meites et al. 2004, Vijayachari et al. 2008).

From a historical perspective, leptospirosis has been primarily reported in tropical and subtropical nations; mainly the developing world (WHO 2003). However, the number of cases reported in temperate regions and in developed nations has been markedly increasing (Bharti et al. 2003, WHO 2003, Lau et al. 2010). This trend can be partially attributed to a growing global human population. This population growth contributes both to urban sprawl and higher population densities. Crowded conditions and the push into previously wild habitats cleared for new construction or agriculture increases the transmission opportunities by increasing the potential for contact with *Leptospira* reservoirs and *Leptospira*-contaminated environments (Guerra 2013). The rising pace of globalization has accelerated the international movement of both human and animal populations. Travel of infected people and animals can introduce *Leptospira* species and animal hosts into new areas. Current and future military conflicts, social upheavals and economic pressures could further expand the distribution of leptospirosis into regions that have been relatively free of this disease (Wallace et al. 2002, Bandara et al. 2014). The rising popularity of adventure sports and races that involve aquatic activities has also been associated with an increase in outbreaks in developed countries. Several outbreaks have been associated with triathlons, where the swimming portion of the race was conducted in a river or lake contaminated with leptospires (Morgan et al. 1998, Brockmann et al. 2006, Radl et al. 2011).

The incidence of leptospirosis is expected to further increase with the predicted rise in global temperatures and associated extreme weather events (Lau et al. 2010). These events include hurricanes, typhoons, and increased flooding. Heavy rainfall and floods are often associated with leptospirosis outbreaks because they increase contact opportunities between humans and domestic animals with *Leptospira*-contaminated surface waters and reservoir hosts (Barcellos and Sabroza 2001, Maskey et al. 2006, Pellizzer et al. 2006, Kawaguchi et al. 2008). Rising temperatures could boost *Leptospira*'s survivability outside of a host in areas where historically lower temperatures had been unfavorable for survival, increasing the time excreted leptospires remain infectious in these areas. In addition, increasing temperatures may expand suitable habitats for reservoir species, further extending *Leptospira*'s potential reach (Chen et al. 2011).

Epidemiological changes in canine leptospirosis has included not only an increase in the number of cases, but also a shift in seroprevalence. Rising case numbers have been reported in the US, Canada, (Rentko et al. 1992, Sykes 2001, Prescott et al. 2002, Ward et al. 2002), Asia Pacific (Victoriano et al. 2009) and parts of Europe (Majetic et al. 2014, Major et al. 2014, Pijnacker et al. 2016). Before 1970, the serovars usually associated with canine clinical disease in the US were Icterohaemorrhagiae and Canicola (Alexander et al. 1957, Hubbert and Shotts 1966). In the early 1970s, a bivalent vaccine against these two serovars was introduced in the US. Following the vaccine's release, the number of reported cases attributed to these serovars decreased (Sykes 2001, Sykes et al. 2011). Around 1990, the serovars

most often associated with canine disease started to change (Rentko et al. 1992, Ward et al. 2004). Increasing numbers of cases were attributed to serovars Grippotyphosa, Bratislava, and Pomona. Case prevalence began to rise once more (Prescott 2002, Ward et al. 2002). Pomona and Grippotyphosa became the predominant serovars linked to canine disease (Prescott et al. 2002, Ward et al. 2004). These serovars do not belong to the same serogroups as the two serovars in the bivalent vaccine. Protection conferred by the bacterins included in the vaccine was generally thought to be serovar/serogroup specific. This meant that dogs that had received the bivalent vaccine were not protected against Pomona, Grippotyphosa, or Bratislava (Ward et al. 2004). This prompted the introduction of a tetravalent canine vaccine in 2001, containing bacterins of serovars Grippotyphosa and Pomona in addition to Canicola and Icterohaemorrhagiae (Alton et al. 2009). Unfortunately, canine leptospirosis cases in the United States have continued to rise (Sessions and Greene 2004, Moore et al. 2006).

In the US, there has been some debate as to whether serovar Autumnalis is responsible for some canine cases (Prescott et al. 2002). In several studies that used the MAT to test canine serum for leptospirosis antibodies, serovar Autumnalis had the highest titer (Moore et al. 2006, Alton et al. 2009, Gautam et al. 2010). However, it has been suggested that Autumnalis cross-reacts considerably with other serovars (Prescott et al. 2002, Moore et al. 2006), and that canine Autumnalis-positive cases should most likely be interpreted as a cross-reaction, rather than evidence that Autumnalis is actually the infecting serovar (Alton et al. 2009).

Could the increase in reported cases be due to a greater awareness of canine leptospirosis among veterinarians? This would increase the likelihood of leptospirosis testing when patient history and symptoms make leptospirosis a differential diagnosis. Though this probably plays a role, it does not appear that increased testing and reporting alone account for the increased number of cases. A retrospective study conducted by Alton et al. (2009) applied the Cochran-Armitage test for trends in proportions to examine the records of 1406 dogs submitted for leptospirosis testing in Ontario. The study found an increasing proportion of leptospirosis infections over an eight-year period, which reflected similar findings in other studies that described an upward trend in prevalence versus increased reporting by veterinarians (Ward et al. 2002, Langston and Heuter 2003, Meites et al. 2004).

Given the predicted increase in cases and changing seroprevalence of canine leptospirosis, it is imperative that regular canine leptospirosis case surveillance be established. Identification of the infecting serovar in clinical cases is necessary for an effective surveillance strategy . Pathogenic serovars are usually adapted to specific animal maintenance hosts (World Organisation for Animal Health (OIE) 2018). Consequently, identification of the infecting serovar can help determine the source of infection. This knowledge can then be used to determine the best strategy for containment and eradication of (Guerra 2013). Serovar identification is also crucial in leptospirosis vaccine design. For a leptospirosis vaccine to

be efficacious, it must protect against the circulating pathogenic serovars that are predominant in the region in which the vaccine will be used. This requires knowledge of the regionally-predominant serovars.

Currently, there is no relatively affordable, rapid and easy-to-implement test that can accurately identify the infecting serovars in clinical cases. Most previous epidemiological studies used the Microscopic Agglutination Test (MAT) for serovar identification. The MAT is the serological reference test for leptospirosis. However, it can only suggest a serogroup with which an animal has come into contact; it cannot identify the infecting serovar. Quantitative Real-Time Polymerase Chain Reaction (qPCR) is the second most commonly used test for leptospirosis diagnosis. Yet, there is no qPCR assay that can differentiate between *Leptospira* serovars at the serovar-level. Since there is currently no method for routine leptospirosis diagnostic testing that can also identify the infecting serovar, there is a need for new methods that can be employed to this end.

2.4 DIAGNOSTICS

Diagnosis of leptospirosis remains a challenge due to, among other factors, the extreme variability in clinical presentation between infected patients (Bharti et al. 2003, Levett 2001). Diagnosis is further complicated by the variability in the time periods during which patients exhibit leptospiremia, leptospiurea and then seroconversion, since diagnostic tests rely on the temporal stage of disease and sample types collected (Ahmad et al. 2005, Musso and La Scola 2013). The incubation period generally lasts from 5 to 14 days but can range from 2-30 days (WHO 2012). The length of this period depends on several factors, including host immune response, infective dose, and virulence of the strain (Sykes et al. 2011). Since the date of infection is rarely known, and the incubation period can vary widely between patients, veterinarians must take a best guess approach in deciding which samples and tests are most appropriate for a particular case. Thus, certain test results taken alone are not definitive (Greenlee et al 2005).

There are three categories of leptospirosis diagnostic tests: those that display the presence of the leptospiral pathogens directly in culture, serological tests that detect anti-leptospiral antibodies, and molecular tests that detect the presence of leptospiral DNA.

Leptospira isolation from an infected patient is considered the definitive diagnosis for leptospirosis (Adler and de la Peña Moctezuma 2010, Goldstein 2010). Leptospires may be isolated for culture from blood samples taken during the first week of illness (Ahmad et al. 2005). Cerebral spinal fluid and dialysate may also harbor leptospires during this stage, and so may likewise be collected for culture. Leptospirurea begins approximately the second week of illness. The amount of time during which leptospires are excreted in the urine varies but can range up to several weeks. Urine is the recommended sample to collect for culture during this period. Some studies have found that urine collected between 14-28 days after infection exhibits the highest concentration of excreted leptospires and; therefore, is most likely to show growth in positive cases when cultured (Bharti et al. 2003). Post-mortem tissue specimens may be collected in fatal cases, with kidney, liver, and brain tissues recommended for culture (Ellis 2015).

As mentioned previously, use of culture on solid or semi-solid media for routine diagnosis of leptospirosis is not a clinically-feasible approach because the bacteria are difficult to culture and positive results may take between 7 days to 13 weeks, depending the strain, culture medium, and inoculation size used (Adler and de la Peña Moctezuma 2010, Cameron 2015). Thus, this method can offer confirmation of diagnosis in retrospect, but cannot guide early disease management decisions (Merien et al. 1992). Additional challenges of *Leptospira* culture include the possible inhibitory effects of blood and urine sample components on leptospiral growth in vitro, contamination with host flora, and false negative results (Cameron 2015).

The most commonly used diagnostic test for leptospirosis is the Microscopic Agglutination Test (MAT) (WHO 2003, Sykes et al. 2011). The MAT was developed shortly after *Leptospira* was first isolated over a century ago (Martin and Pettit 1918). Considered the “gold standard” for *Leptospira* serological testing, the MAT uses a panel of live *Leptospira* serovars to test serial dilutions of patient sera for IgM and IgG agglutinating antibody reactions to *Leptospira* LPS (Ahmad et al. 2005, Miller et al. 2011). Patient sera are incubated with whole cell antigen, after which this mixture is observed using dark-field microscopy for evidence of *Leptospira* agglutination. Positive reactions are attributed to those serovars and dilutions for which at least 50% agglutination occurs (WHO 2003, Sykes et al. 2011). While antibody levels usually become detectable via MAT approximately 7-10 days after the onset of symptoms (Levett 2001, André-Fontaine 2006), some cases have had detectible antibodies as early as two days and as late as four weeks after signs of infection first appear. Due to the temporal variability in antibody production, The American College of Veterinary Internal Medicine Leptospirosis Consensus Statement advises that both acute and convalescent titers be tested, with acute and convalescent samples collected two to four weeks apart (Sykes et al. 2011). The definitive criteria for a diagnosis of leptospirosis by the MAT is seroconversion or a ≧ 4-fold rise in antibody titer against the same serovar between acute and convalescent sera (Levett 2001, Sykes et al. 2011, Greene et al. 2012). A ≧ 4-fold increase in titer is thought to indicate a recent infection (Miller et al. 2011, Sykes et al. 2011). Some dog owners may be disinclined or unable to take their dog in for collection of a convalescent serum sample. In these cases, a presumptive diagnosis of acute leptospirosis can be made with a single positive titer ≧1:800 if the animal also exhibits clinical signs suggestive of leptospirosis (Levett 2001, Sessions and Greene 2004). In such a case, the diagnosis would be presumptive rather than confirmed because a single elevated titer of ≧ 1:800 in dogs can also indicate recent leptospirosis vaccination or prior *Leptospira* exposure (Greene et al. 2012). Additionally, a single negative titer cannot confirm absence of infection because the negative result could be due to patient serum collected prior to seroconversion. The threshold titer for a positive diagnosis of active infection is dependent on the endemicity of leptospirosis in the population. In non-endemic areas, a lower titer may be considered diagnostic whereas in endemic areas, a higher titer may be required to indicate active infection. (Levett 2001, CDC 2020). To account for potential differences in MAT results between vaccinated and non-vaccinated dogs, it was previously suggested that a different

titer threshold be applied for each group when using a single titer for presumptive diagnosis of an active infection. A threshold of ≥ 1:1,600 would be applied to vaccinated dogs with an elevated titer to a MAT panel serovar for which they had been previously inoculated. Non-vaccinated dogs would have a titer threshold of ≥ 1:800 (Langston and Heuter 2003, Miller et al. 2011, Sykes et al. 2011). However, this recommendation is not widely applied. The MAT does not distinguish between vaccination and natural exposure. While post-vaccinal titers in dogs are typically ≤ 1:400, lower than the titer threshold which suggests active infection, some dogs have been recorded to have post-vaccinal titers of ≥ 1:1600 (van de Maele et al. 2008).

The MAT antigen test panel is comprised of serovars that represent different serogroups. Diagnostic laboratories select serovars for panel inclusion based on those thought to be predominant in the relevant geographic region. In the US, the most frequently used serovars in veterinary MAT panels are Autumnalis, Bratislava, Canicola, Grippotyphosa, Hardjo, Icterohaemorrhagiae and Pomona (Sessions and Greene 2004, Davis et al. 2008). The panel should include serovars that represent all serogroups present in a region so that the test will recognize all positive samples that contain anti-*Leptospira* agglutinating antibodies. Since antibodies against serovars that belong to the same serogroup are thought to cross-react, having one representative serovar for each serogroup of interest theoretically should detect samples infected with a serovar absent from the MAT panel but in the same serogroup as another serovar present in the panel. Therefore, the MAT is not a serovar-specific test (Levett 2001), though results are often reported as such. Due to non-serovar-specific shared antigens (Levett 2001), some cross-reactivity between different serogroups also occurs and can potentially be seen as positive reactions to multiple serovar antigens in the MAT. These cross reactions are thought to occur most often during the first six weeks of illness (Miller et al. 2011). Traditionally, the infecting serogroup was thought to be that represented by the serovar with the highest titer. However, studies have shown that there are exceptions to this interpretation. For example, one study found that the MAT accurately predicted the infecting serovar in less than 50% of human cases confirmed by culture (Levett 2003). Frequently, more than one serogroup is tied for the highest titer in the panel, even though the patient is actually infected with only one serovar (Miller et al. 2011). Concurrent infections with more than one serovar do, however, occur. Additionally, “paradoxical reactions,” in which a non-infecting serogroup has a higher, cross reactive titer than the infecting serovar, commonly occurs, particularly in the early stage of infection (Levett 2001). Furthermore, in some cases of canine leptospirosis, MAT serovars with positive titers have been shown to change over the course of infection (Miller et al. 2011). Previous vaccination against leptospirosis may also affect MAT serovar reactivity (Barr et al. 2005). Vaccinated dogs have also been shown to develop the highest titers to serogroups not included in the vaccine. For example, in a study by Barr et al. (2005), dogs inoculated with a vaccine containing bacterins of serovars Grippotyphosa and Pomona developed highest MAT titers to serogroup Autumnalis (Barr et al. 2005). Titer levels due to vaccination usually diminish within three months (Barr et al. 2005, van de Maele et al. 2008), but may be

present even after 12 months for some serovars (Klaasen et al. 2003). In animals that are chronic carriers, antibody titers may fall so low as to be undetectable by the MAT (OIE 2018).

Considering that the MAT relies on the production of patient agglutinating antibodies that are generally not detectable by the MAT until approximately a week post onset of symptoms (Cerqueira and Picardeau 2009), this test is unsuitable for early diagnosis. However, it is important to obtain a diagnosis as early as possible so that appropriate treatment can be started. The longer the infection progresses without treatment, the greater the potential for severe disease (Guerra 2013). Given the potential severity of infection and the possibility of inadvertent spread to naïve humans and animals, early diagnosis is important (Miotto et al. 2018).

Additional serological tests developed for the diagnosis of leptospirosis include various enzyme-linked immunosorbent assays (Adler and de la Peña 2009, Loureiro et al. 2014), the indirect hemagglutination assay (Sulzer et al. 1975), the latex agglutination test (Ramadass et al. 1999), the indirect immunofluorescence method to detect IgG, IgM and IgA antibodies (Appassakij et al. 1995) and the lateral flow assay for IgM antibodies (Smits et al. 2001). While some have proved useful for specific studies, none have so far replaced the MAT as the reference standard for serological testing. Additionally, these tests also lack the ability to identify the infecting serovar (Picardeau 2013).

The Polymerase Chain Reaction (PCR) is the second most common method used for diagnosis of leptospirosis. Unlike the MAT, this and other molecular methods are used to identify leptospires directly in patient samples by detecting the presence of bacterial DNA. Since these methods do not rely on a patient's production of anti-leptospiral antibodies, they can be used for earlier disease detection, prior to seroconversion (Haake and Levett 2015). PCR can be performed on blood sampled during the acute leptospiremic phase (Greenlee et al. 2005), which lasts for approximately a week after the onset of symptoms (Picardeau 2013). With the production of antibodies, the majority of leptospires are cleared from the blood and most tissues and accumulate in immunologically privileged sites such as the proximal renal tubules. The patient then enters a leptospiurea state, during which the organism is shed in the urine, and urine samples can be used for PCR diagnosis (Goldstein 2010, Ellis 2015). Since the 1990s, a variety of PCR assays have been designed for the diagnosis of leptospirosis (Adler and de la Peña Moctezuma 2010). These assays amplify target DNA from either a highly conserved gene, such as *rrs* (16s rRNA gene), or genes found only in pathogenic leptospires, like *lipL32* (Smythe et al. 2002). PCR assays have been found to have a higher sensitivity in detecting leptospires in clinical samples compared to culture and MAT (Merien et al. 1992, Brown et al. 1995, Wagenaar et al. 2000). Quantitative PCR (qPCR) assays are now generally preferred for their faster turnaround time and greater sensitivity (Ahmed et al. 2009, Picardeau 2013).

Quantitative PCR combines the PCR assay with fluorescent reporter molecules to measure the amplification of template product after each cycle of the reaction (Navarro et al. 2015). The intensity of fluorescence reflects the quantity of amplicons in the sample (Kralik and Ricchi 2017). The two most

commonly used fluorescence chemistries in *Leptospira* qPCR assays are the fluorescently labelled TaqMan probe (Navarro et al. 2015) and SYBR Green I fluorescent dye (Levett et al. 2005, Merien et al. 2005, Slack et al. 2006b). In 2005, Monis et al. developed SYTO9, which is a double-stranded DNA intercalating dye that may be used in place of SYBR Green I.

As in PCR, the qPCR assays can be divided into those which detect genes present in all *Leptospira*, including *rrs* (Smythe et al. 2002, Slack et al. 2007), *gyrB* (Slack et al. 2006b, Subharat et al. 2011), and *secY* (Ahmed et al. 2009), and those that detect genes present only in pathogenic *Leptospira,* namely *lipL32* (Levett et al. 2005, Rojas et al. 2010, Villumsen et al. 2012) and *ligA* and *B* (Palaniappan et al. 2005). The limit of detection (LOD) for qPCR assays using human blood/serum ranges from two to 30 genome copies per reaction (Smythe et al. 2002, Levett et al. 2005, Palaniappan et al. 2005, Slack et al. 2007, Ahmed et al. 2009). For assays using human urine samples, the detection limit is roughly 10 genome copies per reaction (Villumsen et al. 2012). Only some of the developed assays have been validated for diagnostic use. Those validated in animals include the Subharat et al. (2011) assay targeting *Leptospira gyrB* in kidney and urine samples collected from deer and the assay developed by Rojas et al. (2010) targeting an amplicon within *Leptospira lipL32* in canine urine samples.

A major factor to keep in mind when using qPCR to test for leptospirosis is that the tested sample must contain a sufficient number of bacteria to provide a positive result. For blood samples, there is only a short window of time within which a sample will contain an adequate number of leptospires (Adler and Faine 2006). While the number of bacteria required for a positive result in urine samples is relatively low compared to that which is usually shed during the leptospiurea phase, the organisms tend to be shed in the urine intermittently by incidental hosts. Thus, there is the potential for false negative results (Greene et al. 2012). Additionally, most PCR-based leptospirosis assays cannot identify the infecting serovar (Picardeau 2013). Methods designed to circumvent this inability include melting curve analysis of amplicons (Merien et al. 2005) direct sequencing of PCR products (Perez and Goarant 2010) and restriction fragment length polymorphism analysis of PCR-amplified DNA (Savio et al. 1994, Brown and Levett 1997). Additional molecular tests for the diagnosis of leptospirosis have been developed, including ribotyping (Perolat et al. 1994) nested PCR (Jouglard et al. 2006) multiplex PCR (Ahmed SA et al. 2012) and loop-mediated isothermal amplification (Koizumi et al. 2012). While useful in research, these methods cannot be used for rapid identification of the infecting serovar at the time of diagnosis.

As additional *Leptospira* genomes are sequenced, validated qPCR assays that allow for serovar differentiation become a greater possibility. One promising assay developed by Cai and colleagues (2010) targets the O-antigen gene (*rfb*) cluster. Studies have found that the epitope for serovar specificity is the LPS O-antigen (Kalambaheti et al. 1999). Variation in the structure of the O-antigen is based on structure variation of the O-antigen gene cluster (de la Peña Moctezuma et al. 1999, de la Peña Moctezuma et al. 2001). Examination of the O-antigen gene clusters of selected strains belonging to several serogroups found that this gene cluster was not conserved, particularly in the 5'-proximal end (He

et al. 2007). Cai et al. (2010) identified six O-antigen specific genes for serogroups Autumnalis, Canicola, Grippotyphosa, Hebdomadis, Icterohaemorrhagiae and Sejroe. They then developed six PCR-based assays using different primer pairs for each tested serogroup. Testing of 40 clinical isolates resulted in PCR products of expected size for each, with the exception of four reference strains belonging to serogroup Sejroe. Cai and colleagues concluded that O-antigen specific genes can be used to differentiate at least some serogroups. However, this is still a research-only technique.

There is clearly a need for a tool that can be used for both leptospirosis diagnostics and identification of the infecting serovar. Recent studies have demonstrated the potential for MALDI-TOF MS to identify bacteria at a subtype level (Seibold et al. 2010, Eddabra et al. 2012, Josten et al. 2013, AlMasoud et al. 2014).

2.5 MALDI-TOF MS

2.5.1 Overview of Mass Spectrometry in Microbial Identification

Mass spectrometry (MS) is an analytical technique that measures the mass-to-charge ratio (*m/z*) of ions generated from a sample to determine its chemical composition, identity, and quantity (Rockwood et al. 2018). Older mass spectrometry systems generally used electron beams for analyte ionization of chemical compounds. In addition to generating ions, this method regularly led to ion fragmentation, and so was characterized as a “hard ionization” technique. Molecules in the analyte were identified by associating known masses with fragment patterns (Hosseini and Martinez-Chapa 2017). While this method was commonly used to analyze small organic molecules, it was not appropriate for evaluating larger biomolecules such as proteins, which could be used to identify microorganisms. A method which could sublimate analyte molecules without destroying them was needed (Hansell 2015). In the mid-1980’s, Karas et al. (1985, 1987) found that certain peptides could be analyzed via MS by embedding the sample in a matrix material and using a laser ionization source. The matrix protected against molecule fragmentation during ionization. In 1987, Tanaka and colleagues devised their own version of this method and showed that large biomolecules could be ionized without fragmentation using the right combination of matrix and laser wavelength (Markides and Gräslund 2002). This technique was characterized as a “soft ionization” method (De Carolis et al. 2014b, Hosseini and Martinez-Chapa 2017) and became known as matrix-assisted laser desorption ionization time-of-flight mass spectrometry (Karas et al. 1985). In 1996, Holland and colleagues reported the accurate identification of bacterial species by comparing MALDI-generated unique spectrum profiles with those of reference standards. This spurred a wave of studies testing MALDI’s ability to differentiate bacteria genera and species (Krishnamurthy and Ross 1996, Vargha et al. 2006, Clark AE et al. 2013, AlMasoud et al. 2014). Studies have shown the ability of whole-cell MALDI to identify most Gram-negative and Gram-positive bacteria, as well as some mycobacteria (Fangous et al. 2014), and some fungi (De Carolis et al. 2014a, Lacroix et al. 2014). MALDI has also shown promise in virus identification (Cobo 2013, Calderaro et al. 2014a). Moreover, MALDI has

successfully identified bacteria below the species level, which illustrates its potential for use in taxonomic characterization and epidemiological research (Croxatto 2012).

2.5.2 MALDI-TOF MS Workflow

In the last few decades, MALDI has become an increasingly important tool in clinical diagnostic laboratories for routine microorganism identification. This is in part due to the relatively rapid turnaround time for test results compared to that of many conventional differentiation methods (Croxatto et al. 2012, Heaton and Patel 2017). Depending on the number of samples to be tested, the time required for sample analysis, ranges from approximately three minutes for one sample, (including laser and vacuum adjustment prior to sample ionization), to 30 minutes for 96 samples, which is the maximum number of samples that can be tested on a ground-steel MALDI target. Faster identification of infecting organisms allows for earlier treatment with appropriate antimicrobials and can improve patient outcomes (Ge et al. 2017). MALDI sample preparation is generally simple, requiring few steps and less prior knowledge of the sample microorganism than many conventional laboratory tests (Heaton and Patel 2017). Also attractive is MALDI's potential for pathogen identification directly from clinical samples without prior culture and isolation of the infecting organisms (De Carolis et al. 2014b). While the initial outlay of cost for the mass spectrometer is high, it is comparable to the investment required for other laboratory equipment (Cobo 2013, Patel 2015). Numerous reports have documented the dramatic decrease in annual laboratory costs due to the reduction in the number of conventional tests and related consumables required for routine diagnostics once MALDI was implemented into the routine diagnostic workflow. It was also found that fewer isolates required 16S ribosomal RNA gene sequencing once MALDI was in use (Seng et al. 2009, Heaton and Patel 2017). As additional species and strains are added to MALDI commercial spectrum libraries, and new techniques and analyses are developed that increase MALDI's discriminatory power, MALDI's use is expected to grow.

MALDI uses a matrix solution that is mixed with the sample to be analyzed. The matrix aids the ionization of sample proteins while protecting the sample from fragmentation and decomposition during the ionization process (Clark AE et al. 2013, Heaton and Patel 2017). Variations of MALDI sample preparation are used depending on the type of sample to be analyzed and the aim of the analysis. In general, the test sample and matrix are mixed together and allowed to co-crystallize on a MALDI target, which is often a metal conductive plate (Croxatto et al. 2012). The target plate is then inserted into the MALDI ionization chamber. A pulsed laser beam, which serves as the ionization source, irradiates the co-crystallized mixture. The matrix absorbs most of the laser's photon energy and in this way, protects the analyte from direct laser damage. The energy absorbed by the matrix leads to the excitation and subsequent sublimation of matrix molecules. Sample molecules are then ionized via proton acquisition during random collisions with matrix ions (Wu and Odom 1998, Hosseini and Martinez-Chapa 2017). While MALDI has been known to generate multiply-charged ions, the typical charge is +1 (Heaton and

Patel 2017, Rockwood et al. 2018). The newly generated ions are accelerated to a time-of-flight (TOF) mass analyzer. This analyzer is a flight (or drift) tube maintained under vacuum through which the ions travel until they collide with an ion detector at the far end. The analyzer separates the ions during their flight based on their velocity, which is inversely proportional to their *m/z* (Heaton and Patel 2017). The "*m*" in *m/z* is the molecular mass of an ion measured in Daltons (Da). Lighter ions travel faster and so reach the detector prior to heavier ions (Mellmann and Muthing 2013). The "*z*" in *m/z* represents the number of charges on the ion in terms of its absolute value (Rockwood et a. 2018). As an ion's charge increases, its *m/z* will be reduced, resulting in a shorter time of flight. The ion detector records the time of flight and signal intensity for each ion, which indicates the relative abundance for each ion at a particular *m/z* (Croxatto et al. 2012). This data is then used to create a mass spectrum for the sample in the form of a graph with ion signal intensities plotted on the y-axis and corresponding ion *m/z* values on the x-axis. In whole-cell MALDI-TOF, spectral peaks are usually recorded in the range of 2,000–20,000 Da (Mellmann and Muthing 2012). Generated spectra act as unique, characteristic profiles for each test isolate (Croxatto et al. 2012, Mellmann and Muthing 2012). Generally, the pattern of distinctive peaks and intensities in these spectra can be used to identify genera, species, and sometimes, strains (Rockwood et al. 2018). Spectra are compared to a library (database) of reference spectra representing well-defined microorganisms (Sauer et al. 2008, Carbonnelle et al. 2011, Heaton and Patel 2017). For each sample spectrum, a score is then generated which indicates the probability that the test isolate matches a known organism in the reference database. MALDI manufacturers each set a possible range of scores separated by threshold values indicating identification at the species level, genus level, or no reliable match (Sauer et al. 2008, De Carolis et al. 2014b).

Proteins detected by whole-cell MALDI are those generally with housekeeping functions that ionize readily, have a low mass and are highly abundant in the microorganism (De Carolis et al. 2014b)**.** Ribosomal subunit proteins are the predominant proteins detected; however, DNA-binding proteins, cold-shock proteins, carbon storage regulators, ribosome modulation factors, and translation initiation factors are also known to contribute to sample mass spectra peaks (Ryzhov and Fenselau 2001, Diekmann and Malorny 2011). While non-protein molecules are also ionized in MALDI, they do not contribute to the generation of mass spectra because their *m/z* falls outside the range detected (Rockwood et al. 2018).

Several factors can affect the quality and type of sample spectra produced and the ability to identify isolates. One factor is matrix choice. A variety of matrices are available, each better suited for different sample types and analytic objectives. The choice of matrix solvents must also be considered. These are involved in the extraction of intracellular proteins and the incorporation of bacterial sample cells within the matrix during the co-crystallization process. Poor crystallization can affect the efficiency of sample ionization (Williams et al. 2003, Vargha et al. 2006). The number of reference spectra in the database can also contribute to MALDI's ability to successfully identify samples. With more spectra representing more organisms, a test isolate is more likely to be identified (De Carolis et al. 2014b). Manufacturers of MALDI

instruments, such as Bruker Daltonics (Bremen, Germany) which produces MALDI Biotyper systems and bioMérieux (Marcy l'Etoile, France) which produces the VITEK® MS system, regularly release reference spectra library updates to increase the number of microorganisms represented in the database. Additionally, some research use only MALDI models allow for the addition of user-created reference spectra. Other factors, no less important, include sample preparation method and the concentration of microorganism in the sample (Williams et al. 2003, Šedo et al. 2011).

2.5.3 Potential for MALDI-TOS MS *Leptospira* Subtyping

The increasing number of cases and changing seroprevalence of canine leptospirosis, along with the zoonotic transmission potential, clearly calls for a rapid and inexpensive serovar typing method that can be applied during routine canine leptospirosis diagnostics. A wide variety of subtyping techniques are available. However, these techniques are generally too costly, time-consuming and laborious to incorporate into the daily diagnostic workflow for the purpose of serovar typing and surveillance (van Belkum et al. 2007, Cerqueira and Picardeau 2009). Given the widespread adoption of MALDI in clinical diagnostic laboratories for routine microbial identification, its potential as a tool for bacterial subtyping has been increasingly explored (Clark AE et al. 2013). Differences in isolate mass spectral profiles have successfully discriminated strains of *Salmonellae* (Dieckmann and Malorny 2011), *Streptococcus agalactiae* (Lartigue et al. 2009), *Staphylococci* (Josten et al. 2013), *Francisella* species (Seibold et al. 2010), and *Acinetobacter baumannii* (Mencacci et al. 2013), among others.

A subtyping technique that could be reasonably integrated into routine diagnostics would need to be cost-effective, easy to perform, repeatable, reproducible, rapid and return results that are easy to interpret (van Belkum et al. 2007). The costs associated with MALDI subtyping, examined in several studies, was found to be competitive with other typing methods (Egli et al. 2015). Approximately two decades of using MALDI for identifying bacteria at the genus and species levels have shown its potential applicability in routine sub-species level identification. The ability to incorporate infecting agent identification into standard routine diagnostic testing offers an opportunity for early outbreak detection and ongoing epidemiological surveillance without added costs or time for additional tests (Spinali et al. 2015).

Several studies have explored MALDI-TOF MS as a tool for *Leptospira* serovar identification. In all studies, the authors created *Leptospira* MSPs in-house, which were then combined into a reference database for testing. Djelouadji *et al*. (2012) created an MSP database representing 19 *Leptospira* species. In a blind-coded trial of the database, all isolates were correctly identified to the species level, except for *L. interrogans* serovar Autumnalis, which was misidentified as *L. kirschneri*. In studies performed by Rettinger *et al.* (2012) and Calderaro et al. (2014b), testing of the custom databases again allowed for identification to the species level. Using Bruker MALDI-TOS MS software to perform additional analyses on reference strain mass spectra, the authors of both studies identified discriminating peaks or peak groups that they suggested might serve as characteristic markers for some of the tested serovars.

While these markers may help to discriminate between strains of these serovars using these same software tools, neither study reported serovar-level identification in real time. In 2018, Karcher et al. created an in-house *Leptospira* MSP database using 31 reference strains. The database was tested using 22 Brazilian *Leptospir*a field isolates, previously identified to the species-level by 16S rRNA sequencing. The authors reported that all test isolates of *L. biflexa, L. interrogans* and *L. santarosai* run against the database returned correct identifications to the species-level. They reported poor differentiation between *L. interrogans* and *L. kirschneri but* did not offer further information regarding database trial results for *L. kirschneri* isolates. Selected serovars of *L. interrogans* and *L. kirschneri* were further analyzed with Bruker's ClinProTools software and IBM® Statistical Package for the Social Sciences software (SPSS) to learn whether specific peak patterns in the mass spectra of these serovars could differentiate between the two species. The authors identified a peak, at approximately 8057 Da, that had a higher peak intensity in the raw spectra of *L. interrogans* serovars relative to that found in *L. kirshneri* serovars. The authors concluded that *Leptospira* species can be distinguished by MALDI using analytical tools included in Bruker's MALDI-TOF MS software. While the studies above used *Leptospira* reference strains from stock cultures and clinical isolates, none used urine samples, which would represent one of the two main types of canine samples submitted for leptospirosis testing. A recent study by Sonthayanon et al. (2019) tested both *Leptospira* isolates from culture and *Leptospira*-spiked urine samples. Both sample types were tested against a custom-created *Leptospira* MSP database. The authors reported that 96 out of 97 samples were correctly identified to the species-level. This study showed that MALDI could be used to detect and identify species of *Leptospira* in urine samples. As in earlier studies, further spectral analysis using ClinProTools found peak patterns that might be unique for some of the serovars used in the study. None of these studies reported accurate identification at the serovar-level in real-time. However, the potentially unique peak patterns for some serovars identified with Bruker software analyses suggests that serovar raw spectra contain sufficient information for differentiation. This suggested the possibility of constructing serovar MSPs that capture these distinct patterns and allow for rapid serovar differentiation during routine diagnostic testing.

Chapter 3: SPECIFICITY

3.1 ABSTRACT

Leptospirosis is recognized as a reemerging zoonotic disease due increasing incidence, widening distribution into new geographic regions, shifting seroprevalence, and more severe forms of disease presenting on a regular basis. Over 250 pathogenic serovars, the etiologic agents of the disease, are recognized. With practically all mammals as potential hosts, there is a critical need for serovar typing and routine surveillance to address this growing threat. Ideally, *Leptospira* typing would be performed as part of routine leptospirosis diagnostics, without requiring additional time, costs, or labor. However, currently available *Leptospira* typing methods are too technically demanding, time-consuming, and expensive to be feasibly incorporated into the routine diagnostic workflow. Here, we propose that MALDI-TOF MS, already used in most diagnostic microbiology laboratories, offers the requisite specificity to identify *Leptospira* serovars, and can reasonably be integrated into routine diagnostics as a complementary diagnostic tool. The aim of this study was to investigate MALDI-TOF MS as a method to rapidly detect and identify selected serovars of *Leptospira interrogans* using a custom *Leptospira* reference library. In pursuit of this objective, numerous methods for optimizing spectra quality and increasing the specificity of a custom *Leptospira* reference library were explored.

The goal was to determine whether MALDI could be used both as a complement to standard canine leptospirosis diagnostic testing and as a method for regular isolate subtyping for epidemiological surveillance. Initially, Main Spectrum Profiles (MSPs) were created for seven *Leptospira* reference serovars using the standard MSP creation guidelines and default parameters. MSPs were tested using serovar sample dilutions. Testing of these original MSPs yielded consistently correct species-level identifications but returned mixed results at the serovar level. To determine whether MSP specificity could be improved, various sample preparation methods, raw spectra and peak data analyses, and modified MSP creation methods were explored. This resulted in the creation of several new MSP types. New MSPs differed from the originals in the number of raw spectra and biological replicates used, the maximum number of MSP peaks allowed, the lower spectrum boundary used for peak picking, and the omission of specific raw spectra from MSP creation based on ClinProTools (CPT) software (v. 3.0 Bruker Daltonics) data analysis. New MSPs were evaluated for serovar specificity. The best-performing MSP for each serovar was then used to create a *Leptospira* reference library. The library was tested in two blind-coded trials using Biotyper Realtime Classification (RTC) software (v.3.1, Bruker Daltonics, Inc.) in the real-time classification workflow. Specificity for serovar MSPs in the first blind trial ranged from 99–100%. In the second blind trial, MSP specificity ranged from 97-100%. This study demonstrated that the MALDI platform can be used to identify *Leptospira* isolates to the serovar-level within the real-time classification workflow by optimizing the creation parameters used to generate custom MSPs.

3.2 INTRODUCTION

Leptospirosis, caused by pathogenic serovars of spirochete bacteria in the genus *Leptospira*, is one of the leading zoonotic causes of morbidity and mortality worldwide, with over 1 million severe cases and an estimated 60,000 human deaths per year (Bharti et al. 2003, Adler and de la Peña Moctezuma 2010, Costa et al. 2015). While recognized for over a century (Weil 1886), it is now considered a re-emerging disease due to a changing epidemiology that includes an increase in the number of reported cases worldwide, an expansion into new geographic areas, and more severe forms of the disease (Bharti et al. 2003, Meites et al 2004, Lau et al. 2010, Truong and Coburn 2012, Vasylieva et al. 2017).

Pathogenic serovars are the etiological agents of leptospirosis and over 250 are currently recognized (Cerqueira and Picardeau 2009, Adler and de la Peña Moctezuma 2010). Serovars are characterized by structural and orientation differences in the O-antigen of the lipopolysaccharide (LPS), as determined by the cross-agglutination absorption test (CAAT) (Adler and de la Peña Moctezuma 2010, Greene et al. 2012). Antigenically related serovars are further organized into serogroups, which are not a recognized taxonomic category, but are useful in serological diagnostics and epidemiological studies (Kmety and Dikken 1993, Sykes et al. 2011).

It has been reported that virtually all mammal species, as well as some reptiles and birds, can be infected with *Leptospira* (WHO 2012, Lunn 2019). While *Leptospira* infection does not cause clinical disease in all infected animals, it can cause morbidity and mortality in many. Both wild and domesticated animals may also serve as maintenance hosts for pathogenic *Leptospira* serovars (Cerqueira and Picardeau 2009, Galloway and Levett 2010). In general, maintenance hosts are chronic, usually asymptomatic, carriers that excrete leptospires in their urine, intermittently or continuously, for extended periods of time, acting as reservoirs of infection (Adler and de la Peña Moctezuma 2010). Particular serovars exhibit characteristic preferences for particular animal maintenance hosts (Ko et al. 2009, Galloway and Levett 2010, Sykes et al. 2011). However, reservoir hosts can maintain one or several particular serovars and these serovar-host associations can vary with geographic location (Levett 2001, Bharti et al. 2003, Klaasen et al. 2003). Incidental (accidental) hosts are those infected with a serovar for which they do not serve as a maintenance host (Levett 2015a). When infected, incidental hosts are likely to experience clinical, often severe, disease (Levett 2015a).

In the US, dogs are the companion animal most commonly affected by leptospirosis (Bowles 2015, Lunn 2019). They are known reservoir hosts for *Leptospira interrogans* serovar Canicola (Levett 2001, Bharti et al. 2003) and often become incidental hosts for different serovars carried by other domestic or wild animals. Infected dogs can intermittently shed leptospires in their urine for extended periods of time without showing clinical signs of disease (Batista et al. 2004, Miotto et al. 2018). Therefore, infected dogs pose a risk to both their owners and other susceptible animals. (Ellis 2015). A recent study by Miotto et al. 2018 found that asymptomatic, chronically-infected dogs may serve as maintenance hosts for *Leptospira* species other than *L. interrogans*. The authors suggest that renal carriage of a variety of *Leptospira*

species in canines is possible and that this may be an under-studied source of leptospirosis transmission.

As in humans, the cases of canine leptospirosis have been increasing and this upward trend is expected to continue (Ward et al. 2002, Moore et al. 2006). This increase can be tied to a variety of factors, one of which is climate change. As climate change contributes to warmer temperatures in temperate latitudes, the number and size of geographic areas which can reasonably sustain leptospires excreted by infected hosts, increases. This results in larger areas over which leptospires can remain infectious in the environment, which in turn, increases transmission opportunities. Moreover, the changing climate is expected to increase the number and scale of extreme weather events, including floods, which create conditions conducive to infection (Lau et al. 2010, Semenza et al. 2012). Considering the growing number of cases, and the frequent severity of this disease, the re-emergence of leptospirosis presents a major challenge to future human and animal health (Lau et al. 2010).

Two factors crucial for effectively addressing the reemergence of leptospirosis are *Leptospira* isolate typing and epidemiological surveillance (Andersen-Ranberg et al. 2016, Ruppitsch 2016). Bacterial typing, (or subtyping), is an essential component for understanding the epidemiology of infectious diseases and implementing effective prevention and control measures (Foxman et al. 2005, van Belkum et al. 2007, Ruppitsch 2016). Typing refers to the identification of isolates at a sub-species level. Phenotypic methods group organisms based on the similarity of physical or biochemical characteristics, the presence and expression of which are due to an organism's genetic makeup under the influence of a particular set of environmental factors (van Belkum et al. 2007). Genotypic methods examine the genomic variation of bacterial isolates based on the presence or absence of clinically or epidemiologically relevant DNA sequences (van Belkum et al. 2007, Goering 2013). Typing of outbreak-associated isolates can help to predict the source of infection and illuminate transmission patterns. Data collected from typing can also be used to study the evolution and distribution of bacterial populations and help to identify disease trends (van Belkum et al. 2007). Given the proclivity for certain associations between *Leptospira* serovars and reservoir hosts, isolate typing to the serovar-level is key to understanding the epidemiology of leptospirosis (Faine et al. 1999, Levett 2001). Effective surveillance for an infectious disease requires routine identification of clinical isolates performed so that outbreaks, shifts in seroprevalence, and the appearance of new hosts can be detected in real-time (van Belkum et al. 2007, Pereira et al. 2017).

One way to achieve this is to integrate bacterial typing into routine diagnostics. A variety of pheno- and genotyping methods have been applied to *Leptospira*; however, no currently available *Leptospira* typing method can feasibly be performed as part of routine diagnostics. These techniques are too labor-intensive and cost-prohibitive for routine use. Additionally, most have displayed limitations in the number of serovars they can distinguish (Cerqueira and Picardeau 2009, Romero et al. 2009, Goering 2013, Miraglia et al. 2013).

Given the predicted increase in the number of leptospirosis cases and the lack of typing methods that allow for routine surveillance, there is a critical need for a new *Leptospira* typing technique that can be

integrated as clinical lab tests. Such a method would allow for surveillance in real time without adding additional work and costs. In addition to elucidating the probable sources of infection (Hartskeerl and Smythe 2015) data gathered through typing and surveillance efforts could be used to establish regional rates of endemicity, define epidemic criteria (van Belkum et al. 2007) and assess and improve disease management strategies (Galloway and Levett 2010, Guerra 2013).

Routine serovar typing is also important for leptospirosis vaccine design and assessment of vaccine efficacy. Immunity against *Leptospira* infection is generally thought to be serogroup specific (Sykes et al. 2011). The canine leptospirosis vaccines currently offered in the United States are based on the bacterins of up to four serovars (Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona), representing four different serogroups. These serovars are presumed to cause the majority of leptospirosis-related canine disease in the US. Differences in regionally predominant circulating serovars may result in different groups of serovars responsible for canine disease. For example, serovar Hebdomadis in Japan (Koizumi et al. 2013), serovar Pyrogenes in the Kerala region of India (Ambily et al. 2012), serovars Patoc and Tarassovi, in the Paraiba and Uberlandia regions, respectively, of Brazil (Batista et al. 2004, Ribeiro de Castro et al. 2011), and serovar Bataviae in the Chaing Mai region of Thailand (Meeyam et al. 2006), are frequently associated with canine disease though are not recognized as causative agents of canine disease in the US.

Seroprevalence can also shift over time. As discussed in the Literature Review, the serovars currently responsible for the majority of canine disease in the US are different than those previously associated with canine leptospirosis prior to the early 1970s, when the canine bivalent vaccine was introduced (Rentko et al. 1992, Ward et al. 2002, Ward et al. 2004). Therefore, serovar-specific *Leptospira* surveillance is important for determining which serogroups are responsible for the majority of disease in a particular geographic area to help ensure continued efficacy of bacterin-based vaccines.

Leptospira have historically been classified using serotyping techniques that detected differences in LPS structure (Faine et al. 1999). Serotyping is a phenotypic method of identification based on reactions that occur between the surface antigens of microorganisms and specific antibodies in patient sera. Until more recently, the Cross-Agglutination Absorption Test (CAAT) was considered the gold standard for *Leptospira* serotyping. This test requires prior culture and isolation of the organism. The fastidious nature of *Leptospira* makes culture and isolation a laborious and time-consuming process. Depending on the serovar, size of inoculum, and growth medium, cultures must be incubated for at least 7-10 days, but possibly up to 13 weeks, before leptospires are detected (Adler and de la Peña Moctezuma 2010, Cameron 2015). Additionally, culture requires specific, enriched growth media (Levett 2001, Cerqueira and Picardeau 2009), and is prone to contamination. Once a sufficient amount of the organism of interest has been obtained, the CAAT is then conducted by comparing the reaction of the unknown isolate with its antiserum to the reactions of reference strains with their corresponding antisera (Musso and La Scola 2013). The hyperimmune sera used in the CAAT is time-consuming to produce, taking approximately 6-

10 weeks before the appropriate antisera titer is achieved (Hartskeerl and Smythe 2015). Additionally, laboratories that perform this test must maintain an extensive range of *Leptospira* reference strains and corresponding live antigens (Moreno et al. 2016). Given the time, resources and expertise required to perform the CAAT, this test is now mainly used for identifying new serovars and is offered by only a few specialized reference laboratories (Cerqueira and Picardeau 2009). For a list of specialized reference laboratories, please see the chapter by Hartskeerl and Smythe entitled “The Role of Leptospirosis Reference Laboratories” in *Leptospira and Leptospirosis* (Hartskeerl and Smythe 2015).

Over the last several decades, additional serological typing methods have been developed with the goal of reducing the challenges posed by the CAAT. One such method involves typing with monoclonal antibodies (mAbs). Isolates of interest (unknown serovars) are tested against panels of mAbs for agglutination. In theory, the mAbs display characteristic patterns of agglutination when tested against different serovars. While this approach offers a more rapid turnaround time than the CAAT and is reported to have successfully typed up to approximately 70% of isolates tested (Hartskeerl and Smythe 2015), there is still room for error. These mAbs are designed to recognize epitopes on *Leptospira* LPS. However, the currently available mAbs recognize only a small number of epitopes and some of these epitopes may be shared by different serovars (Bourhy et al. 2012). Other *Leptospira* serological tests offer limited serovar typing capacity due to serovar cross-reactions (Levett 2001) and changes in serovar LPS driven by horizontal gene transfer (Haake et al. 2004, Morey et al. 2006, van Belkum et al. 2007).

Given the challenges associated with *Leptospira* serotyping, there has been great interest in developing molecular methods which can be used as an alternative to serotyping for serovar identification. Molecular methods have been found to offer greater discriminatory power, faster turnaround times, and greater sensitivity in the identification of many bacteria compared with traditional, phenotypic methods (Adzitey et al. 2013). Molecular genotypic methods are typically based on the detection of genomic variations between isolates. These methods are used primarily to detect variations in single nucleotide polymorphisms (SNPs), insertions, and deletions (Ruppitsch 2016). Various genotypic techniques have been examined for their ability to characterize selected serovars of *Leptospira.* Some methods, such as restriction fragment length polymorphism analysis (RFLP), analyze the size and pattern of genomic fragments produced by restriction enzymes (Perolat et al. 1993). Until recently, pulsed-field gel electrophoresis was considered the current gold standard of molecular typing (Cerqueira and Picardeau 2009). This technique is an RFLP-based method that uses rare cutting restriction endonucleases to create large DNA fragments, which are then separated into a unique pattern of bands via gel electrophoresis. While this method has shown agreement with serotyping results in some *Leptospira* studies, it is very labor intensive, time-consuming, and cannot be used to differentiate between some serovars (Romero et al. 2009, Galloway and Levett 2010, Miraglia et al. 2013). PCR-based genotyping methods have had varying levels of success differentiating between pathogenic and saprophytic serovars, between some *Leptospira* species, and between select groups of serovars. These

methods have included amplified fragment length polymorphisms (AFLPs) and fluorescent-labelled AFLP (Vijayachari et al. 2004), arbitrarily primed PCR (Ralph et al. 1993) and randomly amplified polymorphic DNA (Corney et al. 1993), multi-locus variable number of tandem repeats analysis (Slack et al. 2006a, Zuerner and Alt 2009), and multilocus sequence typing (MLST) (Ahmed et al. 2006, Varni et al. 2014). While PCR-based methods require only small amounts of sample DNA, the number of serovars that can be identified by most assays are limited by the specific primers that need to be created for sub-species level identification (Cai et al. 2010, Ruppitsh 2016). Currently, whole genome sequencing (WGS) is considered the gold standard in for *Leptospira* typing. Since 2003, when the first full *Leptospira* genome was reported by Ren et al. 2003, a growing number of *Leptospira* isolates have been sequenced (Fouts et al. 2016, Llanes et al. 2016, Kurilung et al. 2019). These genomic data offer valuable information that can be used in subsequent analyses that many other forms of typing do not. For example, sequencing data can be used to study the differences between genomes of pathogenic, saprophytic, and intermediate isolates; to elucidate the evolution of virulence factors in pathogenic isolates; and to identify targets for new diagnostic tests and vaccines (Jorge et al 2018). Additionally, WGS does not require the maintenance of live cultures. While WGS offers several advantages over other typing methods, isolate sequencing currently costs approximately $180.00 - $240.00 and the process can take around a week to complete (C. Maddox, personal communication, Nov. 2019). Given the predicted increase in both the incidence and distribution of leptospirosis together with the considerable technical, time and labor requirements of current typing methods, there is an urgent need for an alternative *Leptospira* typing method that can be implemented in the routine diagnostic workflow.

MALDI-TOF MS has been widely adopted by diagnostic microbiology laboratories as a principal component of routine bacterial identification (Freiwald and Sauer 2009, Rodríguez-Sánchez et al. 2014, Heaton & Patel 2017). Whole cell MALDI-TOF MS uses intact cells or cell extracts (Singhal et al. 2015) from a microorganism of interest to generate a unique, molecular fingerprint that can be used for the organism's identification. This fingerprint is a representative mass spectrum created by the MALDI based on the analyte's detected protein profile. For most MALDI systems, identification of the organism then proceeds by using a pattern matching algorithm to compare the analyte's representative spectrum, containing genus- and species-specific peaks, to a commercial database of reference spectra for well-defined microorganisms (Heaton and Patel 2017). The algorithm takes into account a spectrum's peak presence and absence at particular locations along the spectrum's measured mass range, peak intensity, and peak frequency. A score is then generated, which indicates the degree to which a sample's spectrum matches that of the most similar reference spectrum in the database. Different MALDI manufacturers establish their own score interpretation guidelines. In general, the score indicates the probability that the sample resembles one of the reference organisms included in the reference database (Bruker Daltonics, Inc. 2012).

MALDI'S rise in popularity can be attributed to its efficient workflow, relative ease-of-use, low per-test cost, and rapid time-to-result (Seng et al. 2009, Heaton and Patel 2017). Many studies have reported reliable MALDI identification at the genus and species levels (Clark AE et al. 2013, AlMasoud et al. 2014). A growing body of work also demonstrates MALDI's potential for differentiation of organisms at a sub-species level (Seibold et al. 2010, Eddabra et al. 2012, Josten et al. 2013). However, the similarity of organisms at this taxonomic level offers more challenges and often requires optimization of sample preparation, raw spectra acquisition, and spectral peak analysis techniques (Arnold et al. 2006, Murray 2010). This is expected, since closely-related organisms produce spectral profiles with greater similarity compared to spectra produced by distantly-related organisms (Zhang and Sandrin 2016).

Commercial MALDI-TOF MS reference databases do not currently include representative spectra for *Leptospira*. In general, to evaluate MALDI's ability to identify *Leptospira*, users must first create their own *Leptospira* Main Spectrum Profiles (MSPs) and reference library. An MSP is a characteristic mass spectrum and associated peak list created from the average of multiple raw spectra acquired from a particular organism (Bruker Daltonik GmbH 2016). It serves as a mass spectral "fingerprint," and incorporates peak data such as peak mass-to-charge ratios, peak frequency, and peak intensity distributions. Bruker's Microflex™ LT mass spectrometer (Bruker Daltonics, Germany), the MALDI instrument used in this study, offers several options for sample identification via several software programs. Sample raw spectra are first acquired using FlexControl™ (v. 3.0 Bruker Daltonics, Inc.). Next, one of four programs can be used for classification. Biotyper Realtime Classification (RTC) is used for rapid, routine diagnostic classification performed in real time. This program is easy to use and allows multiple samples to be run at once. MALDI Biotyper (MBT) Compass Explorer software (4.1 Bruker Daltonics, Inc.) is used for MSP creation and exploring the relationships among organisms based on their spectra. Identification using this program is considered an offline classification process, requiring additional steps and training to perform. Bruker's FlexAnalysis software allows for the evaluation of individual raw spectra and corresponding peak characteristics as well as comparison of acquired spectra. ClinProTools (CPT) is used for advanced statistical analyses of raw spectra and peak data. Classification can be performed using several tools, but again, is considered an offline process due the time and work involved.

Recently Bruker and the CDC partnered to offer access to a Main Spectrum Profile (MSP) library through MicrobeNet.cdc.gov. After registering on the site and downloading the specific software patches required, users may submit saved spectra for identification against an extended library containing MSPs for some organisms not included in the commercial MSP libraries offered through Bruker alone. The two *Leptospira* species included in the library are: *Leptospira borgpetersenii and L. interrogans.* One *Leptospira* serovar is included: *L. interrogans* serovar Pyrogenes. One strain-level isolate is also included: *L. interrogans* serovar Icterohaemorrhagiae strain RGA. However, "*Leptospira interrogans*" is listed in the notes section next to this entry in the database. It is not clear what this indicates. So, out of four MSP

entries, only two serovar-level MSPs are given. One, for serovar Pyrogenes, is not used in this study and is not known to be associated with canine disease in the United States. The other, for serovar Icterohaemorrhagiae, is given for a different strain than that used in this study. Furthermore, the note next to this entry in the CDC/Bruker database may indicate that this MSP is the same as that given for the *L. interrogans* species. MSPs for most *Leptospira* serovars are not available through this library. Since *Leptospira* from different geographic areas may exhibit some variation in their spectral profiles, additional MSPs may be needed for accurate identification.

A few studies have explored MALDI's ability to detect and identify *Leptospira* (Djelouadji et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Xiao et al. 2015, Karcher et al. 2018, Sonthayanon et al. 2019). In a study conducted by Djelouadji et al. (2012), MSPs were created for 19 *Leptospira* species. Samples were prepared using a modified direct spot method. Six out of six clinical isolates, and 20 out of 21 reference isolates tested against these MSPs were correctly identified to the species level. The one isolate that was identified incorrectly, *Leptospira interrogans* serovar Autumnalis (Akiyami), was matched first to *Leptospira kirshneri.* This was not unexpected, given that only one insertion/deletion and one base differentiate the type strains for these two species (Morey et al. 2006). The authors concluded that MALDI could be used to complement other *Leptospira* typing methods for species-level identification. However, they did not report that serovar-level discrimination was possible. Using a different sample preparation technique, Rettinger et al. (2012) evaluated MALDI's ability to identify 28 *Leptospira* reference strains, representing 12 species and 23 serovars. This group applied an ethanol/formic acid protein extraction method for MALDI sample preparation, which had been reported in some studies (Alatoom et al. 2011, Clark AE et al. 2013) to produce higher-quality spectra and allow for better identification compared to the direct spot method used by Djelouadji and colleagues (2012). MSPs created for the reference strains were tested against both the reference strains and 16 field isolates. All strains were correctly identified to the species level, though one of the *L. kirschneri* field isolate strains matched to both *L. kirshneri* and *L. interrogans*. Results were confirmed via 16S rRNA gene sequencing and MLST. The authors additionally used ClinProTools to evaluate spectra for three of the represented species. The authors reported discriminatory peak combinations detected for seven serovars within these species. It was concluded that MALDI is a reliable method for species-level detection and that further analysis using CPT could potentially differentiate between some serovars. However, they determined it was not possible to differentiate strains below the species level using the MSPs created in-house during real-time identification.

Calderaro et al. (2014b) tested MALDI's ability for both species- and serovar-level identification using some strains tested in the two previous studies and some strains which had not yet been examined. Samples were prepped using an EtOH/FA extraction method similar to that used by Rettinger et al. (2012). MSPs representing six *Leptospira* species were created and tested. Authors reported that all isolates correctly identified at the species- and serovar-levels. However, it was not made clear in the

paper whether the MSPs were tested using the Biotyper Real Time Classification (RTC) (used in real-time identification), or MBT Compass Explorer software (used in offline identification). Raw spectra for two of the six species (*L. interrogans* and *L. borgpetersenii)* were further analyzed in ClinProTools. Like the previous study by Rettinger et al., discriminating peak patterns were detected for some of the serovars represented by these two species. The authors proposed that CPT analysis of isolate raw spectra could be used as an initial method of *Leptospira* detection and, possibly, discrimination of serovars.

MALDI's ability for *Leptospira* identification was further explored by Xiao et al. (2015). Thirty-two out of 33 strains tested against an MSP database created in-house were correctly identified to the species level. The authors additionally used UPLC-MS/MS to detect 108 proteins that were differentially expressed between pathogenic and saprophytic *Leptospira* strains. The authors concluded that MALDI can accurately identify *Leptospira* isolates to the species-level during real-time classification, but that this method cannot be used to differentiate between serogroups.

In a study by Karcher et al. (2018), *Leptospira* MSPs created in-house were used to correctly identify 22 field isolates to the species-level. The majority of the samples were prepared using an EtOH/FA protein extraction method similar to those used in previous studies. However, they reported that a few samples were prepped without extraction, though they do not describe exactly how these samples were prepped. The authors used the two different prep methods to compare the quality of raw spectra obtained with and without extraction. The authors found that extracted samples produced higher quality spectra, containing peaks with higher intensities. A combination of CPT, FlexAnalysis, and SPSS analyses performed on the raw spectra of *L. interrogans and L. kirshneri* found that one peak, located between 8000–8100 Da, could be used to differentiate spectra belonging to these two species. The authors reported that a signal-to-noise ratio (S/N) cut-off value of seven for this peak was required for differentiation. A S/N value below seven indicated that the raw spectrum belonged to *L. interrogans*, while a S/N value above seven suggested that the raw spectrum belonged to *L. kirshneri*. Interestingly, Rettinger et al. had also identified this as a discriminatory peak in their study. However, they found that the peak was present in *L. kirshneri* serovar Grippotyphosa, yet absent in all *L. interrogans* strains tested. As in the previous studies, Karcher and colleagues determined that MALDI could be a valuable tool for initial *Leptospira* identification at the species-level. However, identification at a lower taxonomic level, when obtainable, required additional, offline analyses.

Most recently, Sonthayanon et al. (2019) created a custom in-house *Leptospira* MSP library using a combination of 15 *Leptospira* reference species and 101 clinical isolates from Thailand and Laos identified to the species-level by 16S rRNA gene sequencing. Reference *Leptospira* and isolates were cultured to reach a concentration of 10^8 CFU/mL. Then, 3 mL of each culture was prepped via a protein extraction method that used sinapinic acid matrix. An Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Inc., USA) was used to acquire the mass spectra. The authors used CPT to create three classification models, with the Genetic Algorithm (GA) model returning the highest recognition

capability (RC) and cross validation (CV) values. This model selected 10 discriminatory peaks to differentiate between *L. interrogans, L. borgpetersenii,* and *L. kirschneri*. These peaks included those found in only one or two of the three species and peaks present in all three species, but at different relative intensities. Additionally, a GA model was used to identify groups of discriminating peaks for eight serovars of *L. interrogans*. Again, these peak groups were based on peak presence or absence and relative peak intensity. The *Leptospira* MSP library was tested using 97 clinical isolates in a blind trial. Ninety-six out of the 97 samples were correctly identified to the species-level. While some peak differences were found between serovars of *L. interrogans* using the tools in CPT, the authors do not mention identification of isolates at the serovar-level via real-time classification. The authors also compared the mass spectra of four *Leptospira* species both in media and spiked into urine. They found that, while the spectra were similar, the urine spectra profiles contained some peaks not found in the spectra of media-only samples. However, spiked urine samples were still identified correctly to the species-level. The minimum concentration of *Leptospira* required for accurate MALDI ID was found to be 1×10^6 CFU/mL of culture media. The authors concluded that MALDI holds promise for *Leptospira* species identification, but that further testing with leptospirosis-positive urine samples is needed to learn whether it can be used as an additional tool for routine leptospirosis diagnosis in humans and animals.

To summarize, these six *Leptospira* studies were able to use the MALDI platform and user-created MSPs for species-level identification, yet serovar-level identification proved to be more challenging. The four studies able to achieve serovar-level identification did so using pattern-matching models created within ClinProTools. To use these models for identification, a technician must first acquire and save sample spectra. The spectra must then be imported into CPT, where they can be run against specific, user-created classification models. In contrast, the real-time classification process used in routine diagnostics uses FlexControl to acquire sample spectra while almost simultaneously applying a pattern-matching algorithm to identify the sample via Biotyper RTC. With the Bruker Microflex series MALDI system and sample target plates used in this and five of the previous studies, up to 96 samples can be tested in one real-time run. The average time required for a real-time run, while dependent on the number of samples tested, ranges from approximately 20 seconds for one spot, (once laser and vacuum adjustments are set and the validation check is passed), to approximately 30 minutes for 96 spots. The time required for CPT identification is also partially dependent on the number of samples to be tested. However, since sample spectra acquisition and model classification are performed separately in this offline process, additional steps and time are required. In fact, depending on the type of model and cross-validation method used, this process can take much longer to achieve sample identification. While CPT may be a valuable tool for serovar discrimination in research, the classification models and additional tools used to achieve serovar identification cannot be feasibly incorporated into routine diagnostics. Though this study also used CPT models to identify potentially-important raw spectra peak combinations that could be used for serovar classification, the study's main focus was building serovar MSPs for use in

real-time isolate classification.

The overall goal of this work was to explore the feasibility of incorporating MALDI into routine diagnostics for canine leptospirosis. The work for this study was used to test the hypothesis that MALDI-TOF MS can be used to identify *Leptospira* isolates to the serovar-level within the real-time classification workflow. The aim here was to investigate the ability of MALDI-TOF MS to detect and identify selected serovars of *Leptospira interrogans.* The first objective of this study was to create a custom Main Spectral Profile reference library for seven selected *Leptospira* serovars. The second objective was to test the custom library's specificity in a blind-coded trial.

To evaluate MALDI's ability to identify and differentiate between *Leptospira* serovars, stock cultures of *Leptospira* reference strains were used to construct a Main Spectrum Profile for each serovar using the Bruker MSP default parameters. These original MSPs were combined into a custom in-house *Leptospira* reference library using Bruker library creation guidelines and default settings. The library was then tested using serial dilutions of *Leptospira* stock cultures**.** While these original MSPs returned consistent, accurate identifications to the species-level, they did not have the specificity needed to consistently differentiate between serovars.

To improve MSP specificity, various sample preparation methods, raw spectra and peak data analyses, and modified MSP creation methods, were explored. Spectrum quality plays an important role in both the creation of MSPs and the generation of raw spectra from test samples (Goldstein et al. 2013). In general, spectrum quality refers to the information contained in a raw spectrum and displayed in various characteristics such as the number, intensity, and signal-to-noise ratio of spectrum peaks. However, the specific characteristics that comprise spectra of good quality are often defined by the user based on the particular MALDI system employed and the type of analysis required. Since sample preparation is one variable that can affect both the reproducibility and quality of MALDI raw spectra (Arnold et al. 2006, Goldstein et al. 2013), various preparation techniques were tested to determine which produced spectra with greater numbers of peaks over the widest range of dilutions tested. The technique producing the highest-quality spectra was then used to acquire raw spectra for use in subsequent MSP creation.

A variety of methods was also used to determine whether serovar-specific peak combinations could be identified and applied as selection criteria for raw spectra inclusion in revised MSPs. Initially, the peak lists of each MSP, and each MSP's constituent raw spectra, were compared. Differential weighting of MSP peaks was explored and subtyping MSPs were created. Serovar raw spectra were additionally analyzed using ClinProTools. This program was used to determine similarity between the raw spectra of different serovars, to identify the top serovar-differentiating peaks using both univariate and multivariate statistics, and to identify individual raw spectra that were potentially responsible for lowering an MSP's specificity.

Information obtained using the various Bruker software programs was then applied to the creation of several new, different MSP types for each serovar. New MSPs differed from the originals in the number of raw spectra and biological replicates used, the maximum number of MSP peaks allowed, the lower spectrum boundary used for peak picking, and the omission of specific raw spectra from MSP creation. MALDI identification trials were performed to test new MSP specificity. Trial results were compared to determine which preparation method and MSP creation parameters allowed for the most accurate and consistent serovar identification. The best-performing MSP for each serovar was selected and combined into a final *Leptospira* reference library. This library was then tested in two blind-coded trials using MALDI Biotyper RTC software (Bruker Daltonics, Inc.) for real-time identification. The first trial used *Leptospira* cultures in UHPLC-grade H_2O. The second trial used *Leptospira* cultures pelleted and resuspended in leptospirosis-negative canine urine. Trial results were evaluated using 2 x 2 contingency tables. Specificity for the selected MSPs ranged from 97–100% over both trials.

In this study, we evaluated sample preparation and deposition techniques as well as MSP creation methods to find those optimal for *Leptospira* serovar detection and identification. We found that the EtOH/FA extraction method resulted in the best quality sample spectra, allowing for a greater number of accurate sample identifications. While various types of MSPs were created for each serovar to determine which offered the highest level of specificity, the best-performing MSPs were found to be those which used creation parameters that deviated from the default settings and recommended MSP creation guidelines. Further testing of these MSPs, using additional reference strains and clinical isolates of the same serovars, is needed to determine identification accuracy when challenged with new isolates.

This study demonstrated that the Bruker MALDI-TOF MS Microflex LT™ and user-created *Leptospira* MSPs can be used to identify selected *Leptospira* isolates to the serovar-level in the real time classification workflow. This suggests that MALDI could be used to identify additional serovars. Further work would need to include the creation of serovar MSPs that combine raw spectra from several strains as well as the creation of MSPs for supplemental serovars.

At the time this project was started, one of the objectives was to find a *Leptospira* detection and identification method that could be used for canine urine samples submitted for leptospirosis testing. Work done to determine MALDI's specificity and sensitivity in *Leptospira* detection was performed in parallel. For the first part of the study, serial dilutions of serovar stock culture combined with either UHPLC-water or qPCR-leptospirosis-negative canine urine were used to test the MSPs. However, as will be discussed in the next chapter, testing showed that MALDI does not have the sensitivity needed for direct use on leptospirosis-positive canine urine samples.

Though MALDI cannot be used to identify most leptospirosis-positive canine urine samples, it might still offer a complementary or alternative identification technique in human leptospirosis diagnostics. The concentration of excreted leptospires in leptospirosis-positive human urine, along with typical human urine sample volumes, may allow for a sufficient number of leptospires per MALDI sample spot such that

Leptospira identification is possible. While variation in the predominant circulating pathogenic serovars and strains between geographic regions renders a universal serovar MSP library impractical, this work proposes a method by which diagnostic laboratories can create custom libraries which can be used for serovar identification and epi-surveillance in real-time.

3.3 MATERIALS AND METHODS

3.3.1 Safety Precautions

Leptospira cultures were handled in a Class II biological safety cabinet while wearing the appropriate personnel protective equipment.

3.3.2 *Leptospira* Serovars

Seven *Leptospira* reference strains, representing seven pathogenic serovars, were obtained from the National Veterinary Services Laboratory (NVSL) (Ames, Iowa)**.** The selected serovars are used in the University of Illinois at Urbana-Champaign's Veterinary Diagnostic Laboratory's (VDL) microagglutination titer panel and consist of: *Leptospira interrogans* serovar Autumnalis (Akiyami A), Bratislava (Bratislava Jez), Canicola (Hond Utrecht IV), Grippotyphosa (Andaman), Hardjo (Hardjoprajtino), Copenhageni ($M20^A$, used as a representative member of serogroup Icterohaemorrhagiae), and Pomona (Pomona). Serovars were grown in Polysorbate-80 bovine albumin liquid medium (P-80 BA, NVSL, Ames, Iowa) at 28°C to 30°C and maintained by serial passage every seven days. Cultures were checked for contamination by weekly plating and assessed for purity weekly using control sera for serovar recognition. *Leptospira* numbers were standardized by adjusting the percent transmittance of culture measured in disposable semi-micro cuvettes (Sarstedt Ag & Co. KG, Nümbrecht, Germany) using a Shimadzu BioSpec mini-1240 UV-Vis Spectrophotometer (Shimadzu Corporation, Kyoto, Japan) at 400 nm.

3.3.3 Canine Urine Specimens

Specimens were obtained from those submitted to the VDL for leptospirosis testing. Remaining sample volumes that tested leptospirosis-negative using qPCR (real-time polymerase chain reaction) were stored at 4 °C and used within two weeks of arrival. When a sufficient volume of usable urine had been collected, the specimens were pooled and centrifuged at 600 x *g* for 15 minutes to remove contaminating cells. The supernatant was then run through an enclosed syringe filter for sterilization. Initially, an EMD Millipore™ Millex™ sterile syringe filter with a Durapore™ polyvinylidene difluoride (PVDF) .22 μm membrane (Burlington, MA, USA) was used. However, this type of filter became clogged after approximately 1 mL of urine was passed through it. Since 15+ mL needed to be filtered, and 3 filters had been required to filter 3 mL, the filter type was switched to an EMD Millipore™ Millex™ -GP sterile syringe filter with a .22 μm pore polyethylsulfone (PES) membrane (Burlington, MA, USA).

3.3.4 Spectrophotometric Quantitation

Serovar stock cultures used to make dilutions were adjusted to 25 percent transmittance (%T) using the Shimadzu spectrophotometer mentioned previously, set at 400 nm. The percent transmittance of each dilution created for MALDI specificity runs was also measured. Measurements were performed using one mL of sample in a disposable semi-micro cuvette (10x4x45 mm, REF 67.742, Sarstedt Ag & Co., Germany), with one mL of UHPLC-grade H_2O used as the reference blank. UHPLC-grade H_2O, rather than the polysorbate 80-bovine albumin (P80-BA) *Leptospira* growth medium was chosen for use as the reference blank because initial MALDI tests of serial dilutions created with P80-BA media resulted in poor quality and unidentifiable serovar mass spectra. Since proteins found in culture media may negatively affect MALDI sample spectrum generation and analyte identification (Honoré et al. 2013), subsequent dilutions were instead created with UHPLC-grade H_2O.

3.3.5 Dilutions

Serovar culture serial dilutions were used for both specificity and sensitivity testing of custom-created *Leptospira* MSPs. To prepare the dilutions, culture suspensions were first spun at 600 x *g* for 15 minutes (Sorvall ST 16R centrifuge, Thermo Scientific, Cat. No. 75004240, 23 °C) to pellet dead bacteria. The resulting supernatant was then adjusted to approximately 25 %T (3.01 x 10^9 organisms/mL), prior to dilution creation unless stated otherwise. Two sets of serial 2-fold dilutions, ranging from 3.01 x 10^9 - 1.01 x 10^8 organisms/mL, were created. The first set was used for MALDI-TOF MS trials. The second set were measured to find the concentration of each dilution. Measurements were performed on the Qubit 4.0 Fluorometer using DNA extracted from 1 mL of each dilution. This data was subsequently used to calculate the estimated number of genome equivalents, and by extension, the number of leptospires per μL, for each of the serial 2-fold dilutions used in MALDI testing. A set of serial 10-fold dilutions was also created for use in qPCR. These dilutions consisted of serovar suspensions mixed with either UHPLC-grade H_2O or canine urine ranging from 10^8 -10^1 organisms/mL.

3.3.6 DNA Extraction

DNA template was extracted using the QIAamp® Viral RNA Mini Kit (Qiagen, Germany) and Purification of Viral RNA protocol. This protocol uses Buffer AVL, which is a reagent that inactivates PCR inhibitors found in urine. Although the name of the protocol refers to viral RNA, the manufacturer of this kit recommends this protocol for the extraction of bacterial DNA from urine samples (QIAamp Viral RNA Mini Handbook, March 2018, p. 36). Additionally, this protocol is commonly used by the University of Illinois Urbana-Champaign's Veterinary Diagnostic Laboratory to extract DNA from canine urine samples submitted for leptospirosis qPCR testing. Extraction was performed using 1 mL aliquots of sample.

3.3.7 Determination of Dilution Concentration

DNA template concentration for each serial 2-fold dilution was measured on a Qubit™ 4.0 Fluorometer (Invitrogen, Life Technologies Holdings Pte Ltd, Singapore) using the Qubit™ dsDNA Broad Range Assay Kit. For each set of 2-fold serial dilutions to be measured, a working solution, two standards, and dilution samples were prepped. To create the working solution, 1 μL of Qubit® dsDNA Reagent and 199 μL of Qubit® dsDNA BR Buffer for each of the two standards and per sample to be measured was mixed together in a plastic tube. Next, the solution to be used as Standard 1 was created by adding 10 μL of the Standard 1 kit reagent and 190 μL of working solution to a Qubit™ 0.5-mL PCR tube. The solution to be used as Standard 2 was prepped by similarly using 10 μL of the Standard 2 kit reagent. Dilution samples were prepped by mixing 3 μL of a sample with 197μL of the working solution in a Qubit™ 0.5-mL PCR tube for a final volume of 200 μL per sample tube. Prepped tubes were incubated at room temperature for two minutes. On the Qubit, the dsDNA Broad Range assay type was selected for measurement. Standards 1 and 2 were then measured to calibrate the instrument. After confirmation that the reading given by Standard #2 was at least 10x higher than that of Standard #1, dilution samples were measured in duplicate. Dilution DNA template concentrations were given in ng/μL. Average measurements per dilution were then calculated.

3.3.8 Estimation of the Number of Genome Equivalents per μL for Each Dilution

Average DNA template dilution concentrations were used to calculate the number of genome equivalents per μL using the following equation (Staroscik 2004):

$$\text{Number of copies/}\mu\text{L} = \frac{(\text{amount of DNA in ng/}\mu\text{L}) \times (6.022\times10^{23}\ \text{molecules/mol})^{a}}{(\text{length of dsDNA template in base pairs [bp]})^{b} \times (650\ \text{g/mol})^{c} \times (1\times10^{9})^{d}}$$

[a] Avogadro's number

[b] length of dsDNA template = size of genome = 4,627,366 bp

[c] 650 g/mol = 650 Da = the assumed average weight of a single DNA bp

[d] 1×10^{9} ng/g was used in the equation to convert our calculated number to ng (which allows for units ng to be cancelled, leaving molecules/μL as units for the copy number)

The amount of DNA used in the equation was taken from the average Qubit concentration for each dilution, given in ng/μL. This number was then multiplied by Avogadro's number, 6.022×10^{23} molecules/mole. The equation was calculated using a template length of 4,627,366 base pairs, based on the reported size of *L. interrogans* serovar Copenhageni's genome (Nascimento et al. 2004). The molecular weight of the genome was estimated by multiplying the length of the template by 650 g/mol (Clifford et al. 2012). The inverse of this value was then taken to give the number of moles of DNA

template in 1 g of material (Kemp et al. 2014). Next, this number was multiplied by Avogadro's number, 6.022×10^{23} molecules/mol, to find the number of DNA template molecules per gram. This number was then multiplied by 1×10^{9} to convert the value to molecules/ng. The resulting value was multiplied by the average DNA concentration of each dilution, given in ng/μL, to obtain the number of genome copies per dilution. Calculated template dilution concentrations were then correlated with the sample dilutions used for MALDI testing.

3.3.9 Reagents

Each MALDI-TOF MS sample preparation, regardless of the preparation method being tested, used UHPLC-grade water (for mass spectrometry, CAS Number 7732-18-5, Sigma-Aldrich GmbH), organic solvent (50% ACN and 47.5% UHPLC-grade H_2O, Sigma-Aldrich GmbH; 2.5% Triflouroacetic acid, Honeywell), Bruker's bacterial test standard (BTS), and Bruker's HCCA matrix solution. To reconstitute the HCCA matrix (α-cyano-4-hydroxycinnamic acid, Bruker Daltonics, Inc.), 250 μL of organic solvent was added to a 2.5 ± 0.3 mg pre-portioned tube of matrix, then vortexed to completely dissolve the matrix in the solvent (Bruker Daltonik GmbH 2015). Non-inoculated matrix was used as the negative control. BTS was used for calibration and validation of MALDI runs and as a positive control. It is comprised of *Escherichia coli* DH5 alpha extract spiked with two proteins of high molecular mass, RNase A and Myoglobin. These constituents create characteristic mass spectral peaks that cover the range of mass detection. To prepare the BTS, 50 μL of organic solvent was added to a microcentrifuge tube containing a pre-portioned BTS pellet. Tube contents were then mixed via pipetting until the pellet dissolved. The mixture was incubated for 5 minutes at room temperature, then mixed again via pipette. The mixture was then centrifuged at 15 700 x *g* for 2 minutes. Finally, the mixture was divided into 5 μL aliquots and stored at -18 °C until needed (Bruker Daltonik GmbH 2017). The extended direct transfer and ethanol-formic acid (tube extraction) prep methods also used formic acid (FA) (70% v/v, Sigma-Aldrich GmbH). Tube extraction (TE) additionally required absolute ethanol (EtOH) (# 24102, HPLC grade, Sigma-Aldrich GmbH) and acetonitrile (ACN) (# 34967, HPLC grade, Sigma-Aldrich GmbH).

3.3.10 Sample Preparation Methods

There are three main sample preparation techniques used for whole-cell MALDI-TOF MS bacteria identification: direct transfer of whole cells to the MALDI target plate, extended direct transfer with on-target FA extraction, and ethanol-formic acid extraction (Šedo et al. 2011). In the standard direct transfer method, a single bacterial colony from an organism that has been cultured and isolated on solid media is transferred directly from the growth media to the MALDI target plate. While not employed as often, direct transfer can also be applied to bacterial cells grown in liquid media and collected via centrifugation. The extended direct transfer method involves the addition of 70% FA to the bacterial sample spot prior to adding matrix solution. Ethanol-formic acid (EtOH/FA) extraction is used for those organisms that do not

produce quality spectra using the other two methods. The type of sample preparation method used can affect spectra quality and reproducibility; and therefore, accuracy of sample identification (Williams et al. 2003, Goldstein 2013). The optimal sample preparation method for a particular organism depends on several factors, including the physiochemical properties of the organism, sample type, and the objective of MALDI analysis (Singhal et al. 2015). Bruker stipulates that the ethanol-formic acid extraction method be used for acquiring raw spectra to be used in MSP creation. However, user-determined optimized methods can be used for the acquisition of spectra for real-time or offline identification.

To identify an optimal sample preparation method that would generate unique and reproducible spectral profiles for the *Leptospira* serovars through the highest dilutions tested, both intact cell and protein extraction preparation techniques were explored. For all trials which used a protein extraction preparation method, sample dilutions were spotted onto the MALDI target in duplicate. In trials which used a direct transfer (DT) prep method, sample dilutions were spotted onto the target either once or twice based on the volume of available sample material. After DT prep, samples of higher dilutions often had insufficient material for more than one sample spot. The final preparation step for all methods was the application of 1 μL of matrix solution over each sample spot. This solution was allowed to dry at room temperature prior to target insertion into the MALDI-TOF instrument. Per Bruker's recommendation, Eppendorf-brand microcentrifuge tubes (1.5 mL, Safe-Lock, Eppendorf N. America Inc., Hauppauge, NY, USA), pipettes (Reference® 2 single-channel 0.5-10 μL, 10-100 μL, 100-1000 μL), and tips (epT.I.P.S., 0.1–10 μL, Cat #: 022491504 and 2–200 μL, Cat #: 022491733, Eppendorf N. America Inc., Hauppauge, NY, USA), were used in sample preparation to prevent potential leaching of plastic additives into the sample and subsequent effects on spectra generation (Freiwald and Sauer 2009, Bruker Daltonics, Inc. 2012). Sample preparation trials were generally performed first with stock culture and H_2O serial dilutions from only one serovar. Assays that showed promising results were repeated for all seven pathogenic serovars using H_2O and canine urine culture dilutions. The goal was to identify the method that produced accurate identification results with match scores ≥ 2.30, indicating highly probable species-level identification.

3.3.10.1 The Direct Transfer Method and variations

For the direct transfer sample preparation method, a 1 mL aliquot of each 2-fold serial dilution was transferred to a 1.5 mL microcentrifuge tube and spun at 15 700 x *g* for 2 minutes to pellet the bacteria. The supernatant was removed and a direct spotting (direct transfer) method was used to place the sample on the MALDI target. The standard method uses a pointed wooden applicator stick to collect a small portion of an isolated bacterial colony from a plated culture and then deposit that material as a thin film in a sample spot of the steel target. When used with bacterial colonies grown on solid agar media, this method is fairly straightforward. However, this technique proved to be problematic for use with bacteria pelleted out of liquid media. Applicator sticks were used to transfer a portion of the pellet in each

tube to a target sample spot. For those dilutions which had no pellet but had visible white film on the side of the tube, the stick was used to scrape the film and spread it inside a sample spot. Collecting a sufficient amount of the moistened pellet on the stick presented a significant challenge. Therefore, a larger applicator stick with a blunt end was also tested for sample transfer. When no pellet and no film was visible in the tube, approximately 5-10 μL of supernatant was left in the tube after centrifugation. A pipette was then used to spot 1 μL of the supernatant onto the target. Sample spots were allowed to dry at room temperature prior to matrix overlay.

An extended direct transfer method was also evaluated. For this technique, the second sample spot for each dilution was overlaid with 1 μL of 70% FA and allowed to dry prior to the addition of matrix. This method is most often used for Gram-positve bacteria to assist with bacterial wall penetration and the release of proteins for MALDI measurement. This method was tested to determine if it allowed for the inclusion of additional protein peaks in serovar spectra.

3.3.10.2 Direct Transfer Method: addition of one or two washes using H_2O and EtOH

In an early assay, additional variations of the direct transfer method evaluated the effect of one or two wash steps added after the initial sample centrifugation to pellet bacteria and the disposal of the supernatant. For each wash step, 1 mL of UHPLC-grade H_2O was added to each sample tube. The tubes were then vortexed and spun at 15 700 x *g* for 2 minutes. After disposal of the supernatant, sample tubes proceeded through the remaining standard direct transfer preparation steps as described above. This assay was repeated using 1 mL of EtOH, rather than H_2O, for the wash steps. Serovars Grippotyphosa and Canicola were used for this trial.

3.3.10.3 Ethanol-formic acid extraction (tube extraction) and variations

The ethanol-formic acid protein extraction method and variations thereof were also evaluated. This method is often used with sample organisms that do not generate reproducible quality spectra when processed using a direct transfer method. This is also the recommended technique for the acquisition of spectra used in the creation of in-house reference MSPs. Initially, 1 mL of each sample dilution was placed in a 1.5 mL microcentrifuge tube and spun at 15 700 x *g* for 2 minutes (Eppendorf 5415D Centrifuge, Eppendorf N. America Inc., Hauppauge, NY, USA) to pellet the bacteria. After removal of the supernatant, the pellet was resuspended in 300 μL of UHPLC-grade H_2O and mixed carefully via pipette. Next, 900 μL of EtOH was added to each tube. Suspensions were then vortexed for one minute and centrifugation was repeated. The supernatant was discarded, and the tubes were spun once more at 15 700 x *g* for 1 minute to collect all remaining supernatant at the bottom of the tube for removal via pipette. Tube lids were then left open for approximately five minutes to allow evaporation of any residual moisture. For protein extraction, 25 μL of 70% FA was added to each tube followed by 25 μL of ACN. The tubes were then vortexed and centrifuged at 15 700 x *g* for 2 minutes. The resulting supernatant,

containing the extracted bacterial proteins, was used for MALDI analysis. One μL of each sample supernatant was pipetted, in duplicate, onto the target and allowed to air-dry (Bruker Daltonics, Inc. 2012, Pranada et al. 2016). Each spot was then overlaid by 1 μL of matrix. When the matrix was dry, the target was loaded into the Microflex LT. While Bruker states that the same volume of FA and ACN must be used during this extraction process and recommends using 25 μL of each for well-defined pellets, smaller volumes may be used when very small pellets, or only white film, is visible in the tube. The extraction method above was repeated using volumes of 20 μL, 18 μL, and 15 μL each of FA and ACN.

3.3.10.4 Ethanol-formic acid extraction: addition of one or two washes using H_2O and EtOH

As tested for the direct transfer method, the addition of 1 or 2 washes using either UHPLC-grade H_2O or EtOH was examined for the EtOH/FA extraction method. The washes were added at the same step in the protocol and in the same volume as described above.

3.3.10.5 Sample, FA, and ACN mixed with matrix prior to deposition (Mixed Prep Method)

Some sample spots overlaid by matrix In a few of the previous MALDI runs had dried to produce a clumpy appearance. Homogenous co-crystallization of matrix with analyte is important in achieving good resolution (Hajduk et al. 2016). In a study conducted by AlMasoud et al. (2014), a 'mix method' sample deposition approach was tested in which sample and matrix were mixed prior to placement on the target. The authors reported that this technique had resulted in improved desorption and ionization as compared to several other methods tested. While the type of matrix and sample processing used in the 2014 study differs from that of the current study, this 'mix method' approach was tested for its potential to improve serovar identification. To test this method, serial dilutions of serovar Grippotyphosa and UHPLC-grade H_2O were prepped using the basic EtOH/FA extraction method described above. After the FA and ACN had been added to each sample tube, the tubes were vortexed and centrifuged at 15 700 x *g* for 2 minutes. Five μL of the supernatant from each sample dilution tube was transferred to a microcentrifuge tube containing 5 μL of matrix. These new sample tubes were then vortexed to thoroughly mix the suspension. Next, tubes were spun at 15 700 x *g* for 30 seconds to collect the mixture at the bottom of the tube. One μL of each mixture was then spotted in duplicate on the target.

A slightly different mixed method was tested as a way to potentially get the same results while using fewer microcentrifuge tubes. Serovar Pomona dilutions were used for this trial. In this approach, 5 μL of the centrifuged sample supernatant were deposited as a droplet onto parafilm. Five μL of matrix were then deposited on top of the supernatant. The sample and matrix were mixed via pipette and 1 μL of the mixture was then aspirated and deposited onto the MALDI target. Results from the two mix method approaches were then compared.

3.3.10.6 The Sandwich Method

A study conducted by Kuehl et al. (2011), found that a 'sandwich' preparation method, compared to several others tested, yielded the highest quality mass spectra obtained for *Enterococcus faecalis*. Based on these results, a similar layered matrix and sample deposition method was tested in two trials. The first trial used 2-fold serial dilutions of serovar Autumnalis, the second used dilutions of serovar Pomona. In the first trial, 1 μL of matrix was deposited via pipette onto target sample spots, two per dilution, and allowed to air-dry. Then, a portion of the pelleted material from each dilution was placed over two dried matrix spots using the direct transfer method, described above. Finally, 1 μL of the matrix was overlaid on each sample spot. Once dry, the target was analyzed. The second trial was performed similarly to the first; however, the samples were prepared using the tube extraction method. For sample deposition on the MALDI target, 1 μL of each dilution was spotted in duplicate via pipette over the initial matrix layer. Lastly, 1 μL of the matrix was overlaid on each sample spot.

3.3.11 MALDI-TOF MS Spectra Acquisition

Raw sample spectra were acquired using the Microflex™ LT mass spectrometer running FlexControl™ software. Measurements were performed using a 20 Hz N2 Nitrogen laser, ion source one set to 20 kV, ion source two set to 17.95 kV, and pulsed-ion extraction time set at 140 ns. MSP 96-spot ground steel target plates (Bruker Daltonics, Inc.) were used for sample analysis. Once all samples were loaded onto a sample target plate, it was placed in the MALDI's ionization chamber. A vacuum was then established, and calibration and validation procedures were automatically performed. Spectra were acquired in linear positive mode within an *m/z* range of 2,000-20,000 Da. Each spot was measured in 40-shot steps, for a total of 240 accumulated laser shots per spectrum. Smoothing, baseline adjustment and averaging of the resulting *Leptospira* spectra were performed in automatic mode. Raw spectra acquired for the creation of Main Spectrum Profiles and for serovar identification testing were obtained from serovar subcultures grown under the same culture protocol but collected on different days.

3.3.12 MALDI-TOF MS Real Time Identification

Real time identification was performed on sample spectra using Biotyper (RTC) software. Biotyper RTC uses a pattern-matching algorithm to compare analyte spectra with reference spectra (MSPs) included in the standard Bruker Daltonics database (BDAL) (version 8, 7854 reference spectra). The pattern-matching algorithm looks for similarities in peak positions, frequencies, and intensities between sample and reference spectra (Bruker Daltonics, Inc. 2012). Real-time identification returns a list of the 10 most likely microorganism matches. Each potential match includes the scientific name of the organism and a logarithmic score (LS) indicating the probability of a reliable match at a genus or species level. Other results can include 'No Peaks Found' or 'No Organism ID Possible', depending on sample quality, freshness of reagents, and concentration of bacteria in the sample, among other factors. A result of 'No

Organism ID Possible' indicates that the MALDI was able to acquire a spectrum from the sample, but that the number, mass, and frequency of peaks that make up the spectrum are not similar enough to those of reference spectra in the MALDI database to return a match with any confidence. This result might occur when a sample does not produce a spectrum of good quality, or when a reference spectrum for the sample organisms is not included in the reference database. Scores range from 0.00, (indicating no similarity) to 3.00 (indicating a perfect match). Bruker guides score interpretation as follows: values from 2.30-3.00 indicate highly probable species identification; values from 2.00-2.29 imply a reliable genus and probable species identification, values from 1.70-1.99 indicate probable identification at only the genus level, and scores from 0.00-1.69 signify no reliable identification. User-created MSPs and reference libraries can also be used for identification.

3.3.13 Influence of Sample Concentration on Identification Accuracy

The lowest concentrations that returned accurate serovar- and genus- level MALDI-TOF identifications for *Leptospira* samples was determined using the results of both specificity and sensitivity trials, fluorometric measurements of DNA template to estimate the concentration of each dilution, and qPCR of serovar dilutions to correlate C_T values with dilution concentration. (The reader is directed to the following chapter on Sensitivity for further detail.)

3.3.14 Creation of the Original *Leptospira* Main Spectrum Profile Reference Library

To create a MALDI *Leptospira* mass spectra reference library, Main Spectrum Profiles (MSPs) were created for each serovar. To acquire raw spectra for MSP creation, serovar culture suspensions were adjusted to 25 %T, then prepared using the TE method. The prepped sample for each serovar was spotted on eight positions of the MALDI target, creating eight technical replicates. Spectra were acquired using the AutoXecute feature in Bruker's FlexControl™ software. Each of the eight sample spots was measured in triplicate, yielding 24 raw spectra. These spectra were loaded into FlexAnalysis software (Compass for flexSeries 1.4, v. 3.4 (Build 76), Bruker Daltonik GmbH) to identify those spectra with characteristics that would preclude their use in MSP creation. Spectra were preprocessed using Savitzky-Golay smoothing, Top Hat baseline subtraction, and peak detection. Mass peak lists were generated for each spectrum using a centroid detection algorithm, a signal-to-noise (S/N) threshold of two, and a maximum peak number of 300. These lists identified the most significant peaks for each spectrum based on peak *m/z* value, relative intensity, area, and frequency. Overlaid and stacked spectra were then visually inspected to identify flatline spectra, spectra with few picked peaks, mass peak shifts, hot spots, (peaks common in all raw spectra for a particular serovar that exhibited highly variable signal intensity), and other anomalies. Spectra with these characteristics were noted for exclusion. Spectra were then loaded into MBT Compass Explorer software for MSP creation. Spectra that had been identified as unsuitable for inclusion in MSPs were removed. The remaining spectra were again preprocessed using

the default settings, which included a mass adjustment lower bound of 3000 Da and upper bound of 15000 Da; a resolution of one; a spectra compressing factor of 10, Savitsky-Golay smoothing with a frame size of 25 Da, baseline correction via the Multipolygon method with a search window of 5 Da in two runs, normalization, and peak detection with a signal-to-noise ratio of three. Default settings were also used for the MSP creation, and included: a maximum mass error of 2000 for each single spectrum, a desired MSP mass error of 200, a 25% desired peak frequency minimum, and a maximum desired MSP peak number of 70. A minimum of 18 raw spectra were used to create each serovar's MSP. Each MSP's constituent raw spectra were then tested against it for identification. Per Bruker's recommended MSP creation guidelines, raw spectra should match to their corresponding MSP with a log score of at least 2.75 (Bruker Daltonik GmbH 2016). Raw spectra were also tested against the Bruker reference database to ensure that they did not match with the MSPs of other bacteria. Once an MSP for each serovar had been created, they were combined into a custom *Leptospira* reference library.

3.3.15 Weighted Pattern Matching and Subtyping MSPs

MALDI differentiation of organisms below the species-level usually requires different or additional sample prep methods and spectra analysis beyond that recommended for standard MSP creation (Basile and Mignon 2016, Zhang and Sandrin 2016). Several studies have reported the use of weighted pattern matching to improve custom MSP specificity. In this approach, weights assigned to MSP peaks are manually adjusted so that highly discriminative peaks are assigned a higher weight than that assigned to nonspecific peaks found in multiple MSPs (Freiwald and Sauer 2009, Dieckmann and Malorny 2011). With this in mind, MSP peak lists were compared to identify peaks that could potentially differentiate the serovars. Peak signals found in only one of the seven peak lists were noted as potential biomarkers and for manual weight adjustment. However, when manual editing of peak weights in Compass Explorer was attempted, it was discovered that Bruker had disabled this option.

Since user-defined weighted pattern matching via this program was no longer accessible, subtyping MSPs were created. This is a different method of altering peak weights of reference spectra, albeit with less user-control. Subtyping MSPs are created by applying a weighted scoring algorithm to peaks in at least two selected, standard (parent) MSPs for which greater specificity is desired. Subtyping mass peak lists consist of the same mass peaks found in that of the parent MSPs, but with different weights assigned to distinguishing peaks. Those peaks found in fewer MSPs are assigned a higher weight than peaks common among all selected MSPs (Bruker Daltonik GmbH 2016). Subtyping MSPs were first created for serovars that had misidentified as a different serovar in several previous identification trials. Serovar Icterohaemorrhagiae had identified as serovar Hardjo in several runs. Therefore, the parent MSPs for these two serovars were used to create two subtyping MSPs. The same was done for serovar Grippotyphosa, which had sometimes identified as Bratislava, and for Pomona, which occasionally matched to Canicola. Identification with subtyping MSPs must be performed within Compass Explorer,

and cannot be used for a real-time run, which uses Biotyper RTC software. Therefore, it is recommended that analyte spectra intended for subtyping identification be acquired using the AutoXecute option in FlexControl, which allows the user to save the spectra under a name, and in a location where they can be easily located for subsequent loading into Compass Explorer. Serovar raw spectra saved from two previous trial runs, and which had not been used to create the parent MSPs, were used to test the subtyping MSPs. Once these raw spectra were preprocessed, the identification workflow was run. Match scores were then generated for both standard and subtyping MSP identification. Results were compared between the two types of MSPs to determine if subtyping MSPs improved specificity. A second set of subtyping MSPs, for all seven serovars, were created using all seven parent MSPs. This was done to determine whether using additional standard MSPs to create subtyping MSPs, which would result in different subtyping MSP peak weighting, would improve subtyping MSP specificity. This new set of subtyping MSPs was tested using newly acquired raw spectra for each serovar. Identification results given by parent and subtyping MSPs were then compared.

3.3.16 Manually Editing Peaks in FlexAnalysis

Since differentially weighting MSP peaks had not improved identification results, focus shifted to peaks found in the raw spectra. FlexAnalysis allows manual editing of peaks in raw spectra peak lists. To determine if these peak edits could be used in the creation of new, more specific MSPs, the raw spectra used in serovar Hardjo's original MSP were loaded and their peak lists were compared to that of the MSP. If a potential biomarker present in the MSP list was absent from a raw spectrum's list, that spectrum was visually examined to see if a peak of similar *m/z* was present. Peaks may be visible in a spectrum, but be absent from that spectrum's selected peak list, potentially due to a relative measurement (e.g. lower relative intensity). Peaks of interest, present in a raw spectrum but missing from the spectrum's mass peak list, were manually added to the list and saved. These raw spectra with manually edited lists were then loaded into Compass Explorer, and a new Hardjo MSP was created. The new MSP peak list was reviewed, with particular attention paid to the frequency listed for each potential biomarker. An increase in frequency would indicate that the manual peak changes had been retained and incorporated into the new MSP.

3.3.17 Comparison of MSP and Raw Spectra Peak Lists to Evaluate the Raw Spectra Combinations Used to Create Each MSP

It was found that manual peak edits could not be incorporated in the MSP creation workflow. So, the combination of raw spectra used to create each MSP was then considered. Peak lists for each MSP and its constituent spectra were compared. Raw spectra found to be missing any MSP-specific peaks were noted to determine whether their removal from the MSP would be an appropriate way to increase MSP specificity.

3.3.18 Incorporation of Additional Raw Spectra and Modification of MSP Creation Parameters

Sets of raw spectra for each serovar, those used to create the serovar's MSP and those used to test the MSP, collected on different dates, were visually compared in FlexAnalysis. Each set of spectra belonging to a particular serovar displayed noticeable differences. Some peaks, found in all sets of spectra, showed different levels of intensity between sets. There were also some differences in peak presence and absence between sets. Since these different sets of spectra were acquired using different subcultures, some of these differences were expected. However, it was not known to what degree these differences contributed to serovar misidentifications.

To determine whether including additional sets of spectra in an MSP would improve specificity, three new MSPs, incorporating two new sets of raw spectra, were created for each serovar. Each of the new MSPs was comprised of the initial set of spectra used in the original MSP, plus two new sets of raw spectra, for a total of three sets of raw spectra per MSP. The two new sets of spectra collected for each serovar were acquired from two different subcultures collected on different days. These subcultures were grown using the same protocol as was used for the original culture but used different lots of culture media. Subculture samples were prepped using the same method used for the original samples but used different lots of MALDI reagents. Each new set consisted of 24 raw spectra.

It was predicted that MSPs incorporating these new spectra, acquired across a broader range of conditions, would better tolerate variability in sample spectra and improve identification accuracy. FlexAnalysis was used to examine the new spectra and identify those spectra that should be omitted from MSP creation. Original and new sets of spectra were then combined in Compass Explorer to create new MSPs, each consisting of 67 to 72 raw spectra representing three biological replicates.

The first of the three new MSPs was created using the same parameters as were used for the original MSP. This included a desired maximum peak number of 70. Identification trial results from the original MSP and this new MSP were compared to determine whether an increase in the number of raw spectra sets used in MSP creation alone improved MSP specificity. The second and third new MSPs for each serovar were generated using maximum desired MSP peak numbers of 100 and 125, respectively. Since the original and the first new MSPs were created using a maximum MSP peak number of 70, these MSPs were created to determine whether an increase in both the number of raw spectra sets and the maximum number of MSP peaks would improve MSP specificity. The three new MSPs and original MSP were then tested to compare their ability to correctly identify the corresponding serovar.

An additional MSP was created for serovars Grippotyphosa and Bratislava. Previously-created MSPs for these two serovars had typically resulted in a greater number of misidentifications compared to the results for other serovars. Raw spectra for these two serovars reviewed in FlexAnalysis showed numerous peaks in the 2000–3000 Da range. However, the default preprocessing method within Compass Explorer has a peak picking lower boundary set at 3000 Da. Raw spectra loaded into Compass

Explorer for MSP Creation are therefore preprocessed using a method that does not allow for MSP peak selection below the 3000 Da limit. This potentially excluded peaks important for serovar differentiation. To address this, a new preprocessing method was created that modified the lower boundary to 2000 Da. This method was then used to create a new MSP for Grippotyphosa and Bratislava. These MSPs used all three sets of raw spectra, a lower preprocessing boundary of 2000 Da, and a maximum desired MSP peak number of 100. These MSPs were then tested for serovar identification.

A few additional MSPs were created using other modified preprocessing or MSP creation parameters but were only tested once and eliminated after poor identification results. One such example was the creation of an MSP that used a desired mass error of 100, rather than the default of 200.

3.3.19 MSP Identification Trials

To assess MSP specificity across a range of dilutions, MSPs were tested using six 2-fold serial dilutions of serovar cultures mixed with either UHPLC-grade H_2O or canine urine. The %T of *Leptospira* stock cultures, measured weekly to assess growth, is typically in the range of 45-64 %T. It was found that the initial concentration had to be adjusted to approximately 25 %T so that four to five of the serial dilutions could be detected by the MALDI. Dilutions ranged from roughly 3.01×10^9 - 1.01×10^8 organisms/mL.

MSP specificity trials used *Leptospira* MSP test libraries created in Compass Explorer. These libraries consisted of the new MSP being evaluated and at least one other MSP selected for each of the other serovars. Tests were performed with either the Biotyper RTC or Compass Explorer software, depending on the raw spectra used for the test and the number of MSP combinations being tested, among other factors. MSPs were tested individually and in group trials**.** Individual trials were used to test one particular MSP for a serovar against one MSP selected for each of the other serovars. Group trials were used to test two or more MSPs for a serovar against each other and against one MSP selected for each of the other serovars. The number and combination of MSPs used in each test were chosen based on results of prior MSP trials, the number and type of MSPs that had been created up to the point at which a new MSP was tested, and the need to create additional MSPs with greater specificity for certain serovars. Selected combinations were chosen to reflect the potential group of MSPs that would eventually be chosen for the final *Leptospira* MSP library. The two trial types were used because it was important to evaluate not only how well an MSP identified its corresponding serovar, but how an MSP performed against other MSPs for the same serovar and how it performed as part of the groups of MSPs that would potentially make up the *Leptospira* MSP library. MSP identification results, both log scores and match rank, can change based on the other MSPs against which it is tested. This means, for example, that an MSP's identification score for a correct first match in an individual trial may differ from the score obtained for the same first match in a group trial. It is not uncommon for MSPs of closely-related serovars to misidentify as each other. MSPs created with certain parameters may offer more specificity and reliability, than MSPs created with different

parameters. These performance differences may be emphasized when MSPs for closely-related organisms are combined in an MSP reference library. Evaluation of MSP performance was based on several factors. MSPs that returned the highest number of correct first matches to the serovar-level with the highest identification scores, across the widest range of tested dilutions, and which appeared less frequently as incorrect matches (false negatives) in trials of MSPs for other serovars, were considered high-performing. In individual trials, MSPs that returned correct serovar-level first matches with scores ≥1.7 were selected for further testing. In group trials, MSPs that returned correct serovar-level identifications, with scores both ≥ 1.7 and higher than scores for matches to other serovars, were chosen for subsequent testing. The Biotyper RTC software used for the trials returns a list of up to 10 match identifications with corresponding scores. The first match has the highest score, and subsequent matches have decreasing scores. For example, a group trial for serovar Bratislava could include three different Bratislava MSPs. A Bratislava MSP identification returned as a third match would still be considered a correct match as long as it had a score ≥1.70 and the two matches with higher scores were identified as the other two Bratislava MSPs. This third match, returning an accurate identification of the Bratislava test sample, would still have a higher score than subsequent matches identified as any other serovar. Only one MSP per serovar would be selected for inclusion in the final *Leptospira* reference library. If the Bratislava MSP that returned the third match in this example was selected to represent Bratislava in the reference library, the other two Bratislava MSPs would not be included in the library. If the same Bratislava sample used in the group trial was then run in an individual trial against the *Leptospira* reference library, the Bratislava MSP that had returned a third match, would then return a first match. Removing the two Bratislava MSPs that had returned the first two matches in the group trial would still leave a Bratislava MSP as the third match with a higher score than that for any other serovar, and thus, would equate a correct match. MSPs that did not perform well in individual or group trials were not tested further.

3.3.20 ClinProTools

To identify characteristics of MSP constituent spectra that potentially contributed to the misidentification of some serovars, ClinProTools (CPT) software was used for raw spectra analysis. Some analyses used two subsets of raw spectra per serovar. These subsets, designated as Training and Test subsets, consisted of the raw spectra that had been collected and used for serovar MSP creation. Each subset consisted of 27 raw spectra. Recall that each of the three sets of raw spectra used to create the MSPs, (with the exception of X Original, which used only one set of raw spectra), were collected on different dates (three experimental replicates). Nine spectra, from each of these three dates, were selected for a serovar's Training subset. Nine different spectra, also from each of these three dates, were assigned to each serovar's Test subset. These represented nine technical replicates. When possible, at least one spectrum acquired from each MALDI target sample spot on each of the collection dates, was

assigned to a serovar's Training subset and a different spectrum from each of these spots was assigned to the Test subset. This gave each subset 24 raw spectra. The three additional spectra per subset were selected from the remaining spectra that had not yet been assigned and was based on spectra quality. For example, some spectra contained a higher relative number of peaks compared with that of the other remaining spectra, and so were considered preferable for subset inclusion. When at least two different spectra from each target spot were not available, (for example, due to flatline spectra), raw spectra were assigned such that the same number of spectra from each collection date, and from as many of the eight target sample spots per date as possible, were used for each subset. In this way, the combination of factors present during each collection date that may have affected the raw spectra generated on those dates, would be equally represented in each subset. Spectra in the Training subsets were used for peak calculation to identify statistically significant, serovar-discriminatory peaks and to create CPT classification models for the seven serovar group. Training subset spectra were also used to create the classification models for the Bratislava and Grippotyphosa pair comparison. Test subset spectra were used to test the ability of the generated classification models to separate spectra into the appropriate serovar class. Combinations of both subsets were used for peak calculation in the pairwise comparison and in additional CPT analyses including creation of 2D peak distribution plots, visual examination of raw spectra using virtual gel and spectral views, and Principal Component Analysis (PCA) for both the seven serovar and pairwise comparisons.

After loading the raw spectra into CPT, the spectra were initially prepped to minimize the effect of measurement variations in spectra analysis. Data prep included Top Hat baseline subtraction with a 10% minimum baseline width, normalization of each spectrum against its total ion count, and recalibration with a maximal peak shift of 500 parts per million (ppm) and 30% match to calibrant peaks. The total average spectrum was used for peak picking and peak calculation at a resolution of 800 ppm with a S/N of five. Peak calculations were run twice, once using peak intensities and once using peak areas, to determine which method provided better insight into significant spectra differences between serovars.

Analyses were performed on the group of seven pathogenic serovars and on the serovar pair of Grippotyphosa and Bratislava. For both comparisons, picked peaks were used to calculate an average peak list, corresponding peak univariate statistics, and a Peak Statistic report (PSR). The report included Anderson-Darling p-values (PADs) calculated for each selected peak to determine peak distribution between the serovars. Peaks with calculated PADs ≤ 0.05 indicated a non-normal distribution. These peaks were further evaluated using either the Wilcoxon or Kruskal-Wallis (WKW) test, based on the number of serovar classes being compared. A class is a group of raw spectra, loaded together into CPT, and treated as a separate group from other loaded groups of spectra. In this study, raw spectra for each serovar were loaded into CPT as different groups, which allowed raw spectra data to be compared between serovars. Peaks with PADs > 0.05, which indicated a normal distribution within the loaded serovar classes, were further evaluated using Welch's t-test or ANOVA, depending on the number of

classes compared. In ClinProTools, p-values generated by WKW tests (PWKWs) or by TTA tests (PTTA) indicate the probability that differences in peak areas/intensities between loaded serovar classes can be observed by chance. The lower the p-value, the better the peak's ability to act as a distinguishing marker for class differentiation (Ketterlinus et al. 2005). P-values ≤ 0.05 were interpreted as significant, while p-values ≤ 0.01 were considered highly significant. These values were then used to sort picked peaks according to their ability to discriminate between loaded serovar classes. Given the large number of peaks and corresponding data for each spectrum, this report helped to identify subsets of peaks on which to focus. The top six discriminatory peaks for the serovar group and serovar pair comparisons, for which at least one of the serovars had an average peak area/intensity value of ≥ 1.00, were used for further analysis. These peaks were examined in the raw spectra using spectral and virtual gel views to evaluate peak intensity differences both within and among serovars.

3.3.20.1 Two-dimensional (2D) peak distribution plots

The top discriminatory peaks identified for both the serovar group and serovar pair comparison were further evaluated using two-dimensional (2D) peak distribution plots. This type of plot allows the user to evaluate two peaks at a time, chosen from peaks listed in the Peak Statistic Report. Since peaks used in model creation are selected from peaks included in the PSR, these plots can be used to analyze all discriminatory peaks identified by both the PSR and the models. Peak pairs were plotted for all serovars to display the distribution of the peaks' average intensity values within the serovar raw spectra. Ellipses were used to visualize the standard deviation of a peak's average intensity within each serovar class. Plots were used to identify raw spectra in which peak data overlapped. Convergent peak distributions may contribute to decreased MSP specificity. This is because peaks identified as important for serovar differentiation are less different in raw spectra that display overlap than in raw spectra which show no overlap. For example, plots using the top discriminatory peaks for serovars Bratislava and Grippotyphosa were created because, when these serovars identified incorrectly in specificity trials, they most often were misidentified as each other. These plots were generated using all of the raw spectra in Grippotyphosa and Bratislava MSPs, minus spectra of poor quality. Distributions for the top discriminatory peaks were evaluated. Raw spectra identified with overlapping peak intensity distributions were noted. These spectra were omitted from a new MSP created for each of these two serovars that consisted of a maximum desired MSP peak number of 125, a lower preprocessing boundary of 3000 Da, and the same three sets of raw spectra used in previously-created MSPs. These new MSPs were then tested for specificity. Additional plots were created for these two serovars using the top discriminatory peaks identified by a second Peak Statistic report using a nonparametric sort-order. A new set of raw spectra for Bratislava and Grippotyphosa were identified as having overlapping peak distributions. These spectra were noted and later combined with data from Principal Component Analysis and classification models to create four new MSPs for each of these two serovars. Two of these MSPs used the default lower peak-picking boundary

of 3000 Da. One of these two used a maximum peak number of 70, and the other used a maximum peak number of 100. The third and fourth MSPs of this new set used a lower peak-picking boundary of 2000 Da. Again, one of these MSPs used a maximum peak number of 70, while the other used 100.

3.3.20.2 Principal Component Analysis (PCA)

To further investigate which peaks and raw spectra exhibited the largest influence on a dataset, Principal Component Analysis was performed. The top three principal components (PC1-PC3) explaining the majority of the variance in the loaded serovar classes were examined for clusters, outliers, and patterns within and among the spectra and peaks that could reveal variables important for serovar differentiation. Attention was focused on informative variables which had not been revealed in the previously-performed analyses. Scores plots, which display the distribution of raw spectra for each PC, were searched for the raw spectra with the most influence on each PC. Loadings plots were investigated for peaks with relatively large absolute values, indicating peaks influential for data separation. PCAs were created using the Pareto scaling, which divides the intensity of each peak by the square root of the standard deviation of this peak within the dataset. This allowed for emphasis of lower intensity peaks, some of which had been identified as top discriminatory peaks in the PSR, while partially preserving the intensity structure in the data. PCA was performed on the group of seven serovars and on the serovar pair of Grippotyphosa and Bratislava. Influential variables were assessed for their potential ability to increase MSP specificity via their omission or emphasis in subsequent MSP creation.

3.3.20.3 Model creation

To further evaluate the discriminatory power of selected peaks, three types of pattern recognition models were created in CPT. Models were generated for the serovar group and serovar pair comparisons using the raw spectra assigned to each serovar's training subset. Calculations were based on the average spectrum for each serovar class. Each model selected a list of the top discriminatory peaks for serovar differentiation. The QuickClassifier (QC) model, which uses a univariate peak detection algorithm, predicts the class to which a raw spectrum belongs and calculates a likeliness for membership in each possible class. A peak sort mode can be selected for this model. This mode determines peak ranking and assigned peak weights. The QC model was created using a P-Val WKW sort mode, which uses the p-values calculated by either the Wilcoxon or Kruskal-Wallis test (depending on the number of loaded spectra classes) in the PSR for the loaded spectra. Default settings were used for all other parameters. The Supervised Neural Network (SNN) model, which uses a multivariate algorithm for peak detection, identifies characteristic spectra from each loaded class and uses those spectra as prototypes for class calculations (Bruker Daltonik GmbH 2007). All default settings were used for this model, including automatic detection of prototype number. The Genetic Algorithm (GA) model, which uses a different multivariate peak detection algorithm, is designed to emulate the evolutionary process by iteratively

testing different peak combinations to evaluate their “fitness” for spectra class separation. In each iteration, peak combinations that improve spectra classification over previously-tested combinations are retained and used for the next comparison cycle. Iterations include “random mutation” and “crossover” events that contribute to the selection of peak combinations tested (Pusch and Kostrzewa 2005). The model finally selects the combination that best differentiates between loaded spectra classes.
This model was applied using the default parameters, which included a maximum of 5 peaks in the model, 50 maximum generations, 100 peak combinations, a mutation rate of 0.2, and a crossover rate of 0.5.

Each model selected a different number and group of discriminatory peaks. This was expected, since each model uses a different peak detection algorithm. The number of peaks selected by a model depends on the detection mode (i.e. manual or automatic) used. The type of mode a model can use depends on the model type. The QC and SNN models use the automatic mode. This mode allows the model to determine the best number of peaks up to a maximum peak number of 25 (limited to no more than 25 to maintain reasonable processing times). The GA model always uses manual mode. The default number of peaks for this model is set at five. So, this model will select five discriminatory peaks unless the number of allowable peaks is specifically changed. Since the maximum number of peaks that can be selected for the QC and SNN models is higher than that of the default for the GA model, the QC and SNN models will usually return a higher number of selected peaks than the GA model. Ultimately, the number of discriminating peaks used in a model is not indicative in itself of confidence level, but rather a result of the way these models are designed to work in CPT.

To assess performance, recognition capability (RC) and cross validation (CV) values were calculated for each model. Recognition Capability describes a model’s ability to accurately classify constituent spectra (Bruker Daltonik GmbH 2007). This percentage is calculated by dividing the number of correctly classified component spectra by the total number of spectra used in the model. Cross validation indicates how well a model handles intraclass variability among sample spectra, and so can be used to gauge model reliability (Bruker Daltonik GmbH 2007). Several cross-validation methods are available in CPT. For the purposes of this study, a leave-one-out cross validation method was applied. In this method, one sample from the data set is randomly selected and removed, while the remaining samples are then used to generate a model using the chosen algorithm. The sample that was omitted is then classified against the model. This process is repeated using each of the samples one time. Once all samples in the data set have been classified, the results are averaged, normalized, and returned as a CV value, which reflects the model’s potential predictive capability (Bruker Daltonik GmbH 2007). Bruker recommends this method for data sets containing less than 30 spectra per class. Since 27 spectra were used for each serovar class, the leave one out (LOO) method was selected for cross validation. Together, RC and CV values can be used to predict the future performance of a model (Ketterlinus et al. 2005). These two parameters

were evaluated for each model within the serovar group and serovar pair comparisons. For each comparison, the model with highest RC and CV values was determined to be the most robust.

To narrow down the number of peaks used in subsequent analyses, the set of peaks identified by the most robust model for each serovar comparison was combined with the top six discriminatory peaks identified by the corresponding Peak Statistic report. These peak combinations were explored using 2D Peak Distribution plots and virtual gel and spectral views. In particular, analyses were performed with the goal of increasing MSP differentiation between serovars Grippotyphosa and Bratislava.

3.3.20.4 External validation of classification models

External validation of CPT models is another method of evaluating model predictive capability, including whether a model can correctly predict some classes better than others (Bruker Daltonik GmbH 2007). Validation was performed using the test subsets of raw spectra for each serovar represented in the model. Results showed how each raw spectrum in the test subset was classified by the model. The percent of correctly classified spectra for each serovar class in each model was then determined. These values for each model were averaged to obtain the overall percentage of correctly classified spectra for each model type These average values were used as a supplementary criterion for predicting future model performance.

3.3.21 Comparison of Discriminatory Peaks Selected by the Peak Statistic Report and Three Classification Models to Differentiating Peaks Identified in the Literature

Four studies have identified potentially unique peaks for certain *Leptospira* species and serovars (Rettinger et al. 2012, Calderaro et al. 2014b, Karcher et al. 2018, Sonthayanon et al. 2019). MSP peak lists were reviewed for the unique serovar peaks and associated characteristics noted in these studies. This information was then compared with the discriminatory peaks selected in this study to learn whether serovar-unique peaks and particular characteristics were consistent across studies. If so, this data would be used to evaluate individual raw spectra in each serovar's MSP to determine if certain spectra should be omitted to create more specific MSPs.

3.3.22 Blind Trials

After testing the multiple types of MSPs created for each serovar, a final *Leptospira* MSP reference library was created using the best-performing MSP for each serovar. This library was then tested in two blind trials using serovar subcultures adjusted to 50 %T and prepared via tube extraction. In the first trial, serovar cultures were adjusted to 50 %T, pelleted, then resuspended in 1 mL of HPLC-grade H_2O before tube extraction. For the second trial, adjusted and pelleted serovar cultures were mixed with 1 mL of qPCR-leptospirosis-negative canine urine prior to tube extraction. Prepared sample tubes were given to a second technician, who blind coded the tubes, and returned them to the first technician for deposition on

the MALDI target. A banked sample of *Brachyspira innocens* and a banked sample of *Brachyspira pilosicoli* were cultured, adjusted to 50 %T of HPLC-grade H_2O, then prepped via TE and blind coded for use in the trial. These two species are the organisms most closely-related to *Leptospira* represented in the Bruker Daltonics reference database (BDAL). Their inclusion was used to test whether the *Leptospira* MSPs could differentiate between the two genera. In each trial, one μL of each sample was deposited onto eight spots of two MALDI targets for a total of 16 technical replicates per serovar. An additional two spots on each target were used for BTS (positive control), while one spot per target was used for non-inoculated matrix (negative control). Samples were run against the BDAL commercial databases provided with the Biotyper unit, (representing approximately 7854 organisms), all in-house created MSP libraries (representing approximately 100 organisms), and the *Leptospira* MSP library. Score interpretation followed the recommended guidelines provided by Bruker (Bruker Daltonics, Inc. 2012).

3.3.23 Data Analysis

All raw spectra and peak statistical analyses were performed using MBT Compass Explorer and ClinProTools software. Blind trial results were analyzed for sensitivity and specificity using 2 x 2 contingency tables. For 2 x 2 analysis, the particular MSP being evaluated was designated as the 'in' group and all other samples designated as the 'out' group. Definitions were as follows:

True positive (TP): an 'in' group sample spot identified correctly

False positive (FP): an 'out' group spot identified as the 'in' group

False negative (FN): an 'in' group spot identified as an 'out' group

True negative (TN): an 'out' group spot identified as an 'out' group sample

3.4 RESULTS

3.4.1 Dilutions

The range of serial two-fold sample dilutions noted above, used throughout the majority of this project, were chosen so that at least four of the six dilutions could be detected by the MALDI. In MSP identification trials, the 1:16 and 1:32 dilutions often resulted in no peaks found due to serovar concentrations below that which the MALDI can detect for *Leptospira*.

3.4.2 MALDI-TOF MS Real Time Identification

All real-time identification runs yielded no peaks for uninoculated matrix sample spots, which served as the negative control. Positive control sample spots, using the Bruker bacterial test standard (BTS) were positively identified, yielding expected identification log scores between 2.00 and 2.50.

3.4.3 Sample Preparation Methods

A variety of sample preparation methods were evaluated to determine which technique generated

high quality spectra that produced consistently accurate identification results to the serovar level.

3.4.3.1 The direct transfer method and variations

Leptospira's unique nutritional requirements necessitate special media for culture (Ellinghausen and McCullough 1965, Staneck et al. 1973, Faine et al. 1999). Serovar cultures used as antigen in the microscopic agglutination test must be grown in broth media because agar present in solid and semi-solid media makes accurate interpretation of test results difficult (Turner 1970). Therefore, specific, highly-enriched broth media is most often used for the growth and maintenance of *Leptospira* in veterinary diagnostic labs.

Leptospira grown in liquid media do not readily accommodate the standard direct transfer sample preparation method that can be used for many bacteria grown on solid media. To adapt this method for use with this organism, *Leptospira* subcultures were used to create serial two-fold dilutions ranging from 3.01×10^9 - 1.01×10^8 organisms/mL. Dilutions were centrifuged at 15 700 x *g* for 2 minutes to pellet the bacteria, and pellet material was spotted onto the MALDI target. As in the typical direct transfer process, a wooden applicator stick was used to collect a portion of the sample to spot on the target. However, this approach proved to be impractical. First, all but the most concentrated samples had little or no formed pellet after centrifugation. The higher dilutions tested, which better reflect concentrations most often seen in leptospirosis-positive canine urine samples, did not yield enough pellet material to collect and transfer evenly onto duplicate target sample spots. Second, the pellets produced by more concentrated samples were partially absorbed by the wooden applicator stick and, therefore; resulted in an insufficient amount of material available to disperse onto target spots evenly. Uneven application of sample on the target results in lower-quality spectra with fewer recognized peaks. The MALDI internal camera, which allows the user to see a magnified view of the target sample spot being analyzed during identification runs, emphasized the uneven sample distribution. Laser bombardment of a sample occurs only at particular positions (raster spots) within the sample spot. These positions are chosen based on the pre-defined pattern or random walk movement option selected. This means that during sample analysis, certain positions within a sample spot will not be irradiated by the laser. If a sample is unevenly distributed on a spot, and the laser targets several raster positions with little to no sample, too few sample ions may be generated to yield an accurate, reliable result. Most sample spots in the direct transfer trials did not meet the required number of acceptable peaks for sample identification and resulted in no peaks found. The addition of formic acid to each sample spot in the extended direct transfer method did not improve identification results. In fact, all sample spots using this method resulted in no peaks found. Similarly, incorporation of wash steps using UHPLC-grade H_2O or EtOH failed to improve spectra quality and accuracy. Ultimately, the direct transfer method for *Leptospira* identification yielded few positive identifications, with most sample spots returning no peaks found, regardless of the variations in technique applied. Therefore, any further attempts to optimize this protocol for *Leptospira* were not pursued.

3.4.3.2 Ethanol-formic acid extraction (tube extraction) and variations

Preliminary trials used one serovar to test techniques and develop an optimized sample preparation method for the study. Trials that yielded promising results were repeated using all seven serovars to ensure that the selected method yielded raw spectra of similar quality for each.

The addition of one wash yielded better quality spectra and identification results than the addition of two. UHPLC-grade H_2O, compared to ethanol, was found to be the preferable wash reagent. However, the addition of one wash with UHPLC-grade H_2O did not appreciably improve results over EtOH/FA sample prep with no wash. Therefore, the added wash was not used for sample prep in subsequent trials.

In the first trial of the mixed prep method, aliquots of serovar Grippotyphosa dilutions were mixed with an equivalent amount of matrix in microcentrifuge tubes prior to deposition on the MALDI target. This approach produced results with a greater number of incorrect first matches, and at lower scores, compared to runs using the basic EtOH/FA extraction method. In the second trial using parafilm, both sample spots for the starting concentration had no peaks found. One of two sample spots for the 1:2 dilution yielded a score of 1.50 with no organism identification possible, while the second spot resulted in no peaks found. The 1:4 dilution returned the correct identification for one sample spot with a score of 1.95, while the second spot was classified as no peaks found. All higher dilutions resulted in no peaks found. Given the poor quality of the spectra produced using these approaches and the subsequent lack of serovar identification, it was determined that a mixed prep method was not appropriate for use in whole-cell MALDI-TOF identification of *Leptospira*.

The sandwich preparation method, in which the sample was deposited between two layers of matrix, also yielded poor results. In the first trial using serovar Autumnalis, a correct first match was obtained for one of two undiluted sample spots, with a score of 2.03. However, all other tested dilutions returned a result of no peaks found. The poor results may have been partially due to the direct transfer method used for sample deposition for this trial. As mentioned above, tests of the direct transfer method did not yield good results. Therefore, the outcome of this trial was not unexpected. The second trial, which used the tube extraction method for sample preparation, produced a few additional identifications compared to the first trial. The two undiluted samples correctly identified as serovar Pomona with scores of 2.01 and 1.96. One of two spots for the 1:2 dilution was correctly identified as serovar Pomona with a score of 2.00, while the second spot was incorrectly identified as serovar Canicola with a score of 1.88. All other dilution spots yielded no peaks found. Given the poor performance of this approach for this study, no further trials using the sandwich method were performed.

Overall, results showed that the standard EtOH/FA extraction method consistently produced the highest quality spectra with the greatest number of peaks and resulted in the most consistently correct identifications to the serovar level. All further specificity and sensitivity trials were performed using this method.

3.4.3.3 Additional methods

Several additional sample concentration methods, used mainly for determining the MALDI-TOF MS sensitivity for *Leptospira* organisms, also yielded specificity results because all trials used the in-house created *Leptospira* reference library. The reader is referred to the Results section in Chapter 4 for the outcomes of trials using differential centrifugation and various filters for concentration and desalting.

3.4.4 Creation of the *Leptospira* Main Spectrum Profile Reference Library

Seven *Leptospira* reference strains were used to create the *Leptospira* MSP reference library. To construct the original MSPs, 24 raw spectra per serovar were acquired. After examining the quality of the raw spectra, 18-24 spectra per serovar were chosen for MSP creation. Constituent raw spectra tested against their corresponding MSP returned identification scores ≥ 2.75. Raw spectra tested against the Bruker Daltonics reference database returned a score of "< 0", indicating no match to the reference spectra of other organisms. Serovar MSPs were then combined to form a *Leptospira* library and initial specificity runs were performed.

3.4.5 Serovar MSP Identification Trials

Preliminary tests of the *Leptospira* reference library resulted in consistent, correct identification of *Leptospira* samples to the genus level for all serovars but produced mixed results for serovar-level recognition. As predicted, identification scores decreased with increasing dilutions, with the 1:16 and 1:32 dilutions often resulting in no peaks found. While this inverse correlation is expected, it was not known whether these results were attributed to MSP quality or to the MALDI's sensitivity. As subsequent identification tests were performed, additional, revised MSPs were created with the goal of increasing MSP specificity (Sp). New MSPs were created using variations of the original creation method, such as incorporation of additional raw spectra, modification of both spectra preprocessing and MSP creation method parameters, and selection of raw spectra for MSP inclusion or exclusion based on CPT analysis.

3.4.6 Weighted Pattern Matching and Subtyping MSPs

Leptospira subtyping MSPs were tested to determine whether they offered a higher level of identification accuracy compared to standard MSPs. In a few previous trials, some spectra for serovar Grippotyphosa had identified as Bratislava, some Icterohaemorrhagiae spectra had identified as Hardjo, and some Pomona spectra had identified as Canicola. Therefore, subtyping MSPs were created for these pairs of serovars. Previously saved raw spectra were then used for testing. Two saved Grippotyphosa spectra, acquired from duplicate MALDI target sample spots, were identified correctly by the parent Grippotyphosa MSP with scores of 2.46 and 2.44. The subtyping MSP identified the spectrum from the first sample spot correctly, with a score of 2.08. However, the spectrum from the second sample spot was

identified as Bratislava with a score of 2.06. While Grippotyphosa was the second match for this second spot, it's score was only 1.03, which is too low to be considered a reliable identification. Spectra representing duplicate MALDI target sample spots for Icterohaemorrhagiae were both identified incorrectly as Hardjo by the parent Icterohaemorrhagiae MSP, with Icterohaemorrhagiae being the sixth match for the first spot and seventh match for the second spot. The subtyping MSP also returned Hardjo as the first match for both spots, but scores were lower overall than those for the parent MSP matches. Using serovar Pomona's parent MSP, Pomona sample spectra were incorrectly identified as Canicola with scores of 2.49 and 2.26. Subtyping MSP identification of the first spot also returned a first match of Canicola with a score of 2.01. The second spot was correctly identified as Pomona. However, the match score for this spot was 1.75, which is only slightly above the cut-off value of 1.70, below which a match cannot be considered reliable.

Contrary to the information provided in the software's user manual, the function allowing users to view and manually edit subtyping MSP peak lists was no longer available. Therefore, there was not a way to determine how the peaks were weighted. A new set of seven subtyping MSPs were created to determine if assigned peak weights would change in such a way as to result in improved serovar identification. Newly collected raw spectra for all seven serovars were used for testing. Duplicate sample spots for all serovars except Canicola and Icterohaemorrhagiae identified correctly using the standard MSPs with scores > 2.40. One of two Canicola spots was misidentified as Hardjo, with a score of 2.41. Canicola was the second match for this spot, with a score of 2.39. Subtyping MSPs also incorrectly identified this spot as Hardjo, but with a lower score of 1.98. Again, Canicola was the second match, but with a score of 1.83. Both Icterohaemorrhagiae sample spots were identified as Hardjo by the standard MSPs. Icterohaemorrhagiae came in as the fourth match for both spots, at scores of 2.43 and 2.38. Subsequent identification via subtyping MSPs also returned Hardjo as the first match for both spots.

Results of these two trials showed that subtyping MSPs did not improve specificity. Not only were some serovars still identified incorrectly, match scores were lower overall than those for standard MSPs. Ideally, correct first matches would have scores > 2.30, which is the reference score above which identification results are considered highly probable at the species level. Those serovars most in need of an increase in specificity returned subtyping scores that fell < 2.00. Per Bruker's score interpretation guidelines, scores between 1.70 and 1.99 indicate a probable identification at only the genus level and are not sufficient for reliable species-level identification. By extension, scores in this range would also not be reliable for sub-species-level identification, though there are no guidelines that stipulate this. Since users can no longer view or edit subtyping MSP peak lists, users cannot compare the weights given to particular peaks in parent versus subytyping MSPs, nor can they choose or alter the assigned weights. Given that the subtyping MSPs did not improve specificity, no further trials using these MSPs were performed.

3.4.7 Manually Editing Peaks in FlexAnalysis

Raw spectra used in serovar Hardjo's MSP were loaded into FlexAnalysis. The raw spectra and corresponding mass peak lists were examined for peaks which had been identified as unique in the MSP. If one of these peaks was missing from a raw spectrum's mass peak list, but was present in the spectrum itself, the peak was manually added to the list. These changes were saved, and the raw spectra were loaded into Compass Explorer with the goal of using these edited raw spectra to create a new, more specific, MSP. Once a new MSP was created, the MSP peak lists for the new and original MSPs were compared. No differences in picked peaks or peak frequency could be found. It was determined that manual edits made to raw spectra and saved in FlexAnalysis were only retained within FlexAnalysis and so could not be imported into Compass Explorer and incorporated into a new MSP**.** Therefore, this approach could not be used to increase MSP specificity.

3.4.8 Comparison of MSP and Raw Spectra Peak Lists

Review of serovar MSP peak lists revealed peaks unique to each. The number of characteristic peaks ranged from two (Bratislava) to eight (Grippotyphosa). Peaks of similar *m/z* present in all MSP peak lists were noted. Since MSP peaks are calculated from a combination of corresponding peaks found in serovar raw spectra, MSP peaks were not expected to have the exact *m/z* as their raw spectra counterparts. However, it was expected that peaks of relatively similar *m/z*, within a particular *m/z* range, would be present in both. While the raw spectra peak lists contained some peaks that fit this assumption, many of the peaks did not. In fact, some peaks in the raw spectra spanned a much wider *m/z* range than expected in relation to the *m/z* of the MSP peak with which they were presumably correlated. For example, a unique peak for Bratislava's MSP had been identified at 3213 *m/z*. Review of Bratislava raw spectra peak lists found that the closest peaks ranged from 3177 m/z to 3226 *m/z*. This range of 49 units was larger than expected.

Further comparison of MSP and raw spectra peak lists discovered that some MSP peaks may actually represent more than one peak found in the raw spectra. For instance, a unique MSP peak at 3562 *m/z* for serovar Hardjo appears to represent the average of two peaks found in the raw spectra. The closest peaks found in the raw spectra were either 3538 or 3539 *m/z* (on the low end) or 3615, 3636, or 3637 (on the high end). The average of either the peaks on the low end, or those on the high end, do not result in a peak value of 3562 *m/z*. To get this value, two separate peaks in each raw spectrum that fall within a range of 76, (3539 *m/z*, 3615 *m/z*), to 99, (3538 *m/z*, 3637 *m/z*), units must be combined to form the single peak in the MSP (Figure 3.1). Another example was found in serovar Pomona's MSP peak list, at 5905 *m/z*. Pomona raw spectra peak lists showed that the closest peaks fall within the range of 5781 *m/z*–5944 *m/z* on the low end, and 6013 *m/z*-6014 *m/z* on the high end. Peak 5905 *m/z* appears to be the average of two of these peaks for each spectrum. The number of *m/z* units between these two peaks ranges from 69–232. However, there were other MSP/raw spectra peak correlations in which raw spectra

peaks within a much smaller *m/z* range were not combined to give the correlating peak in the MSP. This shows the difficulty in attempting to correlate peaks in serovar raw spectra with those in MSPs.

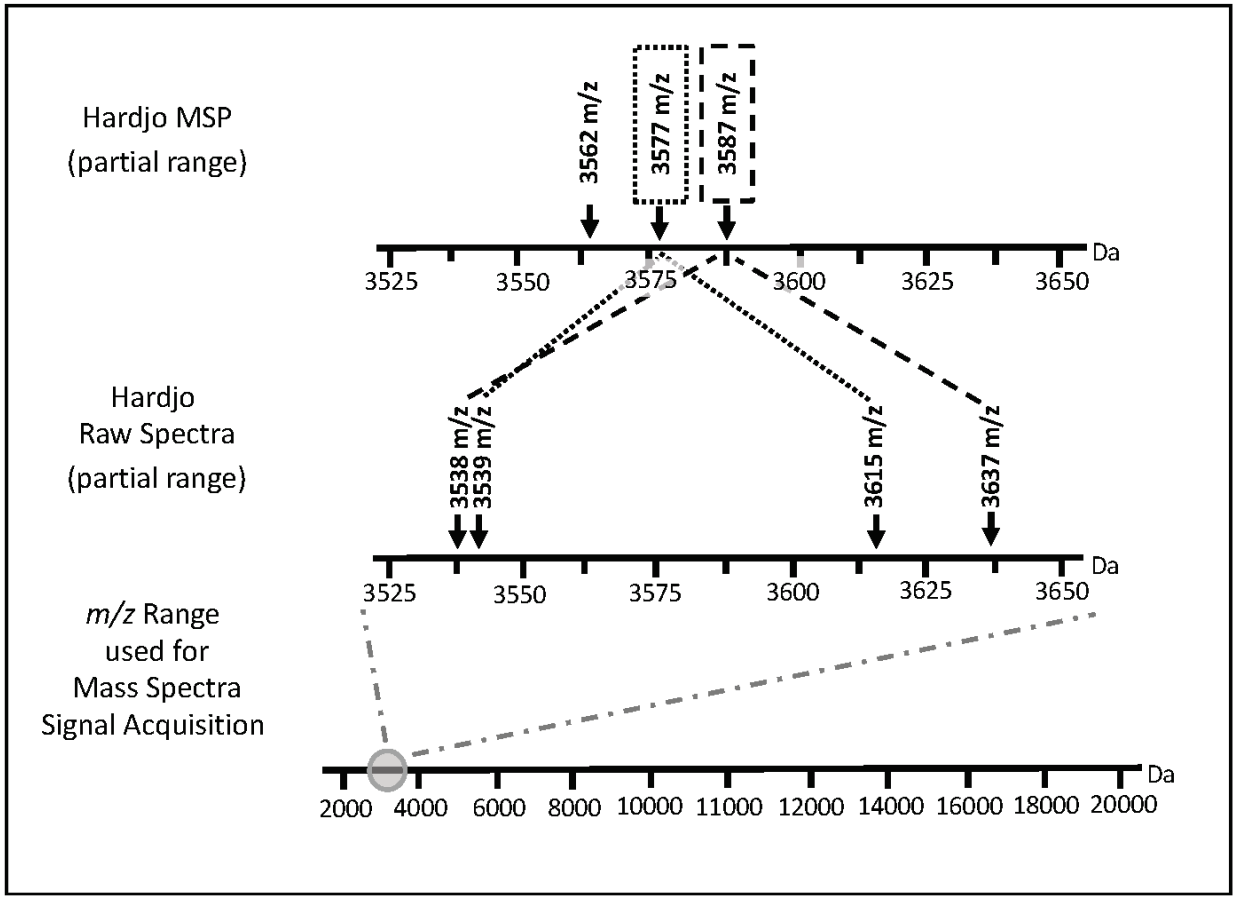

Figure 3.1. One Peak in an MSP May Represent More than One Peak in the MSP's Constituent Raw Spectra. Many peaks in an MSP do not have exact peak matches in their constituent spectra. Instead, some MSP peaks represent the average of two raw spectra peaks. In this example, the top portion of the figure represents a spectral range of 3525 Da to 3650 Da for a Hardjo MSP. The middle portion of the figure represents this spectral range in the MSP's constituent raw spectra. The average of two peak combinations in the raw spectra correspond to one peak found in the MSP. The average of peaks 3538 *m/z* and 3627 *m/z*, denoted by dashed lines, is 3587 *m/z*. The average of peaks 3539 *m/z* and 3615 *m/z*, denoted by dotted lines, is 3577 *m/z*. Corresponding peaks between different raw spectra acquired from a serovar and between a serovar's MSP and constituent raw spectra, can vary over an *m/z* range. Hence the two averages from the peak combinations given here do not exactly equal the MSP peak of 3562 *m/z*. The bottom mass spectral range is that within which sample peak signals are detected. The gray circle indicates the section of this range enlarged to show Hardjo Raw Spectra and MSP peak correlations.

Another interesting observation was that some individual peaks in the raw spectra appeared to be correlated with more than one MSP peak. To illustrate, the peak found at 6013 or 6014 *m/z* in serovar

Pomona's raw spectra correlates with the MSP peak of 6014 *m/z*. However, it also appears to correspond with the MSP peak of 5905 *m/z*, as stated above.

Unexpectedly, some of the MSP peak lists contained peaks with higher *m/z* values than those of the highest peaks found in the raw spectra. For instance, serovar Hardjo's MSP peak list included peaks at 11543 *m/z* and 12021 *m/z*, yet the highest peak found in Hardjo's raw spectra was 11053 *m/z*. The number of peaks in Hardjo raw spectra peak lists ranged from 68-101, while Hardjo's MSP peak list had a total of 70 peaks. Since the peaks in an MSP peak list are generated from the combination of peaks in constituent raw spectra, it was expected that raw spectra peak lists would contain peaks at more extreme *m/z* values, at both the lower and higher ends of the detected spectrum range, compared to peaks in the MSP peak list. The idea here is similar to calculating an average value for a group of individual values. The average represents the central value for the group. It was unexpected that an MSP peak, calculated from one or more raw spectra peaks found in multiple raw spectra, would have a peak *m/z* value that falls outside the *m/z* range of the peaks used for its calculation.

3.4.9 Additional MSPs Constructed Using Modified Creation Parameters

For the reader to more easily grasp the discussion of particular MSPs in this and subsequent chapters, the MSP naming convention is reviewed here first. The first letter used in an MSP name denotes the relevant serovar (e.g. A Original denotes an MSP for serovar Autumnalis.) A first letter followed by the term "Original" refers to the original MSP created for that serovar. As a reminder, the original MSPs were created using one set of raw spectra, a maximum desired peak number of 70, and a lower boundary of 3000 Da for peak detection. A first letter followed by "70" refers to the MSP created using the same parameters as used for the original, but with two additional sets of raw spectra. A first letter followed by "100" or "125" denotes an MSP with the same three sets of raw spectra and the same lower peak detection boundary as the Original and "70" MSPs, but with a maximum desired peak number of 100 or 125, respectively. An X Original, X70, X100, and X125 MSP was created for each of the seven serovars. Two additional MSPs were created for serovars Grippotyphosa and Bratislava. One, named X 2000 M100, includes the three sets of raw spectra used in the previously-mentioned MSPs, a maximum desired peak number of 100, and a lower peak detection boundary of 2000 Da. The second additional MSP for these two serovars was created using data from 2D peak distribution plot analysis and is named X 2D. This MSP was created using the same three sets of raw spectra used in the previous MSPs, a maximum desired peak number of 125, a lower peak picking boundary of 3000 Da, and omitted the particular raw spectra, identified by the 2D plots.

A few additional MSPs, created for serovars Bratislava and Grippotyphosa, will be discussed separately below. A summary of MSP names and main characteristics can be found in Table 3.1.

TABLE 3.1. Main Spectrum Profile Types Created for Each Serovar			
MSP Name	**Sets of Raw Spectra**	**Maximum Desired Peak #**	**Lower Peak m/z Bound for Peak Detection**
X Original	1	70	3000
X 70	3	70	3000
X 100	3	100	3000
X 125	3	125	3000
X 2000 M100	3	100	2000
X 2D	3 (with selected spectra removed)	125	3000

The 'X' in each MSP name above represents the first letter of a serovar's name.

After the X Original MSPs did not yield consistent serovar-level specificity, three new MSPs were created for each serovar by manipulating the following MSP creation parameters: (a) the number of raw spectra used to create an MSP, (b) the number of maximum desired peaks, and (c) the *m/z* range within which spectra peaks were selected for MSP inclusion (lower bound). These parameters were modified to determine their influence on MSP identification accuracy. These new MSPs combined the raw spectra used in the original MSP with two additional sets of raw spectra, each collected from a different subculture. In this way, each new MSP contained three sets of raw spectra, each representing a biological replicate. Incorporation of these supplementary spectra was expected to result in MSPs that could better reflect intraserovar mass spectra variability and would; consequently, increase identification specificity. The first of these new MSPs, named X 70, had a total of 70 picked peaks, like the original MSP. However, while X Original contained 18-24 raw spectra, X 70 contained peak signal data from 67-72 raw spectra. As noted above, the second and third new MSPs, X 100 and X 125, used the same combination of raw spectra used to create a serovar's X 70, but with a higher desired maximum peak number of 100 or 125, respectively. It should be mentioned that changing the maximum desired peak number parameter for MSP creation does not mean that the number of peaks selected for an MSP will equal the maximum desired peak number selected. Changing this parameter just changes the number of peaks that may be selected for an MSP, up to the maximum desired peak number entered.

MSPs created using different parameters and additional raw spectra had peak lists that differed from those of the original MSPs. While many of the same peaks were present in both new and original MSP peak lists, each list contained peaks not found in the other. This was partially attributed to changes in the relative intensity for each peak as more peaks or new spectra were added. Since relative intensity is a contributing factor in MSP peak selection, the presence or absence of a peak or set of peaks in a raw

spectrum alone is not sufficient for determining whether that spectrum should be included in a serovar's MSP for the purpose of maximizing specificity.

3.4.9.1 Effect of additional raw spectra on MSP specificity

To test how incorporation of additional sets of raw spectra affected MSP specificity, trial results for X Original and X 70 MSPs were compared. In serovar Autumnalis individual trials, A Original returned four out of four possible correct first matches for the starting concentration and the 1:4 dilution. It returned three out of four possible correct first matches for the 1:2 and 1:8 dilutions. For the 1:16 dilution, it returned two of four possible correct first matches, and no matches for the 1:32 dilution. A 70 had a total of two possible correct first matches in its individual trials, compared to four for A Original MSP individual trials. A 70 returned two out of two possible correct first matches for the starting concentration, but only one of two possible correct first matches for the 1:2 through 1:8 dilutions. It returned no matches for the 1:16 or 1:32 dilutions. Results from these individual trials suggest that A Original performed slightly better than A 70 since A Original returned a higher number of correct first matches relative to its total number of tested sample spots and returned correct first matches through a higher dilution than A 70. However, the difference in total number of sample spots tested in individual trials for the two MSPs makes comparison somewhat difficult. In Group trials, A Original returned only one correct first match over all sample spots and dilutions tested, which was for the 1:4 dilution. Conversely, A 70 obtained at least one correct first match for all but the 1:4 dilution. It additionally returned two of four possible correct first matches for the 1:32 dilution. Since A Original performed slightly better than A 70 in Individual trials, and A 70 performed better than A Original in Group trials, the addition of MSP constituent spectra to an Autumnalis MSP with a maximum peak number of 70 did not clearly improve MALDI MSP specificity (Table 3.2).

TABLE 3.2. Average scores for Autumnalis MSPs per dilution

Dilution	Average # organisms/mL	**Autumnalis** MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		A Original		**A 70**		**A 100**		**A 125**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind.	Group	Ind.	Group	Ind.	Group	Ind.	Group
		n = 4	**n = 4**	**n = 2**	**n = 4**	**n = 2**	**n = 4**	**n = 2**	**n = 4**
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
SC	2.91×10^9	2.57		2.61	2.54 (1)	2.67	2.60	2.67	
Matched to Other			[*1, 1, 1, 1*]		[*1, 1, 1*]				[*1, 1, 1, 1*]
1:2	1.13×10^9	2.54 (3)		2.58 (1)	2.54 (1)	2.60	2.55	2.60	
Matched to Other		[*2*]	[*1, 1, 1, 1*]	[*2*]	[*1, 1, 1*]				[*1, 1, 1, 1*]
1:4	5.95×10^8	2.51	2.48 (1)	2.56 (1)		2.49	2.48 (3)	2.49	
Matched to Other			[*1, 1, 1*]	[*2*]	[*1, 1, 1, 1*]		[*1*]		[*1, 1, 1, 1*]
1:8	3.10×10^8	2.33 (3)		2.40 (1)	2.27 (1)	2.41 (1)	2.38 (2)	2.41 (1)	
Matched to Other		[*2*]	[*1, 1, 1, 2*]	[*2*]	[*1, 1, 2*]	[*2*]	[*1, 2*]	[*2*]	[*1, 1, 1, 2*]
1:16	1.56×10^8	1.87 (2)			2.16 (1)		2.31 (1)		
Matched to Other		[*2,2*]	[*1, 1, 2, 2*]	[*2,2*]	[*1, 2, 2*]	[*2,2*]	[*1, 2, 2*]	[*2,2*]	[*1, 1, 2, 2*]
1: 32	7.81×10^7				1.92 (1)				
Matched to Other		[*3, 3, 3, 3*]	[*1, 1, 3, 3*]	[*3,3*]	[*3, 3, 4*]	[*3,3*]	[*1, 1, 3, 3*]	[*3,3*]	[*1, 1, 3, 3*]

Average MALDI identification scores are given for four Autumnalis MSPs, (A Original, A 70, A 100, A 125), in both Individual (Ind.) and Group trials. The 2-fold serial dilutions used in the trials are given in the left-most column, starting with "SC" (Starting Concentration). The sixth row in the table above shows "n=" at the top of each column, which indicates the number of sample spots tested for each dilution in the Individual and Group trials. In general, each dilution was tested in duplicate in each trial. Therefore, "n=2" would represent duplicate sample spots for a particular dilution tested in one trial. Some MSPs show that a higher number of sample spots were tested. This is because these MSPs were used in a greater number of trials. The original MSP shows a greater number of sample spots tested in Ind. trials because it was the first MSP created, and was tested prior to the creation of subsequently-created MSPs. MALDI identification score data returned for each dilution is presented in two rows. The first row of score data for each dilution is white, while the second row is gray and is titled "Matched to Other." The first row of data for each dilution gives the average identification scores for each MSP and trial type. In general, the number of scores used to calculate the average score is given by the "n=" number. Scores presented without additional notation (as described below) indicate that all tested sample spots for that trial were matched to the correct serovar. An average score followed by a number in parentheses indicates that fewer than the "n=" number of scores were used to calculate the average. The number in parentheses indicates how many scores were used in the calculation.

TABLE 3.2. ***continued***

Four types of MALDI results caused the scores for some sample spots to be omitted from the average. These four reasons were assigned a number from 1-4 and are defined in the key below.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (possible in Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match may have been to the correct serovar and MSP, or to a different serovar; however, the score was < 1.70, (which is considered unreliable).

For average scores calculated using fewer than the "n=" number of sample spots, the reason(s) why scores from a fewer number of sample spots were used is given in the second row of data per dilution. The number(s) representing the reason(s) for score omission are presented in italics and within brackets. One number per each sample spot/score that was omitted from the average calculation is listed.

Please also note: some first matches returned a tie (same identification score) between two MSPs for the same serovar. In those cases, a first match was counted for each of the tied MSPs, and those scores were included in the average score calculation for the appropriate dilution.

Results for serovar Bratislava's B Original and B 70 were similar. In individual trials, both MSPs returned at least one correct first match for all but the 1:32 dilution. In group trials, B Original returned only one correct first match, for the 1:2 dilution, over all dilutions and sample spots. B 70 returned a total of two correct first matches, one for the starting dilution and one for the 1:4 dilution. As was the case with Autumnalis, the total number of sample spots tested for the B Original MSP was four per dilution, whereas two per dilution were tested for B 70. With Bratislava, results showed that a greater percentage of correct first matches, out of the total available for each dilution, were gained with B 70. Additionally, average scores per dilution were slightly higher for B 70. However, the difference in total sample spots tested for the two MSPs made it difficult to determine a clear improvement in specificity with the addition of constituent spectra to the Bratislava B Original MSP (Table 3.3).

TABLE 3.3. Average scores for Bratislava MSPs per dilution

Dilution	Average # organisms/mL	**Bratislava** MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		B Original		**B 70**		**B 100**		**B 125**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind. n = 4	Group n = 4	Ind. n = 2	Group n = 2	Ind. n = 2	Group n = 2	Ind. n = 2	Group n = 2
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
SD	1.87×10^9	2.61 (2)		2.69	2.75 (1)	2.67	2.64(1)		
Matched to Other		[*2, 2*]	[*1, 1*]		[*1*]			[*2, 2*]	[*1, 1*]
1:2	8.07×10^8	2.47 (1)	2.54 (1)	2.54		2.50 (1)			
Matched to Other		[*2, 2, 2*]	[*1*]		[*1, 1*]	[*2*]	[*1, 1*]	[*2, 2*]	[*1, 1*]
1:4	3.86×10^8	2.42 (3)		2.48 (1)	2.48 (1)	2.44	2.41(1)	2.40	
Matched to Other		[*2*]	[*1, 1*]	[*2*]	[*1*]		[*1*]		[*1, 1*]
1:8	1.75×10^8	1.97 (2)		2.19		2.24		1.97 (1)	
Matched to Other		[*2, 2*]	[*1, 2*]		[*1, 2*]		[*1, 2*]	[*2*]	[*1, 2*]
1:16	8.81×10^7	1.81 (2)		1.80 (1)		1.81 (1)	1.81 (1)		
Matched to Other		[*3, 3*]	[*1,3*]	[*3*]	[*1,3*]	[*3*]	[*3*]	[*2, 3*]	[*1,3*]
1: 32	4.40×10^7								
Matched to Other		[*3, 3, 3, 3*]	[*3, 3*]	[*3, 3*]	[*3, 3*]	[*3, 3*]	[*3, 3*]	[*3, 3*]	[*3, 3*]

Average MALDI identification scores are given for four Bratislava MSPs (B Original, B 70, B 100, B 125) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

In individual trials for serovar Canicola, C Original returned two of four correct first match results for each dilution through 1:8. C 70 returned one of two correct first match results for the starting and 1:8 dilutions, and two of two correct first matches for the 1:2, 1:4, and 1:16 dilutions. Average scores per dilution for C 70 were slightly higher than those for C Original. In group trials, C Original returned no first

matches, while C 70 returned between one and four out of four possible correct first matches for all dilutions. Overall, C 70 offered a greater number of correct first matches over a wider range of dilutions than C Original (Table 3.4).

TABLE 3.4. Average scores for Canicola MSPs per dilution

Dilution	Average # organisms/mL	**Canicola** MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		C Original		**C 70**		**C 100**		**C 125**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind. n = 4	Group n = 4	Ind. n = 2	Group n = 4	Ind. n = 2	Group n = 4	Ind. n = 2	Group n = 4
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
U	2.95 x 10^9	2.44 (2)		2.55 (1)	2.56 (3)	2.49 (1)	2.62 (1)	2.46 (1)	
Matched to Other		[*2, 3*]	[*1, 1, 1, 1*]	[*3*]	[*1*]	[*3*]	[*1, 1, 1*]	[*3*]	[*1, 1, 1, 1*]
1:2	1.46 x 10^9	2.48 (2)		2.55	2.58	2.51		2.47	
Matched to Other		[*2, 2*]	[*1, 1, 1, 1*]				[*1, 1, 1, 1*]		[*1, 1, 1, 1*]
1:4	6.85 x 10^8	2.45 (2)		2.57	2.58 (2)	2.56	2.57 (2)	2.58 (1)	
Matched to Other		[*2, 2*]	[*1, 1, 1, 1*]		[*1, 1*]		[*1, 1*]	[*2*]	[*1, 1, 1, 1*]
1:8	3.26 x 10^8	2.52 (2)		2.48 (1)	2.09 (1)	2.47 (1)	2.61 (2)		
Matched to Other		[*2, 3*]	[*1, 1, 1, 3*]	[*3*]	[*1, 1, 3*]	[*3*]	[*1, 3*]	[*3*]	[*1, 1, 1, 3*]
1:16	1.50 x 10^8			2.33	2.53 (2)	2.30	2.52 (1)	2.28	
Matched to Other		[*2, 2, 3, 3*]	[*1, 1, 1,3*]		[*1, 3*]		[*1, 1, 3*]		[*1, 1, 1, 3*]
1: 32	7.41 x 10^7				2.21 (1)				
Matched to Other		[*3, 3, 3, 3*]	[*1, 2, 3 ,3*]	[*3, 3*]	[*2, 3, 3*]	[*3, 3*]	[*1, 2, 3 ,3*]	[*3, 3*]	[*1, 2, 3 ,3*]

Average MALDI identification scores are given for four Canicola MSPs (C Original, C 70, C 100, C 125) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

Serovar Grippotyphosa's G Original and G 70 both returned correct first matches for the starting concentration through the 1:8 dilution in individual trials. Both MSPs performed poorly in group trials, with G Original returning no first matches and G 70 returning only one correct match, for the 1:2 dilution. Average scores per dilution were similar between the two, with each MSP showing slightly higher scores for different dilutions. The additional raw spectra sets in G 70 did not result in increased Sp for this serovar (Table 3.5).

TABLE 3.5. Average scores for Grippotyphosa MSPs per dilution

Dilution	Average # organisms/mL	**Grippotyphosa** MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		G Original		**G 70**		**G 100**		**G 125**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind. n = 4	Group n = 4	Ind. n = 2	Group n = 4	Ind. n = 2	Group n = 4	Ind. n = 2	Group n = 4
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
SD	1.97×10^9	2.62 (3)		2.69		2.72	2.72	2.72	2.72 (2)
Matched to Other		*[1]*	*[1, 1, 1, 1]*		*[1, 1, 1, 1]*				*[1, 1]*
1:2	1.19×10^9	2.55		2.64	2.65 (1)	2.62	2.73 (2)	2.62	
Matched to Other			*[1, 1, 1, 1]*		*[1, 1, 1]*		*[1, 1]*		*[1, 1, 1, 1]*
1:4	6.65×10^8	2.54 (2)		2.49 (1)		2.56	2.62 (1)	2.56	2.62 (1)
Matched to Other		*[2, 2]*	*[1, 1, 1, 1]*	*[2]*	*[1, 1, 1, 1]*		*[1, 1,1]*		*[1, 1, 1]*
1:8	5.21×10^8	2.44 (2)		2.54 (1)		2.36	2.66 (2)	2.36	2.57 (1)
Matched to Other		*[2, 3]*	*[1, 1, 1, 2]*	*[2]*	*[1, 1, 1, 2]*		*[1, 2]*		*[1, 1, 2]*
1:16	3.50×10^8						2.39 (1)		
Matched to Other		*[2, 3, 3, 3]*	*[1, 3, 3, 3]*	*[3, 3]*	*[1, 3, 3, 3]*	*[3, 3]*	*[3, 3, 3]*	*[3, 3]*	*[1, 3, 3, 3]*
1: 32	1.8s4 x 10^8								
Matched to Other		*[3, 3, 3, 3]*	*[3, 3, 3, 3]*	*[3, 3]*	*[3, 3, 3, 3]*	*[3, 3]*	*[3, 3, 3, 3]*	*[3, 3]*	*[3, 3, 3, 3]*

Average MALDI identification scores are given for four Grippotyphosa MSPs (G Original, G 70, G 100, G 125) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

TABLE 3.5. *continued*

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

Serovar Hardjo's H Original returned correct first matches for the starting sample concentration through the 1:4 dilution in individual trials but had no first matches in group trials. Individual trial results for H 70 showed that this MSP returned valid first match scores only for the starting and 1:2 dilutions. As with H Original, H 70 returned no first matches in group trials. Both of these MSPs performed poorly, and no improvement was seen with the addition of raw spectra to the serovar's H Original MSP (Table 3.6).

TABLE 3.6. Average scores for Hardjo MSPs per dilution

		Hardjo MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		H Original		**H 70**		**H 100**		**H 125**	
Dilution	**Average # organisms/mL**	**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind. n = 4	Group n = 2	Ind. n = 2	Group n = 2	Ind. n = 2	Group n = 2	Ind. n = 2	Group n = 2
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
SD	3.29×10^9	2.54 (3)		2.66 (1)		2.59		2.59	2.56
Matched to Other		[*2*]	[*1, 1*]	[*2*]	[*1, 1*]		[*1, 1*]		
1:2	1.65×10^9	2.45		2.54		2.46		2.43	2.71
Matched to Other			[*1, 1*]		[*1, 1*]		[*1, 1*]		
1:4	7.91×10^8	2.46 (1)							2.67
Matched to Other		[*2, 2, 2*]	[*1, 1*]	[*2, 2*]	[*1, 1*]	[*4, 4*]	[*1, 1*]	[*2, 2*]	
1:8	3.54×10^8								2.62
Matched to Other		[*2, 2, 2, 3*]	[*1, 1*]	[*2, 5*]	[*1, 1*]	[*4, 4*]	[*1, 1*]	[*2, 2*]	
1:16	1.78×10^8								2.28
Matched to Other		[*3, 3, 3, 3*]	[*1, 1*]	[*3, 3*]	[*1, 1*]	[*3, 3*]	[*1, 1*]	[*3, 3*]	
1: 32	8.81×10^7								2.03
Matched to Other		[*3, 3, 3, 3*]	[*1, 1*]	[*3, 3*]	[*1, 1*]	[*3, 3*]	[*1, 1*]	[*3, 3*]	

TABLE 3.6. ***continued***

Average MALDI identification scores are given for four Hardjo MSPs (H Original, H 70, H 100, H 125) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

In serovar Icterohaemorrhagiae individual trials, I Original returned no correct matches for the starting concentration but returned between one and three correct first matches out of four for the 1:2 through 1:16 dilutions. I 70 returned two of two correct first matches for the starting concentration, 1:2 and 1:4 dilutions, and one of two correct first matches for the 1:8 and 1:16 dilutions. Average score per dilution was slightly higher for I 70. Both MSPs performed poorly in group trials. I Original returned only one correct first match at the 1:32 dilution with a score of 1.81. I 70 returned only one correct match at the 1:16 dilution at a score of 2.17. There was no clear improvement in performance for I 70 compared to I Original (Table 3.7).

TABLE 3.7. Average scores for Icterohaemorrhagiae MSPs per dilution										
		Icterohaemorrhagiae MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		I Original		**I 70**		**I 100**		**I 125**	
Dilution	**Average # organisms/mL**	**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind. **n = 4**	**Group** **n = 4**	**Ind.** **n = 2**	**Group** **n = 4**	**Ind.** **n = 2**	**Group** **n = 4**	**Ind.** **n = 2**	**Group** **n = 4**
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
SD	5.05×10^9			2.67		2.67	2.69	2.60	
Matched to Other		[*2, 2, 2, 2*]	[*1, 1, 1, 1*]		[*1, 1, 1, 1*]				[*1, 1, 1, 1*]
1:2	2.65×10^9	2.51 (1)		2.65		2.66	2.67	2.58	
Matched to Other		[*2, 2, 2*]	[*1, 1, 1, 1*]		[*1, 1, 1, 1*]				[*1, 1, 1, 1*]
1:4	1.29×10^9	2.44 (3)		2.58		2.58	2.63 (1)	2.48	
Matched to Other			[*1, 1, 1, 1*]		[*1, 1, 1, 1*]		[*1, 1, 1*]		[*1, 1, 1, 1*]
1:8	6.43×10^8	2.40 (3)		2.47 (1)		2.44 (1)	2.47 (1)	2.36 (1)	
Matched to Other		[*2*]	[*1, 1, 1, 3*]	[*3*]	[*1, 1, 1, 3*]	[*3*]	[*1, 1, 3*]	[*3*]	[*1, 1, 1, 3*]
1:16	3.22×10^8	2.06 (1)		2.18 (1)	2.17 (1)	2.18 (1)		2.16 (1)	
Matched to Other		[*3, 3, 3*]	[*1, 3, 3, 3*]	[*3*]	[*3, 3, 3*]	[*3*]	[*1, 3, 3, 3*]	[*3*]	[*1, 3, 3, 3*]
1: 32	1.60×10^8		1.81 (1)						
Matched to Other		[*3, 3, 3, 3*]	[*3, 3, 3*]	[*3, 3*]	[*1, 3, 3, 3*]	[*3, 3*]	[*1, 3, 3, 3*]	[*3, 3*]	[*1, 3, 3, 3*]

Average MALDI identification scores are given for four Icterohaemorrhagiae MSPs (I Original, I 70, I 100, I 125) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

Comparison of serovar Pomona's P Original and P 70 MSPs found that neither returned correct first match results for any of the dilutions in the group trials. Though P 70 returned two correct first matches for

the 1:32 dilution in the individual trial, the average of these scores was 1.79. This value is very close to the threshold score of 1.70, which separates scores interpreted as probable genus identification from those interpreted as unreliable results. Overall, no appreciable improvement was seen with the added raw spectra of P 70 (Table 3.8).

TABLE 3.8. Average scores for Pomona MSPs per dilution

Dilution	Average # organisms/mL	**Pomona** MSPs - Mean MALDI ID Scores for Correct First Matches							
		MSPs							
		P Original		**P 70**		**P 100**		**P 125**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**		**Tested in an Individual or Group Trial**	
		Ind. n = 4	Group n = 2	Ind. n = 2	Group n = 2	Ind. n = 2	Group n = 2	Ind. n = 2	Group n = 2
		Average Scores		**Average Scores**		**Average Scores**		**Average Scores**	
SD	3.05×10^9	2.59 (2)		2.61		2.58	2.58		
Matched to Other		[*2, 2*]	[*1, 1*]		[*1, 1*]			[*2, 2*]	[*1, 1*]
1:2	1.35×10^9	2.55 (2)		2.57		2.57	2.57	2.51	
Matched to Other		[*2, 2*]	[*1, 1*]		[*1, 1*]				[*1, 1*]
1:4	6.41×10^8	2.45 (3)		2.51		2.52	2.52	2.46	
Matched to Other		[*2*]	[*1, 1*]		[*1, 1*]				[*1, 1*]
1:8	3.20×10^8	2.29 (2)		2.32		2.35	2.35	2.31	
Matched to Other		[*2, 3*]	[*1, 1*]		[*1, 1*]				[*1, 1*]
1:16	1.60×10^8	2.01 (1)		1.86		1.86	1.86	1.85	
Matched to Other		[*2, 2, 3*]	[*1, 1*]		[*1, 1*]				[*1, 1*]
1: 32	8.01×10^7			1.79		1.86 (1)	1.86 (1)	1.82	
Matched to Other		[*2, 2, 3, 3*]	[*1, 1*]		[*1, 1*]	[*4*]	[*4*]		[*1, 1*]

Average MALDI identification scores are given for four Pomona MSPs (P Original, P 70, P 100, P 125) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

TABLE 3.8. ***continued***
For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

3.4.9.2 Effect of an increase in the number of MSP peaks on MSP specificity

To determine if an increase in the number of MSP peaks increased MSP specificity, X 70 trial results were compared to those for X 100 and X 125. For serovar Autumnalis individual trials, A 70, A 100, and A 125 returned at least one of two correct first matches for the starting concentration through the 1:8 dilution. A 100 and A 125 had the same number of first matches and the same average scores per dilution, both of which showed some improvement over the number of correct first matches and average scores for A 70. In group trials, A 100 again had a higher number of correct first matches with higher average scores per dilution for the starting concentration through the 1:16 dilution compared with A 70. At the 1:32 dilution, A 70 had one out of four possible correct first match scores while A 100 had none. However, even with one correct match at the highest dilution tested, A 70 did not perform as well as A100. A125 returned no first matches for the group trials. Therefore, the best results overall were obtained with A 100, followed by A 70.

Results for Bratislava's B 70 and B100 were comparable. Both MSPs returned first matches in individual trials through the 1:16 dilution. Average scores for first matches were higher for B 70 for the starting concentration, 1:2 and 1:4 dilutions, while average first match scores for A 100 were higher for the 1:8 and 1:16 dilutions. The average scores for each dilution were in the same score interpretation range for both A 70 and A 100. Group trial results for these two MSPs were also comparable, though B 100 returned one of two possible first matches for the 1:16 dilution while B 70 returned none. B 125 performed poorly overall, with first matches returned for only the 1:4 and 1:8 dilutions in individual trials, and no first matches in group trials. Based on these results, either B 70 or B 100 could have been chosen to represent Bratislava. However, B 100 was eventually chosen based on the results of trials for some of the other serovars, particularly Grippotyphosa. In general, when serovars other than Bratislava were incorrectly identified as Bratislava, the B 100 MSP was returned as the identification less often than other Bratislava MSPs. Therefore, it was thought that using B 100 as the Bratislava MSP in the final configuration of the *Leptospira* MSP library would result in fewer false negatives for non-Bratislava serovars.

Individual trials for serovar Canicola found that C 70 and C 100 returned similar results, with both MSPs yielding correct first matches from the starting concentration through the 1:16 dilution. However, the average score per dilution was slightly higher for C 70. In group trials, C 70 returned first matches from the starting concentration through the 1:32 dilution, while C 100 returned first matches for only the starting concentration and the 1:4, 1:8, and 1:16 dilutions. However, the one match for C 70 at the 1:32 dilution was not sufficient to claim that C 70 performed better across a wider range of dilutions. Overall,

the number of correct first matches for C 70 was slightly higher than that for C 100. C 125 returned correct first matches for the starting concentration and the 1:2, 1:4, and 1:16 dilutions in individual trials, but returned no first matches in group trials. Due to a fewer number of first matches and a generally lower average first match score compared to C 70 and C 100, C 125 was not chosen to represent serovar Canicola. While C 70 performed slightly better than C 100 when considering results for Canicola samples, it also produced a greater number of false positives in trials for other serovars. Therefore, C 100 was determined to be the best MSP to represent this serovar.

For serovar Grippotyphosa, individual trials showed that G 70, G 100, and G 125 yielded similar results, with all three MSPs returning correct first matches for the starting concentration through the 1:8 dilution. In group trials, G 70 returned only one first match, for the 1:2 dilution. G 100 performed appreciably better, offering correct first matches for the starting concentration through the 1:16 dilution. G 125 returned first matches for only the starting concentration and the 1:4 and 1:8 dilutions. So, while G 125 performed better than G 70, it did not perform as well as G 100.

In individual trials, serovar Hardjo's H 70, H 100, and H 125 returned correct, reliable first matches for the starting concentration and 1:2 dilution. The few additional first matches that H 1000 returned at higher dilutions had average scores ≤ 1.70, which is within the score range interpreted as no reliable identification. In group trials, H 70 and H100 performed poorly returning no first matches for any dilution. H 125 performed better, returning reliable identifications for the starting concentration through the 1:32 dilution. It was determined that H 125 offered the best specificity for serovar Hardjo.

For serovar Icterohaemorrhagiae, performance in individual trials was similar for I 70, I 100 and I 125. All three MSPs returned correct first matches from the starting concentration through the 1:16 dilution. Average scores per dilution were similar for I 70 and I 100, but slightly lower for I 125. In group trials, I 100 showed a marked improvement over I 70. I 100 returned first match scores for the starting concentration through the 1:8 dilution, while I 70 returned only one first match score, at the 1:16 dilution. I 125 returned no first matches in group trials. Of the three MSPs, I 100 showed the greatest specificity, given its ability to outperform I 70 and I 125 in group trials.

For serovar Pomona individual trials, P 70, P 100 returned correct first matches for all dilutions tested, while P 125 returned correct first matches for all but the starting concentration. In group trials, P 100 also returned correct first matches for all dilutions, while P 70 and P 125 returned none. These results showed that P 100 offered the greatest specificity.

Overall, X 125 offered the greatest specificity for serovar Hardjo, while X 100 offered the greatest specificity for all other serovars.

3.4.9.3 Effect of additional MSP constituent raw spectra on MSP peak frequency

The MSP Creation method in Compass Explorer software consists of several parameters, one of which is the desired % frequency that all peaks in an MSP's peak list must meet or exceed. Bruker

recommends that all MSP peaks have a frequency of at least 75%. This means that peaks selected for an MSP must be present in the peak lists of at least 75% of the MSP's constituent raw spectra. However, a review of the peak lists for the new MSPs containing additional raw spectra showed that some of the peaks had a peak frequency < 75%, regardless of whether the MPS had been created with a maximum peak number of 70, 100, or 125. Efforts were made to revise the new MSPs such that all peaks would meet the ≥ 75% frequency. To do this, raw spectra used in each MSP were reviewed in FlexAnalysis to identify spectra that displayed greater peak pattern variability compared to that of the other raw spectra in the group. This had been done previously, in the process of creating the new MSPs. It was performed this second time to select additional spectra to omit in creating revised X 70, X 100, and X 125 MSPs. This process was repeated several times, with additional spectra removed from subsequent revisions of these MSPs. Only when all but 18-20 raw spectra were left in an MSP did all peaks meet the ≥ 75% peak frequency. Additionally, for most of the revised MSPs, reaching a 75% peak frequency meant using raw spectra from only one of the three spectra sets. This negated any improvement in specificity gained by using a greater number of raw spectra and by using raw spectra acquired from different subcultures.

The feasibility of adjusting the recommended 75% peak frequency to a lower frequency for these MSPs was then considered. Bruker's guidelines for MSP creation are written for one set of spectra, not multiple sets of spectra. Additionally, the default setting for this parameter within Compass Explorer is actually 25%. A review of the literature also found several studies in which user-generated MSPs were created using the 25% default setting (Rettinger et al. 2012, Normand et al. 2013, Karcher et al. 2018). Comparison of MSP peak lists for the different types of MSPs showed that the number of peaks with lower frequencies increased as the number of constituent raw spectra in an MSP also increased. This was expected because, as the number of raw spectra in an MSP increases, the number of peaks that are present in all constituent spectra with the requisite characteristics appropriate for MSP selection, (e.g. intensity level, signal-to-noise ratio, and area), decreases. Given this trend, one has to determine how an increase in the number of lower frequency peaks affects an MSP's ability to accurately identify its respective serovar, and whether this effect outweighs any increase in specificity offered by the inclusion of additional raw spectra. It was determined that the ≥ 75% peak frequency guideline may be a best practice when using only one set of raw spectra to create MSPs for species-level identification. However, it may not be feasible when using multiple sets of raw spectra to create MSPs for differentiation at a subtype level. Ultimately, a balance between an optimal number of component raw spectra, an optimal number of peaks, and an acceptable peak frequency, must be found. It was decided that the X 70, X 100, and X 125 MSPs would remain as originally created, with all three sets of spectra and some peak frequencies < 75%. The group of MSPs eventually chosen to represent the serovars in the *Leptospira* MSP library each had from 34 to 45 peaks in the MSP peak list with a frequency < 75%. The lowest peak % frequency for each MSP peak list ranged from 25% to 37.5%.

3.4.9.4 Effect of a lower peak picking boundary on MSP specificity

The X 2000 M100 MSP, created for serovars Grippotyphosa and Bratislava, used a lower peak picking boundary of 2000, rather than 3000, Da. This lower boundary allowed for MSP peak selection from within a lower spectrum range than that used in previously created MSPs. To evaluate whether the lower boundary increased MSP specificity, trial results for X 2000 M100 were compared with those of X 100 since the parameters used to create both MSPs differed only in the lower boundary setting. In serovar Bratislava individual trials, B 100 returned correct first matches for the starting concentration through the 1:16 dilution, whereas B 2000 M100 returned first matches only through the 1:8 dilution. In group trials, the two MSPs returned first matches for different dilutions. B 100 was the first match for the starting concentration and the 1:4 and 1:16 dilutions. B 2000 M100 was the first match for the 1:2 and 1:8 dilutions. Neither MSP returned a first match for the 1:32 dilution. In general, these two MSPs performed similarly. B 100 had a slightly higher average score for two of the tested dilutions, and a few additional correct first matches compared to B 2000 M100. For serovar Bratislava, the lower boundary did not improve specificity. The number of peaks with a < 75% frequency for B 100 was 38, while that for B 2000 M100 was 33 (Table 3.9).

TABLE 3.9. Average scores for Two Additional Bratislava MSPs per dilution

Dilution	Average # organisms/mL	**Bratislava** MSPs - Mean MALDI ID Scores for Correct First Matches			
		MSPs			
		B 2000 M100		**B 2D**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**	
		Ind. **n = 4**	**Group** **n = 2**	**Ind.** **n = 2**	**Group** **n = 2**
		Average Scores		**Average Scores**	
SD	1.87×10^9	2.64		2.64	
Matched to Other			[*1, 1*]		[*1, 1*]
1:2	8.07×10^8	2.56 (1)	2.55 (1)		
Matched to Other		[*2*]	[*1*]	[*2, 2*]	[*1, 1*]
1:4	3.86×10^8	2.44 (1)		2.45 (1)	
Matched to Other		[*2*]	[*1, 1*]	[*2*]	[*1, 1*]
1:8	1.75×10^8	2.07 (1)	2.07 (1)	2.05 (1)	
Matched to Other		[*2*]	[*2*]	[*2*]	[*1, 2*]
1:16	8.81×10^8				
Matched to Other		[*2, 3*]	[*1, 3*]	[*2, 3*]	[*1, 3*]
1: 32	4.40×10^7				
Matched to Other		[*3, 3*]	[*3, 3*]	[*3, 3*]	[*3, 3*]

Average MALDI identification scores are given for two Bratislava MSPs (B 2000 M100 and B 2D) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

In serovar Grippotyphosa individual trials, results for G 100 and G 2000 M100 were similar, with both MSPs returning first matches through the 1:8 dilution. G 100 average scores were slightly higher for three out of four dilutions. In group trials, G 100 returned correct first matches for the starting concentration

through the 1:16 dilution. G 2000 M100 returned correct first matches for only the 1:2, 1:4 and 1:8 dilutions. Given that G 100 returned correct first matches for a few dilutions that G 2000 M100 did not, it was determined that G 100 performed slightly better. The lower boundary did not improve specificity for Grippotyphosa. G 100 had 45 peaks with a frequency < 75%, compared with G 2000 M100, which had 37 (Table 3.10).

TABLE 3.10. Average scores for Two Additional Grippotyphosa MSPs per dilution

Dilution	Average # organisms/mL	**Grippotyphosa** MSPs - Mean MALDI ID Scores for Correct First Matches			
		MSPs			
		G 2000 M100		**G 2D**	
		Tested in an Individual or Group Trial		**Tested in an Individual or Group Trial**	
		Ind. **n = 2**	Group **n = 4**	Ind. **n = 2**	Group **n = 4**
		Average Scores		**Average Scores**	
SD	1.97 x 10^9	2.67		2.69	
Matched to Other			[*1, 1, 1, 1*]		[*1, 1, 1, 1*]
1:2	1.19 x 10^9	2.61	2.64 (1)	2.58	
Matched to Other			[*1, 1, 1*]		[*1, 1, 1, 1*]
1:4	6.65 x 10^8	2.55 (1)	2.65 (3)	2.53	
Matched to Other		[*2*]	[*1*]		[*1, 1, 1, 1*]
1:8	5.21 x 10^8	2.54 (1)	2.58 (1)	2.35	
Matched to Other		[*2*]	[*1, 1, 2*]		[*1, 1, 1, 2*]
1:16	3.50 x 10^8				
Matched to Other		[*3, 3*]	[*1, 3, 3, 3*]	[*3, 3*]	[*1, 3, 3, 3*]
1: 32	1.84 x 10^8				
Matched to Other		[*3, 3*]	[*3, 3, 3, 3*]	[*3, 3*]	[*3, 3, 3, 3*]

Average MALDI identification scores are given for two Grippotyphosa MSPs (G 2000 M100 and G 2D) in both Individual (Ind.) and Group trials.

Key for bracketed numbers in "Matched to Other" row :

1: First match was to a different MSP than the one being tested for a particular serovar (Group Trials)

2: First match was to a different serovar

3: Results returned NPF (no peaks found)

4: First match was to the correct serovar & MSP, but the score was < 1.70 (considered unreliable)

For additional information on the data and terms presented here, please refer to the description for TABLE 3.2.

While the X 2000 M100 MSPs for both Bratislava and Grippotyphosa had fewer peaks with less than the 75% recommended frequency compared to that in the X 100 MSPs, the X 2000 M100 MSPs did not perform as well.

Trial results for the X 2D MSPs, created for serovars Bratislava and Grippotyphosa, will be discussed below.

3.4.10 ClinProTools Analysis

3.4.10.1 Peak statistics

ClinProTools was used to analyze serovar raw spectra to identify unique peak combinations and characteristics that could be used for serovar differentiation. Analyses were performed on the group of seven serovars and on the serovar pair of Bratislava and Grippotyphosa. Peak picking was performed on the total average spectrum for the loaded serovar classes. Peak calculations were performed using maximum peak intensities. An initial analysis for Bratislava and Grippotyphosa used the Welch's t-test to sort peaks in the PSR. The top six peaks in the sorted report were used to generate several 2D Peak Distribution plots for this serovar pair. However, when data from this analysis was not found to be helpful in improving MSP specificity. the PSR for this serovar pair was re-examined. It was then found that the PSRs generated for the both the serovar group and pair analyses showed that Anderson-Darling p-values (PADs) for most peaks indicated non-normal distribution (p-values were ≤ 0.05). Therefore, the Wilcoxon or Kruskal-Wallis test (WKW), depending on the number of serovar classes in the comparison, was chosen for peak sorting and selection. The top six statistically significant peaks for the serovar group and pair comparisons were used for further analyses.

3.4.10.1.1 Peak statistics for seven serovar group

The Peak Statistic report generated for the seven serovar group contained a total of 169 common peaks. Of these, 145 had PADs of ≤ 0.05, indicating a non-normal distribution. Kruskal-Wallis p-values (PWKWs) were then calculated using the average peak areas/intensities for each serovar class Results indicated that 137 peaks exhibited highly significant differences ($p \leq 0.01$) between the serovars. Peaks were sorted based on their statistical separation strength and the top six discriminatory peaks were identified (Table 3.11). Interestingly, a number of peaks with relatively large differences between the maximum and minimum area/intensity values (e.g. difference of 35.65) were designated as having less statistical separation power than some peaks with much smaller differences in their average values (e.g. difference of 0.83).

TABLE 3.11. Top Six Discriminatory Peaks for the Group of Seven Serovars as Determined by the Kruskal-Wallis Test in ClinProTools

	Average (Avg) Peak Intensities and **Standard Deviations (SD)** for **Seven Serovars**													
Serovars	**A**		**B**		**C**		**G**		**H**		**I**		**P**	
Select Peaks (*m/z*)	**Avg**	**SD**	**Avg**	**SD**	**Avg**	**SD**	**Avg**	**SD**	**Avg**	**SD**	**Avg**	**SD**	**Avg**	**SD**
7406.70	3.12	0.21	4.22	0.56	4.88	0.77	4.61	0.50	*9.10*	1.19	2.21	0.39	1.78	0.32
7256.16	*12.8*	4.49	4.70	1.45	0.97	0.29	6.75	1.85	3.69	0.82	2.74	1.10	0.56	0.20
6712.74	6.42	0.72	*9.93*	2.71	2.20	0.86	9.18	0.98	7.93	1.07	2.99	0.49	2.00	0.58
7388.46	6.44	1.12	9.23	1.62	10.5	1.54	12.2	2.98	*13.5*	2.14	2.93	0.92	2.49	0.56
7360.97	21.6	6.25	49.7	13.4	51.3	11.7	41.3	8.44	*63.4*	15.6	1.21	0.33	0.91	0.19
6698.86	1.22	0.20	1.13	0.27	6.87	1.03	1.07	0.40	0.85	0.28	*11.4*	3.26	7.63	2.88

Univariate statistical analysis was performed in ClinProTools to identify the top discriminatory peaks for the seven serovars (the first six of which are given in the table's left-most column). The Anderson-Darling test was used to determine whether average peak intensity data followed a normal, (p-value > 0.05), versus non-normal, (p-value ≤ 0.05), distribution among the loaded serovar classes. The Kruskal-Wallis test was used to determine the top six peaks exhibiting highly significant ($p \leq 0.01$) differences between the seven serovars. The top six peaks given in the table had Anderson-Darling and Kruskal-Wallis test p-values of <0.01, indicating that the average intensities for these peaks had non-normal distributions and that there were significant differences in these average peak intensities between the serovars. Average values are given in arbitrary units.

A: Autumnalis, **B:** Bratislava, **C:** Canicola, **G:** Grippotyphosa, **H:** Hardjo, **I:** Icterohaemorrhagiae, **P:** Pomona.

Avg: peak area / intensity average value

SD: standard deviation for the corresponding average value

3.4.10.1.2 Peak statistics for serovar pair Bratislava and Grippotyphosa

Given the higher number of misidentifications between serovars Bratislava and Grippotyphosa, the decision was made to use all of the raw spectra included in their respective MSPs, rather than just the training subsets of spectra, for analysis. This way, each raw spectrum in these serovars' MSPs could be examined for qualities that might contribute to lower specificity. This information could then be used to create new MSPs with potentially improved specificity. During initial analysis of this serovar pair in ClinProTools, the peaks selected by the Peak Statistic Report were sorted using the Student's *t*-test. The top six peaks of the report were noted and then used for preliminary CPT analyses. A review of the PSR found that 138 peaks exhibited non-normal distribution as indicated by PAD values ≤ 0.05. Since over half of the common peaks for these two serovars had non-normal distributions, peak statistic analysis was repeated using the nonparametric Wilcoxon rank sum test. The new PSR identified a total of 168 common peaks with 86 of these identified as having highly significant area/intensity differences ($p \leq 0.01$) between

the two serovar classes. A new set of top differentiating peaks (Table 3.12) was also identified. The top six discriminatory peaks were selected for use in further analysis.

TABLE 3.12. Top Six Separating Peaks for Serovars Grippotyphosa and Bratislava as Determined by the Wilcoxon Rank Sum Test in ClinProTools

Average (Avg) Peak Intensities and **Standard Deviations (SD) for Serovars Bratislava and Grippotyphosa**				
	B		**G**	
Select Peaks (*m/z*)	**Avg**	**SD**	**Ave**	**SD**
8058.91	26.56	6.75	16.87	3.10
6356.83	12.88	2.71	8.43	1.40
7231.52	3.62	1.85	1.55	0.29
5076.25	2.17	1.32	0.87	0.30
7845.58	2.38	0.38	1.54	0.25
7825.24	1.24	0.19	0.81	0.14

Univariate statistical analysis was performed in ClinProTools using the Wilcoxon test to identify the top six peaks exhibiting highly significant ($p \leq 0.01$) differences between the serovars Grippotyphosa and Bratislava. Average peak intensity for the top six peaks differed between Bratislava and Grippotyphosa by 9.69. All six peaks had a higher average intensity in Bratislava raw spectra compared to the raw spectra of Grippotyphosa.

Key: **B:** Bratislava, **G:** Grippotyphosa. **Avg:** peak area /intensity average value. **SD:** standard deviation for the corresponding average value.

3.4.10.2 Two-dimensional (2D) peak distribution plots

3.4.10.2.1 2D peak distribution – seven serovar group

Peak distribution plots for all seven serovars were created for the top six discriminatory peaks selected in the PSR. While these plots provided a way to initially visualize similarities and differences in peak distribution, ultimately, they did not prove helpful for identifying peaks and raw spectra that potentially lowered MSP specificity due to the sheer number of overlapping data points. This tool was more valuable as a way to focus on two serovars at a time that frequently misidentified as each other (e.g. Bratislava and Grippotyphosa).

3.4.10.2.2 2D peak distribution–serovar pair Bratislava and Grippotyphosa

As mentioned above, the peaks initially chosen for preliminary CPT analyses, including 2D peak distribution analysis, were selected using the Student's *t*-test. The top six discriminatory peaks identified

by this method were used to construct 2D peak distribution plots to identify raw spectra in which intensity of selected peaks exhibited overlap. Out of the three sets of raw spectra loaded for each serovar, 15 spectra for Bratislava and six spectra for Grippotyphosa were found to contain peaks with intensity distributions that overlapped. These spectra were noted and a new MSP, X 2D, was created for each of these serovars with these raw spectra omitted. Bratislava's new MSP, named B 2D, contained a total of 57 raw spectra and 106 peaks, while Grippotyphosa's new MSP, G 2D, consisted of 63 raw spectra and 103 peaks.

The new MSPs were tested in both individual and group trials. Previously-created MSPs for Bratislava and Grippotyphosa were also tested individually using the same set of raw test spectra in the individual B 2D and G 2D trials so that results from all MSPs for each of these serovars could be compared.

In individual trials, serovar Bratislava's B 2D returned correct first matches for the starting concentration and the 1:4 and 1:8 dilutions. The second 1:16 dilution spot and both 1:32 dilution spots returned no peaks found (NPF). The NPF results are thought to be due to the quality of the raw test spectra for these dilutions rather than a reflection of the MSP's ability to identify spectra at higher dilutions. This is because each Bratislava MSP tested using this same set of raw test spectra also returned NPF for these same dilutions. Compared with the other Bratislava MSPs tested, B 2D returned fewer correct first matches than B 70, B 100, and B 2000 M100. In group trials, B 2D returned no first matches (Table 3.9). These results showed that the B 2D MSP did not improve specificity, and so it was not used for further testing.

For serovar Grippotyphosa individual trials, G 2D returned correct first matches for the starting concentration through the 1:8 dilution. Results for the 1:16 and 1:32 dilutions were NPF (Table 3.10). Like results for Bratislava's individual trial, these NPF results were thought to be due to the quality of the raw test spectra. Each Grippotyphosa MSP tested with these spectra also retuned NPF for these two dilutions. G 2D returned a greater number of correct first matches than the G Original, G 70 and G 2000 M100 MSPs. G 2D returned the same number of correct first matches as G 100 and G 125, albeit with a slightly lower average score per dilution was than that for the other two MSPs. In group trials, G 2D returned no first matches. It was also not the 2nd or 3rd match chosen for Grippotyphosa and was often a lower match than an incorrect identification of a Bratislava MSP. Therefore, G 2D was not chosen to represent Grippotyphosa.

Since these two MSPs did not result in improved serovar identification, the Peak Statistic report for Bratislava and Grippotyphosa was reassessed and a new PSR was created using the Wilcoxon test to identify a new set of top differentiating peaks (Table 3.12). Six new peaks with PADs ≤ 0.05, Wilcoxon p-values ≤ 0.01, and peak intensity averages > 1.00 for at least one of the two serovars were chosen. The selected peaks were 5076 *m/z*, 6356 *m/z*, 7231 *m/z*, 7825 *m/z*, 7845 *m/z*, and 8058 *m/z*. Pairs of these peaks were then used to create 2D peak distribution plots (Figure 3.2). Some of the same Bratislava raw

spectra that had previously been chosen for MSP omission due to convergent peak distributions again displayed overlapping distributions for the new set of peaks. However, several new spectra for Bratislava and for Grippotyphosa showed peak distribution overlap. These spectra were noted for potential omission in new MSPs.

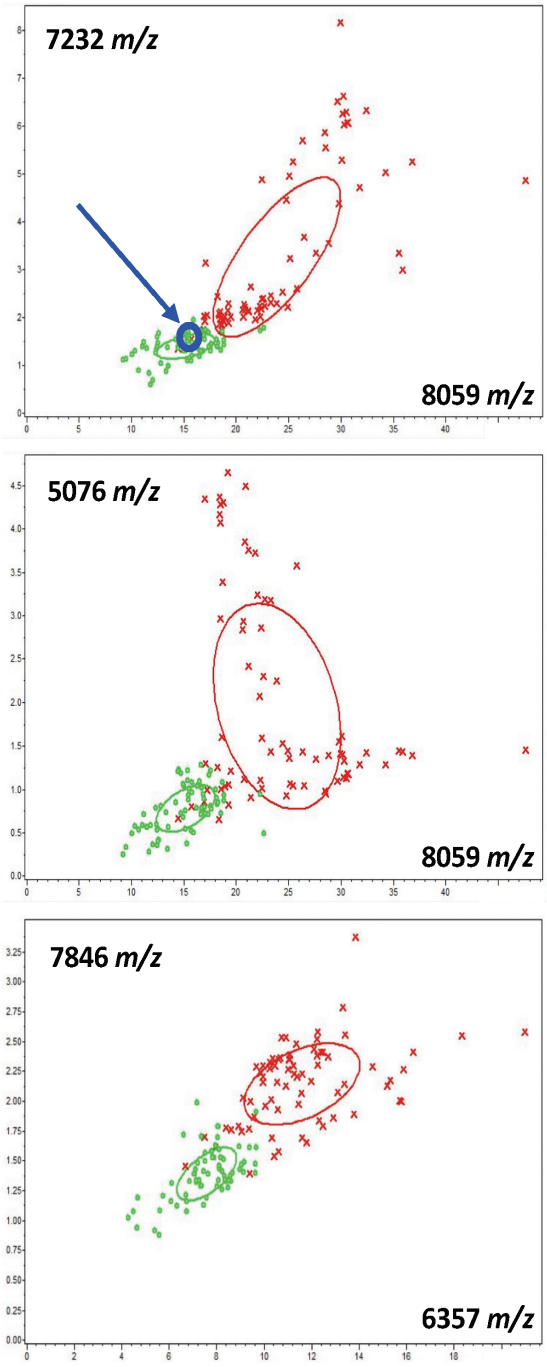

Figure 3.2. ClinProTools' 2D Peak Distribution Plots Displaying the Average Peak Intensity Distributions for Three Pairs of Peaks Within the Raw Spectra of Serovars Bratislava and Grippotyphosa

Each red 'x' represents a Bratislava raw spectrum, while each green circle represents a Grippotyphosa raw spectrum. The x-axis shows the distribution of the area/intensity values for the peak given in the lower right corner. The y-axis shows this distribution data for the peak given in the upper left corner. Values along both axes are given in arbitrary units. Ellipses represent the standard deviation of the average peak intensity for each serovar class. Overlapping markers for the two serovars represent raw spectra which may reduce the specificity of the corresponding serovar's MSPs. Several raw spectra that fell within the opposite serovar's ellipsis, or that fell outside their own ellipsis and overlapped with spectra from the opposite serovar, were identified for removal in newly created test MSPs. An example of such a spectrum is indicated in the top panel, where a red 'X', denoting a Bratislava spectrum, is circled, indicating its located within the ellipsis for Grippotyphosa.

Color key:

- *L. interrogans* Bratislava
- *L. interrogans* Grippotyphosa

3.4.10.3 Principal Component Analysis

PCA was performed on both the group of seven serovars, and for the serovar pair of Bratislava and Grippotyphosa. This method was used to visualize the variability between and among serovar raw spectra and identify potentially influential factors useful for sample differentiation. The 27 raw spectra in each serovar's training subset was used for PCA of the seven serovar group. All spectra used in Bratislava's and Grippotyphosa's MSP were used for the PCA of the serovar pair. Spectra were first normalized, centered, and scaled (Pareto), then used to calculate the relevant PCs. The top three principal components, (PC1-PC3), explaining the majority of the variance in the dataset, were examined in each PCA. Scores and loadings plots were created to assess these PCs.

3.4.10.3.1 PCA for seven serovar group

PCA results for the seven serovar comparison are shown below (Figure 3.3). The top row of the figure displays one 3D and three 2D scores plots. Each plot point represents an individual spectrum, and spectra for each serovar are represented by a different color. The bottom row displays the loadings plots, with each 2D loadings plot corresponding to the scores plot above it. Points in the loadings plots represent individual peaks (*m/z*). These peaks, also termed variables, are chosen based on their contribution to the variance explained by the corresponding principal component. Loadings values range from -1 to 1, with 0 representing no influence and peaks located toward either end of this spectrum considered to be greater contributors to the explained variance in the associated PC (Shao et al. 2012).

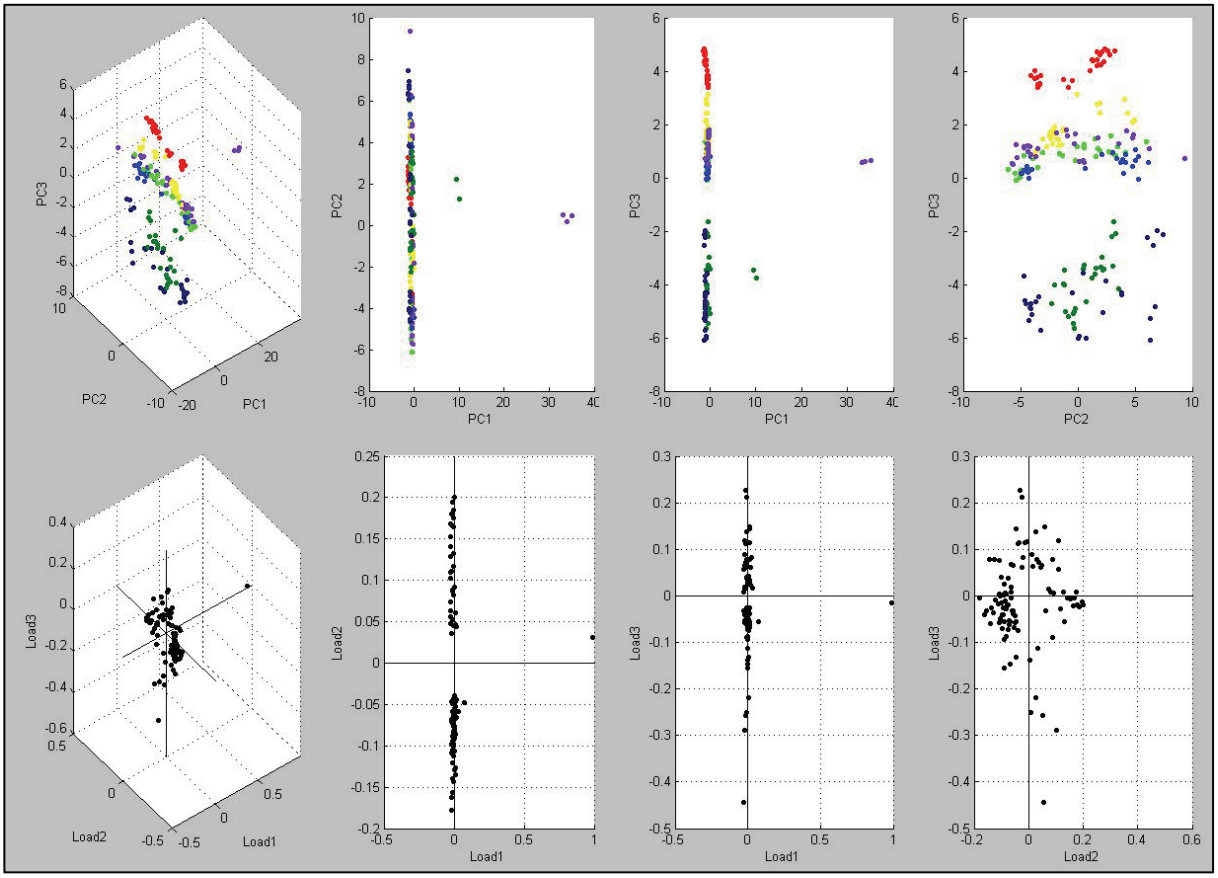

Figure 3.3. Principal Component Analysis (PCA) for Seven *Leptospira* Serovars using Raw Spectra Training Subsets
Twenty-seven raw spectra for each serovar were used to perform the PCA. Each point in the top four Scores plots represents a raw spectrum. Each serovar is represented by a different color: Autumnalis (red), Bratislava (light green), Canicola (light blue), Grippotyphosa (yellow), Hardjo (purple), Icterohaemorrhagiae (dark green), Pomona (dark blue). Each point in the Loadings plots of the second row represents a peak (*m/z* value) from the Peak Statistic report. Loadings values for each point (variable) are based on that variable's contribution to the explained variance in the respective PC. Peaks with greater values (both positive and negative), located farther away from 0, indicate those responsible for a greater portion of the variance for the corresponding PC.

Color key: *L. interrogans* Autumnalis *L. interrogans* Bratislava *L. interrogans* Canicola *L. interrogans* Grippotyphosa *L. interrogans* Hardjo *L. interrogans* Icterohaemorrhagiae *L. interrogans* Pomona

The first 2D scores plot (PC A) showed no clear separation of serovar classes for PC1 or PC2. These two PCs were mainly influenced by a few outlier raw spectra for serovars Hardjo (n=4) and Icterohaemorrhagiae (n=2). These outlier spectra were located away from the main body of plotted points, with the Hardjo outliers having the highest positive values for both PCs. The first loadings plot (Load1 and Load2), which corresponded to PC A, showed only one point, representing peak 3003 *m/z*, located away from the rest of the points for Load1 (Figure 3.4). This point's location indicated that this peak contributed substantially more to the variance represented in PC A compared to that of the other points, all located around 0 for Load1. Per the Peak Statistic report, the average intensity for this peak was highest for

serovars Hardjo and Icterohaemorrhagiae. The raw spectra for these two serovars were subsequently loaded into FlexAnalysis to compare the data for this peak in the outlier versus non-outlier spectra. The peak lists for the three Hardjo outlier spectra located on the right side of the PC A plot, with scores of around 30 for PC1, all contained this peak. However, the peak list for the Hardjo outlier spectrum located at the top of this plot, with a score of 0 for PC1, did not show this peak. The peak lists for all other, non-outlier Hardjo spectra were also missing this peak. Serovar Icterohaemorrhagiae raw spectra were also subsequently reviewed in FlexAnalysis. While peak lists for this serovar's outlier spectra contained this peak, it was missing from the peak lists of all non-outlier Icterohaemorrhagiae spectra. The higher average intensity for this peak among these three Hardjo outlier spectra compared to this peak's average intensity in the Icterohaemorrhagiae outlier spectra, explain why PC1 scores were higher for the Hardjo outliers. This peak was not selected by the GA, SNN, or QC models, nor was it in the top 10 discriminatory peaks of the Peak Statistic report. Per the PWKW sort order used in the PSR, this peak was ranked 154th out of 169 in its ability to discriminate between the serovars, with all serovars, except Hardjo and Icterohaemorrhagiae, having an average intensity for this peak < 1.00.

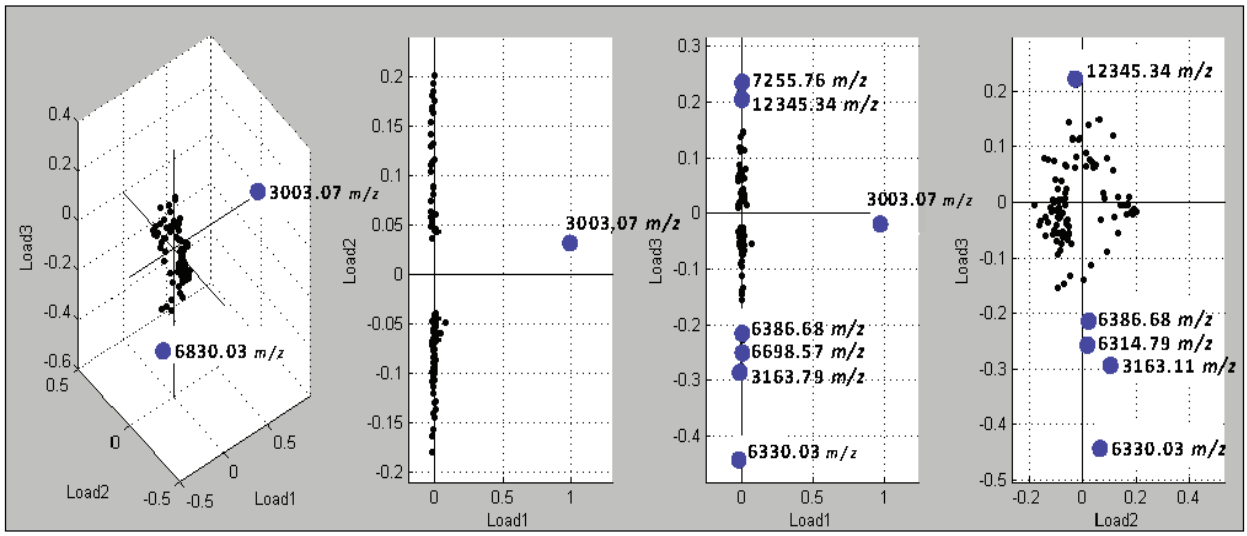

Figure 3.4. Loadings plots from the Principal Component Analysis of seven *Leptospira* serovars with selected points labeled
Four loadings plots, one 3D and three 2D, were created for the PCA. Each of the three 2D loadings plots presents the peaks contributing to the variance explained by its corresponding scores plot (shown in Figure 3.3). Each point in the loadings plots represents the average intensity of a particular spectral peak (*m/z*). The first 2D loadings plot showed peak 3003.07 *m/z*, as being the most influential regarding the variance in PC A (PC1 and PC2). In the second 2D loadings plot, peaks 6386.68, 6698.57, 6330.03, 3163.11 *m/z*, with negative scores in Load3, appeared to be influential. Additionally, peaks 7255.76 and 12345.34 *m/z*, with positive scores for Load3, were revealed as major contributors to the variance explained in PC B (PC1 and PC3). Peak 3003.07 *m/z* was also given again for Load1. In the third loadings plot, peaks 6330.03, 3163.11, 6314.79. 6386.68 and 12345.34 *m/z* were given as influential peaks for Load3.

Review of all Hardjo raw spectra in FlexAnalysis did not reveal a clear reason why the Hardjo outlier spectrum in PC2, at the top of PC A, was positioned as such. A difference in peak presence/absence

between the outlier spectrum and the other spectra did not seem to be the cause for its outlier status. Resolution, S/N, and relative intensity were similar for most peaks amongst the raw spectra. The outlier spectrum did have higher relative intensities for a set of seven peaks, ranging from 3177 *m/z* to 3265 m/z, compared with the relative intensities for these peaks in the other spectra. The intensities for these peaks in the outlier spectrum ranged from 65.1-100. In the non-outlier spectra, the intensities ranged from 22.3-57.5. While these differences in relative intensities may have contributed to the outlier spectrum's position in PC2, this cannot be determined with certainty.

In PC B (PC1 and PC3), five of the same outlier spectra identified along PC1 in PC A were again identified along PC1. Along PC3, the spectra were separated into three different clusters. The first, located at the top of the plot, consisted only of raw spectra for serovar Autumnalis, with no overlap of points representing the spectra of other serovars. The second cluster consisted of raw spectra for serovars Bratislava, Canicola, Grippotyphosa, and Hardjo. PC3 scores for these spectra were positive, yet lower than those for Autumnalis. Many of the spectra in this cluster overlapped, indicating similarity in particular features between these serovars. The third cluster was made up of spectra for serovars Icterohaemorrhagiae and Pomona. These spectra had negative PC3 scores (Figure 3.5). Feature similarity for these two serovars was also suggested by spectra overlap. Raw spectra that are represented by overlapping points may contribute to lower specificity in their respective MSPs.

In the corresponding second Loadings plot (Load1 and Load3), two points were located at the top of the plot, away from the main point cloud. These points had the highest positive values for Load3. These points represented peaks 7255 *m/z* and 12345 *m/z*. Peak 7255 *m/z* was ranked 2nd by the PSR and selected by both the SNN and QC models. Peak 12345 *m/z* was ranked 25th by the PSR and chosen by only the SNN model. Per the PSR, serovar Autumnalis had the highest average intensity for these two peaks. This would explain why Autumnalis raw spectra was clustered separately with the highest scores for PC3. The second Loadings plot also showed four points plotted towards the bottom of the plot, separate from the main point cloud, with negative Load3 values. These points represented peaks 3163 *m/z*, 6330 *m/z*, 6386 *m/z*, and 6698 *m/z*. A review of the Peak Statistic report found that the average intensities for these four peaks were highest for serovars Icterohaemorrhagiae and Pomona**.** This may explain the clustering of these two serovars within the negative scores range of PC3. Peak 3163 *m/z* was ranked 22nd by the PSR, and not chosen by any of the three classification models. Peaks 6330 m/z, and 6698 *m/z* were ranked 13th and 6th, respectively, by the PSR. These two peaks were selected by the SNN and QC models. Lastly, peak 6386 was ranked 8th by the PSR, and selected by both the GA and QC models.

PC C (PC2 and PC3) again showed no serovar differentiation along PC2. PC3 again the same three serovar clusters seen in PC B. The corresponding third Loadings plot (Load2 and Load3), again showed a point towards the top of the plot, representing influential peak 12345 *m/z*. Towards the bottom of the plot, four of five points positioned the furthest away from 0 along Load3 represented the same peaks

identified in the second Loadings plot. An additional point, representing peak 6314 *m/z*, was also plotted in this area. This peak was ranked 15th by the PSR and selected by both the SNN and QC models. Average intensities for this peak were highest for serovars Icterohaemorrhagiae and Pomona. This point contributed to the clustering of these two serovars along PC3 in PC C.

Principal components 1, 2, and 3 contributed to approximately 28%, 17%, and 14% of the overall variance.

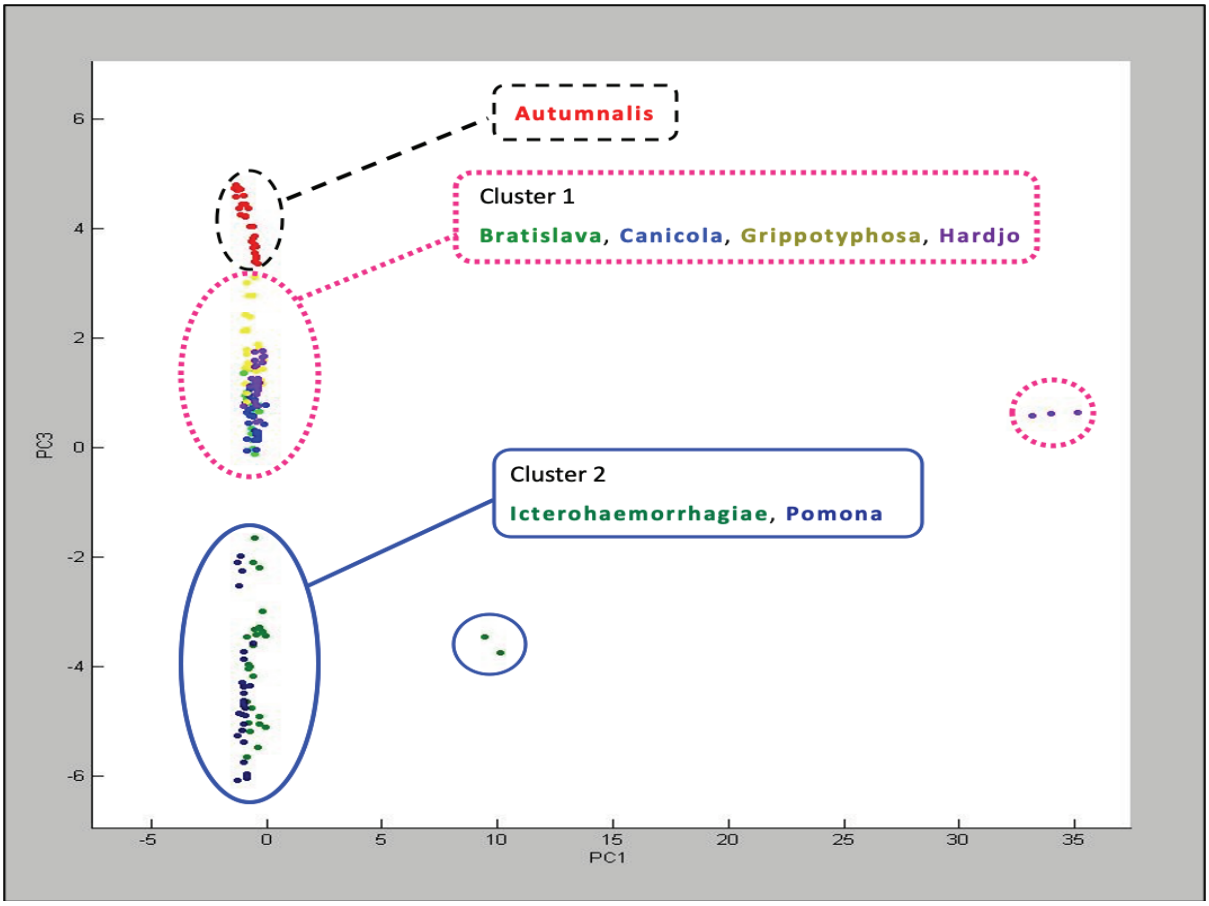

Figure 3.5. Enlarged View of the PC B (PC1 and PC3) Scores Plot for the PCA of Seven Serovars This figure presents an enlarged view of the PC1 and PC3 scores plot, generated for PCA of the seven serovars. The points represent individual raw spectra, with the spectra for each serovar represented by a different color. PC 1 is plotted along the x-axis, while PC 3 is plotted along the y-axis. Raw spectra for serovar Autumnalis, (in red), are grouped together at the high end of PC 3, with no overlap from other serovars. Spectra for serovars Bratislava (light green), Canicola (light blue), Grippotyphosa (yellow), and Hardjo (purple), are mostly grouped together to make up Cluster 1. Three raw spectra for serovar Hardjo are located on the right side of the figure (denoted by the dotted pink circle), away from the main group of spectra. Spectra for serovars Icterohaemorrhagiae (dark green) and Pomona (dark blue) form a second cluster within the region of negative values for PC 3. Two spectra for Icterohaemorrhagiae, surrounded by the blue circle, are located apart from the spectra that make up this second cluster. The outlier spectra, located in the two small circles, contain peak 3003 *m/z* at a higher average intensity compared to that found for this peak in all of the other spectra.

Color key: *L. interrogans* Autumnalis *L. interrogans* Bratislava *L. interrogans* Canicola *L. interrogans* Grippotyphosa *L. interrogans* Hardjo *L. interrogans* Icterohaemorrhagiae *L. interrogans* Pomona

Overall, only three of the 10 peaks identified as influential in these first three PCs were in the top 10 discriminatory peaks selected by the PSR and/or selected by the GA model (which was chosen as the best-performing model for the group of seven serovars because it had the highest RC and CV values) (Table 3.13). The peaks found by PCA were expected to be somewhat different from the group of discriminatory peaks chosen by the PSR and models since each uses a different algorithm. It should be noted that since the PCA was not designed for classification influential peaks identified in the loadings plots are not necessarily the best peaks for differentiating between the loaded serovar classes (Xi et al. 2014). However, further evaluation of peaks found to differ between groups of spectra for the same serovar may offer insight into how sample spectra acquired under different conditions may change and which of these conditions create sample spectra that best represent each serovar.

TABLE 3.13. Peaks Contributing the Most to the Explained Variance in PCs 1-3 for PCA Performed on the Training Subset Spectra for Seven *Leptospira* Serovars

Select Peaks *(m/z)*	Average Peak Intensity per Serovar[a] (a.u.)							Rank in PSR[b]	Selected by Classification Model?		
	A	B	C	G	H	I	P		GA[c]	SNN[d]	QC[e]
3003	0.86	0.99	0.96	0.90	***11.56***	3.36	0.91	154	--	--	--
3163	2.76	5.07	10.83	4.54	7.56	21.24	***38.42***	22	--	X	X
6314	2.24	2.11	2.50	1.85	1.73	7.73	***10.93***	15	--	X	X
6330	6.43	2.24	2.26	1.96	1.96	48.47	***87.28***	13	--	X	X
6386◆	0.53	1.38	1.56	2.08	1.83	4.14	***5.43***	8	X	--	X
6698◆	1.22	1.13	6.87	1.07	0.85	***11.43***	7.63	6	--	X	X
7255◆	***12.76***	4.70	0.97	6.75	3.69	2.74	0.56	2	--	X	X
12345	***2.29***	0.24	0.37	0.21	0.33	0.30	0.28	25	--	X	--

Eight peaks were identified as contributing the most to the explained variance in PCs 1–3 of a PCA performed on the training subset of spectra for all seven *Leptospira* serovars. Average peak intensity for each peak is, given in arbitrary units (a.u.) for each serovar. The highest average intensity for each peak is indicated in bold, italic text. Serovars Pomona and Autumnalis had the highest average intensities for four of the eight, and two of the eight peaks, respectively. Hardjo and Icterohaemorrhagiae each had the highest average intensity for one of the eight peaks. Serovars Bratislava, Canicola, and Grippotyphosa did not have the highest average intensity for any of the peaks. PCA-identified peaks were compared to those selected in the PSR (n=169). PCA-identified peaks covered a range of PSR peak ranks, (e.g. 7255 *m/z,* rank 2 to 3003 *m/z*, rank 154). PCA-identified peaks were also compared to peaks selected by the three classification models. Peaks selected by each model are indicated by the presence of an 'X'.

[a] **A**: Autumnalis, **B**: Bratislava, **C**: Canicola, **G**: Grippotyphosa, **H**: Hardjo, **I:** Icterohaemorrhagiae, **P**: Pomona.

[b] **PSR**: Peak Statistic Report

[c] **GA**: Genetic Algorithm model

TABLE 3.13. ***continued***

[d] **SNN**: Supervised Neural Network model

[e] **QC**: Quick Classifier model

◆ Indicates peak was chosen for further analysis.

3.4.10.3.2 PCA for serovar pair Bratislava and Grippotyphosa

In the PCA performed for serovars Bratislava and Grippotyphosa, scores plots for the top three PCs showed a mixture of separate and overlapping spectra for the two serovar classes (Figure 3.6). In PC A (PC1 and PC2), no clear separation between the two serovars was seen in the majority of the spectra along PC1. However, a cluster of Bratislava spectra had the highest positive scores, and so were plotted to the right of the rest of the spectra. Differentiation between the two serovars was increased along PC2, with most Bratislava spectra assigned negative scores and most Grippotyphosa spectra assigned positive scores.

The first Loadings plot showed that peak 5076 *m/z* appeared to have the greatest influence on the variance represented in PC1. Peak 5076 was ranked fourth in the PSR. This peak's average intensity in Bratislava was 2.17, whereas for Grippotyphosa, it < 1.00. Peak lists for the raw spectra of these serovars were reviewed for this peak in FlexAnalysis. This peak was not present in the peaks lists for Grippotyphosa spectra. This peak was also not selected by any of the three classification models created for this serovar pair. This was not unexpected as the models were built using the training subset of spectra for Bratislava and Grippotyphosa rather than all spectra included in their respective MSPs (Table 3.14). Load2 showed points representing peaks 5420 *m/z*, 5671 *m/z*, 7231 *m/z*, and 11338 *m/z* as having the most influence on the variance explained in PC2. Peak 5420 *m/z* was selected by the QC classification model. It was also ranked seventh in the PSR for these two serovars. Average intensity for this peak was 1.74 for Grippotyphosa, but < 1.00 for Bratislava. The difference in intensity level, and location of this point in the first Loadings plot corresponds to Grippotyphosa, rather than Bratislava, spectra in PC A. Peak 5671 *m/z* was not selected by any of the classification models. It was ranked 13th in the PSR, with an average peak intensity of 8.43 for Grippotyphosa, and 2.24 for Bratislava. Peak 7231 *m/z*, represented by a point towards the bottom of Loading plot 1, was ranked third in the PSR. The average peak intensity was 3.62 for Bratislava and 1.55 for Grippotyphosa. This peak had the lowest loading value for PC2 and corresponded with negative scores for PC2, where the majority the spectra belong to Bratislava. This peak was selected by both the SNN and QC models. Peak 11338 *m/z* had the highest value for Load2. It was ranked ninth in the PSR, with an average peak intensity of 2.97 for Grippotyphosa and an average of < 1.00 for Bratislava. The greater average intensity of this peak in Grippotyphosa corresponds with the higher PC2 scores for this serovar.

In PC B (PC1 and PC3), spectra for Bratislava were arranged into two clusters along PC1. One cluster overlapped with Grippotyphosa spectra, and had scores ranging from around -7 to around 2. The second cluster was made up of only Bratislava, and had scores ranging from approximately 3 to 7.

Spectra for Grippotyphosa also formed two groups along PC1. Scores for one group ranged from around -7 to 2 and overlapped with some of the Bratislava spectra in this range. A second group, with scores ranging from around 3 to 5, was clustered separately. In PC3, a cluster of overlapping Bratislava and Grippotyphosa spectra extended along almost the range of scores for this PC. Again, a separate cluster of Bratislava and of Grippotyphosa spectra were also evident along PC3, falling mainly within positive score values. A closer look at the spectra that made up these clusters found that the separate Bratislava group consisted mainly of spectra acquired in the first set of spectra collected for this serovar. Likewise, the spectra in the separate Grippotyphosa group were mainly those collected in the first set for this serovar. In contrast, most of the Bratislava and Grippotyphosa spectra clustered together were those in the second and third sets of raw spectra acquired for these serovars.

The second Loadings plot (Load1 and Load3) did not yield any new influential peaks for PC1. However, peak 3443 *m/z* had the highest loadings value for Load3. This corresponded to the cluster of Grippotyphosa-only spectra located at the top of the PC B plot. This cluster consisted of only raw spectra from the original set of spectra acquired for this serovar Accordingly, FlexAnalysis showed that this peak was present only in the peak lists for the original set of Grippotyphosa spectra (7). This peak was ranked 75th in the PSR and had a higher average intensity for Grippotyphosa (5.37) than for Bratislava (2.19). It was not selected by any of the classification models as it only served to differentiate one set of Grippotyphosa spectra from the other two, rather than differentiating one serovar from another. Since the three sets of Grippotyphosa raw spectra used in this PCA came from three different subcultures, small variations between each set's mass spectra was not unexpected. In this case, a potential difference in serovar culture age between that used for the original set of raw spectra versus the other two, (up to a week difference), may have contributed to the production of this peak. A variety of other factors, including contamination of the culture or sample preparation reagents, could have also contributed. Since the two spectra sets that don't contain this peak can still be differentiated from the other serovars, this peak is not vital for serovar Grippotyphosa identification and differentiation. However, if further refinement of serovar MSPs were to be done, it would be interesting to compare the original Grippotyphosa raw spectra set with a newly acquired set. If the new set does not contain peak 3443 *m/z,* then it would be recommended that a new MSP be created with the three sets of raw data that do not contain the peak. Then, both the new and the current Grippotyphosa MSPs should be tested to see which offers greater specificity.

In PC C (PC2 and PC3), spectra for the two serovars were again distinguished along PC2, with negative scores for the majority of Bratislava spectra and positive scores for most of the Grippotyphosa spectra. In PC3, spectra for both serovars spanned the range of scores, creating no obvious separation of the two classes. The third Loadings plot revealed no new discriminatory peaks for PC2 or PC3.

PCs 1, 2, and 3 represented approximately 37%, 18%, and 14% of the total explained variance, respectively. In all, 12 overlapping spectra were identified for omission from the spectra used to create four new MSPs for each of these serovars.

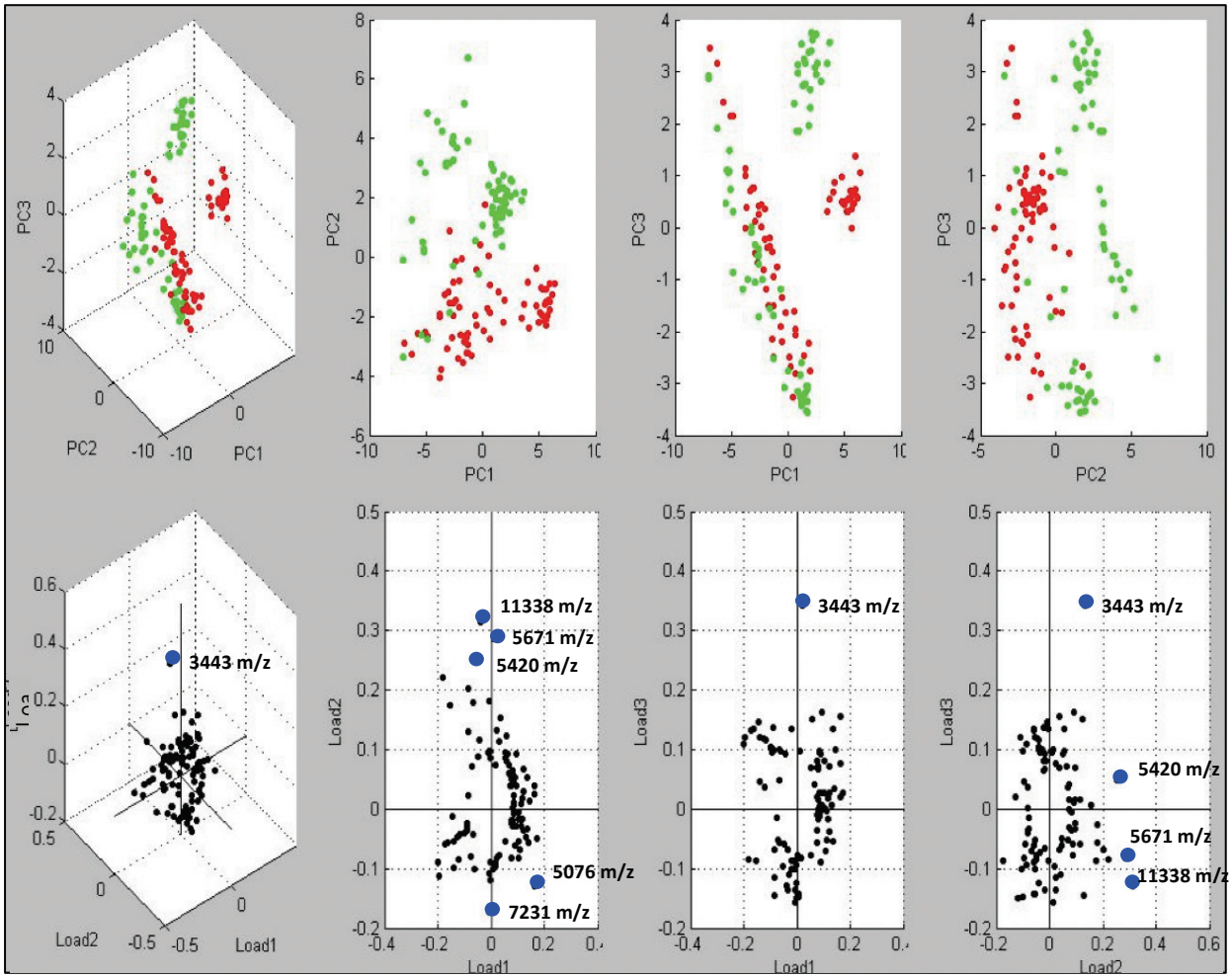

Figure 3.6. Principal Component Analysis (PCA) for Bratislava and Grippotyphosa Raw Spectra All raw spectra used to create MSPs for these two serovars were used in this analysis. Each point in the top four scores plots represents a raw spectrum from one of three different sets of spectra collected for each of these serovars. Bratislava spectra are represented by red points. Grippotyphosa spectra are represented by green points. Spectra of the two serovars that overlap indicate the potential for serovar misidentification and may contribute to lower specificity in their respective MSPs. Points labeled in the bottom four loadings plots denote those peaks (variables) that were identified as contributing the most to the variability explained in the corresponding PC. Samples used in the analysis were prepared from culture suspensions adjusted to approximately 3 x 10^9 cells/mL. Axis values for these plots are in arbitrary units (a.u.).

Color key: ● *L. interrogans* Bratislava ● *L. interrogans* Grippotyphosa

TABLE 3.14. Data on Top PCA-Identified Peaks - Bratislava and Grippotyphosa PCA

Select Peaks (*m/z*)	Average Peak Intensities (a.u.)		Rank in PSR[a]	Selected by Classification Model?		
	Bratislava	Grippotyphosa		GA[b]	SNN[c]	QC[d]
3443	2.19	5.37	75	--	--	--
5076	2.17	0.87	4	--	--	--
5420	0.50	1.74	7	--	--	X
5671	2.24	8.43	13	--	--	--
6386	1.43	2.05	91	--	--	--
7231	3.62	1.55	1	--	X	X
11338	0.67	2.97	9	--	--	--
16648	0.51	0.68	151	--	--	--

Eight main peaks were identified as the top contributors to the explained variance represented by Principal Components 1–3 (PC1–PC3) in the Bratislava and Grippotyphosa PCA.

[a] **PSR**: Peak Statistic Report

[b] **GA**: Genetic Algorithm model

[c] **SNN**: Supervised Neural Network model

[d] **QC**: Quick Classifier model

3.4.10.4 Classification models

The Genetic Algorithm (GA), Quick Classifier (QC) and Supervised Neural Network (SNN) algorithms were used to generate classification models for the serovar group and pair comparisons. Recognition capability (RC) and cross validation (CV) were calculated for all models. Peaks selected by the most robust model(s) for the serovar group and pair comparisons were used for further analysis. Serovar test spectra subsets were classified against the models in external validation. Each spectrum that classified incorrectly was noted and removed from the respective serovar's MSP.

3.4.10.4.1 Models for seven serovar group

Of the three models generated for the comparison of the seven serovar group, the Genetic Algorithm produced the best-performing model, with the highest RC (100%) and the highest CV (98.41%) values. This model selected a combination of five discriminatory peaks, three of which were also found in the top ten discriminatory peaks of the PSR for the group of seven serovars. Compared to the GA model, the Supervised Neural Network model had lower overall RC (98.94%) and CV values (85.71%). Interestingly, the individual CV value for Grippotyphosa was only 18.52%, while the individual CV values for the other serovars ranged from 92.59% to 100%. Therefore, while the SNN model's overall RC and CV values were

relatively high, this model failed to discriminate Grippotyphosa isolates. This model identified 24 discriminatory peaks, six of which were also within the PSR's top 10 selected peaks (Table 3.15). The Quick Classifier model had the lowest RC value, at 90.48%. It's CV value of 85.71% was tied with that of the SNN model (Table 3.16). Eighteen peaks were chosen by the QC model, 10 of which were the same as the top 10 peaks in the PSR. Since the PSR and the QC model both use univariate statistical methods for peak selection, the higher number of top ten peaks common between the PSR and this model was not unexpected. Two peaks were selected by all three models (5490 *m/z* and 7234 *m/z*).

TABLE 3.15. Discriminatory Peaks Selected by the PSR, and by the GA, SNN, and QC Models for Differentiation Between the Seven *Leptospira* Serovars

Rank of Peak in PSR[a]	Peak (*m/z*)	PSR	GA[b]	SNN[c]	QC[d]
104	1997			X	
50	2706			X	
46	3717		X		
95	4927			X	
9	**5490**	X	X	X	X
41	5672			X	
49	5814			X	
15	6315			X	X
13	6330			X	X
7	**6357**	X		X	X
8	**6387**	X	X		X
6	**6698**	X		X	X
3	**6712**	X			X
10	**7234**	X	X	X	X
2	**7256**	X		X	X
5	**7360**	X		X	X
4	**7388**	X		X	X
1	**7406**	X			X
23	7437			X	
17	7825			X	X
16	7846				X
43	7976		X	X	
29	7996			X	
18	8015				X
11	8059			X	X
14	8454				X
45	8717			X	
42	10432			X	
12	10975				X
25	12345			X	

Top 10 peaks in the PSR are listed in bold. Table rows for these 10 peaks are shaded gray. Only two peaks, 5490 *m/z* and 7234 *m/z,* were selected by all three models as well as the PSR.

[a] **PSR** (Peak Statistic report)

[b] **GA** (Genetic Algorithm model)

[c] **SNN** (Supervised Neural Network model)

[d] **QC** (Quick Classifier model)

TABLE 3.16. ClinProTools Classification Models Created for Comparison of Seven Serovars					
Model	**Max # Peaks Used in Model**	**# of Differentiating Peaks Selected by Model**	**RC**[a] (%)	**CV**[b] (%)	**External Validation** (% of correctly classified spectra)
Genetic Algorithm	All Peaks in PSR	5	**100**	**98.41**	**97.88**
Supervised Neural Network	All Peaks in PSR	24	98.94	85.71	97.37
Quick Classifier	All Peaks in PSR	18	90.48	85.71	93.72

Comparison of recognition capability, cross validation, and percent of correctly classified spectra in external validation for the three classification models created for the seven pathogenic serovars. A total of 189 raw spectra, (27 per serovar), were used for external validation. A 20% leave one out cross validation method using 10 iterations was applied. The Genetic Algorithm model was the most robust, with a RC of 100%, a CV of 98.41%, and of 98% of all test subset spectra correctly classified in external validation.

[a] **RC** (recognition capability)

[b] **CV** (cross validation)

Confusion matrices were created to visualize results of external validation for all three models (Table 3.17). In the GA model, only four out of 189 total raw spectra were classified incorrectly. Two of serovar Bratislava's spectra were identified as Grippotyphosa, while two of Grippotyphosa's spectra were classified as Bratislava. This reflected the misclassification occasionally seen between these two serovars. The GA model had the highest percentage of correctly classified spectra for external validation, at 98%. In the SNN model, three of the seven serovars returned incorrect classifications. Serovar Bratislava incorrectly identified twice as Canicola and once as Hardjo. Additionally, one raw spectrum for serovar Canicola, and two for serovar Icterohaemorrhagiae, were misclassified as serovar Hardjo. External validation of the SNN model returned the second highest percent of correctly classified spectra, at 97%. In the QC model, a total of 11 raw spectra (5.8%) were misclassified. Five Bratislava spectra (19.5%) were incorrectly classified: four as Grippotyphosa and one as Canicola. The other misclassified spectra consisted of three Canicola spectra identified as Autumnalis (11.1%), one Hardjo spectrum identified as Grippotyphosa (3.7%) and two Icterohaemorrhagiae spectra identified as Pomona (7.4%). External validation of this model returned the lowest percentage of correctly classified spectra, at 94%. Only one raw spectrum, belonging to serovar Bratislava, was misclassified in all three models. In the GA and QC models, this spectrum was identified as Grippotyphosa. In the SNN model, this spectrum was identified as Hardjo. One additional raw spectrum, again belonging to serovar Bratislava, was misclassified in both the SNN and QC models. It was classified as Canicola in the SNN model, and as Grippotyphosa in the QC model. Overall, all three models had a high percentage of correctly classified

test subset spectra, indicating that these models offer a method by which these seven reference serovars can be correctly differentiated.

TABLE 3.17. External Validation Results for the **Genetic Algorithm, Supervised Neural Network**, and **Quick Classifier Models** created for the **Seven Serovar Group**

A.	GA Model	Predicted Class						
Target Class		**A**	**B**	**C**	**G**	**H**	**I**	**P**
	Autumnalis (**A**)	**27**	0	0	0	0	0	0
	Bratislava (**B**)	0	**25**	0	2	0	0	0
	Canicola (**C**)	0	0	**27**	0	0	0	0
	Grippotyphosa (**G**)	0	**2**	0	**25**	0	0	0
	Hardjo (**H**)	0	0	0	0	**27**	0	0
	Icterohaemorrhagiae (**I**)	0	0	0	0	0	**27**	0
	Pomona (**P**)	0	0	0	0	0	0	**27**

B.	SNN Model	Predicted Class						
Target Class		**A**	**B**	**C**	**G**	**H**	**I**	**P**
	Autumnalis (**A**)	**27**	0	0	0	0	0	0
	Bratislava (**B**)	0	**24**	**2**	0	**1**	0	0
	Canicola (**C**)	0	0	**26**	0	**1**	0	0
	Grippotyphosa (**G**)	0	0	0	**27**	0	0	0
	Hardjo (**H**)	0	0	0	0	**27**	0	0
	Icterohaemorrhagiae (**I**)	0	0	0	0	**2**	**25**	0
	Pomona (**P**)	0	0	0	0	0	0	**27**

C.	QC Model	Predicted Class						
Target Class		**A**	**B**	**C**	**G**	**H**	**I**	**P**
	Autumnalis (**A**)	**27**	0	0	0	0	0	0
	Bratislava (**B**)	0	**22**	**1**	**4**	0	0	0
	Canicola (**C**)	**3**	0	**25**	0	0	0	0
	Grippotyphosa (**G**)	0	0	0	**27**	0	0	0
	Hardjo (**H**)	0	0	0	**1**	**26**	0	0
	Icterohaemorrhagiae (**I**)	0	0	0	0	0	**25**	**2**
	Pomona (**P**)	0	0	0	0	0	0	**27**

TABLE 3.17. *continued*

The three panels show confusion matrices generated by external validation of the three models (GA, SNN, QC) using serovar test spectra subsets for LOO cross validation.

Since the GA model had the highest RC, CV and percent of correctly classified test spectra, the five peaks selected by this model were chosen for further analysis. Three of these peaks were found in the top 10 discriminatory peaks listed for the PSR. Interestingly, the two other peaks selected by the model were ranked in the PSR as the 43rd, (7976 m/z) and the 46th, (3717 m/z), in discriminatory power. The five model-selected peaks plus the top six PSR-selected peaks were chosen for subsequent analyses (Table 3.18).

TABLE 3.18. Discriminatory Peaks Selected for the Seven Serovar Group

Peak Source	Peak 1 (Da)	Peak 2 (Da)	Peak 3 (Da)	Peak 4 (Da)	Peak 5 (Da)	Peak 6 (Da)	Peak 7 (Da)	Peak 8 (Da)	Peak 9 (Da)	Peak 10 (Da)	Peak 11 (Da)
PSR[a]	--	--	--	6698	6712	--	7256	7360	7388	7406	--
GA[b]	3717	5490	6387	--	--	7234	--	--	--	--	7976

The table shows the combination of 11 peaks selected for serovar group analysis. The peaks chosen from the PSR were the top six identified as having the greatest statistical separation strength with an average intensity value of ≥ 1 for at least one of the seven serovars. All five peaks selected by the GA model were chosen for use in serovar group analysis. These peaks had Anderson-Darling p-values ≤ p 0.05, indicating non-normal distribution. Mass (*m/z*) of selected peaks are given in Daltons (Da).

[a] **PSR (**Peak Statistic report)

[b] **GA** (Genetic Algorithm model)

3.4.10.4.2 Models for serovar pair Bratislava and Grippotyphosa

The three models generated for the comparison of serovars Bratislava and Grippotyphosa produced almost identical results. The GA and QC models tied for the highest RC and CV values, which were 100% for both parameters. The SNN model also had an RC of 100%, but a slightly lower CV value of 98.15% (Table 3.19). The number of peaks chosen by each model were 5 (GA), 2 (QC), and 8 (SNN). One of the five peaks selected by the GA model, 7976 *m/z*, was also selected by the SNN model. Likewise, one peak chosen by the SNN model, 7231 *m/z*, was also chosen by the QC model. No peak was found in all three models (Table 3.20).

Table 3.19. ClinProTools Classification Models Created for the Comparison of Serovars Bratislava and Grippotyphosa

Model	Max # Peaks Used in Model	# of Peaks Selected by Model	RC[a] (%)	CV[b] (%)	External Validation (% of correctly classified spectra)
Genetic Algorithm	All Peaks	5	**100**	**100**	96.30
Supervised Neural Network	All Peaks	8	100	98.15	**98.15**
Quick Classifier	All Peaks	2	**100**	**100**	94.45

Comparison of the recognition capability, cross validation, and percent of correctly classified spectra in external validation for the three classification models created for the comparison of serovars Bratislava and Grippotyphosa. A 20% leave one out cross validation method using 10 iterations was applied. The Genetic Algorithm and Quick Classifier models tied for best RC and CV values, at 100% each. The % of correctly classified test subset spectra was slightly higher for the GA model (96.30%) than for the QC model (94.45%). The SNN model had the highest % of correctly classified spectra in external validation, at 98.15%. However, given its lower CV value, the peaks selected by this model were not chosen for further analysis.

[a] **RC** (recognition capability)

[b] **CV** (cross validation)

TABLE 3.20. Discriminatory Peaks Selected by the PSR, and by the GA, SNN, and QC Models for Differentiation Between Serovars Bratislava and Grippotyphosa

Rank of Peak in PSR[a]	Peak *(m/z)*	PSR	GA[b]	SNN[c]	QC[d]
23	2011		X		
15	4926			X	
4	**5076**	X			
7	**5420**	X			X
2	**6356**	X			
3	**7231**	X		X	X
42	7255		X		
6	**7825**	X		X	
5	**7845**	X	X		
26	7975		X	X	
1	**8058**	X		X	
11	8728	X		X	
10	**8742**	X			
8	**9452**	X	X		
9	**11338**	X			
12	12513			X	
36	12931			X	

Top 10 peaks in the PSR are listed in bold. Table rows containing data for these 10 peaks are shaded gray. An 'X' indicates that the classification model given at the top of the column selected the peak in the given row. No peak was chosen by all four peak selection methods.

[a] **PSR** (Peak Statistic report)

[b] **GA** (Genetic Algorithm model)

[c] **SNN** (Supervised Neural Network model)

[d] **QC** (Quick Classifier model)

In external validation of the GA model, all 27 Bratislava spectra identified correctly, while two of the Grippotyphosa spectra misidentified as Bratislava. QC model external validation saw two of 27 Bratislava spectra misidentify as Grippotyphosa and one of 27 Grippotyphosa spectra incorrectly identified as Bratislava. For external validation of the SNN model, one of 27 Bratislava spectra incorrectly matched to Grippotyphosa, while all Grippotyphosa spectra were correctly identified. Though the SNN model had the highest percent of correctly classified spectra in external validation, at 98.15%, it had a slightly lower CV

value than the GA and QC models. The percentages of correctly classified spectra in external validation of the GA and QC models were close, at 96.30% and 94.45%, respectively (Table 3.21).

TABLE 3.21. External Validation Results for the **Genetic Algorithm, Supervised Neural Network**, and **Quick Classifier Models** created for the Comparison of **Serovars Bratislava and Grippotyphosa**

A.	**GA Model**	**Predicted Class**	
Target Class		**B**	**G**
	Bratislava (**B**)	**27**	0
	Grippotyphosa (**G**)	**2**	**25**

B.	**SNN Model**	**Predicted Class**	
Target Class		**B**	**G**
	Bratislava (**B**)	**26**	**1**
	Grippotyphosa (**G**)	0	**27**

C.	**QC Model**	**Predicted Class**	
Target Class		**B**	**G**
	Bratislava (**B**)	**25**	**2**
	Grippotyphosa (**G**)	**1**	**26**

The three panels show confusion matrices generated by external validation of the three models GA, SNN, QC using serovar test spectra subsets for LOO cross validation.

The RC and CV values determined the best-performing model per Bruker's guidelines. Since the highest RC and CV values were tied between the GA and QC models, the peaks selected by both of these models were chosen for further analyses. These peaks were combined with the top six discriminatory peaks from the PSR which had an average intensity value of ≥ 1 for at least one of the two serovars. This resulted in a total of nine peaks used for subsequent analysis for this serovar pair (Table 3.22).

TABLE 3.22. Discriminatory Peaks Selected for Serovars Grippotyphosa and Bratislava

Source	Peak 1 (Da)	Peak 2 (Da)	Peak 3 (Da)	Peak 4 (Da)	Peak 5 (Da)	Peak 6 (Da)	Peak 7 (Da)	Peak 8 (Da)	Peak 9 (Da)	Peak 10 (Da)
PSR[a]	--	5076	--	6356	7231	--	7845	--	8058	--
GA[b]	2011	--	--	--	--	7255	7845	7975	--	9452*
QC[c]	--	--	5420	--	7231	--	--	--	--	--

Ten discriminatory peaks were selected for the comparison of serovars Grippotyphosa and Bratislava. Selected peaks were taken from the GA and QC models, since both had RC and CV values of 100%, and from the Peak Statistic report. Peaks chosen from the PSR were the top six ranked peaks which had an average intensity value of ≥ 1 for at least one of the two serovars. All peaks selected by the classification models are also found in the PSR. This is because the models are created from the peaks selected by the PSR. However, if one if the peaks selected by the GA or QC models noted in the table above is not also listed for the PSR, that indicates that the selected peak was not in the top six ranked PSR peaks or that the peak did not reach an average intensity of ≥ 1 for at least one of the two serovars. All peaks identified by the QC model were used. Four out of five peaks identified by the GA model were used.

*9452 *m/z* – this peak selected by the GA model was omitted because its average intensity value for both serovars in the PSR was < 1.0.

Mass (*m/z*) of selected peaks given in Daltons (Da).

[a] **PSR** (Peak Statistic report)

[b] **GA** (Genetic Algorithm Model)

[c] **QC** (Quick Classifier Model)

3.4.10.5 New MSPs for Bratislava and Grippotyphosa created using the data collected from revised 2D peak distribution plots, PCA, and model validation

Data collected from the revised 2D plots, Principal Component Analysis, and external validation of classification models for serovars Bratislava and Grippotyphosa were combined to select a new set of raw spectra to be omitted from four new MSPs created for each of these two serovars. Each of the four new MSP types was created using a different set of creation parameters. Two of these MSPs used the default lower peak picking boundary of 3000 Da. The first of these, named X Revision_3000_70, used a maximum desired peak number of 70 (the default). The second MSP, X Revision_3000_100 used a peak maximum of 100. The third and fourth new MSPs used a lower peak picking boundary of 2000 Da. These MSPs, named X Revision_2000_70 and X Revision_2000_100, used maximum peak numbers of 70 and 100, respectively (Table 3.23). The raw spectra previously omitted from the B 2D and G 2D MSPs were once again included in the group of raw spectra used to construct the new MSPs since omitting them had not improved MSP specificity.

These eight new MSPs, (four for Bratislava and four for Grippotyphosa), were tested in Compass Explorer in individual trials. The trials used raw spectra collected from culture of the respective serovar. All test spectra were acquired from samples adjusted to a concentration of 7.22 x 10^8 organisms/mL, (50

%T and approximately equivalent to the 1:4 dilution used in serial dilution trials).

Table 3.23. Parameters and Performance of Four New MSPs Created for Serovars Grippotyphosa and Bratislava

MSPs	# Raw Spectra	Max. Desired # Peaks	Lower Peak Picking Boundary (Da)	Correct # 1st Matches* (n = 10)
Bratislava				
B Revision_3000_70	64	70	3000	6
B Revision_3000_100	64	100	3000	6
B Revision_2000_70	64	70	2000	4
B Revision_2000_100	64	100	2000	4
Grippotyphosa				
G Revision_3000_70	59	70	3000	3
G Revision_3000_100	59	100	3000	7
G Revision_2000_70	59	70	2000	2
G Revision_2000_100	59	100	2000	4

Identification scores for 4 new MSPs created for serovars Bratislava and Grippotyphosa using different preprocessing and MSP creation method parameters and the selective exclusion of particular raw spectra based on CPT analyses. *Number of correct first matches does not include correct first matches with scores tied with a second, incorrect match.

For serovar Bratislava, the two MSPs created with a lower bound of 3000 Da yielded better results than the two MSPs created with a lower bound of 2000 Da. B Revision_3000_70 and B Revision_3000_100 both returned correct first matches for six out of 10 sample spots, compared with four of 10 first matches for B Revision_2000_70 and B Revision_2000_100. The two MSPs with a lower number of correct first matches were eliminated from further analysis. Of the two MSPs with a higher number of correct matches, B Revision_3000_100 had a slightly higher score range (1.85–2.56) and median score (2.14) than the score range (1.82–2.51) and median score (2.04) for B Revision_3000_70. B 100, chosen as the best-performing, previously-created MSP for Bratislava, was also tested individually using these same raw test spectra. Results for B Revision_3000_70, B Revision_3000_100, and B 100 were compared. B 100 returned the same number of correct first matches (6) as the two new MSPs and had a slightly higher score range (1.97–2.56). Its median score, of 2.08, was slightly lower than that for B Revision_3000_100. An additional consideration was the number of correct first matches that tied with an incorrect match. As an example, for a particular sample spot, B Revision_3000_70 returned a correct first match with a score of 1.82. However, the second match for this spot, G 100, also had a score of 1.82.

These tied scores mean that the sample spot has the same probability of belonging to Grippotyphosa as it does to Bratislava. This type of result was not counted as a correct first match for Bratislava because it was not specific for this serovar. B Revision_3000_70 and B Revision_3000_100 both had two tied matches, while B 100 had one. All three of these MSPs performed similarly.

Of the four new MSPs created for Grippotyphosa, G Revision_3000_100 yielded the best results. It returned 7 out of 10 correct first matches, with scores ranging from 2.12 to 2.51 and a median score of 2.27. G Revision_3000_70, the other MSP created using a lower bound of 3000 Da, performed poorly. It returned only two of 10 correct first matches. The two new Grippotyphosa MSPs using a lower boundary of 2000 Da also performed poorly. G Revision_2000_70 returned only two of 10 correct matches and G Revision_2000_100 returned only four of 10 correct matches. The three MSPs that performed poorly were omitted from further analysis. G 100, considered to be the best-performing MSP previously created for Grippotyphosa, was also tested using the same test spectra. Out of 10 samples, seven were correctly matched by G 100. Scores ranged from 2.14 to 2.51, with a median score of 2.34. Both G 100 and G Revision_3000_100 returned one correct match that was tied with an incorrect match. In general, results for these two MSPs were similar.

Per the results of this limited run, the new MSPs for both Bratislava and Grippotyphosa did not demonstrate increased specificity compared with that of the two previously-created MSPs, (X 100), against which they were tested. Therefore, these new MSPs were not selected to represent their respective serovars.

3.4.11 Selection of One MSP to Represent Each Serovar in the *Leptospira* MSP Library

Results from MSP identification trials showed that no one type of MSP creation method increased specificity for all serovars. Different modifications worked better for different serovars. This finding emphasized the need for custom MSP library construction to obtain MALDI-TOF MS identification at the serovar level for this species. An MSP's performance was judged not only on how well it represented its corresponding serovar but also on whether it regularly identified as a different serovar. For example, for serovar Bratislava, B 100 yielded the best overall identification results. However, it was also the Bratislava MSP that misidentified as other serovars most often. Though it would be preferable for B 100 to return only correct identifications, these misidentification results alone did not remove this MSP from the pool of MSPs that might potentially be used to create a *Leptospira* reference library. This is because the performance of B 100 depends partially on the combination of other serovar MSPs with which it is used.

Based on MSP trial results, the following MSPs have been selected to represent each serovar: A 100 (serovar Autumnalis), B 100 (serovar Bratislava), C 100 (serovar Canicola), G 100 (serovar Grippotyphosa), H 125 (serovar Hardjo), I 100 (serovar Icterohaemorrhagiae) and P 100 (serovar Pomona). These MSPs, selected as the best-performing MSP for their respective serovars, all contained

three sets of raw spectra and a > 70 maximum desired peak number. With the exception of serovar Hardjo's H 125, which used a maximum desired peak number of 125, all other selected MSPs were created using a maximum peak number of 100.

3.4.12 Comparison of Discriminatory Peaks Selected by the Peak Statistic Report and Three Classification Models to Differentiating Peaks Identified in the Literature

Previous studies were reviewed for reported serovar-unique peaks to compare them to the peaks selected by the PSR and classification models in this study. The objective was to find whether specific peaks with particular identifying characteristics could be identified for each serovar, and if specific selected peaks were consistent across multiple studies. If so, then learning how characteristics of these peaks differed between the serovars would allow for detailed examination of the raw spectra that make up each serovar's MSP to find the right combination that best represents each serovar.

Calderaro et al. (2014b) reported a unique combination of peaks at 3684 Da, 5527 Da, and 11049 Da in the average spectrum for serovar Autumnalis, which they found had higher average areas/intensities compared to corresponding peaks in the total average spectrum of compared serovars. These findings differ from those of the current study, which found that these three peaks were not useful for discriminating serovar Autumnalis spectra from the spectra of other serovars. In this study, all serovars except Pomona had a peak within ± 3 *m/z* of 3684 Da. Additionally, all seven serovars had peaks within ± 3 *m/z* of 5527 Da and within ± 4 *m/z* of 11049 Da. The Peak Statistic report was used to compare the average areas/intensities of these peaks within the raw spectra of the seven study serovars. While a peak at 3684 Da was not included in the report, peaks at 3680 and 3693 Da were. Serovar Hardjo had the highest average area/intensity for both of these peaks. Serovars Autumnalis and Canicola were tied for the highest average area/intensity for the peak at 5527 Da. For the peak at 11049 Da, the highest average area/intensity belonged to serovar Autumnalis. However, serovar Canicola had an average area/intensity for this peak which was only one unit less. Calderaro and colleagues also reported that serovar Bratislava had characteristically lower average areas/intensities for peaks at 5671 Da and 6915 Da. In the current study, the Bratislava MSP peak list included a peak at 5672 Da, but the closest peaks to 6915 Da were 6842 and 6941 Da. The Peak Statistic report showed that Bratislava's average area/intensity for the peak at 5672 Da was approximately the same as that for serovars Autumnalis, Canicola, Icterohaemorrhagiae, and Pomona. The closest peaks to 6915 Da selected in the report were 6840 and 6970 Da.

Rettinger et al. (2012) had also used CPT analyses to identify peak presence and absence combinations potentially characteristic of certain serovars. Two classification models, QC and SNN, were used to detect the most discriminating peaks among the serovars in their study. Based on classification results, 10 peaks were selected for serovars belonging to *L. interrogans* and *L. kirschneri.* A pattern, based on the presence or absence of these peaks for each serovar, was then identified. The Rettinger et

al. study used 28 strains, six of which are the same as those used in this study. In this study, the QC and SNN models selected a combined total of 29 different discriminatory peaks for the seven serovars used. Only two of these were also selected in the Rettinger et al. study. These peaks, at 6330 Da and 7360 Da, corresponded to their peaks at 6327 Da and 7358 Da, respectively. Peak 5526 Da, selected as a discriminatory peak in the Rettinger et al. study, was ranked 148th out of 169 total peaks in the PSR for this study. This peak was present in all seven serovars with an average peak area/intensity ranging from 25.87 to 43.27 a.u. Another difference between the two studies was found for the peak at 6191 Da. In the Rettinger et al. study, this peak was present in all serovars except Grippotyphosa. In this study however, the peak was present in Grippotyphosa. Furthermore, Rettinger et al. identified a peak at 8097 Da, which they proposed differentiated *L. kirshneri* (peak present in serovar Grippotyphosa) from *L. interrogans* (peak absent from *L. interrogans* serovars). No peak of a similar m/z was selected by the three classification models in this study. The closest peak selected in the PSR and two of the three models was at 8059 Da. Model start and end mass integration regions for this peak did not overlap with 8097 Da. Although both previous studies used a Bruker Microflex LT instrument and a similar panel of serovars, the additional serovars included in those studies very likely contributed to differences in the model-selected peaks and corresponding data compared to those of the current study. This seems to suggest that MSP peak combinations, identified as unique to a particular serovar, may vary based on the group of serovars to which it is compared.

Karcher et al. (2018) used a combination of CPT classification models, FlexAnalysis, and SPSS to identify peaks that could discriminate between strains of *L. interrogans* and *L. kirshneri*. While the authors originally identified five peaks, further analysis eliminated four of the five, leaving one peak, at 8057 Da, as the only discriminatory peak found to differentiate between isolates of the two species. This peak correlates with the peak at 8097 Da identified in the Rettinger et al. study, which was also proposed to differentiate between *L. interrogans* and *L. kirshneri*. As mentioned in the previous paragraph, no obvious corresponding peak was selected by the three models in this study nor found in the PSR.

A recent study by Sonthayanon et al. (2019) reported 19 peaks in various distinct presence/absence and level of intensity combinations that could be used to differentiate between eight serovars of *L. interrogans.* Four of the eight *L. interrogans* serovars used in their study: Autumnalis, Canicola, Grippotyphosa, and Pomona, were also used in this study. The 19 identified peaks were compared to the top 10 peaks selected by the PSR and by the three classification models created for the seven serovar group in this study. Four peaks, at 7221 Da, 7435 Da, 7852 Da and 8056 Da, were found to have corresponding selected peaks in the current study. Sonthayanon et al. reported that peaks 7221 Da and 7852 Da were found to be present in in the four serovars common to both studies at an equal or higher intensity compared to the total average spectrum for all study serovars. Peak 7435 Da was noted as having average intensities higher than the total average spectrum for serovars Autumnalis and Canicola, but not for serovars Grippotyphosa and Pomona. For peak 8056 Da, serovars Autumnalis, Canicola, and

Grippotyphosa had higher intensities, while Pomona did not. The peak corresponding to 7221 Da in this study was 7234 Da, which was ranked as the 10th highest discriminatory peak by the PSR and selected by all three classification models. Among all seven serovars in this study, serovar Canicola had the highest average intensity, while serovars Pomona, Autumnalis and Grippotyphosa had the fourth, fifth and sixth highest averages, respectively. Though this peak was chosen as an important discriminatory peak in this study, the way in which it differentiates between the serovar classes differs somewhat from how Sonthayanon et al. used it for differentiation. The average intensity of this peak, when compared across the group of seven serovars used in this study, was not present at an equal or higher intensity than the total average spectrum in all four of the serovars common to both studies, as it was for Sonthayanon et al. The peak corresponding to 7852 Da in this study was 7846 Da, which was selected by the QC model. As seen with the previous peak, when compared across the seven serovars in this study, the average intensity for peak 7846 Da for the four serovars common to both studies was not found to meet or exceed that of the total average spectrum. So again, simply comparing the average peak intensity of each of the serovars in this study to the total average spectrum does not result in the same pattern of higher peak intensities as seen in the Sonthayanon et al. study. Peak 7435 Da, analogous to peak 7437 Da in the current study, was reported as having a higher average intensity in serovars Autumnalis and Canicola than in serovars Grippotyphosa and Pomona. Average peak intensities for 7435 Da, selected by the SNN model in the current study, were also found to be higher for serovar Canicola than for serovar Pomona. However, the average intensity for serovar Grippotyphosa was right below that of Canicola, and higher than that of serovar Autumnalis. So again, comparative peak intensity levels differed for some serovars in the two studies. The last common peak between the two studies was 8056 Da for Sonthayanon et al., corresponding to 8059 Da in the current study. Peak 8056 Da was found to have intensities higher than or equal to the total average spectrum for serovars Autumnalis, Canicola and Grippotyphosa, but not for Pomona. In this study, peak 8059 Da was selected by the QC and SNN models. Among the four common serovars, average intensity for this peak was highest in serovar Autumnalis. However, unlike in the Sonthayanon et al. study, the average peak intensity for serovar Pomona was higher than that for Canicola and Grippotyphosa. While there are some similarities in peak intensity levels between the two studies, the differences do not allow the same relative intensity levels to be used as identifying characteristics among the seven serovars in this study. This is not unexpected. Relative peak intensity levels appear to be related to the group of serovars being compared. If one or more serovars in either of these studies were replaced by different serovars, the relative peak intensities would be expected to change, at least somewhat. Additionally, it is possible that the same serovar, from different geographical regions, may produce slightly different mass spectra. When comparing serovars that are very closely related, small differences in mass spectra may affect MALDI identification.

Given these findings, it was decided that comparison of MSP and raw spectra peaks lists for unique peaks, even with additional relative intensity data, was not an accurate and reliable way to determine which raw spectra best represented a serovar and which should be excluded from an MSP.

3.4.13 Blind Trials

In the first trial, which used serovar culture samples, all but one sample spot on the first target returned correct first matches. This included correct matches for the two *Brachyspira* species tested as genus-specific controls. The one spot that did not yield a correct match belonged to serovar Grippotyphosa. This spot returned a result of no peaks found. Since the seven other spots on this target for this sample returned correct first matches, the NPF result was thought to be due to problems with sample deposition or uneven co-crystallization of the matrix and analyte, rather than a reflection of the MSP's ability to accurately identify the serovar at the concentration tested. The second target had one spot that returned an incorrect first match. This spot, which belonged to serovar Grippotyphosa, returned an incorrect first match of Bratislava, with a score of 2.54. However, the second match was a correct identification, with a score of 2.53. Two additional spots, one belonging to serovar Canicola and the other belonging to Grippotyphosa, resulted in NPF. Again, it was concluded that these two NPF did not reflect MSP specificity or sensitivity, because the other seven spots for each of these serovars on this target returned correct first matches with scores ≥ 2.22. Beyond these three spots, an additional spot for Grippotyphosa returned a tied first match of Grippotyphosa and Bratislava with a score of 2.31. Not including NPF results, for the reason described above, only one sample spot out of the total 112 serovar sample spots on both targets, returned an incorrect first match. While this spot matched incorrectly at the serovar level, this match was correct at the genus level. Overall, 99% (n = 111/112) of the serovar sample spots returned a correct first match. All correct serovar-level first matches for *Leptospira* MSPs returned scores ≥ 2.29. Additionally, the two BTS spots included on each target returned the expected identification of *E. coli*, at scores ≥ 2.19. The uninoculated matrix spot on each target returned no peaks found. Sensitivity and specificity for each MSP was calculated using 2 x 2 contingency tables and is given in Table 3.24. Specificity for MSPs in this first trial ranged from 99 to 100%.

Table 3.24. Sensitivity and Specificity of the Best-Performing MSP per Serovar Based on Results from Blind Trial 1

MSP	Serovar	Sensitivity % TP	Specificity % TN	% FP	% FN
A 100	Autumnalis	100	100	0	0
B 100	Bratislava	100	99.22	0.78	0
C 100	Canicola	93.75	100	0	6.25
G 100	Grippotyphosa	81.25	100	0	18.75
H 125	Hardjo	100	100	0	0
I 100	Icterohaemorrhagiae	100	100	0	0
P 100	Pomona	100	100	0	0

Serovar samples adjusted to 50 %T during MALDI sample prep. Serovar MSP % TP, % TN, % FP, and % FN were calculated using 2 x 2 contingency tables.

In the second trial, which used serovar-spiked canine urine samples, six serovar sample spots returned incorrect first matches on the first target. Three of these were Bratislava spots that returned first matches of Grippotyphosa with scores of 2.65, 2.57, and 2.50. However, Bratislava was the second match for these spots with scores of 2.58, 2.54, and 2.48, respectively. So, for two of these three spots, the difference in score between the first and second matches was ≤ 0.03. One Hardjo spot returned an incorrect first match of Autumnalis, with a score of 2.39 that was tied with a match of Bratislava with the same score. The correct Hardjo MSP was actually the third match for this spot, with a score of 2.34. Lastly, two Grippotyphosa spots returned incorrect first matches of Bratislava with scores of 2.44 and 2.46. Grippotyphosa was the second match for each of these spots, with scores of 2.43 and 2.40, respectively. The difference in scores between the incorrect first matches and correct matches for both Hardjo and Grippotyphosa all differed by ≤ 0.05. As in the first blind trial, *Brachyspira innocens* and *Brachyspira pilosicoli* were included as genus-specific controls. Five of the eight *B. innocens* spots returned 'No Organism ID Possible' with low ID scores ranging from 1.31 to 1.63. These results were not listed as NPF, because the MALDI was able to detect peaks and suggest a list of organisms that might match the sample. However, since the match scores for these suggested organisms were < 1.70, and Bruker score interpretation guidelines state that scores < 1.70 indicate unreliable matches, the suggested matches were considered invalid. A review of the MALDI-suggested matches for the five *B. innocens* sample spots which returned 'No Organism ID Possible' found that *B. innocens* was suggested as the first match for only one of these spots. The rest returned first matches to other organisms. None of the 10 suggested organisms for each of these sample spots were *Leptospira.* The remaining three spots for this

sample returned NPF. Of the eight B. pilosicoli spots, seven returned correct first matches with scores ranging from 1.73 to 1.85. The eighth spot returned No Organism ID Possible with a score of 1.34. However, the first suggested match for this spot was correct.

On the second target, only one of the serovar sample spots returned an incorrect first match. This was a Bratislava sample spot, which returned a first match of Grippotyphosa with a score of 2.64. The second match for this spot was Bratislava, with a score of 2.63. Such small differences between scores, as in this case, means that if this same target was run again, results from the second run may show the first and second matches for this spot switched. As on the first target, five of the eight *B. innocens* spots returned No Organism ID Possible, with scores ranging from 1.35 to 1.48. *B. innocens* was not listed as the first match for any of these spots. Two other *B. innocens* sample spots returned NPF. The last sample spot for this organism returned a correct first match with a score of 1.71. Two of the eight sample spots for *B. pilosicoli* returned No Organism ID Possible with scores of 1.57 and 1.61. However, the suggested first match for both of these spots was correct. The remaining six spots all returned correct first matches, with scores ranging from 1.73 to 1.97. Like the first target, *Leptospira* was not in the list of 10 organisms suggested for any of the *Brachyspira* sample spots. It is not known why most of the *B. innocens* sample spots for the second blind trial returned either No Organism ID Possible or NPF results. Maintaining an even suspension of *Leptospira* and *Brachyspira* for measurement between concentration adjustments can be a challenge. This may have led to minor fluctuations in the %T for the *B. innocens* sample, resulting in a lower sample concentration than that needed for accurate MALDI identification.

Eight spots of urine only were also included on both targets as another negative control. All spots returned NPF. The uninoculated matrix spot on each target also returned no peaks found.

Incorrect first matches for the serovar sample spots were incorrect at the serovar level but correct at the genus level. The two positive control spots on each target returned the correct identification of *E. coli* at scores ≥ 2.36. Overall, seven out of 112 serovar sample spots on both targets returned an incorrect first match, leaving 94% (n = 105/112) serovar spots correctly identified to the serovar-level. All correct serovar-level first matches for *Leptospira* MSPs returned scores ≥ 2.16. Sensitivity and specificity values for the second trial was also calculated using 2 x 2 contingency tables and are given in Table 3.25. MSP specificity for this second trial ranged from 97 to 100%.

Table 3.25. Sensitivity and Specificity of the Best-Performing MSP per Serovar Based on Results from Blind Trial 2

MSP	Serovar	Sensitivity % TP	Specificity % TN	% FP	% FN
A 100	Autumnalis	100	99.22	0.78	0
B 100	Bratislava	75.00	98.61	1.56	25.00
C 100	Canicola	100	100	0	0
G 100	Grippotyphosa	87.50	96.88	3.13	12.50
H 125	Hardjo	93.75	100	0	6.25
I 100	Icterohaemorrhagiae	100	100	0	0
P 100	Pomona	100	100	0	0

Serovar samples adjusted to 50 %T were combined with qPCR-leptospirosis-negative canine urine during MALDI sample prep. Serovar MSP % TP, % TN, % FP, and % FN were calculated using 2 x 2 contingency tables.

Results from the two blind trials supported the hypothesis that MALDI-TOF MS can be used to identify *Leptospira* isolates to the serovar-level within the real-time classification workflow.

3.5 DISCUSSION

The goal of this study was to learn whether whole-cell MALDI-TOF MS has the specificity to discriminate between *Leptospira* serovars and can be used for real-time leptospirosis diagnosis and epidemiological surveillance. In this study, several different MALDI sample preparation techniques and MSP creation parameters were tested to evaluate the influence of these factors on MSP specificity. Finding the optimal combination of preparation technique and creation method allowed for real-time, serovar-level differentiation using a custom *Leptospira* MSP reference library without the need for subsequent ClinProTools analysis.

Though initial serovar MSPs, created using the default parameters and recommended manufacturer guidelines consistently returned accurate genus-level identification, accurate serovar-level results proved more variable. This outcome was expected, given the greater similarity of mass spectral profiles among *Leptospira interrogans* serovars compared to that between *L. interrogans* serovars and organisms belonging to a different species (Zhang and Sandrin 2016). Since substantial genetic overlap occurs between serovars of the same species, the mass spectra of these serovars are expected to be highly similar, having many of the same peaks (Arnold et al. 2006).

One factor that may contribute to the similarity of *Leptospira* protein profiles is horizontal gene transfer (HGT), which has been shown to occur between *Leptospira* species (Haake et al. 2004). In fact,

HGT of genes encoding outer membrane proteins has been proposed as the mechanism by which certain serovars are found in more than one species. Examples include serovar Hardjo, found in both *L. interrogans* and *Leptospira borgpetersenii*, (de la Peña-Moctezuma et al. 1999), and serovars Bulgarica, Grippotyphosa, Mwogolo, and Valbuzzi, found in both *L. interrogans* and *Leptospira kirschneri* (Brenner et al. 1999, Feresu et al. 1999, Levett 2001). It has also been suggested that HGT of ribosomal genes between *Leptospira broomii* with *Leptospira fainei* and *Leptospira inadai* may have occurred, and that this may also have occurred between other *Leptospira* species. (Morey et al. 2006). A more recent study conducted by Xu et al. (2016) using whole genome sequencing found widespread occurrence of HGT in the genomes of pathogenic species, including 32.7% of pathogenic genes in *Leptospira kmetyi* and 39.2% of pathogenic genes in *Leptospira alstoni*.

A review of the literature found that most studies that obtained accurate, MALDI identification below the species level required additional or alternative sample preparation and/or spectra analysis steps beyond the standard required for identification at higher taxonomic levels. Some studies also found that adjustment to the manufacturer's identification score categories was needed (Arnold et al. 2006, Murray 2010, Croxatto et al. 2012, Culebras 2018). The greater level of similarity between spectra of closely-related organisms means that slight variations in sample preparation or MALDI analysis between runs or between laboratories may trigger enough of a difference in a sample's spectrum to negatively affect classification at the subtype-level. These same variations in species-level classification are more likely to have little or no effect (Valentine et al. 2005, Vargha et al. 2006, Dieckmann et al. 2008, Sauer et al. 2008, Singhal et al. 2015). The difficulty of MALDI *Leptospira* identification at the serovar-level is demonstrated by the results of the six previous MALDI/*Leptospira* studies, in which serovar-level identification could not be achieved using a custom MSP library and Biotyper real time classification alone (Djelouadji et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Xiao et al. 2015, Karcher et al. 2018, Sonthayanon et al. 2019).

Serovars are differentiated based on antigenic variation. This is chiefly based on the structural differences in the carbohydrate component of *Leptospira* LPS, determined by the organization of genes within the LPS biosynthetic (*rfb*) locus (de la Peña-Moctezuma et al. 2001, Levett 2001, Bharti et al. 2003). Some studies have reported MALDI detection of outer membrane proteins (OMPs), such as LPS or porins (Zhou et al. 2010, Hu et al. 2015) when combined with an additional form of analysis, such as SDS-PAGE. However, OMPs are not included in the protein subset thought to compose the majority of proteins detected by MALDI-TOF MS alone. Instead, conserved, highly-abundant proteins with housekeeping functions, such as ribosomal proteins, cold- and heat-shock proteins, and nucleic-acid binding proteins, comprise the majority of detected proteins that create a sample's mass spectrum (Ryzhov and Fenselau 2001, De Carolis et al. 2014b, Basile and Mignon 2016).

This then begs the question: Can this type of MALDI platform detect antigenic proteins that differentiate *Leptospira* serovars, given the type of proteins detected within the *m/z* range analyzed? To

answer this question, one needs to consider potential serovar-specific variations in ribosomal proteins, which are considered the major protein type detected (Holland et al.1996 and 1999, Mellmann and Müthing 2012, De Carolis et al. 2014b). Several papers have reported high similarity in the 16S rRNA gene sequences of different *Leptospira* serovars within the same species. A study conducted by Morey and colleagues (2006) compared the 16S rRNA gene sequences of 39 strains, representing 17 *Leptospira* species. The authors found that sequences belonging to different serovars of the same species were either identical or had only an average of 0.2 base pair differences out of approximately 1430 bp total. They also found that sequences for these genes were very similar between some *Leptospira* species. In fact, some intraspecies distances for the 16S rRNA genes were found to be greater than interspecies distances. For example, only a single base, at position 94, and an insertion/deletion at position 784, was found to differentiate the type strains of *L. kirshneri* and *L. interrogans* (Morey et al. 2006). However, other papers have reported that while the amino acid sequences of ribosomal proteins are highly conserved, minor variations in sequence occur even at a subspecies level (Freiwald and Sauer 2009). In particular, ribosomal protein variations, related to differences in antibiotic resistance, have been observed in strains of *E. coli*, albeit via electrospray ionization and MS/MS analyses (Wilcox et al. 2001). While this does not tell us whether linear MALDI-TOF MS can detect differences in ribosomal proteins that can differentiate serovars, it does suggest that detection of such ribosomal protein differences in *Leptospira* is possible.

A search of the literature was performed to learn whether *Leptospira* antigenic components, other than LPS, have molecular weights within the MALDI's standard range of detection and exhibit serovar-specific variations useful for MALDI serovar identification. Several studies have used immunoblotting techniques to detect *Leptospira* antigens. However, the detected antigens were not identified, (Cinco et al. 1992, Doungchawee et al. 2007, Lafetá et al. 2009), and/or the antigen molecular masses fall outside of the mass gate analyzed by the MALDI (Brown et al. 1991, Cinco et al. 1992, Doungchawee et al. 2007). Some studies specifically identified lipoprotein, lipid, and flagellar antigenic proteins, but again, reported molecular masses that are beyond the standard MALDI detection range. (Cullen et al. 2002, Biswas et al. 2005). The studies that used more than one serovar type to explore putative, non-LPS antigens offered limited comparison of antigen differences between serovars (Lessa Aquino et al. 2013), or were limited in scope to the predominant serovars circulating in particular regions and known to cause disease in particular species (Brown et al. 1991, Cinco et al. 1992, Cullen et al. 2002, Biswas et al. 2005, Doungchawee et al. 2007, Lafetá et al. 2009). Attempts to extrapolate these findings to other serovars in different regions is likely to result in inaccuracies. Ultimately, studies reporting *Leptospira* antigens other than LPS that are differentially-expressed between any of the serovars used in this study and that have masses within the detected mass range were not found. As the number of sequenced genomes for *Leptospira* species, serovars, and strains has grown, our understanding of the *Leptospira* proteome has

increased. However, much work is still needed before we can identify antigenic components of *Leptospira* that are both detectable by whole-cell MALDI-TOF MS and that can differentiate serovars.

The six previous studies that examined MALDI's ability to discriminate between *Leptospira* isolates were unable to differentiate between serovars using the *Leptospira* MSPs created in-house within the real-time classification workflow (Djelouadji et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Xiao et al. 2015, Karcher et al. 2018, Sonthayanon et al. 2019). The four studies from this group that reported some serovar-level classification instead used additional software, such as ClinProTools, to analyze raw sample spectra and create classification models to identify discriminatory mass peaks (Rettinger et al. 2012, Calderaro et al. 2014b, Karcher et al. 2018, Sonthayanon et al. 2019). Study authors suggested that the presence or absence of these peaks within the spectra of the respective serovars created unique patterns that could be used to distinguish these serovars within ClinProTools. (The reader is referred back to the Results section of this chapter for details on peaks identified in those studies.)

In contrast**,** peaks reported in previous studies to be characteristic of serovars based on presence or absence and/or relative intensity levels were not found to be characteristic presence/absence markers for the respective serovars in this study. As mentioned previously in the Results section, although the current study used the same MALDI platform, sample preparation method, and CPT classification models as the Rettinger et al. study, only two of 10 discriminatory peaks selected by the QC and SNN models for *L. interrogans* and *L. kirshneri* in that previous study were selected by the QC and SNN models for the *L. interrogans* serovars used in this study. The Peak Statistic report for the seven serovars was also examined for the presence of these 10 peaks. The report revealed some peaks, thought to correspond with the selected peaks from the previous study due to similar mass-to-charge ratios, were present. However, these peaks were ranked low in discriminatory power. For example, in the Rettinger et al. study, peaks at 3206 Da, 3220 Da, and 3234 Da were identified with high intensities in the raw spectra of serovar Pomona, creating a characteristic peak pattern for this serovar. In this study, the peaks with the closest m/z values to these three in the PSR were 3210 Da, 3224 Da, and 3239 Da. These peaks were ranked 82nd, 113th, and 162nd in discriminatory power, respectively, out of 169 total peaks. Since these peaks were ranked low in separation power and were not selected by the CPT models, they were not considered discriminatory peaks in this study. However, given Rettinger et al.'s observation that the three peaks in their study had characteristic higher intensities in Pomona raw spectra, the average areas/intensities of the three corresponding peaks in this study were compared among the seven serovars to learn whether they exhibited higher intensities within the spectra of any particular serovar. If so, these peaks/intensities could serve as a unique peak pattern for that serovar's identification. The PSR showed that serovar Pomona did not have the highest average area/intensity for any of these peaks. In fact, no one serovar had the highest average area/intensity for all three of these peaks.

Another peak, between 8000 and 8100 Da, was identified by both Rettinger et al. and Karcher et al. (2018) as one that could be used to differentiate between the *L. interrogans* and *L. kirshneri* isolates used in their studies. Since no *L. kirshneri* serovars were used in this study, this peak's ability to differentiate between the two species could not be explored. The peaks identified as serovar-specific by Calderaro and colleagues (2014b) were not found to be serovar-specific in this study, in either presence/absence or in average peak area/intensity. Lastly, of the 19 discriminatory peaks reported by Sonthayanon et al. (2019), only four had corresponding selected peaks in this study. The characteristics of these peaks, which the authors used to differentiate between the *L. interrogans* serovars used in their study, were largely absent in the analogous peaks of this study.

Altogether, this seemed to suggest that the top discriminatory peaks selected by CPT analyses appear to be relative to the group of loaded spectral classes, (in this case, serovars). For example, if two sets of raw spectra, representing two different groups of *Leptospira* serovars, were analyzed by CPT, at least some of the top discriminatory peaks selected for each group would be expected to differ. If the two serovar groups included some of the same serovars/strains, then some overlap in peak selection would be expected. However, I would not expect the top peaks selected in the PSR, or the peaks selected by the classification models, to be the same in both situations. Relative peak areas/intensities will change as the group of loaded serovars change. Moreover, selection of differentiating peaks for a group of serovars is partially dependent on relative peak intensities. Therefore, a particular peak or set of peaks at a particular level of intensity cannot be expected to consistently distinguish a specific serovar, regardless of the group of serovars against which it is compared and classified in CPT. As evidenced by the different top discriminatory peaks selected in this versus the previous MALDI/*Leptospira* studies, a particular peak or set of peaks which may differentiate a particular serovar in one study, may not serve to identify the same serovar in a different study. The top peaks selected by both the PSR and the classification models can also change based on CPT parameter settings. For example, the use of different sort orders for PSR calculations will result in some differences in the top discriminatory peaks selected by the PSR. In this study, a PTTA statistical sort order used for the pairwise comparison of Bratislava and Grippotyphosa resulted in a different list of top discriminatory peaks than when the sort order was changed to PWKW. Karcher and colleagues (2018) used a PWKW sort order for the PSR. Rettinger et al. and Calderaro et al. do not mention the sort order used in their studies. Given that the discriminatory peaks identified in the Rettinger et al. study had low PSR ranks in this study, it is likely they used PTTA.

The Peak Statistic report for the analysis of all seven serovars showed that out of 145 peaks with a non-normal distribution, 137 were statistically significant differentiating peaks (PWKW < 0.01). While the high number of significant peaks may seem to indicate that serovar-specific MSP identification is easily achievable, one must keep in mind that these peak differences may occur in only a subset of the serovars. In other words, a particular peak may exhibit a significant difference in average values between, for example, three of the seven serovars compared to the remaining group of four serovars. Therefore,

while this peak's characteristics may be helpful in separating three of the serovars from the other four, by itself it cannot serve to differentiate between all seven serovars. Additionally, the close relationship among these serovars means that all of them likely produce the protein that gave rise to a particular spectral peak, just perhaps in relatively different amounts. Comparison of MSP peak lists created using different parameter settings were compared in Compass Explorer. Picked peaks varied depending on the choice of Preprocessing and MSP Creation Method parameters. This means that MSP peaks are picked in relation to all other picked peaks. This also means that proteins represented by particular peaks may not be included in an MSP peak list, even if that protein is present in the sample. This further supports the idea that the top discriminatory peaks selected by CPT models and the Peak Statistic report are relative to the group of spectral classes loaded.

Overall, this information suggests that, for the type of MALDI platform used in this study, dependence on particular peaks for serovar differentiation is not a reliable method for serovar discrimination. Additionally, reliance on selected biomarkers for differentiation means that serovar identification cannot be performed in real-time, during standard MALDI runs used daily in many diagnostic laboratories. Instead, identification would need to be performed using an offline method, such as CPT analyses, which requires more time, labor, and technician training. This negates some of the advantages of using MALDI for identification, such as a quick turnaround time and ease of use. Therefore, an important aspect of this study was creating serovar MSPs with sufficient specificity such that the MALDI's real-time pattern-matching algorithm could be used for serovar identification.

With this goal in mind, several sample preparation techniques were evaluated to find whether any resulted in raw spectra of higher quality. The idea behind this approach was that new serovar raw spectra, acquired using a different sample preparation technique, might be of higher quality than the raw spectra collected and used in the original MSPs. Higher quality spectra, (e.g. containing more peaks, with higher peak intensities and greater S/N ratios), could then be used to create new MSPs, which might better represent each serovar and allow for greater specificity. Among the variations on direct-transfer (direct-spotting), extended direct-transfer, and protein extraction preparation methods, extraction proved to be the most effective, producing the highest number of identifiable spectra over the widest range of tested dilutions. Since harvesting *Leptospira* samples from solid agar media for MALDI analysis is not practical, samples were prepped from cultures grown in liquid media. This limited the amount of sample pellet material available for direct application on the MALDI target, with more dilute samples having no visible pellet to transfer. Therefore, the lower numbers of good-quality spectra obtained from the direct-transfer methods compared to the extraction methods, was not unexpected. This finding is consistent with that of Karcher et al. (2018), who found that mass spectra for *L. interrogans* and *L. borgpetersenii,* obtained via a direct-transfer method, resulted in peaks with low intensities compared with spectra obtained via a protein extraction method. Other groups studying different organisms have also found the protein extraction method to be preferable for identification of bacterial subtypes, reporting an increase in

peak signal reproducibility and range within which biomarkers are detected (Alatoom et al. 2011, Croxatto et al. 2012, Clark AE et al. 2013, Lartigue 2013). Bruker recommends that MSPs be created with spectra acquired from samples prepped with the protein extraction method (Bruker Daltonik GmbH 2014) , and no alternative sample prep method resulted in higher-quality raw spectra. Therefore, all spectra used in MSP creation were acquired from samples prepped with the EtOH/FA extraction method. Since this sample prep method was the same one used to create the original MSPs, and the original MSPs did not offer the specificity needed for serovar identification, focus shifted to testing alternate methods of MSP creation.

Original MSPs were created using the default creation parameters and recommended manufacturer guidelines. However, these MSPs did not allow for serovar differentiation using the Biotyper Realtime Classification software's pattern matching algorithm used for MALDI species-level diagnostics. Several MSP creation parameter modifications were tested to assess whether specificity was improved. At least four different MSPs were created for each serovar. To determine the MSP that best represented each serovar, accurate identification results from both individual and group specificity trials were considered. This is because a serovar's MSP must be able to return accurate matches not only when tested by itself, but also within the context of all other MSPs in the *Leptospira* library. Modifications to the standard MSP creation parameters resulted in MSPs that could be used for serovar differentiation within Biotyper RTC.

One modification was incorporation of raw spectra from three different subcultures into each serovar's MSP. The idea to use spectra from different subcultures came after noticing that the spectral profiles of some serovars showed a notable degree of intraserovar variability between raw spectra collected from subcultures on different dates. These variations in peak intensity and peak shift were most clearly seen using the spectra overlay feature in the FlexAnalysis software program. Such differences can occur for a variety of reasons, including alterations in sample preparation and bacterial adaptation to growth conditions (Valentine et al. 2005, Goldstein et al. 2013, Josten et al. 2013 Sauget et al. 2017). Some variation may also occur due to uneven distribution of bacterial colony material on the MALDI target and slight differences in the amount of ionized protein generated by each group of laser shots on a target spot. We may even see that a particular peak present in one raw spectrum of an isolate may be absent in a different spectrum of lower quality from the same isolate (Ostergaard et al. 2015). Given the proclivity for slight variations in an isolate's raw spectra, it was thought that including raw spectra with such variations in serovar MSPs would result in MSPs that better represented each serovar.

The second modification was an increase in the maximum desired number of MSP peaks. The high level of similarity between the serovars means that serovar MSP peak lists have many of the same peaks. MSPs created using the standard maximum desired peak number of 70 did not have the specificity needed for real-time serovar differentiation. It was thought that allowing a greater number of peaks in each MSP peak list would potentially increase the number of serovar-differentiating peaks included in each serovar's MSP.

Ultimately, it was found that serovar MSPs created with additional raw spectra obtained from multiple subcultures and with a higher maximum desired peak number of 125, (Hardjo), or 100, (the six other serovars), had the greatest specificity.

MSP inclusion of both additional peaks and raw spectra from several subcultures meant that the number of unique peaks found in every raw spectrum used in a serovar's MSP, and the number of those peaks that met the requirements for MSP selection, decreased. Therefore, the default guideline that all MSP peaks have ≥ 75% peak frequency, was not met. The decision to accept MSPs that contained some peaks with frequencies below 75% was based on the fact that the guidelines were not created for MSPs using 64+ raw spectra from different subcultures. Additionally, the default parameter setting for the desired peak frequency in Compass Explorer is 25%.

Deviation from recommended Bruker guidelines is not unprecedented. A review of the literature found previous studies that have proposed modifications to the MSP creation guidelines and to the interpretation of sample identification log scores (Bourassa and Butler-Wu 2015, Pérez-Sancho et al. 2015, Pranada et al. 2016, Zboromyrska et al. 2018). For example, while Bruker recommends that MSPs be created using a minimum peak frequency of 75%, multiple studies have reported the use of a 25% minimum peak frequency instead, including those performed by Rettinger et al. (2012), Normand et al. (2013), Karcher et al. (2018), and Boyer et al. (2019). Bruker's recommended MSP creation guidelines (Bruker Daltonik GmbH 2016) also pertain to only one set of raw spectra, acquired during one run. However, Normand and colleagues (2013) found that the number of accurate identifications improved significantly when raw spectra from distinct subcultures of a particular strain were included in that strain's MSP.

A call to the Bruker representative for our lab confirmed that they do not yet offer *Leptospira* MSPs in their commercial database (M. Santino, personal communication, November 25, 2018). Given the importance of leptospirosis as a re-emerging infectious disease, it seems plausible that Bruker would want to add *Leptospira* MSPs to their commercial reference library. The fact that they have yet to do so may indicate that their current guidelines are not conducive to the creation of *Leptospira* MSPs and may need to be modified to obtain MSPs for this genus. While further testing of the modified MSP creation approach used in this study is needed to confirm accurate identification over a wider set of serovars and strains, the modifications used may offer a way to use MALDI for serovar identification.

Specificity of the custom *Leptospira* MSP library was tested in two blind trials. In the first trial, which used only serovar culture adjusted to 50 %T with UHPLC-grade water, 99% (n = 111/112) of the serovar sample spots correctly identified to the serovar level. MSP specificity ranged from 99 to 100%. In the second trial, which combined adjusted and pelleted serovar culture with 1 mL of canine urine before sample prep, 105 out of 112 (94%) serovar sample spots were correctly identified to the serovar-level. MSP specificity ranged from 97 to 100%. Results for both blind trials were similar, with only a 3% difference in the number of correct first serovar-level matches. There are a variety of possible reasons

why the second trial resulted in more incorrect first matches than the first. These include not only the addition of urine to the sample pellets in the second trial, but also potentially uneven sample deposition on some target spots, or uneven co-crystallization of the matrix and sample across a sample spot. For all spots that returned an incorrect first match, the difference in score between the first match and the correct match was ≤ 0.07.

One should keep in mind that if a sample spot returns matches with very similar scores in a MALDI run, a rerun of that sample spot may result in a change in the ranks of those matches. For example, in the second blind trial, on the second target, a sample spot for serovar Bratislava initially returned an incorrect first match of Grippotyphosa, with a score of 2.41. The second match was Bratislava, with a score of 2.37. When this spot was rerun, the first and second matches switched ranks. Bratislava was then the first match, with a score of 2.66, while Grippotyphosa was the second match, with a score of 2.62. This switch in identification match ranks is not uncommon in closely related organisms. To increase MSP specificity, which would translate to larger differences in the scores of subsequent matches, further work on each MSP would be needed.

There were several limitations in this study. The first was the small number of replicates, both within individual trials and of individual trials, in which both H_2O and urine serial dilutions were used for MALDI testing, qPCR, and Qubit concentration measurement. Several factors contributed to lower replicate numbers. Trials in which urine dilutions were used were limited due to the availability of this resource. Urine used for this project was collected from the remaining canine specimens that had been submitted to the VDL for testing. Specimens that were qPCR-leptospirosis-positive, exhibited bacterial growth in culture, or that contained visible blood, sediment, or other particles that might complicate MALDI interpretation, were not used. Suitable specimens were generally available only in small volumes after VDL testing. Six sets of dilutions (three with H_2O and three with urine), were required to run a full trial (MALDI, qPCR, Qubit). Therefore, it sometimes took several weeks to collect the volume of urine needed for a complete run. Obtaining the required volume of *Leptospira* culture needed for trials could also be challenging. Cultures required growth for approximately seven days before use. Multiple cultures per serovar had to be used to obtain the needed volume and starting concentration. Also contributing to the limited number of usable technical replicates was not knowing at the start of this project that MALDI identification scores and match level were relevant to the combination of MSPs included in an MSP library. This means that during early trials, when only one or two MSP types had been created, it was not evident that results from group and individual trials testing the same serovar MSP could not be considered technical replicates. Therefore, results from the two trial types could not be compared to each other statistically. Therefore, while many trials were run, few individual trial replicates can be directly compared to each other and used for statistical analysis. Group trial results provided valuable insight into the performance of different MSP types. However, since trends in group trial results were consistent with those of individual trials and only one MSP would be chosen to represent each serovar in the *Leptospira*

library, data from individual trials, rather than group trials, was ultimately used for specificity testing. Had the association between ID scores, match levels, and loaded serovar classes been discovered earlier in the study, testing could have been planned such that fewer group and more individual trials would have been run.

Another limitation was the lack of MSP testing using *Leptospira* strains other than those used to create the MSPs. Since positive patient samples rarely have the bacterial concentration needed for MALDI detection, positive canine samples submitted to the VDL were not an adequate source for additional test isolates. To determine if *Leptospira* MSPs, created using these alternate parameters, can be useful for leptospirosis research, further testing with additional strains is necessary. Not only would this testing allow us to learn whether various strains within a serovar can be recognized by the serovar's MSP, but whether uncultivated strains produce mass spectra with enough similarity to that of their cultivated counterparts such that both can be identified by the MSP.

To be a reliable tool for the identification of locally-circulating *Leptospira* serovars, the specificity of this MSP library would need to be further tested and improved to reduce the number of misidentifications and increase the differences between subsequent MALDI ID scores. Numerous variations of these MSPs, based on the combination of constituent raw spectra, have yet to be tried. For example, the results of CPT raw spectra and peak analyses informed the creation of the X 2D and four X Revision MSPs for serovars Bratislava and Grippotyphosa. However, these MSPs used only a portion of the data found that could lead to greater MSP specificity. The X 2D MSPs were created using the same three sets of raw spectra used for previous MSPs, with the omission of a group of overlapping spectra identified in CPT 2D peak distribution plots for the top six discriminatory peaks identified in the Bratislava and Grippotyphosa PSR after PTTA sorting. The group of four X Revision MSPs used the same three sets of raw spectra but omitted a different group of spectra than was omitted from the X 2D MSPs. The X Revision MSPs omitted a group of overlapping spectra selected from a combination of 2D peak distribution plots created for the top PSR discriminatory peaks using PWKW sorting, the PCA performed for these two serovars, and some of the raw spectra incorrectly identified in external validation of the classification models for this serovar pair. However, not all of the identified overlapping spectra were removed. The spectra identified by the 2D and PCA plots chosen for MSP omission were selected based on the degree of overlap exhibited by their location in the plots. The omission of other overlapping spectra, in different combinations, might result in more specific MSPs. The creation of numerous other MSPs, each with a different group of identified overlapping or outlier raw spectra omitted for Bratislava and Grippotyphosa, were not pursued as time and resources did not allow. It should also be noted that removing selected raw spectra from MSPs risks inadvertently removing the number of representative unique MSP peaks that differentiate one serovar from another. When manipulating the group of constituent raw spectra that make up a serovar's MSP, potential effects of the manipulation on the performance of all serovar MSPs included in the MSP library must be considered and checked.

As mentioned in previously, it is not yet known whether clinical isolates of the same serovars used in this study would be identified by these MSPs. To learn this, clinical isolates of locally-circulating serovars need to be tested against these MSPs. Raw spectra acquired from such isolates could also be added to or replace some of the constituent raw spectra in these MSPs as another method of increasing specificity and expanding MSP coverage for various strains of local serovars.

In summary, this study evaluated the feasibility of using MALDI-TOF MS for rapid and accurate *Leptospira* serovar identification. Testing of several sample preparation methods revealed that direct-spotting methods produced more inconsistent and poorer-quality spectra than the EtOH/FA protein extraction method. The standard TE method, without added wash steps, was found to produce the best quality mass spectra for the study serovars. The use of MSP creation parameters that differed from manufacturer's suggested guidelines and default settings, led to improved serovar identification, demonstrated by MSP specificities ranging from 97 to 100 % in two blind trials. These two trials differed in that the first used only serovar cultures, while the second used pelleted cultures suspended in canine urine. Resulting specificity for all MSPs were similar between the two trials, suggesting that the custom *Leptospira* MSP library can reliably differentiate between study serovars in pure culture and urine samples. Results from the blind trials, run in Biotyper Realtime Classification (RTC) software, showed that it is possible to use custom MSPs to differentiate between *Leptospira* serovars in real-time. This finding supports the hypothesis that linear, whole-cell MALDI-TOF MS has the specificity to accurately identify *Leptospira* isolates to the serovar level. This is the first study, to the author's knowledge, that has reported MALDI-TOF MS *Leptospira* serovar-lever identification within the real-time classification workflow. The modified MSP creation method used in this study may further offer a way to increase custom MSP specificity for other bacterial subtypes not yet included in commercial MSP libraries and for which standard MSP creation parameters result in MSPs with specificities insufficient for subtype identification.

Chapter 4: SENSITIVITY

4.1 ABSTRACT

Leptospirosis is a leading cause of zoonotic morbidity and mortality worldwide. It has been recognized as a re-emerging infectious disease due to a shifting epidemiology demonstrated, in part, by increased incidence across the globe. The causative agents of leptospirosis are pathogenic serovars of the genus *Leptospira,* of which over 250 are currently recognized. These serovars display characteristic associations with certain animal reservoir hosts. As such, the identification of serovars associated with clinical cases is critical for epidemiological surveillance and determination of appropriate and effective outbreak response and prevention strategies. However, a tool for rapid, economical, serovar typing is still lacking. Matrix-Assisted Laser Desorption Ionization Time-of-Flight mass spectrometry (MALDI-TOF MS), has become an increasingly-important tool for fast, cost-efficient, and accurate microorganism identification in diagnostic microbiology laboratories. To explore the potential use of the MALDI platform for epidemiological surveillance of leptospirosis, MALDI's ability to identify *Leptospira* to the serovar level was tested (please see the previous chapter for further details). Having obtained promising MALDI specificity results, MALDI sensitivity then needed to be assessed. Specifically, we needed to learn whether MALDI could be used to detect *Leptospira* directly from patient samples, without a prior culture and isolation step. The objective of this study was to discover the MALDI's sensitivity for *Leptospira* serovar detection and identification.

To do this, serial dilutions of *Leptospira* culture were used to create a standard curve plotting qPCR cycle threshold (C_T) values against the corresponding $\log_{10}$ of the genome equivalents per mL for each dilution. The dilutions were then tested in MALDI sensitivity assays against the previously-created *Leptospira* Main Spectrum Profile reference library, (see Chapter 3). The highest dilutions which returned accurate identifications were noted and correlated with corresponding dilution concentrations and estimated C_T values to determine sensitivity. For this study, sensitivity was determined for accurate serovar-level and genus-level identification. The average C_T value representing the sensitivity was then compared to the average C_T value of qPCR-leptospirosis-positive canine urine samples submitted to the University of Illinois Urbana-Champaign's Veterinary Diagnostic Laboratory (VDL) from the last approximately 2.5 years. A difference of approximately 5 $\log_{10}$ leptospires/mL was found between the concentration required for accurate MALDI serovar identification and the concentration of typical leptospirosis-positive canine urine samples received by the VDL for leptospirosis testing (3.55×10^8 organisms/mL vs 1.18×10^3 organisms/mL, respectively). This means that the average concentration of leptospirosis-positive samples submitted to the VDL is much too dilute to be directly detected by MALDI. To determine if samples at these dilute concentrations could be concentrated enough for MALDI detection, various filters, differential centrifugation protocols, and MALDI sample preparation methods

were tested. None of these resulted in sufficient concentration for direct detection. Given the average concentration and volume of leptospirosis-positive canine urine samples, results showed that MALDI identification of *Leptospira* serovars directly from canine urine samples during routine diagnostic testing is not feasible.

4.2 INTRODUCTION

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has become an increasingly important and transformative tool for the identification of microbial species in diagnostic microbiology laboratories (Seng et al. 2009, Clark AE et al. 2013). Several commercially available MALDI-TOF MS platforms are available for use in whole-cell microorganism identification, including the Bruker MALDI Biotyper® (Bruker Daltonics, Germany) and VITEK® MS (bioMérieux, France). While the organism identification process varies between platforms, generally the workflow entails the acquisition of a raw mass spectrum from an unknown isolated organism and the subsequent comparison of this spectrum to a library of reference spectra to find a match. The spectrum collected from the isolated organism is a reflection of its proteome. The reference library to which the sample spectrum is compared consists of representative spectral profiles for many well-characterized bacteria and fungi. The comparison is performed in a pattern-matching process that takes into account the location, intensity, and frequency of peaks in the spectra.

A large and growing body of research describes the various sample preparation protocols, MALDI platforms, software programs for spectra acquisition and analysis, and variety of organisms that have been accurately identified (Croxatto et al. 2012, Heaton and Patel 2017). Studies have found that integration of whole-cell MALDI-TOF MS identification into routine diagnostic workflows results in faster time to results and a decrease in the cost of labor and consumables per test isolate (Croxatto et al. 2012, Ge M-C et al. 2017). Given the MALDI's record of success in bacterial identification at the genus- and species-levels, its growing role in routine diagnostic testing, and the high cost and labor burdens of existing microbial typing methods, there is considerable interest in the ability of this technology to identify pathogens at a sub-species level.

Bacterial typing at this lower taxonomic level is valuable for a number of key reasons. Subtyping is an essential tool for understanding infectious disease epidemiology. It allows for phenotypic and genotypic characterization of different bacterial populations (Leone et al. 2008, Ruppitsch 2016). Subtypes often exhibit host nidality, exhibiting characteristic associations with particular reservoir (maintenance) hosts (Ko et al. 2009, Galloway and Levett 2010, Sykes et al. 2011). Host-adapted strains usually cause only sub-clinical or mild forms of disease in their reservoir hosts, making recognition of carriers based on clinical signs alone, difficult (Sessions and Greene 2004). Bacterial typing can help to identify and elucidate these pathogen-host relationships. This data can then be used to determine the probable source of infection in clinical cases, and predict likely transmission routes (Foxman et al. 2005, Wang et

al. 2015). This, in turn, can help control the spread of disease by guiding the design and implementation of appropriate measures to reduce transmission opportunities (Guerra 2013). Applied to clinical isolates during routine diagnostic testing, bacterial typing can be used for real-time epidemiological surveillance. Data obtained through surveillance can be used to identify disease trends, providing baseline information that will be helpful in recognizing outbreaks. Early recognition of changes to baseline disease prevalence allows for faster disease management response (Adzitey et al. 2013).

Subtype identification can also be crucial in vaccine design. This is true for leptospirosis. Current US canine leptospirosis vaccines are bacterin-based. These protect against a targeted subset of serovars presumed to be the predominant circulating serovars responsible for causing the majority of canine leptospirosis in the US. However, seroprevalence may shift over time (Bharti et al. 2003, Moore et al. 2006, Gautam et al. 2010), and the efficacy of these vaccines may be reduced (Sykes et al. 2011). Regular bacterial typing and epidemiological surveillance can help to identify shifts in strain prevalence that may warrant a public/veterinary health response or vaccine redesign. In addition, real-time surveillance allows for a more rapid response compared to prevalence shifts identified only in retrospective epidemiological analyses.

Studies examining MALDI's ability to identify microorganism subtypes have met with varying success. Some studies have sought to identify unique spectral peaks (biomarkers) in bacterial sample spectra which could be used to differentiate strains. However, this approach by itself has not led to consistent identification (Sandrin et al. 2013). Subtyping often requires optimization of sample preparation, including supplementary preparation steps and reagents. Additionally, adjustment of spectra acquisition parameters and extended spectra analyses are often required (Murray 2010). Even with these challenges, a number of studies have found that MALDI may be a promising alternative to standard subtyping methods. MALDI has been used to successfully subtype various Gram-negative bacteria, including serovars of *Salmonella enterica* (Leuschner et al. 2003, Dieckmann et al. 2008, Dieckmann and Malorny 2011), strains of *Erwinia* (Sauer et al. 2008), and pathotypes of *Escherichia coli* (Barbuddhe et al. 2008, Clark GC et al. 2013). Gram-positive bacteria that have been subtyped by MALDI include *Listeria monocytogenes* serotypes 4a and 4c (Barbuddhe et al. 2008), *Streptococcus equi* subsp. *equi* and *S. equi* subsp. *zooepidemicus* (Mani et al. 2017), and *Clostridia difficile* ribotypes 001, 027, 126 and 078 (Reil et al. 2011).

There is also interest in determining whether MALDI can be used for the detection and typing of leptospirosis. With an increasing number of reported cases and outbreaks across the globe, and more severe clinical presentations in both humans and other animals, leptospirosis has become an increasingly important public and veterinary health concern (Ward et al. 2002, Moore et al. 2006, Costa et al. 2015, Haake and Levett 2015). The increase in leptospirosis case numbers and expansion into previously non-endemic areas is predicted to continue (Lau et al. 2010, Chen et al. 2011, Guerra 2013). Additionally, future shifts in *Leptospira* seroprevalence are considered likely due to climate change and an increasing

and increasingly globalized population (Lau et al. 2010, Bandara et al. 2014, Haake and Levett 2015, Day et al. 2016). These shifts may negatively affect the efficacy of leptospirosis vaccines allowing for an increase in leptospirosis-caused morbidity and mortality. Given the potential for severe disease and death, it is important that changes in leptospirosis epidemiology that may impact human and animal health, including outbreaks, be detected as early as possible. Regularly performed epidemiological surveillance, which would include typing *Leptospira* isolates in patient samples, is necessary for early detection. A variety of typing methods have been applied to *Leptospira*, with variable success. However, currently available methods are too laborious, costly, and time-consuming to incorporate into routine leptospirosis testing (Cerqueira and Picardeau 2009, Moreno et al. 2016). Therefore, there is a critical need for a detection and typing method that can be used in daily diagnostics.

A review of the literature found only a few studies that have tested the MALDI's ability to identify *Leptospira* at any level (Djelouadji et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Xiao et al. 2015). Currently, commercial MALDI platforms do not include *Leptospira* representative spectral profiles in their respective reference libraries, with one exception. Recently the CDC and Bruker have collaborated to offer access to a Main Spectrum Profile (MSP) library through MicrobeNet.cdc.gov. This site allows submission of user spectra for identification against a library that contains MSPs for some organisms not included in the MSP libraries offered through Bruker alone. MSPs for several *Leptospira* species are included. However, it currently lists MSPs for only two serovars (*L. interrogans* serovar Pyrogenes and *L. interrogans* serovar Icterohaemorrhagiae). So, MSPs for most *Leptospira* serovars are not available through this library. Additionally, *Leptospira* from different geographic areas may differ somewhat in their spectral profiles, requiring additional MSPs for accurate identification. Researchers who previously tested MALDI's ability to identify *Leptospira* created their own reference libraries using a selected group of representative *Leptospira* isolates (Djelouadji et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Xiao et al. 2015, Karcher et al. 2018, Sonthayanon et al. 2019). One reason for the lack of *Leptospira* spectral profiles in commercial reference libraries could be the fastidious nature of this organism. The majority of pathogenic Gram-negative bacteria processed by veterinary diagnostic laboratories can be cultured on solid agar media within 24–48 hours for isolation and testing. In contrast, *Leptospira* do not grow on conventional solid media. Growth on specially-enriched agar media, if successful, takes a considerable amount of time, and is often not apparent until 4–8 weeks after inoculation (Levett 2001, Adler and de la Peña Moctezuma 2010). *Leptospira* cultures needed for serological diagnostic testing, such as whole-cell *Leptospira* antigen used for the MAT, are instead grown in enriched, liquid media (OIE 2018). While this type of media is suitable for the MAT, it presents its own challenges for MALDI testing.

The challenges of culturing *Leptospira* may also be why there are relatively few studies examining the use of MALDI for *Leptospira* identification. In general, whole-cell MALDI-TOF MS bacterial identification involves the culture and isolation of bacteria on solid agar media prior to analysis. This helps to ensure a

monomicrobial sample and a sufficient number of bacteria available for successful MALDI detection. The majority of Gram-negative bacteria tested on the MALDI are directly spotted onto the MALDI target from an isolated colony grown on solid media. *Leptospira* grown in liquid or semi-solid media do not offer colonies that can be directly collected from the media and deposited onto the target. Therefore, a different sample preparation method must be used. Most studies that have tested MALDI identification of *Leptospira* serovars have used variations of an ethanol-formic acid extraction (TE) method as part of the sample preparation (Rettinger et al. 2012, Calderaro et al. 2014b, Xiao et al. 2015, Karcher et al. 2018). This results in a suspension of bacteria in formic acid and acetonitrile. After centrifugation, the resulting supernatant is deposited onto the target via pipette (Bruker Daltonics, Inc. 2012).

In a study conducted by Djelouadji et al. (2012), a modified direct spot sample preparation method was used to test the MALDI's ability to detect and identify 10-fold serial dilutions of *Leptospira* reference and clinical isolates. The lower limit of detection for the correct identification to the species level was reported to be 10^5 leptospires/mL. In another study, Rettinger et al. (2012) used an ethanol-formic acid (EtOH/FA) extraction sample prep method on a panel of 28 *Leptospira* reference strains. While all of the strains tested, including saprophytic and intermediate strains, were correctly identified to the species level, they found that their sample prep method required a concentration of at least 10^6 leptospires/mL. In 2014(a), Calderaro et al. tested a group of *Leptospira* reference strains circulating in Italy. Like Rettinger and colleagues, they employed an EtOH/FA extraction method and found that a minimal concentration of 10^6 leptospires/mL was required for species-level identification. The study performed by Xiao et al. (2015) similarly used an EtOH/FA sample prep. Thirty-two of 33 strains tested using their *Leptospira* reference library were correctly identified to the species level and all strains were accurately classified as either pathogenic or non-pathogenic. However, the authors did not provide the leptospire concentration(s) of the strains tested. Therefore, a minimum required concentration was not noted. In 2018, Karcher and colleagues (2018) used a group of well-described Brazilian *Leptospira* strains to test MALDI's *Leptospira* detection ability. Isolates were standardized to a concentration of 1×10^8 organisms per mL prior to sample preparation via EtOH/FA extraction. The reference library used in this study consisted of 31 *Leptospira* strains, representing six species. The library was tested using these same 31 strains as well as 22 field isolates belonging to *L. biflexa, L. interrogans*, and *L. borgpetersenii*. All 22 field isolates were correctly identified to the species-level. These authors reported some difficulty using their reference library to distinguish isolates of *L. interrogans* from those of *L. kirshneri*. Further analyses in ClinProTools and SPSS were needed to obtain reproducible differentiation between the spectra for these two species.

The most recent study published on this subject was performed by Sonthayanon et al. (2019). This group used a combination of *Leptospira* reference species and clinical isolates to create an in-house MSP library used for testing. Serovars of *L. interrogans, L. borgpetersenii, L. kirschneri*, *L. meyeri, L. noguchii*, and *L. weilii* were included. Ninety-six out of 97 samples blind-tested against the library were correctly identified to the species-level. The minimum concentration of *Leptospira* required for accurate MALDI ID

was found to be 1 x 10^6 CFU/mL of culture media. Samples of *Leptospira*-spiked urine were also tested. Samples were created by adding 3 μL of *Leptospira* culture grown to 10^8 CFU/mL to 3 mL of urine. Results were reported as promising, but the authors stated that further work was needed to determine the lowest *Leptospira* concentration needed for MALDI ID in urine samples.

These six studies demonstrated MALDI's ability to correctly identify *Leptospira* at the species-level, with two of the six (Calderaro et al. 2014a, Sonthayanon et al. 2019) also differentiating between certain serovars using additional software, such as ClinProTools. These results suggest that quality sample spectra obtained from *Leptospira* isolates likely contain sufficient information for MALDI serovar differentiation. However, can this spectral data be harnessed for serovar identification via an efficient and reproducible protocol that can practically be applied to real-time typing of patient samples?

To use MALDI for routine leptospirosis surveillance, the average *Leptospira* concentration in positive patient samples submitted for leptospirosis testing would need to fall within the MALDI's sensitivity range. Additionally, average concentration of patient specimens partially depends on the type of patient (e.g. human versus canine) (Barragan et al. 2017). If a MALDI protocol can accurately identify serovars, but only at concentrations higher than that typically found in positive patient samples, then MALDI cannot be feasibly applied to patient samples unless MALDI sensitivity can be increased and/or patient samples can be sufficiently concentrated (van Belkum et al. 2017). Natural amplification by culture is not practical for *Leptospira* diagnosis and routine surveillance due to the amount of time required to culture *Leptospira* and the potential for contaminant overgrowth (Brown et al. 1995, Ahmad et al. 2005). Hence, the need for a protocol that allows MALDI testing of patient samples directly without a culture step.

The companion animal most often diagnosed with leptospirosis in the US is the dog (Bowles 2015, Lunn 2019). Canine leptospirosis can range from asymptomatic infection in canine reservoir hosts, to severe clinical illness and death. While confirmed cases of leptospirosis transmission from pet dogs to their owners are rare (White et al. 2017), infected dogs do pose a risk of zoonotic transmission (Klaasen and Adler 2015, Pijnacker et al. 2016). As seen in humans, the epidemiology of canine leptospirosis has also been changing, with an upsurge in reported cases and new syndromes causing increased morbidity (Ward et al. 2002, Sykes et al. 2011, Pijnacker et al. 2016, White et al. 2017) as well as shifts in seroprevalence (Rentko et al. 1992, Ward et al. 2002, Ward et al. 2004).

The sample types most commonly submitted for canine leptospirosis testing are urine, serum, and tissues (the latter used for postmortem testing). Of these three sample types, urine allows for the earliest detection of leptospirosis infection (IDEXX Reference Laboratories 2014). Therefore, a MALDI protocol used for routine typing of leptospirosis-positive samples would ideally allow for direct testing of urine specimens. To determine whether urine direct testing is feasible, it must first be ascertained whether constituents in urine samples interfere with MALDI testing (e.g. requiring a higher concentration of *Leptospira* in the sample for detection compared to the concentration required for pure cultures). Patient urine samples often contain red and white blood cells, as well as other contaminants. These can hinder

the ability of MALDI to accurately identify the pathogen of interest (DeMarco and Burnham 2014). Identification can also be negatively affected by low numbers of bacteria in dilute samples. Therefore, the sample preparation method used must address the removal of contaminants and concentration of specimen bacteria such that the sample deposited onto the MALDI target contains the lowest amount of contaminants and the highest number of bacteria possible. In five of the aforementioned *Leptospira* studies, test samples consisted of either *Leptospira* cultures alone or cultures mixed with sterile water, deionized water, or saline solution buffered with Sorensen's solution. These sample types are not representative of patient samples (DeMarco and Burnham 2014), and so the results of these studies cannot be extrapolated to clinical specimens.

A growing number of studies have explored various sample preparation methods that allow for MALDI direct testing of patient specimens without a culture step. For example, Ferreira et al. (2010) evaluated MALDI's ability to identify bacteria directly from clinical urine samples by analyzing 220 urine specimens from patients with urinary tract infections. Four mL of each sample were initially prepped for MALDI analysis using a differential centrifugation method. Samples that did not yield reliable identification were then prepped using an ethanol-formic acid extraction protocol. An Autoflex III (Bruker Daltonics, Inc., Germany) MALDI-TOF MS was used for sample analysis. Of the samples that exhibited bacterial growth $> 10^5$ CFU/mL, MALDI correctly identified 92.7% of the infecting bacteria to the genus level, and 91.8% to the species level. The authors found the lower limit of MALDI detection to be 8×10^4 CFU/mL for *Esherichia coli*. Other bacteria were identified only at higher concentrations, such as *Enterococcus faecalis*, which required a bacterial count of 5×10^5 CFU/mL. The authors determined that, in general, concentrations equal to $< 10^5$ CFU/mL result in low confidence scores or no identification. In 2012, Köhling et al. aimed to develop a MALDI sample prep method that could filter, collect, and identify bacteria from patient urine samples even when bacterial concentration was $< 10^5$ CFU/mL. This team used a Voyager DE STR MALDI-TOF mass spectrometer (Applied Biosystems) to analyze 107 patient urine samples. Specimen preparation included filtration via a Millipore Microfil S Filtration Device and several centrifugation steps. Sixty-two of the samples were identified by MALDI. The authors reported that in 22 of 26 samples that were not identified, resultant spectra contained three intense peaks which correspond to the mass-to-charge-ratios of human α-defensins 1, 2, and 3 (Zhang et al. 2002). They suggest that the presence of these defensins may have suppressed the ionization of the infecting bacteria, resulting in spectra without the same identifying peaks that would have been present had the defensins been absent from the samples. This study reported MALDI identification of some samples at concentrations below the lower LOD found by Ferreira et al. (2010). Six out of 14 samples with a concentration $\geq 10^3$ CFU/mL were identified. Only three of these identifications had a confidence level of 99.8-80.0%, while the remaining three had lower confidence levels. While Köhling et al. lists the bacterial species identified, it is not clear which species were identified at lower versus higher concentration levels. DeMarco and Burnham (2014) tried yet another approach using Amicon Ultra-15 Centrifugal Filter Units

(Millipore) to desalt, fractionate, and concentrate patient urine specimens prior to MALDI testing. Fifteen mL of each urine specimen was processed using this diafiltration method. Samples were then analyzed by a Bruker Biotyper Microflex LT mass spectrometer. The lower LOD for Gram-negative bacteria was found to be 10^5 CFU/mL, while that for Gram-positive bacteria was 10^6 CFU/mL. The authors additionally used *E. coli*-spiked sterile urine specimens to evaluate and compare the limits of detection achieved by their sample prep method, and those described by Ferreira et al. and Köhling et al. DeMarco and Burnham found that the method used by Ferreira and colleagues required 10^6 CFU/mL, while that detailed by Köhling et al. required 10^7 CFU/mL. The different LODs reported by these studies may be due to a variety of factors, including the use of different MALDI-TOF MS models and software, different sample preparation steps, and different bacteria in the urine samples. Some studies, which have reported successful MALDI species-level identification directly from urine samples, found that the additional steps and analyses performed to obtain the ID were not readily-translatable for the direct ID of different microorganisms. (Croxatto et al. 2012, Íñigo et al. 2016, Zboromyrska et al. 2016).

While studies that have investigated MALDI direct testing of urine samples have reported species-level identification when sample bacterial concentration was at or above 10^5 CFU/mL (Ferreira et al. 2010, Croxatto et al. 2012, Köhling et al. 2012, DeMarco and Burnham 2014, Kim et al. 2015), success at a sub-species-level appears to be more elusive. In fact, a review of the literature could find no published reports of MALDI identification of bacterial subtypes directly from urine samples. This was not unexpected. MALDI identification at the sub-species level, directly from patient urine samples is a task that combines the challenges of both direct MALDI ID from urine samples and MALDI ID at a sub-species level. Successful direct subtyping involves the satisfactory removal of contaminating substances from a patient sample, sufficient concentration of the sample's infecting pathogen(s), acquisition of high-quality sample spectra that contain enough information for discrimination between similar organisms at a sub-species level, and the creation of in-house MSPs for subtypes that are not yet included in a lab's MALDI-TOF MS MSP library.

Successful identification is partially dependent on the concentration of the target bacteria in a specimen. The direct identification studies described above were all performed using human urine samples. Canine urine samples; however, are generally smaller in volume than human samples and so may yield lower bacterial numbers. Additionally, while canine and human urine composition is similar, there can be differences in urine constituents between the two species (Syme et al. 2007). Therefore, sample preparation methods developed for use with human urine may not be appropriate for use with canine urine. A review of the literature found no published studies that have examined MALDI direct testing on canine urine samples**.** The dearth of published work in this area, the changing epidemiology of leptospirosis, and the lack of a practical tool for routine surveillance, inspired this project.

In this study, we evaluated the sensitivity of whole-cell MALDI-TOF MS for *Leptospira* detection. The hypothesis was that whole-cell MALDI-TOF MS has the sensitivity to detect *Leptospira* serovars in the

urine of clinically affected canines. The overall goal of this study was to determine whether whole-cell MALDI-TOF MS could be used as a rapid and relatively inexpensive method for the detection of clinical canine leptospirosis serovars. This included ascertaining whether typical concentrations allowed for genus versus serovar-level detection. The first objective in achieving this aim was to create a standard curve to estimate the concentration of *Leptospira* in samples of unknown concentration. The second objective was to identify the average C_T value and corresponding *Leptospira* concentration for VDL qPCR-positive canine leptospirosis cases from a 2.5-year period. This average would be used for comparison with the lowest average number of leptospires/mL found to be required for accurate serovar-level identification. The third objective was to conduct MALDI-TOF MS sensitivity trials against the custom *Leptospira* MSP library using serial 2-fold dilutions of serovar cultures and serovar-spiked canine urine samples. Results from these trials could then be evaluated to learn whether urine matrix affects the *Leptospira* concentration required for MALDI detection. If it was found that the MALDI could only detect *Leptospira* at concentrations higher than the average concentration found in positive canine samples submitted for testing, various MALDI sample preparation protocols would be tested to discover whether a particular technique could be used to sufficiently concentrate positive patient samples for accurate MALDI identification.

To assess the MALDI's detection range for *Leptospira*, serial dilutions of serovar cultures were tested against a custom MALDI *Leptospira* Main Spectrum Profile (MSP) reference library created in-house (described in the previous chapter). The results of MALDI runs were used to learn whether the presence of urine affects MALDI sensitivity for *Leptospira* and to determine the lowest concentration that accurately detected test organisms at the serovar- and genus-levels. To determine dilution concentrations, two dilution sets, one consisting of serovar culture plus ultrahigh-performance liquid chromatography (UHPLC)-grade water, and one consisting of culture plus leptospirosis-negative canine urine, were used to perform qPCR. Data from qPCR runs were evaluated to find whether the presence of urine affects *Leptospira* qPCR assay results and to create a standard curve correlating C_T value with the $\log_{10}$ number of genome equivalents.

To get an idea of typical sample concentrations seen in qPCR-leptospirosis-positive canine urine samples submitted to the University of Illinois at Urbana-Champaign's (UIUC) Veterinary Diagnostic Laboratory (VDL), leptospirosis qPCR results from the past 2.5 years were reviewed to determine the average C_T value associated with positive tests. A cycle threshold (C_T) value is the first cycle of an qPCR reaction at which the generated fluorescence crosses a specific threshold (Cepheid, Smart Cycler II Operator Manual, 1999-2005). This specific threshold, (different than the cycle threshold), is a user-defined level of fluorescence signal that distinguishes amplification signal from background noise. The C_T value can often be used to determine the initial DNA template copy number of a sample because it is inversely related to a sample's starting amount of template DNA. A sample with a higher starting concentration of the target DNA template will have a lower C_T value than a sample with a lower starting

concentration (Life Technologies Corporation 2012). The VDL uses the following diagnostic interpretation of C_T values for canine leptospirosis qPCR urine testing:

- C_T values from 0 to 38 are considered positive.
- C_T values > 38 up through 42 are considered suspect and confirmed by gel electrophoresis detection of an 87 base pair PCR product.
- C_T values ≥ 43 are interpreted as negative.

The standard curve was then used to gauge the bacterial concentrations in leptospirosis-positive samples previously received by the VDL and to compare the C_T values and concentrations of the previous VDL samples with those of the study samples.

When study results showed that MALDI did not detect *Leptospira* in samples at concentrations typically received for testing, several different sample concentration and desalting methods were tested to determine whether they could improve MALDI sensitivity. These additional preparation steps did not allow MALDI detection of samples at lower concentrations than that which could be detected by the standard ethanol-formic acid extraction method. Sensitivity of the MALDI for *Leptospira* was found to be approximately 5 $\log_{10}$ less sensitive than qPCR. Given the difference between the concentrations at which MALDI can detect and accurately identify *Leptospira* versus the typical concentrations of leptospirosis-positive patient samples received by the VDL, it was concluded that the MALDI cannot practically be used for the routine diagnosis or epidemiological surveillance of canine leptospirosis. The main limiting factors are the low concentrations of *Leptospira* in typical leptospirosis-positive canine urine samples and the low volume of urine per sample.

Given the large number of pathogenic serovars, and variation in the predominant serovars responsible for disease from region to region, attempting to create a universal reference library that can identify all serovars is impractical. However, the work presented here describes a method by which diagnostic laboratories that use MALDI for routine diagnostics may create *Leptospira* MSPs for use in real-time identification. Moreover, MALDI identification of *Leptospira* could be used for studies that would normally require *Leptospira* culture and isolation and for diagnostic laboratory serovar culture contamination checks. With additional work, this method could be extended as an alternative identification technique in human leptospirosis diagnostics.

4.3 MATERIALS AND METHODS

4.3.1 Safety Precautions

MALDI-TOF MS *Leptospira* sensitivity (Se) trials followed the same safety precautions as that used for specificity (Sp) trials.

4.3.2 *Leptospira* Serovars

The same seven *Leptospira* reference strains used in specificity testing were also used for sensitivity

trials. This group of pathogenic strains was comprised of *Leptospira interrogans* serovar Autumnalis (Akiyami A), serovar Bratislava (Bratislava Jez), serovar Canicola (Hond Utrecht IV), serovar Grippotyphosa (Andaman), serovar Hardjo (Hardjoprajtino), serovar Copenhageni (M20[A], used as a representative member of serogroup Icterohaemorrhagiae), and serovar Pomona (Pomona). Strains, procured from the National Veterinary Services Laboratory (NVSL, Ames, Iowa), were grown in P80 media to a standard density as described in Chapter 3.

4.3.3 Canine Urine Samples

Urine samples were collected from the unused portions of canine urine specimens submitted to the VDL for leptospirosis testing. Only those samples which had tested leptospirosis qPCR-negative were used. Specimens were stored no longer than two weeks at 4 °C before being pooled to obtain the needed volume. Pooled urine was spun at 600 x *g* for 15 minutes to remove larger particles and contaminating cells. The supernatant was then filtered for sterilization using EMD Millipore™ Millex™ -GP sterile syringe filters with a .22 μm pore polyethylsulfone (PES) membrane, as described in Chapter 3.

4.3.4 Spectrophotometric Quantitation

The methods and protocols used for serovar dilution creation, spectrophotometric quantitation of serovar dilutions and DNA template, and DNA extraction were the same as described in the previous chapter on specificity.

4.3.5 Dilutions

Three different types of serial dilutions were used in this sensitivity study. Prior to dilution creation, suspensions of serovar cultures were centrifuged at 600 x *g* for 15 minutes (Sorvall ST 16R centrifuge, Thermo Scientific, Cat. No. 75004240, 23°C) to pellet dead cells for removal. The supernatant was then adjusted to a standardized percent transmittance (%T) of approximately 25, (3.01×10^9 organisms per mL), unless otherwise noted. Early attempts at lower concentrations demonstrated the need for 25 %T to attain 4-5 serial two-fold linear dilutions in range of MALDI detection limits. Adjustments and %T measurements were performed on a Shimadzu BioSpec mini-1240 UV-Vis Spectrophotometer (Shimadzu Corporation) set to a wavelength of 400 nm. One mL aliquots of culture supernatant were measured in disposable semi-micro cuvettes (Sarstedt Ag & Co., Germany). One mL of UHPLC-grade H_2O was used as the reference blank. Dilutions were created using the adjusted serovar cultures. Serial two-fold dilutions, ranging from 3.01×10^9 to 1.01×10^8 organisms per mL, and consisting of serovar cultures mixed with either UHPLC-grade H_2O or canine urine, were used for MALDI testing. A second set of two-fold serial dilutions was created with DNA extracted from serovar cultures adjusted to 25 %T. Extracted DNA was combined with UHPLC-grade water to create 2-fold dilutions through 1:32, the same range of dilutions as those used for MALDI testing. These dilutions were then measured for DNA concentration.

The third dilution set consisted of ten-fold serial dilutions, ranging from 10^8 to 10^1 organisms per mL**,** also consisting of culture and either UHPLC-grade H_2O or canine urine, were used for qPCR.

Various trials called for different sets and numbers of dilutions. Some trials required a substantial volume of *Leptospira* culture not only due to the number of dilutions needed, but because the culture first had to be concentrated to the standardized 25 %T prior to dilution creation. While serovar cultures were created as needed for specific trials, there were times when the volume of culture needed for a trial surpassed the amount grown specifically for this project. In those situations, the appropriate serovar cultures leftover from VDL leptospirosis microscopic agglutination testing were combined to acquire the needed volume, so as not to delay the trial while new cultures were grown and to save on the amount of culture media used. VDL cultures that were used for this study were of similar age as the study cultures with which they were combined.

4.3.6 DNA Extraction

DNA was extracted from 1 mL aliquots of each dilution sample using the QIAamp® Viral RNA Mini Kit (Qiagen, Germany) and Purification of Viral RNA protocol (QIAamp Viral RNA Mini Handbook, 2018, p. 27). This kit and protocol are recommended by the manufacturer for the extraction of bacterial DNA from urine samples to aid in the removal of qPCR inhibitors potentially present in the urine. Moreover, the University of Illinois Urbana-Champaign's Veterinary Diagnostic Laboratory commonly uses this protocol to extract DNA from canine urine samples submitted for leptospirosis qPCR testing.

4.3.7 Measurement of DNA Template Concentration and Estimation of Genome Equivalents per Dilution

DNA template concentrations for serial two-fold dilutions were measured using a Qubit™ 4.0 Fluorometer (Invitrogen, Life Technologies Holdings Pte Ltd, Singapore) with the Qubit™ dsDNA Broad Range Assay Kit as previously described in Chapter 3. Briefly, for each sample dilution to be measured, 3 μL of the sample and 197μL of the working solution were combined in a clear 0.5 mL PCR tube and incubated for two minutes. After calibrating the Qubit with the two prepared Broad Range assay standards, the dsDNA assay setting was used to measure each sample in duplicate. Average measurements per dilution were then calculated.

Calculation of genome equivalents per μL for each dilution was also performed as previously described (Chapter 3). In brief, the average Qubit concentration measurement for each dilution (DNA ng per μL) was used in the following equation (Staroscik 2004):

$$\text{Number of copies } /\mu\text{L} = \frac{(\text{amount of DNA in ng}/\mu\text{L}) \times (6.022\times10^{23}\ \text{molecules/mol})^{a}}{(\text{length of dsDNA template in base pairs [bp]})^{b} \times (650\ \text{g/mol})^{c} \times (1\times10^{9})^{d}}$$

[a] Avogadro's number

[b] length of dsDNA template = size of genome = 4,627,366 bp

[c] 650 g/mol = 650 Da = the assumed average weight of a single DNA bp

[d] $1x10^9$ ng/g was used in the equation to convert our calculated number to ng (which allows for units ng to be cancelled, leaving molecules/μL as units for the copy number)

DNA template, not qPCR product, was measured for concentration. Therefore, this equation was calculated using a template length of 4,627,366 bp, which is the reported size of *L. interrogans* serovar Copenhageni's genome (Nascimento et al. 2004). The calculated number of copies per μL were then converted to mL.

4.3.8 qPCR Assay

Real-time PCR was performed using the DNA extracted from the serial 10-fold dilutions, described above. The qPCR assay used was based on previous work by Smythe et al. (2002). Originally designed for use with human samples, this assay was modified and validated for use with veterinary samples by our lab (Borst and Maddox 2004). Primers Lepto F2 (5'^{171} CCCGCGTCCGATTAG 3') and Lepto R2 (5'^{258} TCCATTGTGGCCGR$^{A/G}$ACAC 3') were used to amplify an 87 bp fragment of the 16S rRNA (*rrs*) gene found in pathogenic *Leptospira* species. This assay uses a TaqMan probe, [5'^{205}(FAM) CTCACCAAGGCGACGATCGGTAGC^{228} 3' (TAMRA)] labeled at the 5' end with the fluorescent reporter dye, 6-carboxy-fluorescein (FAM). The 3' end contains the quencher, 6-carboxy-tetramethyl-rhodamine (TAMRA) (LGC Biosearch Technologies, Inc., Petaluma, CA). The reaction mixture was prepared using OmniMix™ HS reagent beads (Cepheid). Each bead contains reagent amounts for two 25 μL qPCR reactions. Reagent amounts per bead, when reconstituted to 50 μL, are as follows: 200 μM dNTPs, 4 mM $MgCl_2$, 25 mM HEPES buffer (pH 8.0 ± 0.1), 3 U hot-start *Taq* polymerase. Additional reagents and corresponding volumes making up the Master Mix for two reactions were as follows: 1 μL of 10 μM Lepto F2 and Lepto R2 primers (Integrated DNA Technologies), 2 μL of 10 μM probe, 5 μL of DNA template, and 45 μL of nuclease-free H_2O in a total volume of 50 μL. Each 25 μL reaction consisted of 22.5 μL of master mix plus 2.5 μL genomic DNA. The positive amplification control consisted of 22.5 μL Master Mix plus 2.5 μL of a 1:1000 dilution of genomic DNA, (1.97×10^{-2} ng/μL, 3.94×10^3 leptospires/μL), from *L. interrogans* serovar Grippotyphosa (original source: NVSL ICL-020). Master mix sans template served as the negative template control. Amplification was performed using a Cepheid SmartCycler®. The thermal cycling protocol consisted of an initial denaturation step at 95 °C for 120 seconds followed by 50 cycles consisting of denaturation at 95 °C for 15 seconds and extension at 60 °C for 30 seconds. The baseline threshold was set at 30 fluorescent units. C_T values, concentrations, and calculated numbers of template copies per mL were correlated for each dilution.

4.3.9 Determining the Average C_T Value of Previous qPCR-Leptospirosis-Positive VDL Cases

The Vetstar Animal Disease Diagnostic System (VADDS) was used to mine VDL records for positive canine leptospirosis cases diagnosed via qPCR of urine for the time period of January 1, 2016 through July 1, 2018. qPCR C_T values for these cases were averaged to obtain the mean C_T value for positive canine urine samples.

4.3.10 Creation of a Standard Curve

To determine whether *Leptospira* dilutions detected and correctly identified by the MALDI during sensitivity testing fell within the concentration range of typical leptospirosis-positive patient samples, a standard curve was created. Serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona were chosen to create the curve. The seven serovar strains used in this study all belong to the same species, *L. interrogans*. Therefore, we anticipated that dilution concentrations and corresponding C_T values would be similar for all seven. In the interest of saving time and resources, we chose to use four instead of all seven serovars for curve creation. These four serovars were selected because they are considered the most common serovars that infect canine companions in the US.

One set of serial 10-fold dilutions, consisting of serovar culture plus UHPLC-grade H_2O as mentioned previously, was created for each of these four serovars. A second set of serial 10-fold dilutions, consisting of serovar culture plus urine, was created for serovars Grippotyphosa and Pomona. DNA extraction was performed on these dilutions and the resulting DNA template was used for qPCR. The urine dilutions were used to mimic patient urine samples and determine whether urine adversely affects the sensitivity of the qPCR assay. Once the qPCR runs were performed, the C_T values for each H_2O dilution were averaged, as were the C_T values for each urine dilution. DNA template dilution concentrations were also calculated and averaged. These averages were then used to create two standard curves, one containing data from the H_2O dilutions and one using data from the urine dilutions. Curves were constructed in Microsoft Excel by plotting the logarithm of the average initial number of leptospires per mL for each dilution along the x-axis, and the average corresponding C_T values along the y-axis.

Linear regression analysis was performed and the coefficients of determination (R^2) were calculated. The resulting regression equations gave the slope of the standard curve for the respective set of dilution data, which was used to estimate the efficiency of the qPCR assay. The efficiency is the rate at which an amplicon is produced during the exponential phase of an qPCR reaction. Ideal efficiency for an assay is 100%, which theoretically translates to a 2-fold increase in amplicon quantity with each cycle, and a 10-fold increase in amplicon quantity for every 3.32 cycles (Kralik and Ricchi 2017). Therefore, a standard curve slope of -3.32 indicates 100% efficiency (Kralik and Ricchi 2017). Slopes less than –3.32 indicate reactions with less than 100% efficiency, while slopes greater than –3.32 may indicate problems with pipetting or with sample quality (Applied Biosystems 2008).

The two curves were compared for differences in the average C_T values and quantity of

leptospires/mL per dilution. Once it was found that differences in the C_T values for urine versus H_2O dilutions were negligible, the H_2O dilution standard curve was used for subsequent analysis.

4.3.11 MALDI-TOF MS Sensitivity Testing

Reagents used for MALDI-TOF MS sensitivity testing were the same as those used for specificity testing. The reader is directed to the previous chapter's Materials and Methods section for detailed information. The procedures used for sample spectra acquisition and MALDI real-time identification were also the same as described in the previous chapter. To briefly recap, two-fold serial sample dilutions were prepared using an ethanol/formic acid protein extraction method. For each MALDI run, one μL of each sample to be tested was spotted in duplicate onto an MSP 96-spot ground steel target plate. Once dry, each sample spot was overlaid with one μL of HCCA matrix solution. Each MALDI run included two target spots of Bruker's bacterial test standard (BTS) used for calibration and validation of the run and as a positive control. One target spot per run was also used for non-inoculated matrix, which served as the negative control. After deposition of samples and reagents on the MALDI target, spots were allowed to air dry prior to MALDI analysis. This allowed for proper sample/matrix co-crystallization required for spectra acquisition. The prepped target was then inserted into the MALDI instrument. Sample spectra acquisition and analyses was performed on a Microflex™ LT mass spectrometer running FlexControl™ software and equipped with a 20 Hz N2 Nitrogen laser. Spectra were acquired in linear positive mode within a mass range of 2000–20,000 Da. Real-time identification was executed using Biotyper Realtime Classification (RTC) software. Samples were tested against the Bruker Daltonics database Main Spectrum Profiles created in-house for additional microorganisms not included in the BDAL database, and the custom *Leptospira* reference library, (the creation of which was discussed in Chapter 3). Interpretation of identification results followed the scoring guidelines given by Bruker, also discussed in Chapter 3. As a brief review, scores are given on a logarithmic scale and range from 0.00 to 3.00. Scores below 1.70 indicate unreliable identification results. Scores at or above 1.70 are divided into three ranges that indicate increasing confidence levels of genus and species identification.

To determine whether urine matrix affects the MALDI's sensitivity for detection and identification of *Leptospira* serovars, two sensitivity trials were performed for each serovar. These trials tested raw spectra from two-fold serial H_2O and urine dilution sample sets against a *Leptospira* MSP test library containing one MSP for each of the seven serovars. For each trial, the *Leptospira* concentration corresponding to the highest H_2O and urine dilutions that returned the correct serovar-level and genus-level identifications with a score ≥ 1.7, were noted. Concentrations recorded for the two dilution sets were then compared.

4.3.12 Sample Concentration Method Development

After discovering that the majority of qPCR-leptospirosis-positive canine urines samples submitted to

the VDL had *Leptospira* concentrations below that which could be detected by the MALDI**,** several sample concentration methods were tested. This was done to determine whether leptospirosis-positive samples could be concentrated such that MALDI could detect and differentiate between *Leptospira* serovars in positive patient specimens.

4.3.12.1 Filtration

One method explored for sample concentration was filtration, of which, the following methods were tested.

4.3.12.2 Enclosed syringe filters

The first method examined used enclosed syringe filters to concentrate serovar Grippotyphosa stock culture. Serovar culture, previously grown for seven days, was initially centrifuged at 600 x *g* for 15 minutes to remove dead cells. The supernatant was then adjusted to a 39.4 %T. An EMD Millipore™ Millex™ -GP sterile syringe filter with a .22 μm pore PES membrane (Burlington, MA, USA) was attached to a 3 mL syringe. A 1.2 mL aliquot of the culture was run through the filter and collected into a 1.5 mL microcentrifuge tube (Evergreen Scientific, Rancho Dominguez, CA). The %T of the collected filtrate was measured to determine if it differed from the initial 39.4%T. A lower %T could indicate an increase in the *Leptospira* concentration of the filtrate, while a higher %T could indicate a decrease in *Leptospira* concentration, compared to that of the starting culture suspension. The %T of the filtrate was recorded, and the filtrate was discarded. Another 1.2 mL aliquot of the serovar stock culture was then run through a different EMD Millipore™ Millex™ Sterile Syringe Filter. This filter had a Durapore™ polyvinylidene difluoride (PVDF) membrane with a .22 μm pore size (Burlington, MA, USA)**.** The filtrate was collected into a new microcentrifuge tube and measured for %T. The %T was recorded, and the filtrate was discarded. Finally, a new 1.2 mL aliquot of the culture was run through a PALL Acrodisc® Syringe Filter with a Versapor® membrane and a .45 μm pore size (Port Washington, NY, USA) and collected into a new microcentrifuge tube. Once more, the %T of this filtrate was measured, recorded, and the filtrate was discarded. All %T measurements were then compared to determine if any of the three filters tested appeared to either increase or decrease *Leptospira* concentration. If the filtrate from a particular filter had a higher %T than that of the initial culture suspension, that would suggest the filter had trapped some of the *Leptospira*. The purpose of testing filtration was to find a way to concentrate the *Leptospira* in a sample rather than remove it. Therefore, since these filters produced filtrate with percent transmittance measurements indicative of trapping the *Leptospira*, they were not tested further.

4.3.12.3 Membrane filters

Two membrane filters with two different pore sizes were tested next. This trial used serovar Icterohaemorrhagiae stock culture adjusted to 39.8 %T. An EMD Millipore Durapore™ PVDF membrane

filter with a 0.22 μm pore size was placed in a 25 mm Swinnex Filter holder and then attached to a 3 mL syringe. Two mL of the stock culture was run through the filter into a snap cap tube. The %T of the filtrate was then measured. A new 25 mm Swinnex Filter Holder was fitted with an EMD Millipore Durapore™ PVDF membrane filter with a 0.45 μm pore size and attached to a new syringe. Two mL of the unfiltered stock culture were run through this filter with the filtrate collected into a new snap cap tube. The %T of this filtrate was measured. To determine if the trapped particles in the filter could be flushed out of the filter with only a few additional steps, 2 mL of sterile water were run through the same filter into a new snap cap tube. Again, the %T of the filtrate was measured. Since this transmittance appeared to indicate that no bacteria had been flushed out with the water, the filter holder was opened, and the membrane was gently scraped with a disposable inoculating loop. The loop was then swirled in the filtrate collected from the prior run with water. This scraping and swirling procedure was repeated numerous times to collect material adherent to the membrane's surface that could be collected with this technique. The filtrate was then mixed thoroughly and measured for %T. Filtrate %T and starting stock culture %T were compared to determine if these filters trapped the *Leptospira.*

4.3.12.4 Amicon centrifugal filters

Based on the previous study by DeMarco and Burnham (2014) mentioned above, Amicon centrifugal filters (Amicon Ultra-4 Centrifugal Filter Unit with Ultracel-10 membrane, Millipore) were chosen for the next filtration test. Serovar Icterohaemorrhagiae stock culture was adjusted to 31.6 %T, (4.34×10^9 organisms/mL), using the Shimadzu spectrophotometer, as described earlier. A portion of the culture was then combined with canine urine to create a 1:2 (2.17×10^9 organisms/mL), and a 1:4 (1.09×10^9 organisms/mL) dilution. Starting concentration and dilutions were based on data from initial standard curve trials and MALDI runs (data not shown) performed to determine what, if any, overlap exists for the qPCR versus MALDI detection range. Two mL of the starting concentration and of each dilution were placed in a centrifugal filter. The three filters were then centrifuged at 4000 x *g* for 25 minutes at 4°C. Flow-through was discarded and 1 mL of UHPLC-grade H_2O was added to each filter's reservoir. The filters were again spun 4000 x *g* for 25 minutes at 4°C. After this second spin, the concentrate from each filter's reservoir was collected using a 20-200 μL pipettor and tips. The pipette tip was gently scraped along the sides of each reservoir to dislodge any remaining *Leptospira*. The concentrate from each filter was then transferred to a 1.5 mL microcentrifuge tube. The tubes were spun at 14,000 x *g* for 3 minutes to pellet the bacteria and the supernatant was discarded. All three tubes had well-formed pellets, which were resuspended in 1 mL of UHPLC-grade H_2O and centrifuged again at 14,000 x *g* for 3 minutes. The supernatant from each tube was carefully removed via pipette. Next, samples were prepared for MALDI analysis using the ethanol-formic acid (EtOH/FA), or tube extraction, method described in the previous Specificity chapter. After the addition of H_2O and EtOH to each sample tube and subsequent centrifugation, no bacterial pellets were observed, but all three tubes had cloudy spots present on one

side. The tubes were again spun at 15,700 x *g* for 1 minute to collect the remaining supernatant, which was removed via pipette and discarded. After the tubes were allowed to dry, 25 μL of both 70% FA and ACN were added to each tube and tube contents were mixed via pipette. The tubes were then vortexed and centrifuged at 15,700 x *g* for 2 minutes. After this, 1 μL from each tube was deposited in duplicate on the MALDI target plate**.** On visual inspection of the target to ensure all sample spots were dry before adding the matrix, one of the two 1:4 dilution spots appeared to have little or no material. So, an additional spot of this dilution was deposited on the target to ensure that at least two spots had sufficient sample material.

A second trial using Amicon filters and additional dilutions was performed. This trial also tested the direct transfer method of sample deposition. Serovar Icterohaemorrhagiae culture stock was again used, this time adjusted to 27.5 %T. The stock was then combined with canine urine to create four dilutions at 1:2000, 1:8000, 1:12,000, and 1:16,000. These dilutions were chosen not only to fall within the linear range of the qPCR assay, but also in an attempt to more closely resemble the concentrations of leptospirosis-positive patient samples usually received by the VDL. The filters were tested to see if they would allow sufficient concentration of patient samples so that they could be detected by the MALDI. As in the previous trial, two mL of each dilution were transferred into an Amicon filter. The subsequent prep and processing steps were also the same as those described above for the first trial. Filtration again resulted in approximately 100 μL of concentrate collected for each dilution. This meant that the resulting filtrate for each dilution had been concentrated approximately 20 times. The filtrate concentrations for each dilution were as follows: 1:2000 dilution (1.01 x 10^8 organisms/mL), 1:8000 dilution (2.52 x10^7 organisms/mL), 1:12,000 dilution (1.90 x 10^7 organisms/mL), and 1:16,000 dilution (1.27 x 10^6 organisms/mL). Concentrates were transferred to microcentrifuge tubes for EtOH/FA extraction. After extraction, pellets were visible in all but the 1:8000 dilution tube. The supernatant in the 1:2000, 1:12,000 and 1:16,000 dilution tubes was removed via pipette and discarded. For the 1:8000 dilution, all but approximately 5 μL of supernatant was removed. This residual supernatant was kept to aid in the transfer of sample material to the MALDI target.

The direct transfer method was used to deposit a portion of each sample onto the target. First, a rounded wooden applicator stick was used to collect part of the pellet from the 1:2000 dilution tube and place it onto a sample spot on the target plate. However, no pellet material could be seen on the target. The stick appeared to have absorbed the moist sample. Next, MALDI transfer sticks with a pointed end were used to try and collect a portion of the pellet from each of the remaining tubes. Since no pellet was seen in the 1:8000 dilution tube, the stick was used to pick up a portion of the remaining supernatant at the bottom of the tube. Only one sample spot was used for each dilution so that some of the sample would be left for tube extraction. Sample material from the 1:12,000 and 1:16,000 dilutions, but not from the 1:2000 or 1:8000 dilutions, could be seen on the target. The remaining sample in each tube was then used for TE so that both direct transfer and extraction sample prep methods could be tested with the

Amicon filter concentrates. Tube extraction was performed as previously described. After the addition of matrix to the target, samples were analyzed.

4.3.12.5 Differential centrifugation and extended direct transfer sample preparation

Differential centrifugation was also tested as a method for sample concentration. Serovar Autumnalis culture stock was adjusted to 32.6 %T. Canine leptospirosis-negative urine samples, which had been previously centrifuged and filtered, was spiked with the culture stock to create two-fold serial dilutions ranging from 1:2 to 1:32 (approximately 1.10×10^9 organisms/mL to 6.81×10^7 organisms/mL). One mL of each dilution was placed in 1.5 mL microcentrifuge tube and spun at 15,700 x *g* for 2 minutes to pellet any remaining larger particles and blood cells present in the urine. The supernatant was transferred to new microcentrifuge tubes and spun at 15,700 x *g* for 4 minutes to pellet the bacteria. The amount of visible pellet in each tube after centrifugation varied. The supernatant was removed from each tube via pipette. Pointed wooden MALDI transfer sticks were then used to deposit a portion of the pelleted material from each tube onto the MALDI target in duplicate. The second spot for each dilution was overlaid with 1 μL of FA to test the extended direct transfer sample preparation method. After drying, target preparation was completed with the addition of BTS and matrix.

4.3.12.6 Other sample preparation and deposition methods

Several other sample preparation and spotting methods were also evaluated to determine what, if any, effect they had on MALDI sensitivity. These methods entailed variations on the direct transfer and TE methods that included one and two added washes using both UHPLC-grade H_2O and EtOH, mixed prep methods in which the sample and matrix were combined prior to deposition on the target, and a sandwich method in which the sample was placed on the target between two layers of matrix. For further description of these methods, the reader is referred to the Direct Transfer and EtOH/FA Method and Variations subsections within the Materials and Methods section in the previous Specificity chapter.

4.3.13 Data Analysis

Lowest concentrations recorded for MALDI sensitivity trials were broken into four sets of 14. Two sets contained the lowest concentrations that returned accurate serovar-level identifications; one for H_2O dilutions and one for urine dilutions. The other two sets contained the lowest concentrations that returned accurate genus-level identifications for the H_2O and urine dilutions. The distribution of each of these sets were evaluated for normality using histograms, kurtosis, skewness, and the Shapiro-Wilk test. The median, range and 10% and 90% percentiles were determined for non-normally distributed data. The two-tailed, Mann-Whitney U test was used to compare the concentrations for serovar-level and genus-level identifications between the H_2O and urine dilution sets. Statistical analyses were performed using the

XLSTAT 365 software version 2.1 (Addinsoft, 2017, USA) in Microsoft Excel, at a significance level of 0.05.

Next, results from MALDI sensitivity trials were reviewed to determine whether MALDI sensitivity allows for the detection and identification of *Leptospira* within the typical concentration range of leptospirosis-positive canine urine samples submitted to the VDL. The standard curve was used to estimate the C_T values that correspond to the lowest *Leptospira* concentrations that returned accurate serovar- and genus-level identifications. These C_T values were then compared to the average C_T value and C_T value range for qPCR-leptospirosis-positive canine cases submitted to the VDL over a 2.5-year time period.

4.4 RESULTS

4.4.1 Percent Transmittance as a Measurement of *Leptospira* Numbers

The Shimadzu spectrophotometer was used to measure the %T of *Leptospira* culture two-fold serial dilutions for each serovar. Individual graphs plotting the %T versus the number of *Leptospira* copies per mL were created. Data from serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona were averaged and used to create a representative graph below (Figure 4.1) (see Appendix A for individual curves for these four serovars).

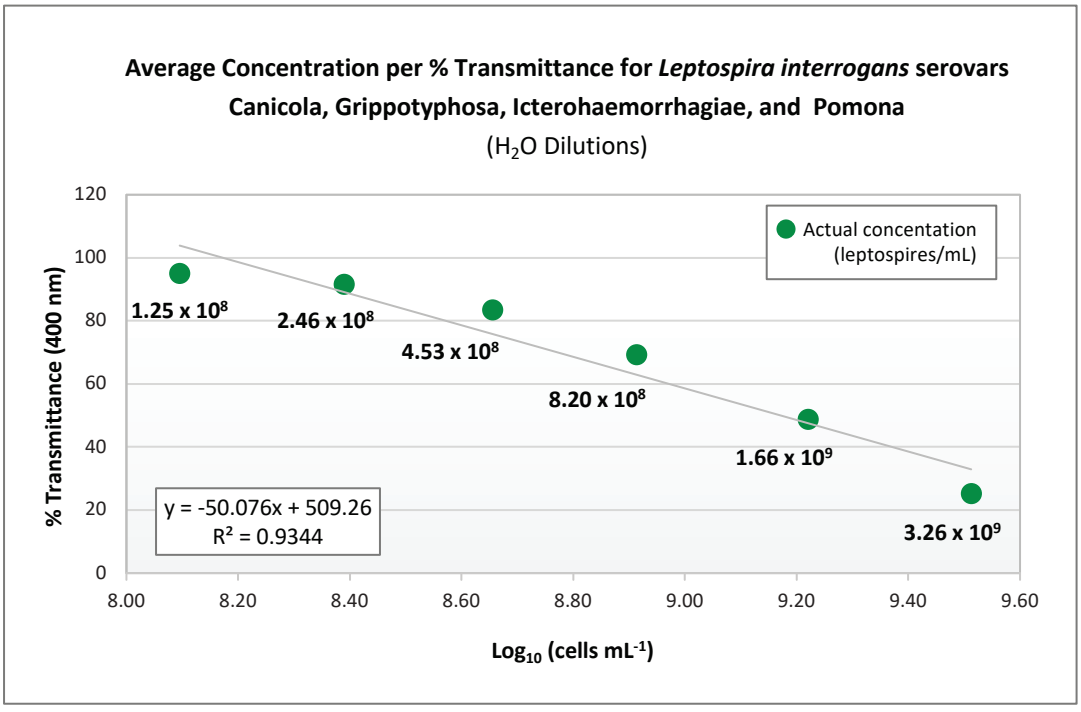

Figure 4.1. Each data point represents the average percent transmittance for one of five sample concentrations tested for serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona. Concentrations corresponded to a series of two-fold serial dilutions created from serovar cultures and UHPLC-water. The graph shows an inverse linear relationship between percent transmittance and *Leptospira* concentration.

4.4.2 Standard Curves

Two standard curves were created using 10-fold serial dilutions. The first curve used data from serovar culture and UHPLC-grade H_2O dilutions (Figure 4.2). The data set for this curve was obtained by averaging qPCR, DNA concentration and corresponding genome equivalents data from serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona (Appendix B). The second curve was generated with data from dilutions of canine urine spiked with serovar culture (Figure 4.3). Data for this this curve was acquired by averaging qPCR, DNA concentration and corresponding genome equivalents data from serovars Grippotyphosa and Pomona (Appendix C). Log quantity values were obtained by taking the log of the calculated number of genome equivalents per mL from the measurements of DNA template concentration per dilution. The linear regression line and regression coefficient (R^2) were calculated for each curve. The standard curve created for the H_2O dilutions had a regression line of $y = -3.495x + 46.299$ with an R^2 of 0.998. The slope, -3.495, was used to calculate a reaction efficiency of 93.252%. The curve for the urine dilutions had a calculated regression line of $y = -3.495x + 47.000$, an R^2 of 0.998, and also showed a reaction efficiency of 93.252%. The data for both curves displayed an inverse, linear relationship, in which the C_T values increased as the number of template DNA copies decreased.

The two curves were then compared to determine whether urine inhibits qPCR sensitivity. If inhibition was detected, then separate H_2O and urine dilution curves would be needed to evaluate qPCR and DNA concentration data. Both curves had similar slopes, differing by approximately .153, and similar regression coefficients, differing by only .003. Given these results, urine did not appear to affect the sensitivity of the qPCR assay. Considering the similar data points for both curves, and the limited urine available for the study, it was determined that the H_2O standard curve would be used for estimating dilution concentration, corresponding C_T values and number of genome copies for subsequent work.

To estimate the number of genome equivalents per mL in leptospirosis-positive samples of unknown concentration, ten-fold serial dilutions of serovar culture and UHPLC-H_2O, ranging from 10^8 to 10^1, were created and used to perform qPCR. Data actually used for analysis fell within the range of 10^7–10^3 organisms/mL because it was found that data for concentrations extending beyond this range fell outside of the linear range.

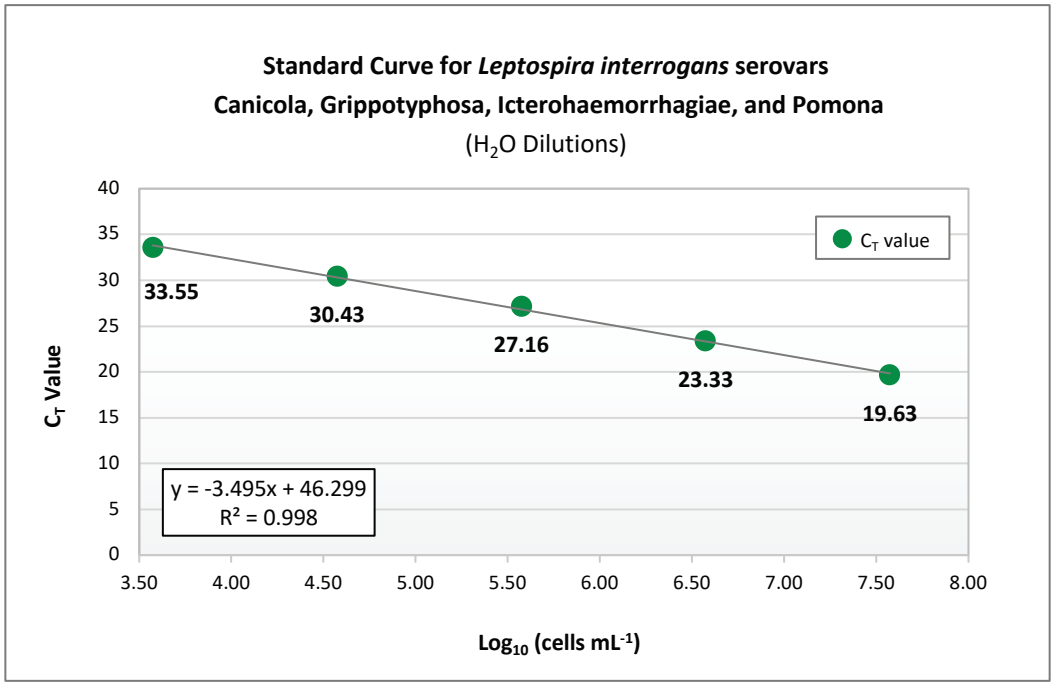

Figure 4.2. A series of 10-fold serial dilutions, ranging from 10^7 to 10^3 leptospires/mL, for serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona were used to perform *Leptospira* qPCR assays. Each data point in the curve is the average C_T value per the log concentration of a sample dilution for the four serovars. This standard curve was used to estimate the average concentration of leptospirosis-positive canine urine samples submitted to the VDL for diagnostics over a 2.5-year period and to estimate the concentrations of samples used during method creation.

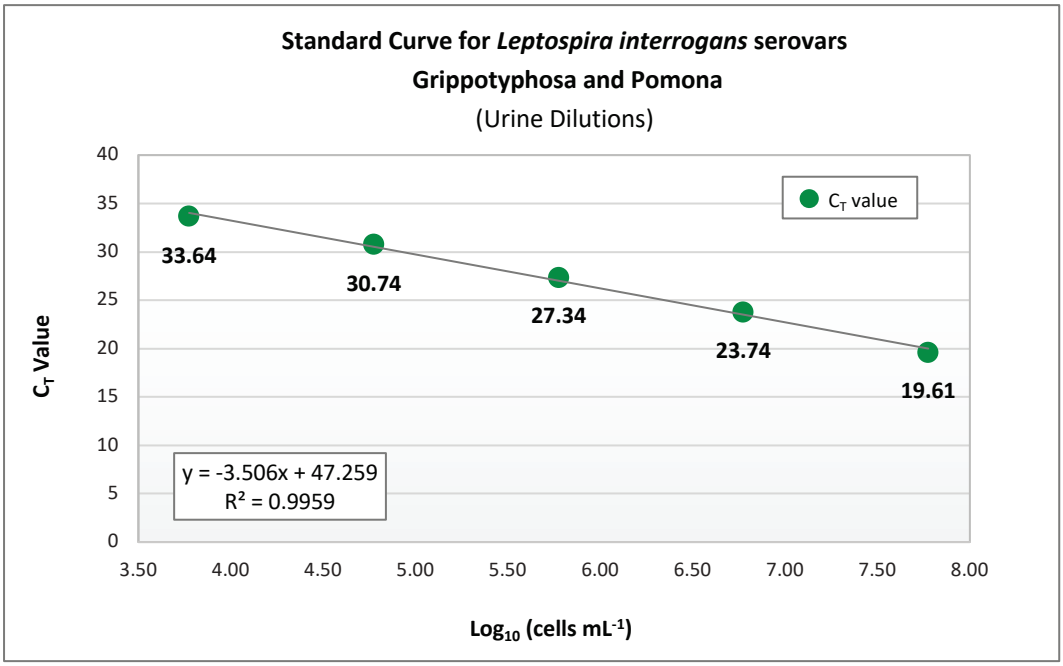

Figure 4.3. A set of ten-fold serial dilutions consisting of serovar culture plus leptospirosis-negative canine urine were used to perform *Leptospira* qPCR assays. Dilutions ranged from 10^7 to 10^3. One assay each was performed for serovars Grippotyphosa and Pomona. Each data point represents the average C_T value per log *Leptospira* target template concentration for these two serovars.

4.4.3 MALDI-TOF MS Sensitivity for *Leptospira*

MALDI sensitivity was assessed from the results of two trials per serovar in which both H_2O and urine dilution test spectra were run against the final *Leptospira* MSP library. The lowest average concentration of serovar culture + H_2O found to return an accurate, MALDI identification over all serovars to the serovar-level was 3.55 x 10^8 organisms/mL and to the genus-level was 3.33 x 10^8 organisms/mL. For dilutions of culture and urine, the lowest average concentration returning accurate MALDI serovar-level identification was 3.43 x 10^8 organisms/mL, while 3.26 x 10^8 organisms/mL was the lowest average concentration for accurate genus-level ID. As expected, lower concentrations and fewer leptospires/mL were required for accurate genus-level ID compared to those required for accurate serovar-level ID.

To determine whether urine significantly affected MALDI detection of *Leptospira* serovars, the lowest concentrations that return serovar- and genus-level identifications in the MALDI sensitivity trials were compared between the H_2O and urine dilution sets. The goal was to learn whether one set of dilutions allowed for *Leptospira* detection at higher or lower concentrations than the other. If no significant difference was found between the concentrations in the two dilution sets, then it would be assumed that results from subsequent testing with H_2O dilutions could be extrapolated to urine dilutions, with the caveat that the hypothetical urine dilutions used urine of similar quality to that which had been used in the sensitivity trials. For the two dilution sets, the lower of the two lowest concentrations that returned accurate serovar-level and genus-level identifications for each serovar was recorded (Table 4.1).

Table 4.1. Lowest concentrations (*Leptospira*/mL) per serovar that returned a correct **serovar-level** and a correct **genus-level** MALDI identification for H_2O and urine dilution sets

	Serovar - Level ID		**Genus -** Level ID	
	H_2O dilutions	**Urine dilutions**	**H_2O dilutions**	**Urine dilutions**
Serovar	**# Template copies/mL**	**# Template copies / mL**	**# Template copies / mL**	**# Template copies / mL**
Autumnalis	3.10×10^8*	1.56×10^8	1.56×10^8*	1.56×10^8
Bratislava	8.81×10^7	8.81×10^7	8.81×10^7	8.81×10^7
Canicola	7.41×10^7	7.41×10^7	7.41×10^7	7.41×10^7
Grippotyphosa	1.84×10^8	7.39×10^8*	1.84×10^8	6.93×10^8*
Hardjo	7.91×10^8	3.54×10^8	7.91×10^8	3.54×10^8
Icterohaemorrhagiae	3.22×10^8	6.43×10^8	3.22×10^8	6.43×10^8
Pomona	8.01×10^7	8.21×10^7	8.01×10^7	8.21×10^7

Two individual trials, testing both an H_2O and a urine dilution set, were performed for each serovar. The lowest concentrations that returned accurate serovar- and genus-level identifications in each dilution set, for each serovar over both trials, is given above. Only two of the lowest concentrations listed, for both dilution sets, differed between serovar- and genus-level identifications (indicated by an *).

The distribution of concentration data for both the H_2O and urine dilution sets at the serovar identification level was found to be non-normal (Shapiro-Wilk, $p < 0.05$). Concentrations in the H_2O dilution set had a median of 2.47×10^8, a range of 7.41×10^7–1.65×10^9, and 10th and 90th percentiles of 7.71×10^7–1.22×10^9 (all units are in organisms/mL). Concentrations in the urine dilution set had a median of 3.18×10^8, a range of 7.41×10^7–7.79×10^8, and 10th and 90th percentiles of 7.81×10^7–7.59×10^8. Likewise, the concentration data sets recorded for genus-level IDs were found to be inconsistent with a normal distribution (Shapiro-Wilk, $p < 0.05$). H_2O dilution set concentrations had a median of 1.56×10^8, but the same range, and 10th and 90th percentiles as the H_2O dilution concentrations for the serovar-level. Urine dilution set concentrations had a median of 2.51×10^8, a range of 7.41×10^7–7.39×10^8, and 10th and 90th percentiles of 7.81×10^7–7.16×10^8. The Mann-Whitney U test showed no significant difference between the lowest concentrations for the H_2O and urine dilution sets at either the serovar-level (U: 84.5, $p > 0.05$) or the genus-level (U: 82.5, $p > 0.05$).

4.4.4 *Leptospira* Concentration Needed for MALDI Detection and Identification Compared to the *Leptospir*a Concentration in Typical Leptospirosis-Positive Canine Urine Samples Received by the VDL

VDL leptospirosis qPCR test results were reviewed for the period ranging from January 1, 2016 to July 1, 2018. During this period, a total of 142 qPCR-positive canine leptospirosis cases diagnosed from urine specimens were identified. C_T values for these cases ranged from 17.68 to 44.27, with the majority falling in the range of 30 to 40. The overall average value was 35.18 (SD: 5.11). The standard curve was used to estimate an average concentration of 1.18 x 10^3 organisms/mL, which corresponded to the 35.18 C_T value. Recall that the average concentrations needed for accurate MALDI identification to the serovar- and genus-levels using H_2O dilution values were 3.55 x 10^8 organisms/mL and 3.33 x 10^8 organisms/mL, respectively. The C_T value estimated to correspond to these concentrations would fall within the range of 16.44–19.63. This is much lower than the average C_T value found for positive canine samples submitted to the VDL. This means that most positive canine samples received by the VDL will not have the concentration required for MALDI detection. Therefore, it was determined that MALDI cannot feasibly be used for the detection and identification of *Leptospira* serovars directly from canine urine as there is a difference of 5 $\log_{10}$ organisms/mL between the lowest concentration needed for MALDI serovar detection and that of typical leptospirosis-positive canine urine samples received for testing. There is also a difference of 5 $\log_{10}$ organisms/mL between the lowest concentration needed for MALDI *Leptospira* genus-level detection and that of typical leptospirosis-positive VDL samples. This indicates that MALDI could also not be used for *Leptospira* detection alone directly from leptospirosis-positive samples. Consequently, MALDI measurement of leptospirosis-positive canine urine samples is not an appropriate method to use for routine leptospirosis epi-surveillance and diagnosis.

4.4.5 Filtration

4.4.5.1 Enclosed syringe filters

There is not a consensus on filter pore size required to trap *Leptospira*. A .22 μm pore size is reported to trap most bacteria (Machtejevas and Unger 2008). While some authors report that this also holds true for *Leptospira* (Zuerner 2005), others have reported that *Leptospira* can actually pass through this pore size (Ausubel et al. 1992, Fain 1994, Walker 2008). It has generally been thought that leptospires could easily pass through a.45 μm diameter pore due to their diameter, reported as approximately 0.1 μm (Levett 2001, Mohammed et al. 2011). However, a few studies have noted that some filters with a .45 μm pore size have been found to trap *Leptospira* (Hawkins 2007, Walker 2008).

Three enclosed syringe filters (an EMD Millipore™ Millex™ filter with a .22 μm pore PES membrane, an EMD Millipore™ Millex™ filter with a 22 μm pore Durapore™ PVDF membrane, and a PALL Acrodisc® Filter with a 45 μm pore Versapor® membrane), were tested to learn whether they could be used to concentrate leptospires in patient urine samples with similar results. Percent transmittance (%T)

measurements were used to compare bacterial concentration between the filtrates produced by each filter type. Before filtration, serovar Grippotyphosa stock culture was adjusted to a 39.4 %T. The adjusted culture was run through the PES membrane filter with a .22 μm pore size, which produced a filtrate with an 89 %T. The higher %T of the filtrate compared to that of the adjusted stock culture suggested that a large portion of *Leptospira* in the had been trapped by the filter. The Durapore™ PVDF Membrane filter, with a .22 μm pore size, produced a filtrate with an 87.2 %T. This result, similar to that of the PES membrane filter, suggested that it, too, had trapped a large portion of the *Leptospira.* The Versapor® membrane filter, with a larger pore size of .45 μm, produced a filtrate with an 83.8 %T. The lower %T of this filtrate was expected since the larger pore size meant that fewer *Leptospira* should be trapped in the filter and a larger portion of the bacteria should be found in the filtrate. However, this %T was still much higher than expected. Since this type of filter is sealed, recovery of *Leptospira* trapped by the membrane is difficult to impossible. Overall, these results indicated that none of the three filters offered a good way to concentrate *Leptospira,* and they were not tested further.

4.4.5.2 Membrane filters

Two Millipore membrane filters, each with a different pore size, were tested to determine their potential for concentrating leptospires from patient samples. The filters were tested using serovar Icterohaemorrhagiae stock culture adjusted to 39.8 %T. The membrane filter with a .22 μm pore size produced filtrate with an 88.5 %T. This value suggested that the filter had trapped a large portion of the *Leptospira* in the culture. Filtrate produced by the membrane with a 0.45 μm pore size had an 85.3 %T. This value indicated that the increase in pore size did not allow an appreciably larger number of leptospires to pass through. Sterile water run through this same filter yielded a filtrate with a 93.5 %T. This result implied that flushing the filter with water via this method would not dislodge a sufficient number of trapped bacteria needed for MALDI detection. An attempt to collect *Leptospira* trapped by the filter via manually scraping the filter with a loop was unsuccessful. Spectrophotometric measurements of these scrapings mixed with filtrate from the prior run had a 97.2 %T, which indicated that this was not an effective method for collecting these *Leptospira*. Given these results, it was determined that use of these membrane filters was not a feasible way to concentrate *Leptospira*.

4.4.5.3 Amicon centrifugal filters

Amicon Ultra-4 Centrifugal Filters with an Ultracel-10 membrane were tested for their ability to concentrate *Leptospira* in two trials. The first trial was performed using three samples of serovar Icterohaemorrhagiae culture stock, each at a different concentration. The first sample was serovar culture adjusted to a starting concentration equivalent to 31.6 %T (4.34 x 10^9 organisms/mL). The second and third samples were a 1:2 dilution (2.17x 10^9 organisms/mL) and a 1:4 dilution, (1.09 x 10^9 organisms/mL), created from the same starting concentration. After samples were processed, 1 μL of each concentration

was deposited onto the MALDI target in duplicate. Both target spots for the first sample produced spectra that correctly matched to the MSP for serovar Icterohaemorrhagiae. Identification scores were 2.40 and 2.53, which indicated a secure genus identification and a highly probable species identification (Bruker Daltonics, Inc. 2012). While one of the sample spots for the 1:2 dilution was correctly identified as serovar Icterohaemorrhagiae with a score of 2.31, the other sample spot matched incorrectly to serovar Hardjo, with a score of 2.39. However, both of these scores also fall within the range, from 2.30 to 3.00, which indicates a highly probable species match. Two of the three sample spots for the 1:4 dilution correctly matched to serovar Icterohaemorrhagiae with identification scores of 2.21 and 2.30. The third spot incorrectly identified as serovar Hardjo, with a score of 2.40. Interestingly, the spots for the 1:2 and 1:4 dilutions that were correctly matched had lower identification scores than the spots that were incorrectly identified as serovar Hardjo. The 1:2 dilution spot that was correctly identified had a score of 2.21, which falls within a lower score classification range, from 2.00 to 2.29. This range indicates a secure genus, and probable species, identification. While it would not be surprising to see identification scores decrease with higher dilutions, no clear inverse relationship was seen between these two variables. Samples were also used for qPCR to determine the C_T value for each dilution. The starting concentration had a C_T value of 15.40, the 1:2 dilution had a C_T value of 15.62, and the 1:4 dilution had a C_T value of 17.10. All three C_T values fell outside the linear range for this qPCR assay.

Recall that the final concentrations for the second Amicon filter trial reflect the original concentration used in each filter x 20. Results for this trial were disappointing. Direct transfer sample spots for 3 of the 4 dilutions resulted in 'No Organism ID Possible.' The sample spot that did result in an identification, for the 1:12,000 dilution was found to be invalid. The C_T value for this dilution was lower, indicating a higher concentration, than the C_T values for the 1:2000 (1.01×10^8 organisms/mL) and 1:8000 dilutions (2.52×10^7 organisms/mL). This revealed that the 1:12,000 dilution had been created incorrectly and all results from this dilution were not accurate. The lack of MALDI identification for the 1:2000 and 1:8000 dilution TE sample spots may have been due to insufficient sample used for the extraction prep. There could have been so little sample left in the tube after direct spotting that the amount of material left in the tubes for these dilutions was inadequate for MALDI detection. The average concentration for leptospirosis-positive canine urine specimens submitted to the VDL is 1.18×10^3 organisms/mL. If one of these Amicon filters was used to concentrate 2 mL of a sample with this concentration, the resulting filtrate would have a concentration of approximately 2.36×10^4 organisms/mL. This concentration is still several logs below that needed for accurate MALDI *Leptospira* detection and identification (3.55×10^8 organisms/mL). The volume and concentration of the filtrate produced by these filters is dependent on the volume of the sample, the amount of time the filter is centrifuged, and the volume of wash reagent used. A longer centrifuge time results in a smaller volume of filtrate. Though a shorter centrifugation time than that used in this trial might be used, the volume and concentration of the filtrate would still be limited by typical canine urine specimen volumes submitted for leptospirosis testing. The volume available for use in these

filters would most often fall below 2 mL. Considering the cost of the filters, longer sample preparation time, and the insufficient volume and concentration of the filtrate, the decision was made to forego any further trials with these filters.

4.4.6 Differential Centrifugation and Extended Direct Transfer Sample Preparation

The two-fold serial dilutions created for the differential centrifugation trial ranged in concentration from 1.05×10^9 organisms/mL to 3.40×10^7 organisms/mL. Two sample spots per dilution were deposited onto the MALDI target using the direct spot technique. The two spots for the starting dilution returned correct MALDI serovar-level identifications at scores of 2.53 and 2.42. One of the 1:2 dilution spots returned a correct Autumnalis ID with a score of 2.01, while the second spot returned No Peaks Found. Similarly, one of the two 1:4 dilution spots returned a correct ID with a score of 2.23, while the second spot returned No Peaks Found. The 1:8 through 1:32 sample spots all returned No Peaks Found. To note, the second sample spot for each dilution used an extended direct transfer sample prep, which involved depositing formic acid directly over this second sample spot. Since the first sample spots for the 1:2 and 1:4 dilutions returned the correct ID, and the second sample spots did not, this could suggest that the extended direct transfer approach yields fewer correct identifications and lower sensitivity than the standard direct transfer approach. However, further testing would be needed to draw this conclusion.

Two additional spots for each of the starting, 1:2, and 1:4 dilutions were spotted onto the target using the sandwich deposition technique. The first spot for the starting sample identified correctly with a score of 2.03, while the second spot ID'd as *L. bratislava,* with a score of 1.98. Both spots for the 1:2 and 1:4 dilutions returned No Peaks Found. These six spots are not enough to draw a general conclusion about this deposition technique. However, given the poor results obtained with this method, it was not used in further trials.

This trial was conducted early in the study. It tested not only a differential centrifugation sample processing technique, but also tested the direct transfer and extended direct transfer deposition techniques. In retrospect, a differential centrifugation sample process should have been retested using the EtOH/FA sample prep method once it was known that the EtOH/FA method yielded superior MALDI ID results compared to direct transfer. Results from this trial cannot be used to draw conclusions because we cannot know if the lack of additional MALDI IDs was due to insufficient concentration by the differential centrifugation process or due to direct transfer deposition of the samples.

4.4.7 Overall Outcome of *Leptospira* Sample Concentration Efforts via Filtration and Differential Centrifugation

Ultimately, the filtration and differential centrifugation experiments did not result in successful concentration of *Leptospira* samples. The types and sizes of the syringe and membrane filters tested trapped most of the *Leptospira* in the samples and did not produce filtrate that could be used for further

MALDI testing. The Amicon filter assay was found to require more steps, time, and resources, than would prove beneficial for real-time *Leptospira* testing. Additionally, the cost of using these filters would partially negate be low cost-per-test benefit of using the MALDI. Future work may discover an optimum filter type or set of differential centrifugation steps that would allow for sufficient *Leptospira* sample concentration, but that is beyond the scope of this study.

4.4.8 Other Sample Preparation and Deposition Methods

Results for the direct transfer and ethanol/formic acid extraction sample preparation methods and their variations were given in the Results section of the previous Specificity chapter. Briefly, the direct transfer method often yielded sample spots which resulted in no MALDI peaks found, and therefore; no identification. Variations on this method did not improve MALDI ID. Variations on the standard EtOH/FA prep, which included the addition of one or two washes, tried with both UHPLC-grade H_2O and ethanol, and with mixed prep and sandwich deposition techniques, resulted in fewer identifiable sample spots, compared to the sample spots prepped with the standard EtOH/FA method. Once this was discovered, the standard EtOH/FA method was chosen to prep samples for the remainder of the MALDI trials.

4.5 DISCUSSION

With a lack of practical tools for active leptospirosis surveillance and a predicted increase in case number, there is an urgent need for the development of new, rapid, and economical diagnostic and surveillance tools for the prevention and control of this disease. This study investigated MALDI-TOF MS as a potential tool for use in the detection and identification of *Leptospira* in leptospirosis-positive canine urine samples. We hypothesized that MALDI had the required sensitivity to detect *Leptospira* in clinical canine urine samples at concentrations typically seen in positive cases. An average of 3.33×10^8 organisms per mL was found to be required for MALDI detection at the genus-level. Comparably, the qPCR assay used by the VDL for leptospirosis diagnostics requires a minimum of approximately 37 organisms per mL. The minimum average number increased to 3.55×10^8 organisms per mL for accurate MALDI serovar-level detection. Since all serovars tested belonged to the same species, a minimum number of *Leptospira* required for correct species-level detection was not investigated.

The minimum *Leptospira* concentration found to return an accurate ID in this study was higher than that found in several previous studies (Djelouadji Z, et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Sonthayanon et al. 2019). In the study conducted by Djelouadji Z, et al. (2012) *Leptospira* were quantified using a BioSigma Fast-read 102 apparatus. The authors reported a minimum required concentration of 1×10^5 organisms/mL specifically for *L. interrogans* serovar Copenhageni and *L. biflexa* serovar Patoc. These two serovars were not used in this study, so it is not known whether a lower minimum concentration for these two serovars would have also been found using the methods in the current study. Rettinger et al. (2012) reported a higher minimum required concentration of 1×10^6

organisms/mL for sample preparation, and quantified *Leptospira* using a Petroff-Hausser counting chamber with darkfield microscopy. Calderaro et al. (2014b) also reported a minimum concentration of 1 x 10^6 organisms/mL and quantified *Leptospira* concentration via microscopy count. This year, Sonthayanon et al. (2019) reported that a minimum concentration of 1 x 10^6 CFU/mL was needed to generate detectable and identifiable sample spectra. The authors were able to use colony counts to quantify *Leptospira* because they used LVW agar, an experimental solid agar medium, to grow the *Leptospira* cultures used in their study.

The minimum concentration needed in this study did correspond with that reported by Karcher et al. (2018). In their study, Karcher and colleagues found that a minimum concentration of 1 x 10^8 organisms/mL was needed for successful MALDI sample preparation, which means lower concentrations did not produce sample spectra that allowed for accurate identification. *Leptospira* quantification was performed via Petroff-Hausser counting chamber and darkfield microscopy. It is unknown why the minimal concentration in this study was higher than that of several other studies. Perhaps the differences may be partially due to the method used for *Leptospira* quantification. Unlike the previous studies, the current study quantified *Leptospira* concentration via Qubit measurement of DNA extracted from dilution sample aliquots. However, this difference in quantification method did not result in a difference in minimum concentration between this study and that found by Karcher et al. (2018).

The aim of this work was to determine whether MALDI could detect *Leptospira*-positive samples at the concentrations typically seen in positive canine samples submitted to the VDL. Therefore, *Leptospira*-spiked canine urine samples were included in MALDI testing. Since canine urine samples appropriate for use in this study were limited, the results obtained from serovar + H_2O dilution testing were assessed to determine if they could substitute for results obtained with serovar + urine dilutions. Specifically, results from *Leptospira* qPCR runs and MALDI sensitivity trials were examined for significant differences between the two dilution sets. Negligible differences between results of the two sets would be interpreted to mean that results obtained with H_2O dilution sets could be used to estimate results given by spiked urine sets. The use of H_2O in place of urine dilutions would allow for more MALDI testing that would not be available if testing were constrained to the limited available volume of urine. The C_T value versus log concentration standard curves created for the two types of dilutions showed a no significant difference in qPCR assay efficiency between the two. Therefore, the H_2O dilution curve was used to estimate the concentration for all test dilutions. In MALDI sensitivity trials, accurate identification and ID scores were also similar for the two dilution sets. It is possible that this lack of difference was due to the urine selected for use in the study. As mentioned previously, canine urine samples were selected from leftover VDL specimens that were qPCR-leptospirosis-negative, showed no bacterial growth on culture, contained no visible blood or discoloration, and had little, if any, particulate. The decision was made to use only those samples which met these requirements to help control potential factors that might confound MALDI results for urine dilutions. The idea was to use urine samples that were absent of potentially interfering

constituents or characteristics for initial urine dilution testing. This was to try to exclude factors other than the urine matrix that could account for any differences seen in trial results between the two dilution sets. If specificity and sensitivity trial results suggested that MALDI could be used for *Leptospira* detection in patient samples during routine diagnostic testing, further trials would be run using urine samples that varied in color, turbidity, presence of blood cells or particulate, and presence of other bacteria or fungi, to more closely mimic the types of patient samples submitted for testing.

The results of this study showed that MALDI requires a concentration approximately 10^5-fold higher than that of typical positive samples received by the VDL. As such, most samples that are positive via qPCR will not exhibit a detectable signal via MALDI. To detect *Leptospira* in these samples, specimens would need to be sufficiently concentrated such that the sample material deposited onto the MALDI target contained a high-enough number of leptospires or MALDI detection and/or analyses parameters would need to be adjusted to allow for detection at lower concentrations. Various concentration methods were explored to determine if they allowed for sufficient sample concentration. Based on literature that reported *Leptospira*'s small diameter may allow it to pass through filters with a .22 μm pore size (Ausubel et al. 1992, Walker 2008), tests using enclosed syringe and membrane filters with a .22 μm pore size were performed. The goal was to learn whether these filters would trap any larger bacteria and contaminating particles present in a sample while allowing *Leptospira* to pass through and be collected in the filtrate. In this way, sample constituents that could interfere with MALDI analysis would be removed. The filtrate could then be centrifuged to concentrate and pellet *Leptospira*. After removal of the supernatant, the pellet material could be deposited directly onto the MALDI target. However, percent transmittance measurements of the filtrate in these tests indicated that the majority of the *Leptospira* in tested samples were trapped by the filter. Filters with a .45 μm pore size were also tested because several sources had reported using this pore size to separate *Leptospira* from other, larger bacteria (Zuerner 2005, Haake and Zückert 2015). Percent transmittance readings of the filtrates from these tests showed that these filters also trapped most of the *Leptospira* in the samples, despite their larger pore size. Further review of the literature found that these results were not completely unexpected. Several sources have found that *Leptospira* recovery from filters with a 0.45 μm pore size can be low, depending on the type of filter material used (Hawkins 2007, Walker 2008, Kaboosi et al. 2010). In this study, the three filter materials tested, (PES, Durapore™ PVDF, and Versapor®), all retained the majority of *Leptospira* in each sample. It has been suggested that the spiracle shape of leptospires might make them more likely to be retained by the filter and so harder to recover via filtrate (Walker 2008, Hawkins 2013). Attempts to dislodge *Leptospira* trapped in the filters by backflushing, (syringe and membrane filters), and scraping the surface of the filters, (membrane filters), proved unsuccessful. Amicon centrifugal filters were also tested as a method to concentrate *Leptospira*. While these filters were able to increase dilution concentrations by 20x, (when starting with 2 mL of sample), this was still below the concentration needed for MALDI detection. Given the typical volume of urine submitted for canine leptospirosis testing and the need to

allocate part of this volume for other tests, 2 mL is on the high end of the sample volume we could expect to have for use in this type of filter. Sample volume affects the volume of filtrate produced, which in turn affects the concentration factor. Therefore, sample concentration with these filters would be limited to an approximate concentration factor of 20x, which would not increase the concentration of most samples enough for MALDI detection.

The differential centrifugation protocol tested was also unable to concentrate samples sufficiently for MADI detection. Perhaps a different combination of centrifugation speeds and time spun might yield more promising results. However, given the level to which typical leptospirosis-positive patient samples must be concentrated for MALDI detection, it is unlikely to be achieved via differential centrifugation alone. Since these methods could not sufficiently concentrate *Leptospira* for MALDI detection, further testing with a wider variety of urine specimens was not pursued. The future development of a method by which patient samples could be sufficiently concentrated for MALDI detection would present the possibility of incorporating MALDI into routine leptospirosis diagnostics and surveillance. To explore this option, additional urine sample testing would be required to learn how samples that span the spectrum of specimen variability might affect MALDI results.

A few studies have explored the average concentrations of *Leptospira* shed in the urine of infected dogs. A study by Rojas et al. (2010), examined the average number of leptospires shed per mL of urine in canine samples submitted to the University Veterinary Hospital (UVT) in University College Dublin, and in samples collected via free catch from local dog sanctuaries. Thirty-four out of the 498 UVT samples were positive for leptospirosis and had an average concentration of 6.22×10^4 leptospires per mL. Three out of 27 samples collected from the sanctuaries were leptospirosis-positive, with an average concentration of 2.3×10^4 leptospires per mL. Both averages fall below the minimum detectable concentrations found in this study and reported in previous studies. More recently, Barragan and colleagues (2017) reported that infected dogs shed an average of 1.4×10^2 leptospires per mL of urine. Again, this suggests that the average concentration of leptospires in positive canine patient samples is too low for detection without additional sample processing or software analysis.

While the various sample preparation and concentration techniques applied in this study did not allow for MALDI detection of *Leptospira* at typical concentrations seen with positive canine samples, other techniques may offer improved detection. Several studies have used affinity-based methods for preconcentration and enhanced detection of selected bacteria in samples (Lin et al. 2005, Wu et al. 2012, Chiu et al. 2014). AnchorChip technology has also been used to achieve a more homogenous deposition of the sample and matrix on the target and improve spectral profile reproducibility (Zhang et al. 2004). Optimization of spectrum processing parameters may also offer improved detection, which, in turn, could boost sensitivity. Whether a combination of these or other techniques can improve sensitivity enough to detect the relatively low concentrations of *Leptospira* in positive samples, and whether such methods can be practically incorporated into routine diagnostics, remains to be seen. Though preliminary results have

found that incorporating MALDI into routine diagnostics for canine leptospirosis testing is not practical, a MALDI leptospirosis reference library can still prove useful in research examining epidemiological aspects of this reemerging zoonotic disease.

Chapter 5: SUMMARY AND CONCLUSIONS

5.1 OVERVIEW

The studies described in this thesis explored the sensitivity and specificity of the MALDI-TOF MS platform for the detection and identification of select *Leptospira* strains to the serovar level. The purpose of this work was to learn whether MALDI could serve as a compliment to routine canine leptospirosis diagnostics and as a method for regular serovar typing and epi-surveillance. Results showed the potential for serovar-level specificity using custom Main Spectrum Profiles designed with alternative MSP creation parameters. However, MALDI was found to lack the necessary sensitivity to detect *Leptospira* at concentrations typically found in leptospirosis-positive patient urine specimens.

5.2 SPECIFICITY

Specificity work for this thesis examined MALDI sample preparation and deposition techniques, alternative methods for creating serovar MSPs, and blind-coded trials of custom MSPs. The standard EtOH/FA protein extraction sample prep method was, unsurprisingly, found to produce the highest-quality sample spectra, resulting in the highest number of accurate identifications.

Serovar MSPs created using the manufacturer's default settings and recommended MSP creation guidelines were found to lack sufficient specificity. Deviation from these conventional guidelines when creating subsequent MSPs resulted in greater MSP specificity and allowed for serovar-level identification during real-time classification**.** Remember that real-time classification is performed using Bruker RTC software, in which the acquisition of sample spectra and identification of the sample occur almost simultaneously. This is the MALDI workflow process used most often in diagnostic microbiology laboratories. The two divergent creation parameters that offered the greatest MSP specificity were inclusion of raw spectra from more than one subculture in each MSP, and the alteration of MSP Maximum Desired Peak Number to > 70. It has been reported that factors such as the age at which an organism is analyzed, the environmental conditions present (e.g. humidity) during sample prep and analysis, and analysis performed at different times in the maintenance cycle of the MALDI system can cause some variability in an organism's mass spectra (Arnold et al. 1999, Garbis et al. 2005, Semmes et al. 2005, West-Norager et al. 2007, Croxatto et al. 2012, Ueda et al. 2015, Sauget et al. 2017). It was thought that taking these factors into account when creating the MSPs would result in more robust MSPs that could tolerate intraserovar mass spectra variability. Therefore, raw spectra from three different subcultures were included in each MSP. These spectra represented three biological replicates and included *Leptospira* grown in different lots of growth media, harvested at different growth phases (from approximately 1 week to 2 weeks after passing), prepped for MALDI analysis using different lots of reagents, grown and

prepped during different seasons with accompanying different atmospheric conditions, and acquired and saved to the MALDI at various time spans after a laser adjustment/cleaning.

Final MSPs selected to represent each serovar also contained some peaks that had < 75% frequency. Bruker guidelines recommends that all MSP peaks have ≥ 75% peak frequency. These guidelines were written for the creation of MSPs used to identify organisms to the genus- and species- levels. For these studies, it was concluded that the identification of *Leptospira* at the sub-species level in real-time may require *Leptospira* MSPs to include peaks with < 75 % frequency.

Examination of MSP peak lists found that characteristic peak patterns of presence/absence and intensity levels reported for certain serovars in previous studies, (Rettinger et al. 2012, Calderaro et al. 2014b, Karcher et al. 2018, Sonthayanon et al. 2019), generally did not apply to those same serovars in this study. These peak patterns appear to be relative to the combination of serovars and strains used to create a *Leptospira* MSP library. While identified peak patterns may be useful for particular labs, it is thought that these patterns cannot be widely applied for serovar identification purposes by other labs which use a different combination of serovars and strains for MSP database creation.

5.2.1 CPT Analysis

It was found that CPT classification models could accurately differentiate between *Leptospira* serovars with recognition capabilities (RC) ranging from 90.48 to 100% and with cross validations (CVs) ranging from 85.71 to 98.41%.

The blind-coded trials of the custom *Leptospira* MSP library, performed using the Biotyper Realtime Classification (RTC) software, found that MSP specificity at the serovar-level ranged from 99 to 100% for the H_2O set, and from 97 to 100% for the urine dilution set. These results supported the hypothesis that MALDI-TOF MS can be used to identify *Leptospira* isolates to the serovar-level in real-time. The average score for correct first matches for the H_2O dilution set was 2.58, while that for the urine dilution set was 2.48. Slightly lower scores for the urine set was not unexpected. Though the urine specimens selected to create the dilutions had no visually obvious characteristics that could interfere with MALDI identification and were additionally filtered prior to use, it is possible that some specimens retained constituent particles that could have affected the scores. Regardless, these two average scores are within the same score range interpreted as highly reliable to the species level. These findings supported the hypothesis that MALDI identification specificity for serovar-spiked urine specimens will be the same as that for culture-only specimens.

There are a few factors that should be kept in mind regarding the assessment of MSP performance in this study. The first is that Bruker guidance for score interpretation covers only genus and species levels. This thesis focused on MALDI identification at the serovar level, for which Bruker provides no guidance. As a quick review, the Bruker guidelines create three tiers of scores which indicate different levels of identification specificity. In the first tier, scores between 1.70 and 1.99 indicate a probable genus-level

identification. The second tier contains scores that range between 2.00 and 2.39 and indicate a secure genus ID and probable species ID. Scores that range between 2.30 and 3.00, in the third tier, indicates a probable species-level ID. Scores ≤ 1.69 are considered unreliable (Bruker Daltonics, Inc. 2012). In these guidelines, higher scores reflect a higher probability that a sample has been accurately identified to the species-level. So, what range of scores would indicate that a MALDI identification is probably accurate to the serovar-level? Contrary to what some may think, confidence in an identification at this lower taxonomic level would not necessarily require scores to fall within the high end of the third tier. Instead, it is recommended that score interpretation guidelines should be reassessed and that different score ranges may be needed for particular organisms. This is because, for identification at a lower taxonomic level, peaks present in only one or some members of a species become more important in the pattern matching process (Culebras et al. 2018). These peaks may represent proteins that are not highly conserved among all members of a species. As mentioned previously in the Specificity chapter, several studies have proposed less restrictive score interpretation for subtype-level identification using the Bruker Microflex LT system (Fedorko et al. 2012, Szabados et al. 2012, Theel et al. 2012, Pranada et al. 2016). The main concern for lowering the score tier thresholds for particular organisms is a potential increase in the number of false positives of these organisms. However, the validation process a clinical lab performs to incorporate a new diagnostic test into routine use can also be used to determine whether lower score thresholds increase the number of true positives without increasing the number of false positives.

The second factor to consider is that sample spot scores can differ from run to run, as mentioned in the Discussion section of Chapter 3. This has been observed in both real-time runs using the Biotyper RTC software and in off-line runs with saved raw test spectra using the MBT Compass Explorer software. These different scores for the same spot are generally very similar and usually in the same score interpretation tier. When trying to identify a sample at the genus or species levels, this does not pose a problem. However, when trying to differentiate between closely related organisms at a sub-species level, such as serovars Bratislava and Grippotyphosa, this can prove challenging and lead to incorrect serovar identification. To counteract this, further work to improve MSP specificity is needed so to increase the differences between identification scores for the first, second, and subsequent serovar matches. Another difference in MALDI identification results that may be seen for sample spots in duplicate runs is a result of No Peaks Found in a first run, and a correct first match on the second run, or vice versa. So, a sample spot that appears to lack sufficient analyte information for spectrum creation in one run, based on an NPF result, may return a correct result in a second run. In these cases, the NPF result obtained in the first run is not a reflection of MALDI sensitivity nor is it indicative of an MSP's ability to identify a particular organism at a particular concentration. It is also not an incorrect match. The change in identification can occur for several reasons. The primary reason is the uneven co-crystallization of the matrix and analyte within target sample spots can create areas referred to as "hot spots," which produce higher ion yields upon laser irradiation and consequently, stronger signals. Likewise, "cold spots" may form, which

generate lower numbers of ions upon irradiation and, therefore; create a weaker signal (Basile and Mignon 2016, O'Rourke et al. 2018). When a sample is analyzed, the MALDI laser strikes various areas within the sample spot, desorbing and ionizing the matrix and analyte in those areas. The ions generated in each laser pulse/strike are recorded. Acquisition parameters dictate peak criteria required for a peak to be recorded. For example, peaks must meet or exceed a certain signal intensity to be accepted. Furthermore, a certain number of accepted peaks per sample spot must be met for a mass spectrum to be produced for that spot. If the required number of accepted peaks for a sample spot is not met within a set number of laser shots, the result for that sample spot will be No Peaks Found, and the MALDI will move on to analyze the next sample spot. Thus, if the laser happens to strike several "cold spots" within a sample spot that do not produce the required number of acceptable peaks, no spectrum will be produced. If that sample spot is run again, and the laser happens to strike "hot spots" that do produce the required number of acceptable peaks, then a sample spectrum will be created and the pattern matching process can return a list of possible identifications.

The third factor to be considered is that sample spots that return NPF results both initially and on re-run may also not reflect MALDI sensitivity. NPF results can be due to a variety of reasons, including uneven matrix/analyte co-crystallization as mentioned above, user error in sample prep or deposition onto the MALDI target, issues with the reagents or supplies used in sample prep, the need for MALDI laser recalibration or system cleaning, choice of matrix not ideal for sample type, and parameter settings (Garbis et al. 2005, Semmes et al. 2005, West-Norager et al. 2007, Croxatto et al. 2012, Ueda et al. 2015, Sauget et al. 2017). In the MALDI sensitivity trials for this project, raw test spectra for each serovar were acquired and saved to be used in various trials. A few sets of test spectra returned NPF for lower dilutions (higher concentrations) than expected, regardless of the number of re-runs performed or the particular combination of serovar MSPs being tested. If these test spectral sets for particular serovars were the only spectra used to test the MSPs created for those particular serovars, I might have concluded that the MALDI required a higher concentration of those serovars for detection. However, subsequent sets of saved test spectra for those same serovars returned accurate genus or serovar-level IDs for some of the dilutions which had yielded NPF results for the first test spectra sets. When assessing MALDI sensitivity for particular organisms, NPF results must be considered in context.

The limitations of and recommendations suggested for the specificity study were previously addressed in the Discussion section of Chapter 3. Briefly, the two study limitations were the small number of replicates used and the limited generalizability of study results. A greater number of replicates may have found that some MSPs other than those chosen for inclusion in the final *Leptospira* MSP library may better represent the serovars. The number of replicates may also not have been sufficient to detect a significant difference results given by H_2O versus urine dilution sets. Neither *Leptospira* serovars and strains other than those used for MSP creation nor clinical isolates were used for testing. These were not available for this study. Since the *Leptospira* MSPs produced for this study were created with serially-

passaged laboratory strains, these MSPs may lack peaks representing certain proteins that are only generated in vivo during an active infection. Therefore, it is unknown whether these MSPs contain a sufficient number of unique peaks that will allow for the accurate identification of clinical isolates. To address this issue, it is recommended that additional testing of the MSP library be performed using local clinical isolates and additional serovars and strains. Additionally, raw spectra acquired from local clinical isolates of the same serovar and strain types as those used to create the MSP library should be added to or replace some of the current raw spectra the respective MSPs. This would most likely result in more robust and specific MSPs. Additional testing and editing of revised MSPs would then be needed to increase MSP specificity such that the MALDI identification scores for first, second, and subsequent matches differ enough such that multiple runs of the same sample spots return the same identifications at the same ranks.

5.3 SENSITIVITY

The overall goal of the sensitivity work for this project was to determine whether MALDI could detect *Leptospira* in typical concentrations found in leptospirosis-positive canine urine specimens, based on estimated concentrations of positive specimens submitted to the VDL over a 2.5-year time period. The sensitivity of the MALDI platform was found to be much lower than that required for *Leptospira* detection. As discussed in the previous chapter, the minimum concentrations for the H_2O dilution set that returned accurate serovar- and genus- level IDs were 3.55 x 10^8 organisms/mL and 3.33 x 10^8 organisms/mL, respectively. These concentrations are higher than that reported for several previous studies (Djelouadji et al. 2012, Rettinger et al. 2012, Calderaro et al. 2014b, Sonthayanon et al. 2019) that found minimum required concentrations between 10^5 and 10^6 organisms/mL. The use of different *Leptospira* quantitation methods between these previous studies and the current one may have contributed to this concentration difference. However, Karcher et al. (2018) also used a *Leptospira* quantitation method that differed from the one used in this study yet reported a similar minimum concentration of 10^8 organisms/mL. This suggests that factors other than quantitation method contributed to the difference in minimum concentration found. MALDI type was not a contributing factor because the current and previous studies all used the same Bruker MALDI-TOF MS model. However, potential variations in MALDI software settings used for spectra acquisition and detection may have played a role. Differences in sample preparation methods may also be a contributing factor. While all previous studies used a protein extraction sample prep method similar to that used in this study, the methods were not exactly the same. Variations included reagent volumes, centrifugation speeds and times applied to specimens, and in the Sonthayanon et al. (2019) study, the use of a different matrix. A recent study performed by Mörtelmaier et al. (2019) also reported a higher minimum bacterial concentration needed for the same MALDI-TOF MS model used in in this study. They found that *Parabacteroides distasonis* and *Staphylococcus aureus* samples required a minimum concentration of 10^8 viable cell counts/mL, while *Enterococcus faecalis* and

Bifidobacterium breve samples required 10^9 viable cell counts/mL to consistently return accurate MALDI identifications with scores > 1.70. These organisms were cultured on agar media and quantified via colony counts. While the Mörtelmaier et al. (2019) study shows that a minimum concentration of ≥ 10^8 cells/mL has been found to be required for the accurate MALDI detection of some other organisms, differences between the organisms and methods used in that and the current study offers no further insight into the minimum concentration difference between this and prior *Leptospira*/MALDI studies.

Analysis of MALDI sensitivity trial results found no significant difference in the lowest average concentrations that returned accurate serovar- and genus-level identifications between H_2O and urine dilution sets. This was interpreted to mean that urine matrix of the quality used in this study's dilution sets does not appear to lower or increase MALDI sensitivity for *Leptospira* detection.

Comparison of the estimated C_T value corresponding to the lowest *Leptospira* concentration required for MALDI detection and the average C_T value of leptospirosis-positive canine urine specimens submitted to the VDL revealed that MALDI sensitivity is too low to be used for routine canine leptospirosis diagnostics & epi-surveillance. Therefore, the hypothesis that whole-cell MALDI-TOF MS has the sensitivity to detect *Leptospira* serovars in the urine of clinically affected canines was not supported by this study. The various filtration and differential centrifugation concentration techniques and sample preparation methods tested did not result in sufficient sample concentration. However, concentration techniques, modifications to MALDI software parameters, or sample preparation methods not addressed in these studies may yet offer a method that can close this sensitivity gap.

The limitations of the sensitivity study overlap those of the specificity study. This includes the need to test additional *Leptospira* serovars and strains as well as clinical isolates against the *Leptospira* MSP library. As mentioned above, some of the proteins produced and detected in serially-passaged serovars may differ in abundance from the proteins produced and detected in wild-type serovars from clinical cases. These differences may be reflected in the mass spectra of these organisms and therefore affect the specificity and sensitivity of the MSP library.

Variations in the quality of acquired and saved serovar mass spectra used for testing were seen to affect both specificity and sensitivity results. Some of the initially-acquired mass test spectra returned more NPF results over a greater range of dilutions tested than subsequently-acquired test spectra. This was most likely due to the refinement of the MALDI sample prep process and workflow over time, and an increasing familiarity with the Bruker MALDI software (e.g. quality assessment of acquired test spectra, repeat analysis of target spots as appropriate to obtain the spectral data from "hot spots"). While the acquisition and testing of new test spectra against the current *Leptospira* MSP library is not expected to return wildly different results from that found in the current study, additional results obtained from new spectra would increase the sample size and may further refine the minimum concentration found to be required for accurate MALDI *Leptospira* ID.

Convenience-based acquisition of urine submitted to the VDL for other diagnostic testing means that the urine used was not representative of the range of physiological states that can be found in the total canine population. Therefore, the non-significant difference found between the results produced by the H_2O versus urine dilution sets may not apply to results produced by urine dilution sets created with urine specimens that are more representative of the canine population as a whole. If future work using more representative urine specimens did find a significant difference between results from H_2O and urine dilution sets, conclusions drawn from H_2O results could not be accurately extrapolated to urine specimens. Acquisition of a more representative pool of urine specimens, including from dogs of different ages, different breeds, and from a variety of geographic locations around the US is recommended to address this issue.

Future work on this subject may yet discover a method(s) for sufficiently increasing MALDI sensitivity for *Leptospira* in serovar suspensions such that typical *Leptospira* concentrations in positive patient specimens could be detected but warrants further investigation. Different matrices may offer increased sensitivity. The physicochemical properties of a matrix determines the degree of co-crystallization that occurs with a sample, the amount of sample/matrix desorption when exposed to the MALDI laser and the extent of sample ionization (Zenobi and Knochenmuss 1998, Dreisewerd 2003). These factors influence a MALDI's limit of detection (Dreisewerd 2014). In addition to commonly used matrices, such as α-cyano-4-hydroxycinnamic acid (HCCA), sinapinic acid (SA) and 2,5-dihydroxybenzoic acid (DHB) (Dreisewerd 2014), a growing variety of matrices have been developed to improve the detection and identification of analytes that have been difficult to evaluate using a common matrix. These include not only other organic acids, but inorganic matrices (Sunner et al. 1995, Fleith et al. 2014), ionic liquids matrices (Li and Gross 2004), and proton stripping matrices (Sun et al. 2007), the depth and breadth of which is beyond the scope of this thesis.

Bruker Daltonics, Inc. has introduced a variety of systems and kits which might allow for sufficient concentration of extracted Leptospira proteins such that identification of leptospirosis-positive canine urine specimens is possible. However, these options are not currently available for the Bruker Microflex LT system and steel targets used in this study. Bruker offers three magnetic bead kits that can be used in protein and peptide isolation and concentration, with each kit enriching a different set of proteins and peptides (Bruker Daltonics, Inc. n.d.). However, use of these kits requires either a manual magnet separator or Bruker's ClinProtTM robotic platform, and a unique combination of MALDI target type, matrix type, and calibration, validation, and acquisition parameter changes. These kits were designed for biomarker discovery and so it is not known whether they would prove useful for whole-cell organism identification and differentiation. Without knowing the proteins and the corresponding masses that can be used to ID and differentiate *Leptospira* serovars, choosing the right combination of bead kit, target, matrix, solvent and additional reagent types would be difficult. Additionally, one cannot use saved spectra and peak lists acquired from previous serovar MALDI runs to determine the proteins and corresponding mass

range that would be most useful. This is because MALDI-TOF MS alone cannot be used to identify proteins based on peak presence/absence. Few peaks in whole-cell MALDI mass spectra can be attributed to just one protein. Often, multiple proteins contribute to spectral peaks and can overlap with peaks from sodium and potassium adducts (Liyanage and Lay Jr. 2006). These or other affinity or chromatographic separation techniques prior to MALDI analysis would need be labor and cost efficient as well as rapid so as not to negate the advantages offered by using MALDI for routine typing.

5.4 CONCLUDING REMARKS

The changing epidemiology of leptospirosis exhibited by a growth in the number of recorded cases, an expansion in geographical distribution, and an increase in the severity of clinical presentations, requires an effective detection and prevention strategy. An important component of any such strategy is an efficient and robust surveillance system. Many countries currently do not have official monitoring programs, and the majority of those that do are limited to human cases reported retrospectively (European Centre for Disease Prevention and Control 2017, Schneider et al. 2017, CDC 2020). This renders any measures taken to address the disease reactive rather than proactive. Implementing routine epi-surveillance that incorporates serovar typing would allow for earlier detection of outbreaks, determination of the probable source, and identification of regionally-predominant serovars. The data collected through surveillance would provide valuable information useful in determining appropriate prevention and control measures. Such data would also prove invaluable for leptospirosis vaccine design.

The advantages MALDI-TOF MS offers in microbial diagnostics to the genus- and species-levels and its widespread use clinical diagnostic laboratories would make it an ideal platform for routine microbial typing. This work has shown that MALDI *Leptospira* serovar-level identification in real-time is possible. While MALDI was found to lack the sensitivity needed to identify *Leptospira* at concentrations found in typical patient specimens, there may yet be a sample concentration or MALDI acquisition method that would allow typing of leptospirosis-positive patient specimens. This work provides a foundation for future studies of real-time *Leptospira* typing by MALDI-TOF MS.

REFERENCES

Adler B, Ballard SA, Miller SJ, Faine S. Monoclonal antibodies reacting with serogroup and serovar specific epitopes on different lipopolysaccharide subunits of *Leptospira interrogans* serovar *pomona*. *FEMS Microbiol Lett*. 1989;47(4):213-218. doi:10.1111/j.1574-6968.1989.tb02385.x

Adler B, de la Peña Moctezuma A. *Leptospira* and leptospirosis. *Vet Microbiol*. 2010;140(3-4):287-296. doi:10.1016/j.vetmic.2009.03.012

Adler B, Faine S. The Genus *Leptospira*. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, eds. *The Prokaryotes*. New York, NY: Springer New York; 2006:294-317. doi:10.1007/0-387-30747-8_10

Adzitey F, Huda N, Ali GRR. Molecular techniques for detecting and typing of bacteria, advantages and application to foodborne pathogens isolated from ducks. *3 Biotech*. 2013;3(2):97-107. doi:10.1007/s13205-012-0074-4

Ahmad F, Shah S, Ahmad FMH. Laboratory diagnosis of leptospirosis. *J Postgrad Med*. 2005;51(3):195-2005.

Ahmed A, Engelberts MFM, Boer KR, et al. Development and validation of a real-time PCR for detection of pathogenic *Leptospira* species in clinical materials. Bereswill S, ed. *PLoS ONE*. 2009;4(9):e7093. doi:10.1371/journal.pone.0007093

Ahmed N, Devi SM, Machang'u RS, et al. Multilocus sequence typing method for identification and genotypic classification of pathogenic *Leptospira* species. *Ann Clin Microbiol Antimicrob*. 2006:10.

Ahmed SA, Sandai DA, Musa S, et al. Rapid diagnosis of leptospirosis by multiplex PCR. *Malays J Med Sci*. 2012;19(3):9-16.

Alatoom AA, Cunningham SA, Ihde SM, et al. Comparison of direct colony method versus extraction method for identification of gram-positive cocci by use of Bruker Biotyper matrix-assisted desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2011;49(8):2868-2873. doi:10.1128/JCM.00506-11

Alexander AD, Gleiser CA, Malnati P, and Yoder H. Observations on the prevalence of leptospirosis in canine populations of the United States. *Am J Hyg*. 1957;65:43-56.

Allan KJ, Biggs HM, Halliday JEB, et al. Epidemiology of leptospirosis in Africa: A systematic review of a neglected zoonosis and a paradigm for 'One Health' in Africa. Zinsstag J, ed. *PLoS Negl Trop Dis*. 2015;9(9):e0003899. doi:10.1371/journal.pntd.0003899

AlMasoud N, Xu Y, Nicolaou N, and Goodacre R. Optimization of matrix assisted desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS) for the characterization of *Bacillus* and *Brevibacillus* species. *Analytica Chimica Acta*. 2014;840:49-57. doi:10.1016/j.aca.2014.06.032

Alton GD, Berke O, Reid-Smith R, et al. Increase in seroprevalence of canine leptospirosis and its risk factors, Ontario 1998–2006. *Can J Vet Res*. 2009;73:167-175.

Ambily R, Mini M, Joseph S, et al. Canine leptospirosis; a seroprevalence study from Kerala, India. *Vet World*. 2013;6(1):42-44. doi:10.5455/vetworld.2013.42-44

Andersen-Ranberg EU, Pipper C, Jensen PM. Global patterns of *Leptospira* prevalence in vertebrate reservoir hosts. *Journal of Wildlife Diseases*. 2016;52(3):468-477. doi:10.7589/2014-10-245

André-Fontaine G. Canine leptospirosis—do we have a problem? *Vet Microbiol*. 2006;117(1):19-24. doi:10.1016/j.vetmic.2006.04.005

André-Fontaine G, Hernandez J. Leptospirosis in dogs. *Point Veterinaire*. 2008;39(284):47+.

Appassakj H, Silpapojakul K, Wansit R, and Woodtayakorn J. Evaluation of the immunofluorescent antibody test for the diagnosis of human leptospirosis. *Am J Trop Med Hyg*. 1995;52:340-343.

Applied Biosystems. *Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR*.; 2008.

Arnold RJ, Karty JA, Ellington AD, and Reilly JP. Monitoring the growth of a bacteria culture by MALDI-MS of whole cells. *Anal Chem*. 1999;71(10):1990-1996. doi:10.1021/ac981196c

Arnold RJ, Karty JA, Reilly JP. Bacterial strain differentiation by mass spectrometry. In: Wilkins CL, Lay Jr. JO, eds. *Identification of Microorganisms by Mass Spectrometry*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2006:181-201. doi:10.1002/0471748641.ch9

Ausubel FM, Brent R, Kingston RE, et al. *Short Protocols in Molecular Biology*. second. New York: John Wiley and Sons; 1992.

Bandara M, Ananda M, Wickramage K, et al. Globalization of leptospirosis through travel and migration. *Global Health*. 2014;10(1):61. doi:10.1186/s12992-014-0061-0

Barbosa AS, Abreu PAE, Vasconcellos SA, et al. Immune evasion of *Leptospira* species by acquisition of human complement regulator C4BP. *Infect Immun*. 2009;77(3):1137-1143. doi:10.1128/IAI.01310-08

Barbuddhe SB, Maier T, Schwarz G, et al. Rapid identification and typing of *Listeria* species by matrix-assisted saser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol*. 2008;74(17):5402-5407. doi:10.1128/AEM.02689-07

Barcellos C, Sabroza PC. The place behind the case: leptospirosis risks and associated environmental conditions in a flood-related outbreak in Rio de Janeiro. *Cad Saúde Pública*. 2001;17(suppl):S59-S67. doi:10.1590/S0102-311X2001000700014

Barr SC, McDonough PL, Scipioni-Ball RL, and Starr JK. Serologic responses of dogs given a commercial vaccine against *Leptospira interrogans* serovar Pomona and *Leptospira kirschneri* serovar Grippotyphosa. *Am J Vet Res*. 2005;66(10):1780-1784. doi:10.2460/ajvr.2005.66.1780

Barragan V, Olivas S, Keim P, and Pearson T. Critical knowledge gaps in our understanding of environmental cycling and transmission of *Leptospira spp.* Schaffner DW, ed. *Appl Environ Microbiol*. 2017;83(19):e01190-17, /aem/83/19/e01190-17.atom. doi:10.1128/AEM.01190-17

Basile F, Mignon RK. Methods and instrumentation in mass spectrometry for the differentiation of closely related microorganisms. In: Demirev P, Sandrin TR, eds. *Applications of Mass Spectrometry in Microbiology*. Cham: Springer International Publishing; 2016:13-50. doi:10.1007/978-3-319-26070-9_2

Batista C, Azevedo S, Alves C, et al. Seroprevalence of leptospirosis in stray dogs from Patos City, state of Paraíba, Brazil. *Braz J Vet Res Anim Sci*. 2004;41:131-136.

Bharti AR, Nally JE, Ricaldi JN, et al. Leptospirosis: a zoonotic disease of global importance. *Lancet Infect Dis*. 2003;3(12):757-771. doi:10.1016/S1473-3099(03)00830-2

Biswas D, Roy S, Vijayachari P, Sugunan AP, Natarajaseenivasan K, Sehgal SC. Comparison of immunoreactive proteins of commonly circulating serogroups of Leptospira in Andaman Islands, India. *Indian J Med Res*. 2005:8.

Blackmore DK, Hathaway SC. The nidality of zoonoses. In: *Proceedings of the Second International Symposium on Veterinary Epidemiology and Economics*. ; 1979:207-213.

Bourassa L, Butler-Wu SM. MALDI-TOF mass spectrometry for microorganism identification. In: *Methods in Microbiology*. Vol 42. Elsevier; 2015:37-85. doi:10.1016/bs.mim.2015.07.003

Bourhy P, Collet L, Lernout T, et al. Human *Leptospira* isolates circulating in mayotte (Indian Ocean) have unique serological and molecular features. *J Clin Microbiol*. 2012;50(2):307-311. doi:10.1128/JCM.05931-11

Bowles M. CVC Highlight: Canine leptospirosis: When to suspect it, how to manage it. http://veterinarymedicine.dvm360.com/cvc-highlight-canine-leptospirosis-when-suspect-it-how-manage-it. Published 2015.

Boyer PH, Almeras L, Plantard O, et al. Identification of closely related Ixodes species by protein profiling with MALDI-TOF mass spectrometry. Munderloh UG, ed. *PLoS ONE*. 2019;14(10):e0223735. doi:10.1371/journal.pone.0223735

Brenner DJ, Kaufmann AF, Sulzer KR, et al. Further determination of DNA relatedness between serogroups and serovars in the family *Leptospiraceae* with a proposal for *Leptospira alexanderi* sp. nov. and four new *Leptospira* genomospecies. *Int J Syst Evol Microbiol*. 1999;49(2):839-858. doi:10.1099/00207713-49-2-839

Brockmann S, Piechotowski I, Bock-Hensley O, et al. Outbreak of leptospirosis among triathlon participants in Germany, 2006. *BMC Infect Dis*. 2010;10(1):91. doi:10.1186/1471-2334-10-91

Brown JA, LeFebvre RB, Pan MJ. Protein and antigen profiles of prevalent serovars of *Leptospira interrogans*. *Infect Immun*. 1991;59(5):1772-1777. doi:10.1128/IAI.59.5.1772-1777.1991

Brown PD, Gravekamp C, Carrington DG, et al. Evaluation of the polymerase chain reaction for early diagnosis of leptospirosis. *J Med Microbiol*. 1995;43(2):110-114. doi:10.1099/00222615-43-2-110

Brown PD, Levett PN. Differentiation of *Leptospira* species and serovars by PCR-restriction endonuclease analysis, arbitrarily primed PCR and low-stringency PCR. *J Med Microbiol*. 1997;46(2):173-181. doi:10.1099/00222615-46-2-173

Bruker Daltonics. Instructions for Use Profiling Kit MB-WCX. https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/InstructionForUse/IFU_223983_223987_MB-WCX_Rev1.pdf. Published n.d.

Bruker Daltonics, Inc. *MALDI Biotyper 3.1 User Manual Rev. 1*.; 2012.

Bruker Daltonik GmbH. *ClinProtools User Manual, Version 2.2.*; 2007.

Bruker Daltonik GmbH. *Instructions for Use Bruker Matrix HCCA, Portioned Rev. C.*; 2015.

Bruker Daltonik GmbH. *MBT Explorer Module User Manual, Rev. A.*; 2016.

Bruker Daltonik GmbH. *Instructions for Use Bruker Bacterial Test Standard Rev. D.*; 2017.

Bruslind L, Oregon State University. *Allied Health Microbiology*. Open Oregon State; 2017. http://library.open.oregonstate.edu/microbiology/chapter/bacteria-cell-walls/.

Cai C-S, Zhu Y-Z, Zhong Y, et al. Development of O-antigen gene cluster-specific PCRs for rapid typing six epidemic serogroups of *Leptospira* in China. *BMC Microbiol*. 2010;10(67). doi:10.1186/1471-2180-10-67

Calderaro A, Arcangeletti M-C, Rodighiero I, et al. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry applied to virus identification. *Sci Rep*. 2014a;4(1):6803. doi:10.1038/srep06803

Calderaro A, Piccolo G, Gorrini C, et al. *Leptospira* species and serovars identified by MALDI-TOF mass spectrometry after database implementation. *BMC Res Notes*. 2014b;7(1):330. doi:10.1186/1756-0500-7-330

Cameron CE. Leptospiral Structure, Physiology, and Metabolism. In: *Leptospira and Leptospirosis*. Vol 387. Current Topics in Microbiology and Immunology. Berlin: Springer; 2015.

Carbonnelle E, Mesquita C, Bille E, et al. MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clin Biochem*. 2011;44(1):104-109. doi:10.1016/j.clinbiochem.2010.06.017

Centers for Disease Control and Prevention. Healthcare Workers, Technical Information for Leptospirosis. https://www.cdc.gov/leptospirosis/health_care_workers/index.html. Published January 7, 2020. Accessed January 29, 2020.

Cepheid USA. *Cepheid Smart Cycler II, Operator Manual.*; 1999.

Cerqueira GM, Picardeau M. A century of *Leptospira* strain typing. *Infect Genet Evol*. 2009;9(5):760-768. doi:10.1016/j.meegid.2009.06.009

Chen SP. Leptospirosis. Case Based Pediatrics for Medical Students and Residents. https://www.hawaii.edu/medicine/pediatrics/pedtext/s06c28.html. Published 2002. Accessed November 21, 2019.

Chen I-C, Hill JK, Ohlemuller R, et al. Rapid range shifts of species associated with high levels of climate warming. *Science*. 2011;333(6045):1024-1026. doi:10.1126/science.1206432

Chiu T-C. Recent advances in bacteria identification by matrix-assisted laser desorption/ionization mass spectrometry using nanomaterials as affinity probes. *Int J Mol Sci*. 2014;15(5):7266-7280. doi:10.3390/ijms15057266

Chow JC, Young DW, Golenbock DT, et al. Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. *J Biol Chem*. 1999;274(16):10689-10692. doi:10.1074/jbc.274.16.10689

Cinco M, Delneri D, Banfi E. Immunodominant antigens recognized by the human immune response to infection by organisms of the species *Leptospira interrogans* serogroup Australis. *FEMS Microbiol Lett*. 1992;89(5):287-298. doi:10.1111/j.1574-6968.1992.tb05008.x

Clark AE, Kaleta EJ, Arora A, and Wolk DM. Matrix-assisted laser desorption ionization-time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin Microbiol Rev*. 2013;26(3):547-603. doi:10.1128/CMR.00072-12

Clark CG, Kruczkiewicz P, Guan C, et al. Evaluation of MALDI-TOF mass spectroscopy methods for determination of *Escherichia coli* pathotypes. *J Microbial Methods*. 2013;94(3):180-191. doi:10.1016/j.mimet.2013.06.020

Clifford RJ, Milillo M, Prestwood J, et al. Detection of bacterial 16S rRNA and identification of four clinically important bacteria by real-time PCR. Willson RC, ed. *PLoS ONE*. 2012;7(11):e48558. doi:10.1371/journal.pone.0048558

Cobo F. Application of MALDI-TOF mass spectrometry in clinical virology: A review. *Open Virol J*. 2013;7(1):84-90. doi:10.2174/1874357920130927003

Corney BG, Colley J, Djordjevic SP, et al. Rapid identification of some *Leptospira* isolates from cattle by random amplified polymorphic DNA fingerprinting. *J Clin Microbiol*. 1993;31(11):2927-2932. doi:10.1128/JCM.31.11.2927-2932.1993

Costa F, Hagan JE, Calcagno J, et al. Global morbidity and mortality of leptospirosis: A systematic review. Small PLC, ed. *PLoS Negl Trop Dis*. 2015;9(9):e0003898. doi:10.1371/journal.pntd.0003898

Croxatto A, Prod'hom G, Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev*. 2012;36(2):380-407. doi:10.1111/j.1574-6976.2011.00298.x

Culebras DE. Application of MALDI-TOF MS in Bacterial Strain Typing and Taxonomy. In: *The Use of Mass Spectrometry Technology (MALDI-TOF) in Clinical Microbiology*. Elsevier; 2018:213-233. doi:10.1016/B978-0-12-814451-0.00015-0

Cullen PA. Global analysis of outer membrane proteins from *Leptospira interrogans* serovar Lai. *Infect Immun*. 2002;70(5):2311-2318. doi:10.1128/IAI.70.5.2311-2318.2002

Cullen PA, Haake DA, Bulach DM, Zuerner RL, Adler B. LipL21 is a novel surface-exposed lipoprotein of pathogenic *Leptospira* species. *Infect Immun*. 2003;71(5):2414-2421. doi:10.1128/IAI.71.5.2414-2421.2003

Cullen PA, Haake DA, Adler B. Outer membrane proteins of pathogenic spirochetes. *FEMS Microbiol Rev*. 2004;28(3):291-318. doi:10.1016/j.femsre.2003.10.004

Davis MA, Evermann JF, Petersen CR, et al. Serological survey for antibodies to *Leptospira* in dogs and raccoons in Washington state. *Zoonoses Public Hlth*. July 2008:436-442. doi:10.1111/j.1863-2378.2008.01137.x

Day MJ, Horzinek MC, Schultz RD, and Squires RA. WSAVA Guidelines for the vaccination of dogs and cats. *J Small Anim Pract*. 2016;57(1):4-8. doi:10.1111/jsap.12431

De Carolis E, Vella A, Vaccaro L, et al. Development and validation of an in-house database for matrix-assisted laser desorption ionization-time of flight mass spectrometry-based yeast identification

ssing a fast protein extraction procedure. *J Clin Microbiol*. 2014a;52(5):1453-1458. doi:10.1128/JCM.03355-13

De Carolis E, Vella A, Vaccaro L, et al. Application of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *J Infect Dev Ctries*. 2014b;8(09):1081-1088. doi:10.3855/jidc.3623

de Castro JR, Salaberry SRS, Souza III MA de, and Lima-Ribeiro AMC. Sorovares de *Leptospira spp.* predominantes em exames sorológicos de caninos e humanos no município de Uberlândia, Estado de Minas Gerais. *Rev Soc Bras Med Trop*. 2011;44(2):217-222. doi:10.1590/S0037-86822011005000012

de la Peña Moctezuma A, Bulach DM, Kalambaheti T, and Adler B. Comparative analysis of the LPS biosynthetic loci of the genetic subtypes of serovar Hardjo: *Leptospira interrogans* subtype Hardjoprajitno and *Leptospira borgpetersenii* subtype Hardjobovis. *FEMS Microbiol Lett*. 1999;177(2):319-326. doi:10.1111/j.1574-6968.1999.tb13749.x

de la Peña Moctezuma A, Bulach DM, Adler B. Genetic differences among the LPS biosynthetic loci of serovars of *Leptospira interrogans* and *Leptospira borgpetersenii*. *FEMS Immunol Med Microbiol*. 2001;31(1):73-81. doi:10.1111/j.1574-695X.2001.tb01589.x

DeMarco ML, Burnham C-AD. Diafiltration MALDI-TOF mass spectrometry method for culture-independent detection and identification of pathogens directly from urine specimens. *Am J Clin Pathol*. 2014;141(2):204-212. doi:10.1309/AJCPQYW3B6JLKILC

Dieckmann R, Helmuth R, Erhard M, and Malorny B. Rapid classification and identification of *Salmonellae* at the species and subspecies levels by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol*. 2008;74(24):7767-7778. doi:10.1128/AEM.01402-08

Dieckmann R, Malorny B. Rapid screening of epidemiologically important *Salmonella enterica* subsp. enterica serovars by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol*. 2011;77(12):4136-4146. doi:10.1128/AEM.02418-10

Dikken H, Kmety E. Chapter VIII Serological Typing Methods of Leptospires. In: *Methods in Microbiology*. Vol 11. Elsevier; 1978:259-307. doi:10.1016/S0580-9517(08)70493-8

Djelouadji Z, Roux V, Raoult D, et al. Rapid MALDI-TOF mass spectrometry identification of *Leptospira* organisms. *Vet Microbiol*. 2012;158(1-2):142-146. doi:10.1016/j.vetmic.2012.01.028

Doungchawee G, Sirawaraporn W, Icksang-Ko A, et al. Use of immunoblotting as an alternative method for serogrouping *Leptospira*. *J Med Microbiol*. 2007;56(5):587-592. doi:10.1099/jmm.0.47143-0

Dreisewerd K. The desorption process in MALDI. *Chem Rev*. 2003;103(2):395-426. doi:10.1021/cr010375i

Dreisewerd K. Recent methodological advances in MALDI mass spectrometry. *Anal Bioanal Chem*. 2014;406(9-10):2261-2278. doi:10.1007/s00216-014-7646-6

Eddabra R, Prevost G, Scheftel JM. Rapid discrimination of environmental *Vibrio* by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Microbiol Res*. 2012;167:226-230. doi:10.1016/j.micres.2011.09.002

Egli A, Tschudin-Sutter S, Oberle M, et al. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) based typing of extended-spectrum β-Lactamase producing *E. coli* – a novel tool for real-time outbreak investigation. Proost P, ed. *PLoS ONE*. 2015;10(4):e0120624. doi:10.1371/journal.pone.0120624

Ellinghausen Jr. HC, McCullough WG. Nutrition of *Leptospira* pomona and growth of 13 other serotypes: fractionation of oleic albumin complex and a medium of bovine albumin and polysorbate 80. *Am J Vet Res*. 1965;26:45-51.

Ellis WA, Hovind-Hougen K, Moller S, and Birch-Andresen A. Morphological changes upon subculturing of freshly isolated strains of *Leptospira interrogans* serovar Hardjo. *Zentralbl Bakteriol Mikrobiol Hyg A*. 1983;255:323-335.

Ellis WA. Animal Leptospirosis. In: Adler B, ed. *Leptospira and Leptospirosis*. Vol 387. Current Topics in Microbiology and Immunology. New York: Springer; 2015:111-112.

Espinal P, Seifert H, Dijkshoorn L, et al. Rapid and accurate identification of genomic species from the *Acinetobacter baumannii* (Ab) group by MALDI-TOF MS. *Clin Microbiol Infec*. 2012;18(11):1097-1103. doi:10.1111/j.1469-0691.2011.03696.x

European Centre for Disease Prevention and Control. *Leptospirosis-Annual Epidemiological Report*.; 2016. https://www.ecdc.europa.eu/en/publications-data/leptospirosis-annual-epidemiological-report-2016-2014-data. Accessed January 30, 2017.

Evangelista KV, Coburn J. *Leptospira* as an emerging pathogen: a review of its biology, pathogenesis and host immune responses. *Future Microbiol*. 2010;5(9):1413-1425. doi:10.2217/fmb.10.102

Everard COR, Carrington D, Korver H, and Everard JD. Leptospires in the marine toad (*Bufo marinus*) on Barbados. *J Wildl Dis*. 1988;24(2):334-338. doi:10.7589/0090-3558-24.2.334

Faine S, Adler B, Bolin C, and Perolat P, eds. *<Leptospira</i> and Leptospirosis*. 2nd ed. Melbourne, Australia: Medisci; 1999.

Faisal SM, Yan W, Chen C-S, et al. Evaluation of protective immunity of *Leptospira* immunoglobulin like protein A (LigA) DNA vaccine against challenge in hamsters. *Vaccine*. 2008;26(2):277-287. doi:10.1016/j.vaccine.2007.10.029

Fangous M-S, Mougari F, Gouriou S, et al. Classification algorithm for subspecies identification within the *Mycobacterium abscessus* species, based on matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Land GA, ed. *J Clin Microbiol*. 2014;52(9):3362-3369. doi:10.1128/JCM.00788-14

Farr RW. Leptospirosis. *Clin Infect Dis*. 1995;21(1):1-6.

Fedorko DP, Drake SK, Stock F, and Murray PR. Identification of clinical isolates of anaerobic bacteria using matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Eur J Clin Microbiol Infect Dis*. 2012;31(9):2257-2262. doi:10.1007/s10096-012-1563-4

Feigin RD. Human leptospirosis from immunized dogs. *Ann Intern Med*. 1973;79(6):777. doi:10.7326/0003-4819-79-6-777

Feresu SB, Steigerwalt AG, Brenner DJ. DNA relatedness of *Leptospira* strains isolated from beef cattle in Zimbabwe. *Int J Syst Bact*. 1999;49(3):1111-1117. doi:10.1099/00207713-49-3-1111

Ferreira L, Sanchez-Juanes F, Gonzalez-Avila M, et al. Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2010;48(6):2110-2115. doi:10.1128/JCM.02215-09

Field CB, Barros V, Stocker TF, et al., eds. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press; 2012. doi:10.1017/CBO9781139177245

Fleith C, Cantel S, Subra G, et al. Laser desorption ionization mass spectrometry of peptides on a hybrid CHCA organic–inorganic matrix. *Analyst*. 2014;139(15):3748-3754. doi:10.1039/C3AN01465G

Fouts DE, Matthias MA, Adhikarla H, et al. What makes a bacterial species pathogenic? : Comparative genomic analysis of the genus *Leptospira*. Small PLC, ed. *PLoS Negl Trop Dis*. 2016;10(2):e0004403. doi:10.1371/journal.pntd.0004403

Foxman B, Zhang L, Koopman JS, et al. Choosing an appropriate bacterial typing technique for epidemiologic studies. *Epidemiol Perspect Innov*. 2005;2(1):10. doi:10.1186/1742-5573-2-10

Freiwald A, Sauer S. Phylogenetic classification and identification of bacteria by mass spectrometry. *Nat Protoc*. 2009;4(5):732-742. doi:10.1038/nprot.2009.37

Galloway RL, Levett PN. Application and validation of PFGE for serovar identification of *Leptospira* clinical isolates. Lukehart S, ed. *PLoS Negl Trop Dis*. 2010;4(9):e824. doi:10.1371/journal.pntd.0000824

Garbis S, Lubec G, Fountoulakis M. Limitations of current proteomics technologies. *J Chromatogr A*. 2005;1077(1):1-18. doi:10.1016/j.chroma.2005.04.059

Gautam R, Wu C-C, Guptill LF, et al. Detection of antibodies against *Leptospira* serovars via microscopic agglutination tests in dogs in the United States, 2000–2007. *JAVMA Sci Rep*. 2010;237(3):293-298.

Ge M-C, Kuo A-J, Liu K-L, et al. Routine identification of microorganisms by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: Success rate, economic analysis, and clinical outcome. *J Microbiol Immunol*. 2017;50(5):662-668. doi:10.1016/j.jmii.2016.06.002

Gendron K, Christe A, Walter S, et al. Serial CT features of pulmonary leptospirosis in 10 dogs. *Veterinary Record*. 2014;174(7):169-169. doi:10.1136/vr.102046

Goering RV. Molecular Typing Techniques: State of the Art. In: Tang Y-W, Stratton CW, eds. *Advanced Techniques in Diagnostic Microbiology*. Boston, MA: Springer US; 2013:239-261. doi:10.1007/978-1-4614-3970-7_13

Goldstein JE, Zhang L, Borror CM, et al. Culture conditions and sample preparation methods affect spectrum quality and reproducibility during profiling of *Staphylococcus aureus* with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Lett Appl Microbiol*. 2013;57(2):144-150. doi:10.1111/lam.12092

Goldstein RE. Canine leptospirosis. *Vet Clin N Am-Small*. 2010;40(6):1091-1101. doi:10.1016/j.cvsm.2010.07.008

Goris MGA, Boer KR, Duarte TATE, et al. Human leptospirosis trends, the Netherlands, 1925–2008. *Emerg Infect Dis*. 2013;19(3):371-378. doi:10.3201/eid1903.111260

Greene CE, Sykes JE, Moore GE, et al. Leptospirosis. In: Greene CE, ed. *Infectious Diseases of the Dog and Cat*. 4th ed. St. Louis: Saunders Elsevier; 2012:431-447.

Greenlee JJ, Alt DP, Bolin CA, et al. Experimental canine leptospirosis caused by *Leptospira interrogans* serovars pomona and bratislava. *Am J Vet Res*. 2005;66(10):1816-1822. doi:10.2460/ajvr.2005.66.1816

Guerra MA. Leptospirosis. *JAVMA*. 2009;234(4).

Guerra MA. Leptospirosis: Public health perspectives. *Biologicals*. 2013;41(5):295-297. doi:10.1016/j.biologicals.2013.06.010

Haake DA, Chao G, Zuerner RL, et al. The leptospiral major outer membrane protein LipL32 is a lipoprotein expressed during mammalian infection. *Infect Immun*. 2000;68(4):2276-2285. doi:10.1128/IAI.68.4.2276-2285.2000

Haake DA, Suchard MA, Kelley MM, et al. Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer. *J Bacteriol*. 2004;186(9):2818-2828. doi:10.1128/JB.186.9.2818-2828.2004

Haake DA, Matsunaga J. *Leptospira*: a spirochaete with a hybrid outer membrane: Leptospiral outer membrane. *Mol Microbiol*. July 2010:no-no. doi:10.1111/j.1365-2958.2010.07262.x

Haake DA, Levett PN. Leptospirosis in Humans. In: Adler B, ed. *Leptospira and Leptospirosis*. Vol 387. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015:65-97. doi:10.1007/978-3-662-45059-8_5

Haake DA, Zückert WR. The Leptospiral Outer Membrane. In: Adler B, ed. *Leptospira and Leptospirosis*. Vol 387. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015:187-221. doi:10.1007/978-3-662-45059-8_8

Hajduk J, Matysiak J, Kokot ZJ. Challenges in biomarker discovery with MALDI-TOF MS. *Clinica Chimica Acta*. 2016;458:84-98. doi:10.1016/j.cca.2016.04.033

Hansell C. Enter the matrix. *Nature Milestones in Mass Spectrometry*. 2015;12(S1):17-17. doi:10.1038/nmeth.3527

Harkin KR, Gartrell CL. Canine leptospirosis in New Jersey and Michigan: 17 cases (1990-1995). *J Am Anim Hosp Assoc*. 1996;32(6):495-501.

Hartskeerl PA, Terpstra WJ. Leptospirosis in wild animals. *Vet Q*. 1996;18(sup3):149-150. doi:10.1080/01652176.1996.9694722

Hartskeerl RA, Smythe LD. The Role of Leptospirosis Reference Laboratories. In: Adler B, ed. *Leptospira and Leptospirosis*. Vol 387. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015:273-288. doi:10.1007/978-3-662-45059-8_11

Hauk P, Guzzo CR, Ramos HR, et al. Structure and calcium-binding activity of LipL32, the major surface antigen of pathogenic *Leptospira* sp. *J Mol Bio*. 2009;390(4):722-736. doi:10.1016/j.jmb.2009.05.034

Hawkins IR. Detecting *Leptospira* in Water: Evaluation of a Proposed Method [thesis]. University of Hawai'i; 2007.

Hawkins I, Walker M, Evensen C, et al. Detecting *Leptospira* in Water: Preliminary Results from a Regional Collaboration. In: ; 2013:1. http://cals.arizona.edu/region9wq/pdf/LeptoConferencePoster_Final.pdf.

He P, Sheng Y-Y, Shi Y-Z, et al. Genetic diversity among major endemic strains of *Leptospira interrogans* in China. *BMC Genomics*. 2007;8(1):204. doi:10.1186/1471-2164-8-204

Heaton P, Patel R. Mass Spectrometry Applications in Infectious Disease and Pathogens Identification. In: *Principles and Applications of Clinical Mass Spectrometry*. Elsevier; 2017:93-114. doi:10.1016/B978-0-12-816063-3.00004-9

Hensley T. Leptospirosis in Cattle. https://www.texvetpets.org/article/leptospirosis-in-cattle/. Published November 2016. Accessed November 26, 2019.

Hoffmeister B, Peyerl-Hoffman G, Pischke S, et al. Differences in clinical manifestations of imported versus autochthonous leptospirosis in Austria and Germany. *Am J Trop Med Hyg*. 2010;83(2):326-335. doi:10.4269/ajtmh.2010.10-0040

Holland D, Wilkes JG, Ralli F, et al. Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption ionization with time-of-flight mass spectrometry. *Rapid Commun Mass Sp*. 1996;10(10):1227-1232.

Holland RD, Duffy CR, Rafii F, et al. Identification of bacterial proteins observed in MALDI TOF mass spectra from whole cells. *Anal Chem*. 1999;71(15):3226-3230. doi:10.1021/ac990175v

Honoré AH, Thorsen M, Skov T. Liquid chromatography–mass spectrometry for metabolic footprinting of co-cultures of lactic and propionic acid bacteria. *Anal Bioanal Chem*. 2013;405(25):8151-8170. doi:10.1007/s00216-013-7269-3

Hosseini S, Martinez-Chapa SO. Principles and Mechanism of MALDI-ToF-MS Analysis. In: *Fundamentals of MALDI-ToF-MS Analysis*. Applied Sciences and Technology. Singapore: Springer Singapore; 2017:1-19. doi:10.1007/978-981-10-2356-9_1

Hsu S-H, Lo Y-Y, Tung J-Y, et al. Leptospiral outer membrane lipoprotein LipL32 binding on Toll-like receptor 2 of renal cells as determined with an atomic force microscope. *Biochem*. 2010;49(26):5408-5417. doi:10.1021/bi100058w

Hu Y-Y, Cai J-C, Zhou H-W, et al. Rapid detection of porins by matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Front Microbiol*. 2015;6. doi:10.3389/fmicb.2015.00784

Hubbert WT, Schotts EB. Leptospirosis in kennel dogs. *JAVMA*. 1966;148:1152-1159.

Íñigo M, Coello A, Fernández-Rivas G, et al. Direct identification of urinary tract pathogens from urine samples, combining urine screening methods and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Munson E, ed. *J Clin Microbiol*. 2016;54(4):988-993. doi:10.1128/JCM.02832-15

Johnson RC, Harris VG. Differentiation of pathogenic and saprophytic leptospires. *J Bacteriol*. 1967;94(1):27-31. doi:10.1128/JB.94.1.27-31.1967

Jorge S, Kremer FS, Oliveira NR de, et al. Whole-genome sequencing of *Leptospira interrogans* from southern Brazil: genetic features of a highly virulent strain. *Mem Inst Oswaldo Cruz*. 2018;113(2):80-86. doi:10.1590/0074-02760170130

Josten M, Reif M, Szekat C, et al. Analysis of the matrix-assisted laser desorption ionization time-of-flight mass spectrom of *Staphylococcus aureus* identifies mutations that allow differentiation of the main clonal lineages. *J Clin Microbiol*. 2013;51(6):1809-1817. doi:10.1128/JCM.00518-13

Jouglard SDD, Simionatto S, Seixas FK, et al. Nested polymerase chain reaction for detection of pathogenic leptospires. *Can J Microbiol*. 2006;52(8):747-752. doi:10.1139/w06-027

Kaboosi H, Razavi MR, al Sadat Noohi A. Efficiency of filtration technique for isolation of leptospires from surface waters: Role of different membranes with different pore size and materials. *Afr J Microbiol Res*. 2010;4(9):671-676.

Kalambaheti T, Bulach DM, Rajakumar K, and Adler B. Genetic organization of the lipopolysaccharide O-antigen biosynthetic locus of *Leptospira borgpetersenii* serovar Hardjobovis. *Microb Pathogenesis*. 1999;27(2):105-117. doi:10.1006/mpat.1999.0285

Karas M, Bachmann D, Hillenkamp F. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal Chem*. 1985;57(14):2935-2939. doi:10.1021/ac00291a042

Karas M, Bachmann D, Bahr U, and Hillenkamp F. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int J Mass Spectrom Ion Processes*. 1987;78:53-68. doi:10.1016/0168-1176(87)87041-6

Karcher D, Grenfell RC, Moreno AM, et al. Identification of pathogenic and nonpathogenic *Leptospira* species of Brazilian isolates by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Braz J Microbiol*. 2018;49(4):900-908. doi:10.1016/j.bjm.2018.03.005

Kawaguchi L, Sengkeopraseuth B, Tsuyuoka R, et al. Seroprevalence of leptospirosis and risk factor analysis in flood-prone rural areas in Lao PDR. *Am J Trop Med Hyg*. 2008;78(6):957-961. doi:10.4269/ajtmh.2008.78.957

Kemp BM, Winters M, Monroe C, Barta JL. How much DNA is lost? Measuring DNA loss of STR length fragments targeted by the PowerPlex 16 System using the Qiagen MinElute Purification kit. *Human Biology Open Access Pre-Prints*. 2014. http://digitalcommons.wayne.edu/humbiol_preprints/61.

Ketterlinus R, Hsieh S-Y, Teng S-H, et al. Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. *BioTechniques*. 2005;38:S37-S40.

Kim Y, Park KG, Lee K, and Park Y-J. Direct identification of urinary tract pathogens from urine samples using the Vitek MS system based on matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Ann Lab Med*. 2015;35(4):416-422. doi:10.3343/alm.2015.35.4.416

Klaasen HLBM, Molkenboer MJCH, Vrijenhoek MP, and Kaashoek MJ. Duration of immunity in dogs vaccinated against leptospirosis with a bivalent inactivated vaccine. *Vet Microbiol*. 2003;95(1-2):121-132. doi:10.1016/S0378-1135(03)00152-4

Klaasen HLBM, Adler B. Recent advances in canine leptospirosis: focus on vaccine development. *VMRR*. June 2015:245. doi:10.2147/VMRR.S59521

Kmety E, Dikken H. *Classification of the Species Leptospira interrogans and History of Its Serovars*. Groningen, the Netherlands: University Press Groningen; 1993.

Ko AI, Goarant C, Picardeau M. *Leptospira:* the dawn of the molecular genetics era for an emerging zoonotic pathogen. *Nat Rev Microbiol*. 2009;7(10):736-748. doi:10.1038/nrmicro2208

Köhling HL, Bittner A, Muller K-D, et al. Direct identification of bacteria in urine samples by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and relevance of defensins as interfering factors. *J Med Microbiol*. 2012;61(3):339-344. doi:10.1099/jmm.0.032284-0

Koizumi N, Nakajima C, Harunari T, et al. A new loop-mediated isothermal amplification method for rapid, simple, and sensitive detection of *Leptospira spp.* in urine. *J Clin Microbiol*. 2012;50(6):2072-2074. doi:10.1128/JCM.00481-12

Koizumi N, Muto MM, Akachi S, et al. Molecular and serological investigation of *Leptospira* and leptospirosis in dogs in Japan. *J Med Microbiol*. 2013;62(Pt_4):630-636. doi:10.1099/jmm.0.050039-0

Kralik P, Ricchi M. A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. *Front Microbiol*. 2017;8. doi:10.3389/fmicb.2017.00108

Krishnamurthy T, Ross PL. Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells. *Rapid Commun Mass Sp*. 1996;10:1992-1996.

Kuehl B, Marten S-M, Bischoff Y, et al. MALDI-ToF mass spectrometry–multivariate data analysis as a tool for classification of reactivation and non-culturable states of bacteria. *Anal Bioanal Chem*. 2011;401(5):1593-1600. doi:10.1007/s00216-011-5227-5

Lacroix C, Gicquel A, Sendid B, et al. Evaluation of two matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) systems for the identification of *Candida* species. *Clin Microbiol Infec*. 2014;20(2):153-158. doi:10.1111/1469-0691.12210

Lafetá BN, de Castro EC, da Silva N. Protein and antigen profiles of *Leptospira interrogans* serovar Hardjo. *Cienc Rural*. 2009;39(9):2539-2543. doi:10.1590/S0103-84782009000900024

Langston CE, Heuter KJ. Leptospirosis. *Vet Clin N Am-Small*. 2003;33(4):791-807. doi:10.1016/S0195-5616(03)00026-3

LaRocque RC, Breiman RF, Ari MD, et al. Leptospirosis during Dengue Outbreak, Bangladesh. *Emerg Infect Dis*. 2005;11(5):766-769. doi:10.3201/eid1105.041212

Lartigue M-F, Hery-Arnaud G, Haguenoer E, et al. Identification of *Streptococcus agalactiae* Isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2009;47(7):2284-2287. doi:10.1128/JCM.00175-09

Lartigue M-F. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry for bacterial strain characterization. *Infect Genet Evol*. 2013;13:230-235. doi:10.1016/j.meegid.2012.10.012

Lau CL, Smythe LD, Craig SB, and Weinstein P. Climate change, flooding, urbanization and leptospirosis: fueling the fire? *Trans R Soc Trop Med Hyg*. 2010;104(10):631-638. doi:10.1016/j.trstmh.2010.07.002

Leone I, Chirillo MG, Raso T, et al. Phenotypic and genotypic characterization of *Pseudomonas aeruginosa* from cystic fibrosis patients. *Eur J Clin Microbiol Infect Dis*. 2008;27(11):1093-1099. doi:10.1007/s10096-008-0551-1

Leshem E, Segal G, Barnea A, et al. Travel-related leptospirosis in Israel: A nationwide study. *Am J Trop Med Hyg*. 2010;82(3):459-463. doi:10.4269/ajtmh.2010.09-0239

Leuschner RGK, Beresford-Jones N, Robinson C. Difference and consensus of whole cell *Salmonella enterica* subsp. enterica serovars matrix-assisted laser desorption/ionization time-of-flight mass spectrometry spectra. *Lett Appl Microbiol*. 2003;38(1):24-31. doi:10.1046/j.1472-765X.2003.01436.x

Levett PN. Leptospirosis. *Clin Microbiol Rev*. 2001;14(2):296-326. doi:10.1128/CMR.14.2.296-326.2001

Levett PN. Usefulness of serologic analysis as a predictor of the infecting serovar in patients with severe leptospirosis. *Clin Infect Dis*. 2003;36(4):447-452. doi:10.1086/346208

Levett PN. Detection of pathogenic leptospires by real-time quantitative PCR. *J Med Microbiol*. 2005;54(1):45-49. doi:10.1099/jmm.0.45860-0

Levett PN. *Leptospira*. In: Jorgensen JH, ed. *Manual of Clinical Microbiology*. 11th ed. Washington, DC: ASM Press; 2015a:916-923.

Li W, Sun E, Wang Y, et al. Rapid identification and antimicrobial susceptibility testing for urinary tract pathogens by direct analysis of urine samples using a MALDI-TOF MS-based combined protocol. *Front Microbiol*. 2019;10:1182. doi:10.3389/fmicb.2019.01182

Li YL, Gross ML. Ionic-liquid matrices for quantitative analysis by MALDI-TOF mass spectrometry. *J Am Soc Mass Spectrom*. 2004;15(12):1833-1837. doi:10.1016/j.jasms.2004.08.011

Life Technologies Corporation. *Real-Time PCR Handbook_2012.Pdf*. Life Technologies Corporation; 2012.

Lin Y-S, Tsai P-J, Weng M-F, Chen Y-C. Affinity capture using vancomycin-bound magnetic nanoparticles for the MALDI-MS analysis of bacteria. *Anal Chem*. 2005;77(6):1753-1760. doi:10.1021/ac048990k

Lin X, Zhao J, Qian J, et al. Identification of immunodominant B- and T-cell combined epitopes in outer membrane lipoproteins LipL32 and LipL21 of *Leptospira interrogans*. *Clin Vaccine Immunol*. 2010;17(5):778-783. doi:10.1128/CVI.00405-09

Lin X, Sun A, Ruan P, et al. Characterization of conserved combined T and B cell epitopes in *Leptospira interrogans* major outer membrane proteins OmpL1 and LipL41. *BMC Microbiol*. 2011;11(1):21. doi:10.1186/1471-2180-11-21

Lindtner-Knific R, Vergles-Rataj A, Vlahović K, et al. Prevalence of antibodies against *Leptospira sp.* in snakes, lizards and turtles in Slovenia. *Acta Vet Scand*. 2013;55(1):65. doi:10.1186/1751-0147-55-65

Liyanage R, Lay Jr. JO. An Introduction to MALDI-TOF MS. In: Wilkins CL, Lay JO, eds. *Identification of Microorganisms by Mass Spectrometry*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2006:39-60. doi:10.1002/0471748641.ch3

Llanes A, Restrepo CM, Rajeev S. Whole genome sequencing allows better understanding of the evolutionary history of *Leptospira interrogans* serovar Hardjo. Stevenson B, ed. *PLoS ONE*. 2016;11(7):e0159387. doi:10.1371/journal.pone.0159387

Loureiro AP, Martins G, Thomé S, and Lilenbaum W. Laboratorial diagnosis of animal leptospirosis* Diagnóstico laboratorial de leptospirose animal. *RBCV*. 2013;20(3):119-126. doi:10.4322/rbcv.2014.071

Lunn KF. Overview of Leptospirosis. In: *The Merck Veterinary Manual*. Kenilworth, NJ, USA: Merck & Co., Inc.; 2019. https://www.merckvetmanual.com/generalized- conditions/leptospirosis/leptospirosis-in-dogs.

Luo D, Xue F, Ojcius DM, et al. Protein typing of major outer membrane lipoproteins from Chinese pathogenic *Leptospira spp.* and characterization of their immunogenicity. *Vaccine*. 2009;28(1):243-255. doi:10.1016/j.vaccine.2009.09.089

Machtejevas E, Unger KK. Chromatography. In: von Hagen J, ed. *Proteomics Sample Preparation*. Weinheim, Germany: Wiley-VCH Verlag GmbH and Co. KGaA; 2008:484.

Majetic ZS, Galloway R, Ruzic Sabljic E, et al. Epizootiological survey of small mammals as *Leptospira spp.* reservoirs in Eastern Croatia. *Acta Tropica*. 2014;131:111-116. doi:10.1016/j.actatropica.2013.12.009

Major A, Schweighauser A, Francey T. Increasing Incidence of Canine Leptospirosis in Switzerland. *Int J Environ Res Public Health*. 2014;11(7):7242-7260. doi:10.3390/ijerph110707242

Malmström J, Beck M, Schmidt A, et al. Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature*. 2009;460(7256):762-765. doi:10.1038/nature08184

Mani RJ, Thachil AJ, Ramachandran A. Discrimination of *Streptococcus equi* subsp. *equi* and *Streptococcus equi* subsp. *zooepidemicus* using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J VET Diagn Invest*. 2017;29(5):622-627. doi:10.1177/1040638717702687

Markides K, Graslund A. *Advanced Information on the Nobel Prize in Chemistry 2002*. The Royal Swedish Academy of Sciences; 2002:13. https://www.nobelprize.org/uploads/2018/06/advanced-chemistryprize2002.pdf. Accessed February 5, 2020.

Maroun E, Kushawaha A, El-Charabaty E, et al. Fulminant leptospirosis (Weil's disease) in an urban setting as an overlooked cause of multiorgan failure: a case report. *J Med Case Reports*. 2011;5(1):7. doi:10.1186/1752-1947-5-7

Martin L, Pettit A. Sero-diagnostic de la spirochaetose icterohaemorrhagique. *Bull Mem Soc Med Hop Paris*. 1918;42:672-675.

Maskey M, Shastri JS, Saraswathi K, et al. Leptospirosis in Mumbai: Post-deluge outbreak 2005. *Indian J Med Microbiol*. 2006;24(4):S337-S338. doi:10.4103/0255-0857.29413

Matsunaga J, Barocchi MA, Croda J, et al. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily: Leptospiral Ig-like protein family. *Mol Microbiol*. 2003;49(4):929-946. doi:10.1046/j.1365-2958.2003.03619.x

Matsunaga J, Werneid K, Zuerner RL, et al. LipL46 is a novel surface-exposed lipoprotein expressed during leptospiral dissemination in the mammalian host. *Microbiology*. 2006;152(12):3777-3786. doi:10.1099/mic.0.29162-0

McMichael AJ. Globalization, climate change, and human health. *N Engl J Med*. 2013;368(14):1335-1343. doi:10.1056/NEJMra1109341

Meeyam T, Tablerk P, Petchanok B, et al. Seroprevalence and risk factors associated with leptospirosis in dogs. *Southeast Asian J Trop Med Public Health*. 2006;37(1):148-153.

Meites E, Jay MT, Deresinski S, et al. Reemerging leptospirosis, California. *Emerg Infect Dis*. 2004;10(3):7.

Mellmann A, Müthing J. MALDI-TOF Mass Spectrometry-Based Microbial Identification. In: Tang Y-W, Stratton CW, eds. *Advanced Techniques in Diagnostic Microbiology*. Boston, MA: Springer US; 2013:187-207. doi:10.1007/978-1-4614-3970-7_10

Mencacci A, Monari C, Leli C, et al. Typing of nosocomial outbreaks of *Acinetobacter baumannii* by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*. 2013;51(2):603-606. doi:10.1128/JCM.01811-12

Meri T, Murgia R, Stefanel P, et al. Regulation of complement activation at the C3-level by serum resistant leptospires. *Microb Pathogenesis*. 2005;39(4):139-147. doi:10.1016/j.micpath.2005.07.003

Merien F, Amouriaux P, Perolat P, et al. Polymerase chain reaction for detection of *Leptospira spp.* in clinical samples. *J Clin Microbiol*. 1992;30(9):2219-2224. doi:10.1128/JCM.30.9.2219-2224.1992

Merien F, Portnoi D, Bourhy P, et al. A rapid and quantitative method for the detection of *Leptospira* species in human leptospirosis. *FEMS Microbiol Lett*. 2005;249(1):139-147. doi:10.1016/j.femsle.2005.06.011

Mgode GF, Machang'u RS, Mhamphi GG, et al. *Leptospira* serovars for diagnosis of leptospirosis in humans and animals in Africa: Common *Leptospira* isolates and reservoir hosts. Zinsstag J, ed. *PLoS Negl Trop Dis*. 2015;9(12):e0004251. doi:10.1371/journal.pntd.0004251

Miller MD, Annis KM, Lappin MR, and Lunn KF. Variability in results of the Microscopic Agglutination Test in dogs with clinical leptospirosis and dogs vaccinated against leptospirosis: Canine leptospirosis MAT variability. *J Vet Intern Med*. 2011;25(3):426-432. doi:10.1111/j.1939-1676.2011.0704.x

Minter A, Diggle PJ, Costa F, et al. A model for leptospire dynamics and control in the Norway rat (*Rattus norvegicus*) the reservoir host in urban slum environments. *Epidemics*. 2018;25:26-34. doi:10.1016/j.epidem.2018.05.002

Miotto BA, Guilloux AGA, Tozzi BF, et al. Prospective study of canine leptospirosis in shelter and stray dog populations: Identification of chronic carriers and different *Leptospira* species infecting dogs. Chang Y-F, ed. *PLoS ONE*. 2018;13(7):e0200384. doi:10.1371/journal.pone.0200384

Miraglia F, Matsuo M, Morais ZM, et al. Molecular characterization, serotyping, and antibiotic susceptibility profile of *Leptospira interrogans* serovar Copenhageni isolates from Brazil. *Diagn Microbiol Infect Dis*. 2013;77(3):195-199. doi:10.1016/j.diagmicrobio.2013.08.003

Mohammed H, Nozha C, Hakim K, and Abdelaziz F. *Leptospira:* Morphology, classification and pathogenesis. *J Bacteriol Parasitol*. 2011;02(06). doi:10.4172/2155-9597.1000120

Monis PT, Giglio S, Saint CP. Comparison of SYTO9 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of dye concentration on amplification and DNA melting curve analysis. *Anal Biochem*. 2005;340(1):24-34. doi:10.1016/j.ab.2005.01.046

Moore GE, Guptill LF, Glickman NW, et al. Canine Leptospirosis, United States, 2002–2004. *Emerg Infect Dis*. 2006;12(3):501-503. doi:10.3201/eid1203.050809

Moreno LZ, Miraglia F, Lilenbaum W, et al. Profiling of *Leptospira interrogans* , *L. santarosai* , *L. meyeri* and *L. borgpetersenii* by SE-AFLP, PFGE and susceptibility testing—a continuous attempt at species and serovar differentiation. *Emerg Microbes Infec*. 2016;5(1):1-7. doi:10.1038/emi.2016.16

Morey RE, Galloway RL, Bragg SL, et al. Species-specific identification of *Leptospiraceae* by 16S rRNA gene sequencing. *J Clin Microbiol*. 2006;44(10):3510-3516. doi:10.1128/JCM.00670-06

Morgan J, Bornstein SL, Karpati AM, et al. Outbreak of leptospirosis among triathlon participants and community residents in Springfield, Illinois, 1998. *Clin Infect Dis*. 2002;34(12):1593-1599. doi:10.1086/340615

Mörtelmaier C, Panda S, Robertson I, et al. Identification performance of MALDI-ToF-MS upon mono- and bi-microbial cultures is cell number and culture proportion dependent. *Anal Bioanal Chem*. 2019;411(26):7027-7038. doi:10.1007/s00216-019-02080-x

Muñoz-Zanzi C, Mason M, Encina C, et al. *Leptospira* contamination in household and environmental water in rural communities in southern Chile. *Int J Environ Res Public Health*. 2014;11:6666–6680.

Murray GL. The Molecular Basis of Leptospiral Pathogenesis. In: Adler B, ed. *Leptospira and Leptospirosis*. Vol 387. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015:139-185. doi:10.1007/978-3-662-45059-8_7

Murray PR. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: usefulness for taxonomy and epidemiology. *Clin Microbiol Infec*. 2010;16(11):1626-1630. doi:10.1111/j.1469-0691.2010.03364.x

Musso D, La Scola B. Laboratory diagnosis of leptospirosis: A challenge. *J Microbiol Immunol*. 2013;46(4):245-252. doi:10.1016/j.jmii.2013.03.001

Nahori M-A, Fournié-Amazouz E, Que-Gewirth NS, et al. Differential TLR Recognition of leptospiral Lipid A and lipopolysaccharide in murine and human cells. *J Immunol*. 2005;175(9):6022-6031. doi:10.4049/jimmunol.175.9.6022

Nascimento ALTO, Verjovski-Almeida S, Van Sluys MA, et al. Genome features of *Leptospira interrogans* serovar Copenhageni. *Braz J Med Biol Res*. 2004;37(4):459-477. doi:10.1590/S0100-879X2004000400003

Navarro E, Serrano-Heras G, Castaño MJ, Solera J. Real-time PCR detection chemistry. *Clinica Chimica Acta*. 2015;439:231-250. doi:10.1016/j.cca.2014.10.017

Normand A-C, Cassagne C, Ranque S, et al. Assessment of various parameters to improve MALDI-TOF MS reference spectra libraries constructed for the routine identification of filamentous fungi. *BMC Microbiol*. 2013;13(1):76. doi:10.1186/1471-2180-13-76

O'Rourke MB, Djordjevic SP, Padula MP. The quest for improved reproducibility in MALDI mass spectrometry: Improving reproducibility in MALDI MS. *Mass Spec Rev*. 2018;37(2):217-228. doi:10.1002/mas.21515

Østergaard C, Hansen SGK, Møller JK. Rapid first-line discrimination of methicillin resistant *Staphylococcus aureus* strains using MALDI-TOF MS. *Int J Med Microbiol*. 2015;305(8):838-847. doi:10.1016/j.ijmm.2015.08.002

Palaniappan RUM, Chang Y-F, Chang C-F, et al. Evaluation of lig-based conventional and real time PCR for the detection of pathogenic leptospires. *Mol Cell Probe*. 2005;19(2):111-117. doi:10.1016/j.mcp.2004.10.002

Patel R. MALDI-TOF MS for the diagnosis of infectious diseases. *Clin Chem*. 2015;61(1):100-111. doi:10.1373/clinchem.2014.221770

Patra KP, Choudhury B, Matthias MM, et al. Comparative analysis of lipopolysaccharides of pathogenic and intermediately pathogenic *Leptospira* species. *BMC Microbiol*. 2015;15(1):244. doi:10.1186/s12866-015-0581-7

Pellizzer P, Todescato A, Benedetti P, et al. Leptospirosis following a flood in the Veneto area, north-east Italy. *Ann Ig*. 2006;18(5):453-456.

Pereira MM, Schneider MC, Munoz-Zanzi C, et al. A road map for leptospirosis research and health policies based on country needs in Latin America. *Rev Panam Salud Publica*. 2017;41:1-9. doi:10.26633/RPSP.2017.131

Perez J, Goarant C. Rapid *Leptospira* identification by direct sequencing of the diagnostic PCR products in New Caledonia. *BMC Microbiol*. 2010;10(1):325. doi:10.1186/1471-2180-10-325

Pérez-Flores J, Charruau P, Cedeño-Vázquez R, and Atilano D. Evidence for wild crocodiles as a risk for human leptospirosis, Mexico. *EcoHealth*. 2017;14(1):58-68. doi:10.1007/s10393-016-1196-7

Pérez-Sancho M, Vela AI, García-Seco T, et al. Assessment of MALDI-TOF MS as alternative tool for *Streptococcus suis* identification. *Front Public Health*. 2015;3. doi:10.3389/fpubh.2015.00202

Perolat P, Lecuyer I, Postic D, and Baranton G. Diversity of ribosomal DNA fingerprints of *Leptospira* serovars provides a database for subtyping and species assignation. *Res Microbiol*. 1993;144(1):5-15. doi:10.1016/0923-2508(93)90210-S

Perolat P, Merien F, Ellis WA, and Baranton G. Characterization of *Leptospira* isolates from serovar hardjo by ribotyping, arbitrarily primed PCR, and mapped restriction site polymorphisms. *J Clin Microbiol*. 1994;32(8):1949-1957. doi:10.1128/JCM.32.8.1949-1957.1994

Picardeau M. Diagnosis and epidemiology of leptospirosis. *Med Maladies Infect*. 2013;43(1):1-9. doi:10.1016/j.medmal.2012.11.005

Picardeau M, Bertherat E, Jancloes M, et al. Rapid tests for diagnosis of leptospirosis: Current tools and emerging technologies. *Diagn Microbiol Infect Dis*. 2014;78(1):1-8. doi:10.1016/j.diagmicrobio.2013.09.012

Pijnacker R, Goris MGA, te Wierik MJM, et al. Marked increase in leptospirosis infections in humans and dogs in the Netherlands, 2014. *Euro Surveill*. 2016;21(17):30211. doi:10.2807/1560-7917.ES.2016.21.17.30211

Pranada AB, Schwarz G, Kostrzewa M. MALDI Biotyping for Microorganism Identification in Clinical Microbiology. In: Cramer R, ed. *Advances in MALDI and Laser-Induced Soft Ionization Mass Spectrometry*. Cham: Springer International Publishing; 2016:197-225. doi:10.1007/978-3-319-04819-2_11

Prescott JF, McEwen B, Taylor J, et al. Resurgence of leptospirosis in dogs in Ontario: recent findings. *Can Vet J*. 2002;43:955-961.

Pusch W, Kostrzewa M. Application of MALDI-TOF mass spectrometry in screening and diagnostic research. *Curr Pharm Des*. 2005;11(20):2577-2591. doi:10.2174/1381612054546932

Qiagen. *QIAamp Viral RNA Mini Handbook.Pdf*.; 2018:49.

Que-Gewirth NLS, Ribeiro AA, Kalb SR, et al. A methylated phosphate group and four amide-linked acyl chains in *Leptospira interrogans* Lipid A: The membrane anchor of an unusual lipopolysaccharide that activates TLR2. *J Biol Chem*. 2004;279(24):25420-25429. doi:10.1074/jbc.M400598200

Radl C, Müller M, Revilla-Fernandez S, et al. Outbreak of leptospirosis among triathlon participants in Langau, Austria, 2010. *Wien Klin Wochenschr*. 2011;123(23-24):751-755. doi:10.1007/s00508-011-0100-2

Ralph D, McClelland M, Welsh J, et al. *Leptospira* species categorized by arbitrarily primed polymerase chain reaction (PCR) and by mapped restriction polymorphisms in PCR-amplified rRNA genes. *J Bacteriol*. 1993;175(4):973-981. doi:10.1128/JB.175.4.973-981.1993

Ramadass P, Samuel B, Nachimuthu K. A rapid latex agglutination test for detection of leptospiral antibodies. *Veterinary Microbiology*. 1999;70(1-2):137-140. doi:10.1016/S0378-1135(99)00133-9

Reagan KL, Sykes JE. Diagnosis of canine leptospirosis. *Vet Clin N Am-Small*. 2019;49(4):719-731. doi:10.1016/j.cvsm.2019.02.008

Reil M, Erhard M, Kuijper EJ, et al. Recognition of *Clostridium difficile* PCR-ribotypes 001, 027 and 126/078 using an extended MALDI-TOF MS system. *Eur J Clin Microbiol Infect Dis*. 2011;30(11):1431. doi:10.1007/s10096-011-1238-6

Ren S-X, Fu G, Jiang X-G, et al. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature*. 2003;422(6934):888-893. doi:10.1038/nature01597

Rentko VT, Clark N, Ross LA, Schelling SH. Canine leptospirosis: A retrospective study of 17 cases. *J Vet Intern Med*. 1992;6(4):235-244. doi:10.1111/j.1939-1676.1992.tb00345.x

Rettinger A, Krupka I, Grünwald K, et al. *Leptospira spp.* strain identification by MALDI TOF MS is an equivalent tool to 16S rRNA gene sequencing and multi locus sequence typing (MLST). *BMC Microbiol*. 2012;12(1):185. doi:10.1186/1471-2180-12-185

Robinson AJ, Ramadass P, Lee A, Marshall RB. Differentiation of subtypes within *Leptosira interrogans* serovars Hardjo, Balcanica and Tarassovi, by bacterial restriction-endonuclease DNA analysis (BRENDA. *J Med Microbiol*. 1982;15:331-338.

Rockwood AL, Kushnir MM, Clarke NJ. Rockwood Kushner & Clark_Chapter 17_Mass Spectrometry.pdf. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier; 2018:295-323.e5.

Rodrigues TCS, Santos ALQ, Lima AMC, et al. Anti-*Leptospira spp.* antibodies in *Crotalus durissus collilineatus* kept in captivity and its zoonotic relevance. *Acta Tropica*. 2016;158:39-42. doi:10.1016/j.actatropica.2016.02.006

Rodríguez-Sánchez B, Marín M, Sánchez-Carrillo C, et al. Improvement of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry identification of difficult-to-identify bacteria and its impact in the workflow of a clinical microbiology laboratory. *Diagn Microbiol Infect Dis*. 2014;79(1):1-6. doi:10.1016/j.diagmicrobio.2014.01.021

Rojas P, Monahan AM, Schuller S, et al. Detection and quantification of leptospires in urine of dogs: a maintenance host for the zoonotic disease leptospirosis. *Eur J Clin Microbiol Infect Dis*. 2010;29(10):1305-1309. doi:10.1007/s10096-010-0991-2

Romero EC, Bernardo CC da M, Yasuda PH. Human leptospirosis: a twenty-nine-year serological study in São Paulo, Brazil. *Rev Inst Med trop S Paulo*. 2003;45(5):245-248. doi:10.1590/S0036-46652003000500002

Romero EC, Blanco RM, Galloway RL. Application of pulsed-field gel electrophoresis for the discrimination of leptospiral isolates in Brazil. *Lett Appl Microbiol*. 2009;48(5):623-627. doi:10.1111/j.1472-765X.2009.02580.x

Roqueplo C, Kodjo A, Demoncheaux J, et al. Leptospirosis, one neglected disease in rural Senegal. *J Vet Med Sci*. 2019;5(4):536-544. doi:10.1002/vms3.186

Rosselló GAM, Rodríguez MPG, de Lejarazu Leonardo RO, Domingo AO, Pérez MÁB. New procedure for rapid identification of microorganisms causing urinary tract infection from urine samples by mass spectrometry (MALDI-TOF). *Enferm Infec Micr Cl*. 2015;33(2):89-94. doi:10.1016/j.eimc.2014.02.022

Ruppitsch W. Molecular typing of bacteria for epidemiological surveillance and outbreak investigation / Molekulare Typisierung von Bakterien für die epidemiologische Überwachung und Ausbruchsabklärung. *Die Bodenkultur: Journal of Land Management, Food and Environment*. 2016;67(4):199-224. doi:10.1515/boku-2016-0017

Ryzhov V, Fenselau C. Characterization of the protein subset desorbed by MALDI from whole bacterial cells. *Anal Chem*. 2001;73(4):746-750. doi:10.1021/ac0008791

Sandrin TR, Goldstein JE, Schumaker S. MALDI TOF MS profiling of bacteria at the strain level: A review. *Mass Spectrom Rev*. 2013;32(3):188-217. doi:10.1002/mas.21359

Sauer S, Freiwald A, Maier T, et al. Classification and identification of bacteria by mass spectrometry and computational analysis. Yang C-H, ed. *PLoS ONE*. 2008;3(7):e2843. doi:10.1371/journal.pone.0002843

Sauget M, Valot B, Bertrand X, and Hocquet D. Can MALDI-TOF mass spectrometry reasonably type bacteria? *Trends Microbiol*. 2017;25(6):447-455. doi:10.1016/j.tim.2016.12.006

Savio ML, Rossi C, Fusi P, et al. Detection and identification of *Leptospira interrogans* serovars by PCR coupled with restriction endonuclease analysis of amplified DNA. *J Clin Microbiol*. 1994;32(4):935-941. doi:10.1128/JCM.32.4.935-941.1994

Schneider MC, Leonel DG, Hamrick PN, et al. Leptospirosis in Latin America: exploring the first set of regional data. *Rev Panam Salud Publica*. 2017;41:1. doi:10.26633/RPSP.2017.81

Schuller S, Callanan JJ, Worrall S, et al. Immunohistochemical detection of IgM and IgG in lung tissue of dogs with leptospiral pulmonary haemorrhage syndrome (LPHS). *Comp Immunol Microbiol Infect Dis*. 2015;40:47-53. doi:10.1016/j.cimid.2015.04.002

Šedo O, Sedláček I, Zdráhal Z. Sample preparation methods for MALDI-MS profiling of bacteria. *Mass Spectrom Rev*. 2011;30(3):417-434. doi:10.1002/mas.20287

Seibold E, Maier T, Kostrzewa M, et al. Identification of *Francisella tularensis* by whole-cell of matrix-assisted laser desorption ionization-time of flight mass spectrometry : Fast, reliable, robust, and cost-effective differentiation on species and subspecies levels. *J Clin Microbiol*. 2010;48(4):1061-1069. doi:10.1128/JCM.01953-09

Semenza JC, Suk JE, Estevez V, et al. Mapping climate change vulnerabilities to infectious diseases in Europe. *Environ Health Perspect*. 2012;120(3):385-392. doi:10.1289/ehp.1103805

Semmes OJ, Feng Z, Adam B-L, et al. Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer. *Clin Chem*. 2005;51(1):102-112. doi:10.1373/clinchem.2004.038950

Seng P, Drancourt M, Gouriet F, et al. Ongoing revolution in bacteriology: Routine identification of bacteria by matrix-assisted laser desorption ionization- time of flight mass spectrometry. *Clin Infect Dis*. 2009;49(4):543-551. doi:10.1086/600885

Sessions JK, Greene CE. Canine leptospirosis: Epidemiology, pathogenesis, and diagnosis. *Infect Dis Comp*. 2004;26(8):11.

Shao C, Tian Y, Dong Z, et al. The use of principal component analysis in MALDI-TOF MS - a powerful tool for establishing a mini-optimized proteomic profile. *Am J Biomed Sci*. January 2012:85-101. doi:10.5099/aj120100085

Shimadzu Corporation. *Shimadzu BioSpec Mini-1240 UV-Vis Spectrophotometer User's System Guide*. Kyoto, Japan; 1994.

Silva ÉF, Medeiros MA, McBride AJA, et al. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. *Vaccine*. 2007;25(33):6277-6286. doi:10.1016/j.vaccine.2007.05.053

Singhal N, Kumar M, Kanaujia PK, and Virdi JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front Microbiol*. 2015;6. doi:10.3389/fmicb.2015.00791

Slack A, Symonds M, Dohnt M, and Smythe L. An improved multiple-locus variable number of tandem repeats analysis for *Leptospira interrogans* serovar Australis: a comparison with fluorescent

amplified fragment length polymorphism analysis and its use to redefine the molecular epidemiology of this serovar in Queensland, Australia. *J Med Microbiol*. 2006a;55(11):1549-1557. doi:10.1099/jmm.0.46779-0

Slack AT, Symonds ML, Dohnt MF, Smythe LD. Identification of pathogenic *Leptospira* species by conventional or real-time PCR and sequencing of the DNA gyrase subunit B encoding gene. *BMC Microbiol*. 2006b;6(1):95. doi:10.1186/1471-2180-6-95

Slack A, Symonds M, Dohnt M, et al. Evaluation of a modified Taqman assay detecting pathogenic *Leptospira spp.* against culture and *Leptospira*-specific IgM enzyme-linked immunosorbent assay in a clinical environment. *Diagn Microbiol Infect Dis*. 2007;57(4):361-366. doi:10.1016/j.diagmicrobio.2006.10.004

Smits HL, Eapen CK, Sugathan S, et al. Lateral-flow assay for rapid serodiagnosis of human leptospirosis. *Clin Vaccine Immunol*. 2001;8(1):166-169. doi:10.1128/CDLI.8.1.166-169.2001

Smythe LD, Smith IL, Smith GA, et al. A quantitative PCR (TaqMan) assay for pathogenic *Leptospira spp. BMC Infect Dis*. 2002;2(1):13. doi:10.1186/1471-2334-2-13

Smythe LD, Wuthiekanun V, Chierakul W, et al. Short report: The Microscopic Agglutination Test (MAT) is an unreliable predictor of infecting *Leptospira* serovar in Thailand. *Am J Trop Med Hyg*. 2009;81(4):695-697. doi:10.4269/ajtmh.2009.09-0252

Smythe L, Adler B, Hartskeerl RA, et al. Classification of *Leptospira* genomospecies 1, 3, 4 and 5 as *Leptospira alstonii* sp. nov., *Leptospira vanthielii* sp. nov., *Leptospira terpstrae* sp. nov. and *Leptospira yanagawae* sp. nov., respectively. *Int J Syst Evol Microbiol*. 2013;63(Pt 5):1859-1862. doi:10.1099/ijs.0.047324-0

Sonthayanon P, Jaresitthikunchai J, Mangmee S, et al. Whole cell matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) for identification of *Leptospira spp.* in Thailand and Lao PDR. Lin T, ed. *PLoS Negl Trop Dis*. 2019;13(4):e0007232. doi:10.1371/journal.pntd.0007232

Spinali S, van Belkum A, Goering RV, et al. Microbial typing by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: Do we need guidance for data interpretation? Doern GV, ed. *J Clin Microbiol*. 2015;53(3):760-765. doi:10.1128/JCM.01635-14

Staneck JL, Henneberry RC, Cox CD. Growth requirements of pathogenic *Leptospira*. *Infect Immun*. 1973;7(6):886-897. doi:10.1128/IAI.7.6.886-897.1973

Staroscik A. Calculator for determining the number of copies of a template. https://web.uri.edu/gsc/dsdna-calculator/. Published January 2004.

Subharat S, Wilson PR, Heuer C, and Collins-Emerson JM. Evaluation of a SYTO9 real-time polymerase chain reaction assay to detect and identify pathogenic *Leptospira* species in kidney tissue and urine of New Zealand farmed deer. *J VET Diagn Invest*. 2011;23(4):743-752. doi:10.1177/1040638711407892

Sulzer CR, Glosser JW, Rogers F, and Jones WL. Evaluation of an indirect hemagglutination test for the diagnosis of human leptospirosis. *J Clin Microbiol*. 1975;2:218-221.

Sun G, Yang K, Zhao Z, et al. Shotgun metabolomics approach for the analysis of negatively charged water-soluble cellular metabolites from mouse heart tissue. *Anal Chem*. 2007;79(17):6629-6640.

Sunner Jan, Dratz Edward, Chen Y-Chie. Graphite surface-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides and proteins from liquid solutions. *Anal Chem*. 1995;67(23):4335-4342. doi:10.1021/ac00119a021

Sykes JE. Canine leptospirosis - a re-emerging disease. In: *Proceedings of the 19th Annual Meeting of the ACVIM*. ; 2001:763-765.

Sykes JE, Hartmann K, Lunn KF, et al. 2010 ACVIM Small Animal Consensus Statement on Leptospirosis: Diagnosis, Epidemiology, Treatment, and Prevention: Leptospirosis. *J Vet Intern Med*. 2011;25(1):1-13. doi:10.1111/j.1939-1676.2010.0654.x

Sykes JE. Leptospirosis. In: Sykes JE, Saunders WB, eds. *Canine and Feline Infectious Diseases*. Elsevier; 2014:474-486. doi:10.1016/B978-1-4377-0795-3.00050-8

Syme HM, Fletcher MGR, Bailey SR, and Elliott J. Measurement of aldosterone in feline, canine and human urine. *J Small Anim Pract*. 2007;48(4):202-208. doi:10.1111/j.1748-5827.2006.00264.x

Szabados F, Tix H, Anders A, et al. Evaluation of species-specific score cutoff values of routinely isolated clinically relevant bacteria using a direct smear preparation for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry-based bacterial identification. *Eur J Clin Microbiol Infect Dis*. 2012;31(6):1109-1119. doi:10.1007/s10096-011-1415-7

Takeuchi O, Hoshino K, Kawai T, et al. Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. *Immunity*. 1999;11(4):443-451. doi:10.1016/S1074-7613(00)80119-3

Tanaka K, Ido Y, Akita S, et al. Detection of High Mass Molecules by Laser Desorption Time-of-Flight Mass Spectrometry. In: *Proceedings of the 2nd Japan-China Joint Symposium on Mass Spectrometry*. Osaka, Japan; 1987.

Theel ES, Schmitt BH, Hall L, et al. Formic acid-based direct, on-plate testing of yeast and *Corynebacterium* species by Bruker Biotyper matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J Clin Microbiol*. 2012;50(9):3093-3095. doi:10.1128/JCM.01045-12

Truong KN, Coburn J. The emergence of severe pulmonary hemorrhagic leptospirosis: Questions to consider. *Front Cell Inf Microbio*. 2012;1. doi:10.3389/fcimb.2011.00024

Turner LH. Leptospirosis III. *Trans R Soc Trop Med Hyg*. 1970;64(4):623-646.

Ueda O, Tanaka S, Nagasawa Z, et al. Development of a novel matrix-assisted laser desorption/ionization time-of-flight mass spectrum (MALDI-TOF-MS)-based typing method to identify methicillin-resistant *Staphylococcus aureus* clones. *J Hosp Infect*. 2015;90(2):147-155. doi:10.1016/j.jhin.2014.11.025

United Nations, Department of Economic and Social Affairs, Population Division. World Urbanization Prospects: The 2014 Revision. 2015. https://esa.un.org/unpd/wup/publications/files/wup2014-report.pdf.

Valentine N, Wunschel S, Wunschel D, et al. Effect of culture conditions on microorganism identification by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol*. 2005;71(1):58-64. doi:10.1128/AEM.71.1.58-64.2005

van Belkum A, Tassios PT, Dijkshoorn L, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infec*. 2007;13 (Suppl. 3):1-46. doi:10.1111/j.1469-0691.2007.01786.x

van Belkum A, Welker M, Pincus D, et al. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry in clinical microbiology: What are the current issues? *Ann Lab Med*. 2017;37(6):475. doi:10.3343/alm.2017.37.6.475

van de Maele I, Claus A, Haesebrouck F, and Daminet S. Leptospirosis in dogs: a review with emphasis on clinical aspects. *Vet Rec*. 2008;163(14):409-413. doi:10.1136/vr.163.14.409

Vargha M, Takáts Z, Konopka A, and Nakatsu CH. Optimization of MALDI-TOF MS for strain level differentiation of Arthrobacter isolates. *J Microbial Methods*. 2006;66(3):399-409. doi:10.1016/j.mimet.2006.01.006

Varni V, Ruybal P, Lauthier JJ, et al. Reassessment of MLST schemes for *Leptospira spp.* typing worldwide. *Infect Genet Evol*. 2014;22:216-222. doi:10.1016/j.meegid.2013.08.002

Vasylieva N, Andreychyn M, Kravchuk Y, et al. Changes in leptospirosis etiology in animals and humans. *Ann Agric Environ Med*. 2017;24(4):671-675. doi:10.26444/aaem/78031

Victoriano AFB, Smythe LD, Gloriani-Barzaga N, et al. Leptospirosis in the Asia Pacific region. *BMC Infect Dis*. 2009;9(1):147. doi:10.1186/1471-2334-9-147

Vijayachari P, Ahmed N, Sugunan AP, et al. Use of fluorescent amplified fragment length polymorphism for molecular epidemiology of leptospirosis in India. *J Clin Microbiol*. 2004;42(8):3575-3580. doi:10.1128/JCM.42.8.3575-3580.2004

Vijayachari P, Sugunan AP, Shriram AN. Leptospirosis: an emerging global public health problem. *J Biosci*. 2008;33(4):557-569. doi:10.1007/s12038-008-0074-z

Villumsen S, Pedersen R, Borre MB, et al. Novel TaqMan® PCR for detection of *Leptospira* species in urine and blood: Pit-falls of in silico validation. *J Microbial Methods*. 2012;91(1):184-190. doi:10.1016/j.mimet.2012.06.009

Wagenaar J, Zuerner RL, Alt D, and Bolin CA. Comparison of polymerase chain reaction assays with bacteriologic culture, immunofluorescence, and nucleic acid hybridization for detection of *Leptospira borgpetersenii* serovar Hardjo in urine of cattle. *Am J Vet Res*. 2000;61(3):316-320. doi:10.2460/ajvr.2000.61.316

Walker M. *Method Development and Preliminary Applications for* Leptospira *Spirochetes in Water Samples*. Environmental Protection Agency, Office of Research and Development; 2008:30.

Wallace MR, Hale BR, Utz GC, et al. Endemic infectious diseases of Afghanistan. *Clin Infect Dis*. 2002;34(Supplement_5):S171-S207. doi:10.1086/340704

Wang Y, Zhuang X, Zhong Y, et al. Distribution of plasmids in distinct *Leptospira* pathogenic species. Picardeau M, ed. *PLoS Negl Trop Dis*. 2015;9(11):e0004220. doi:10.1371/journal.pntd.0004220

Ward MP, Glickman LT, Guptill LF. Prevalence of and risk factors for leptospirosis among dogs in the United States and Canada: 677 cases (1970-1998). *JAVMA*. 2002;220(1):53-58. doi:10.2460/javma.2002.220.53

Ward MP, Guptill LF, Prahl A, and Ching Wu C. Serovar-specific prevalence and risk factors for leptospirosis among dogs: 90 cases (1997–2002). *JAVMA*. 2004;224(12):1958-1963. doi:10.2460/javma.2004.224.1958

Weil A. Ueber eine eigenthümliche, mit Milztumor, Icterus and Nephritis einhergehende, acute Infectionskrankheit. In: *Deutsches Archiv Für Klinische Medicin*. 1886:209-232. https://ia800704.us.archive.org/0/items/deutschesarchiv125unkngoog/deutschesarchiv125unkngoog.pdf.

Werts C, Tapping RI, Mathison JC, et al. Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nat Immunol*. 2001;2(4):346-352. doi:10.1038/86354

West-Nørager M, Kelstrup CD, Schou C, et al. Unravelling in vitro variables of major importance for the outcome of mass spectrometry-based serum proteomics. *J Chromatogra B*. 2007;847(1):30-37. doi:10.1016/j.jchromb.2006.09.048

White AM, Zambrana-Torrelio C, Allen T, et al. Hotspots of canine leptospirosis in the United States of America. *Vet J*. 2017;222:29-35. doi:10.1016/j.tvjl.2017.02.009

Wilcox SK, Cavey GS, Pearson JD. Single ribosomal protein mutations in antibiotic-resistant bacteria analyzed by mass spectrometry. *Antimicrob Agents Chemother*. 2001;45(11):3046-3055. doi:10.1128/AAC.45.11.3046-3055.2001

Williams TL, Andrzejewski D, Lay JO, and Musser SM. Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. *J Am Soc Mass Spectrom*. 2003;14(4):342-351. doi:10.1016/S1044-0305(03)00065-5

World Health Organization. *Human Leptospirosis: Guidance for Diagnosis, Surveillance and Control*. Geneva, Switzerland; 2003. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0036-46652003000500015&lng=en&tlng=en. Accessed February 5, 2020.

World Health Organization. Leptospirosis Fact Sheet. http://www.wpro.who.int/mediacentre/factsheets/fs_13082012_leptospirosis/en/. Published 2012.

World Organisation for Animal Health (OIE). Leptospirosis. OIE Terrestrial Manual 2018. https://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.01.12_LEPTO.pdf. Published 2018. Accessed February 5, 2020.

Wu H-F, Gopal J, Manikandan M. Future perspective of nanoparticle interaction-assisted laser desorption/ionization mass spectrometry for rapid, simple, direct and sensitive detection of microorganisms: Future of NPs in MALDI-MS of microbes. *J Mass Spectrom*. 2012;47(3):355-363. doi:10.1002/jms.2962

Wu KJ, Odom RW. Peer reviewed: Characterizing synthetic polymers by MALDI MS. *Anal Chem*. 1998;70(13):456A-461A. doi:10.1021/ac981910q

Wuthiekanun V, Amornchai P, Paris DH, et al. Rapid isolation and susceptibility testing of *Leptospira spp.* using a new solid medium, LVW agar. *Antimicrob Agents Chemother*. 2013;57(1):297-302. doi:10.1128/AAC.01812-12

Xi B, Gu H, Baniasadi H, and Raftery D. Statistical Analysis and Modeling of Mass Spectrometry-Based Metabolomics Data. In: Raftery D, ed. *Mass Spectrometry in Metabolomics*. Vol 1198. New York, NY: Springer New York; 2014:333-353. doi:10.1007/978-1-4939-1258-2_22

Xiao D, Zhang C, Zhang H, et al. A novel approach for differentiating pathogenic and non-pathogenic *Leptospira* based on molecular fingerprinting. *J Proteomics*. 2015;119:1-9. doi:10.1016/j.jprot.2014.10.023

Xu Y, Zhu Y, Wang Y, et al. Whole genome sequencing revealed host adaptation-focused genomic plasticity of pathogenic *Leptospira*. *Sci Rep*. 2016;6(1):20020. doi:10.1038/srep20020

Zboromyrska Y, Rubio E, Alejo I, et al. Development of a new protocol for rapid bacterial identification and susceptibility testing directly from urine samples. *Clin Microbiol Infec*. 2016;22(6):561.e1-561.e6. doi:10.1016/j.cmi.2016.01.025

Zboromyrska Y, Bosch J, Aramburu J, et al. A multicentre study investigating parameters which influence direct bacterial identification from urine. Rohde H, ed. *PLoS ONE*. 2018;13(12):e0207822. doi:10.1371/journal.pone.0207822

Zenobi R, Knochenmuss R. Ion formation in MALDI mass spectrometry. *Mass Spec Rev*. 1998;17:337-366.

Zhang L, Caffrey RE, Dalmasso EA, et al. Contribution of human a-Defensin 1, 2, and 3 to the anti-HIV-1 activity of CD8 antiviral factor. *Science*. 2002;298:995-1000.

Zhang L, Sandrin TR. Maximizing the Taxonomic Resolution of MALDI-TOF-MS-Based Approaches to Bacterial Characterization: From Culture Conditions Through Data Analysis. In: Demirev P, Sandrin TR, eds. *Applications of Mass Spectrometry in Microbiology*. Cham: Springer International Publishing; 2016:147-181. doi:10.1007/978-3-319-26070-9_6

Zhang X, Leung S-M, Morris CR, and Shigenaga MK. Evaluation of a novel, integrated approach using functionalized magnetic beads, bench-top MALDI-TOF-MS with prestructured sample supports, and pattern recognition software for profiling potential biomarkers in human plasma. *J Biomolecular Techniques*. 2004;15(3):167-175.

Zhou P, Altman E, Perry MB, Li J. Study of matrix additives for sensitive analysis of lipid A by matrix-assisted laser desorption ionization mass spectrometry. *Appl Environ Microbiol*. 2010;76(11):3437-3443. doi:10.1128/AEM.03082-09

Zuerner RL. Laboratory Maintenance of Pathogenic *Leptospira*. In: Coico R, Kowalik T, Quarles J, Stevenson B, Taylor R, eds. *Current Protocols in Microbiology*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2005:mc12e01s00. doi:10.1002/9780471729259.mc12e01s00

Zuerner RL, Alt DP. Variable nucleotide tandem-repeat analysis revealing a unique group of *Leptospira interrogans* serovar Pomona isolates associated with California sea lions. *J Clin Microbiol*. 2009;47(4):1202-1205. doi:10.1128/JCM.01639-08

Zuerner RL. Host Response to *Leptospira* Infection. In: Adler B, ed. *Leptospira and Leptospirosis*. Vol 387. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015:223-250. doi:10.1007/978-3-662-45059-8_9

APPENDIX A: Concentration per %T Curves for *Leptospira interrogans* serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona

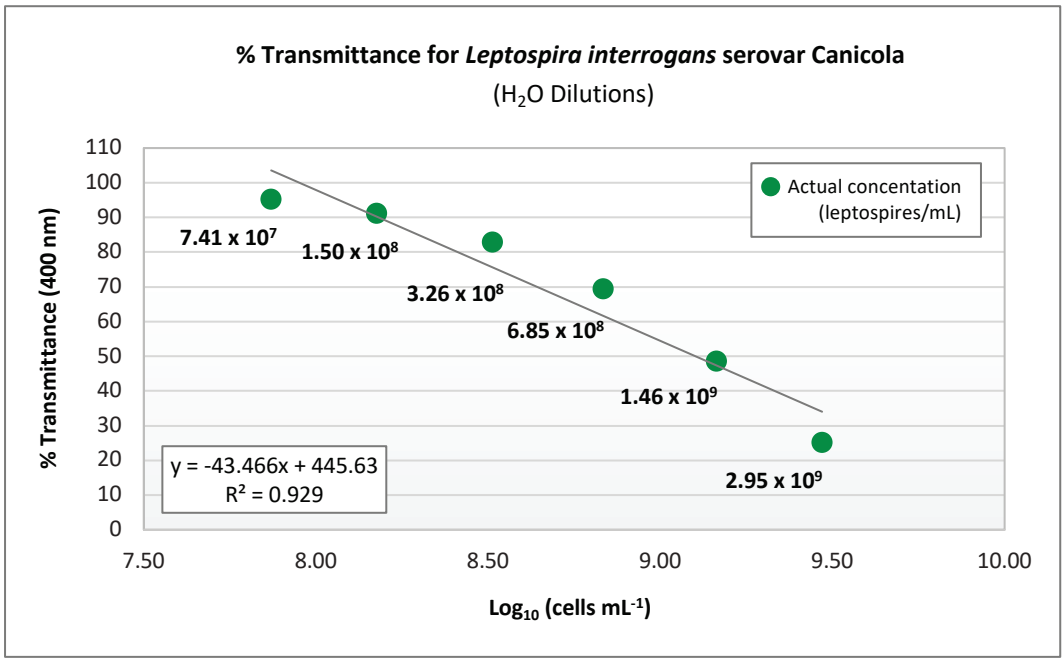

Figure A.1. Percent transmittance values per log concentration for two-fold serial dilutions of *Leptospira interrogans* serovar Canicola.

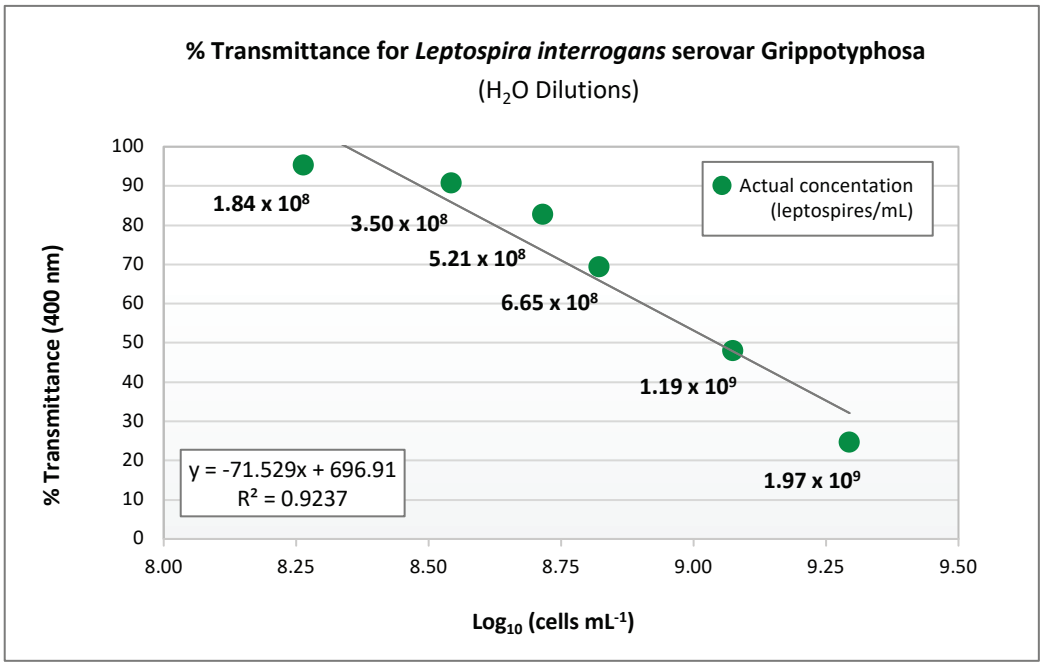

Figure A.2. Percent transmittance values per concentration for two-fold serial dilutions of *Leptospira interrogans* serovar Grippotyphosa.

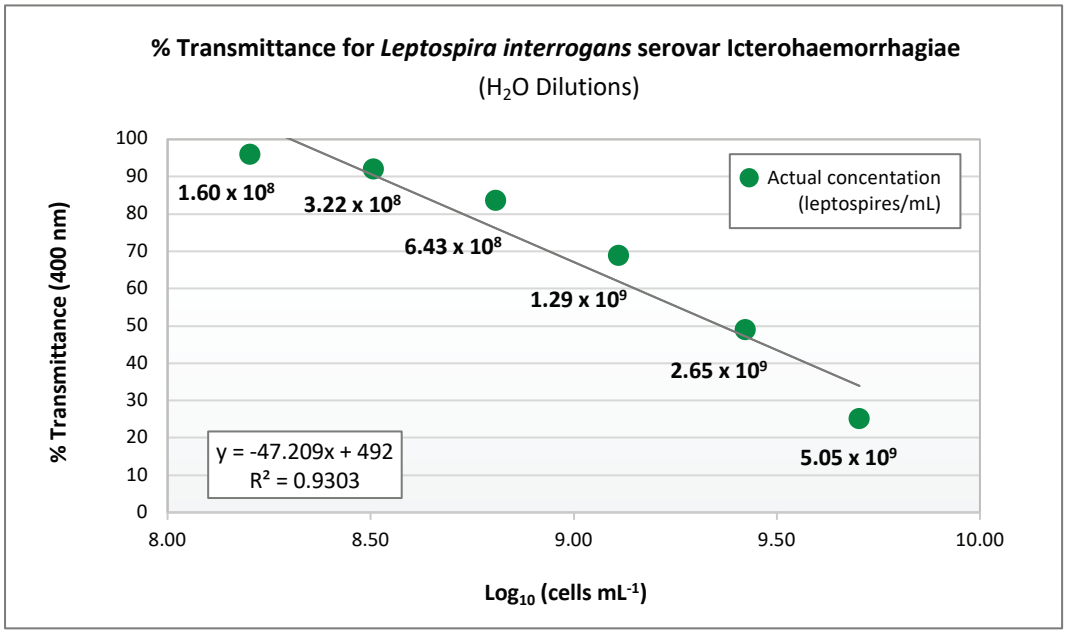

Figure A.3. Percent transmittance values per concentration for two-fold serial dilutions of *Leptospira interrogans* serovar Icterohaemorrhagiae.

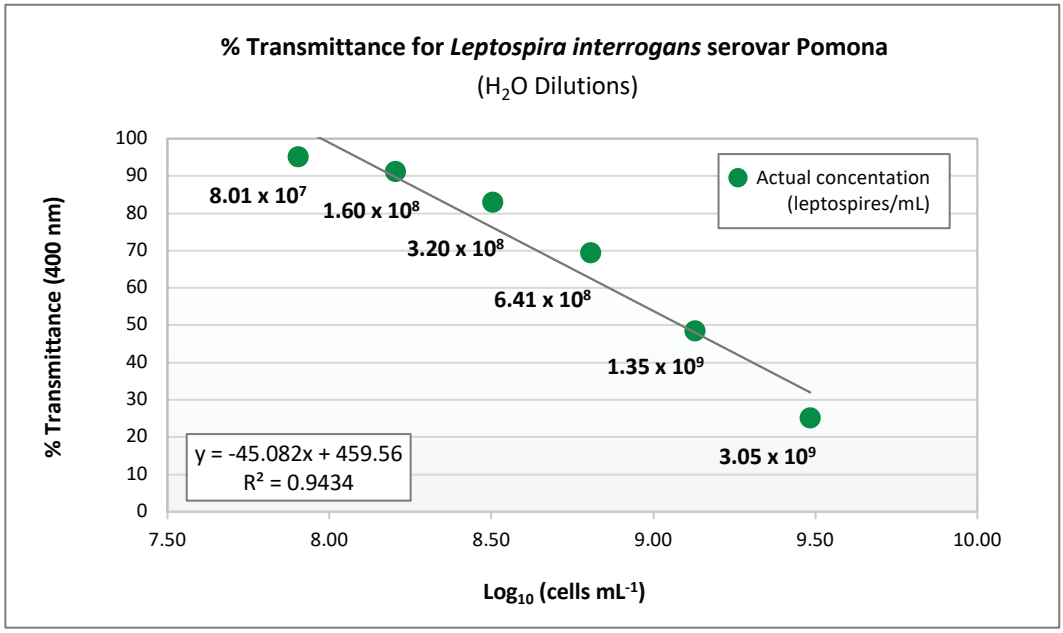

Figure A.4. Percent transmittance values per log concentration for two-fold serial dilutions of *Leptospira interrogans* serovar Pomona.

APPENDIX B: Individual Standard Curves for *Leptospira interrogans* serovars Canicola, Grippotyphosa, Icterohaemorrhagiae, and Pomona

(H_2O Dilutions)

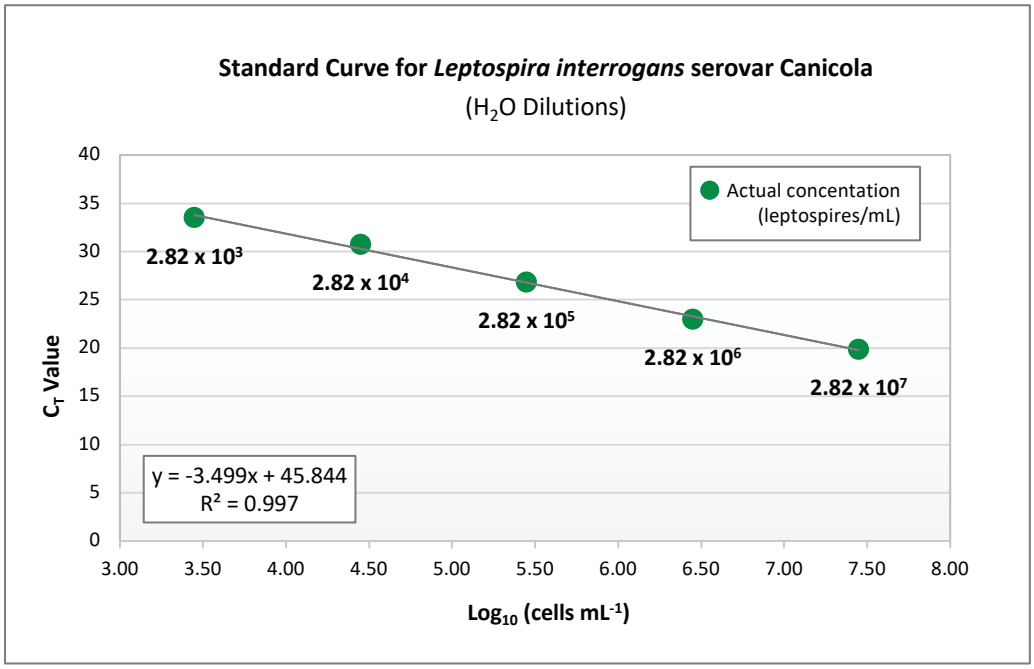

Figure B.1. Cycle threshold value per log concentration for ten-fold serial dilutions of *Leptospira interrogans* serovar Canicola.

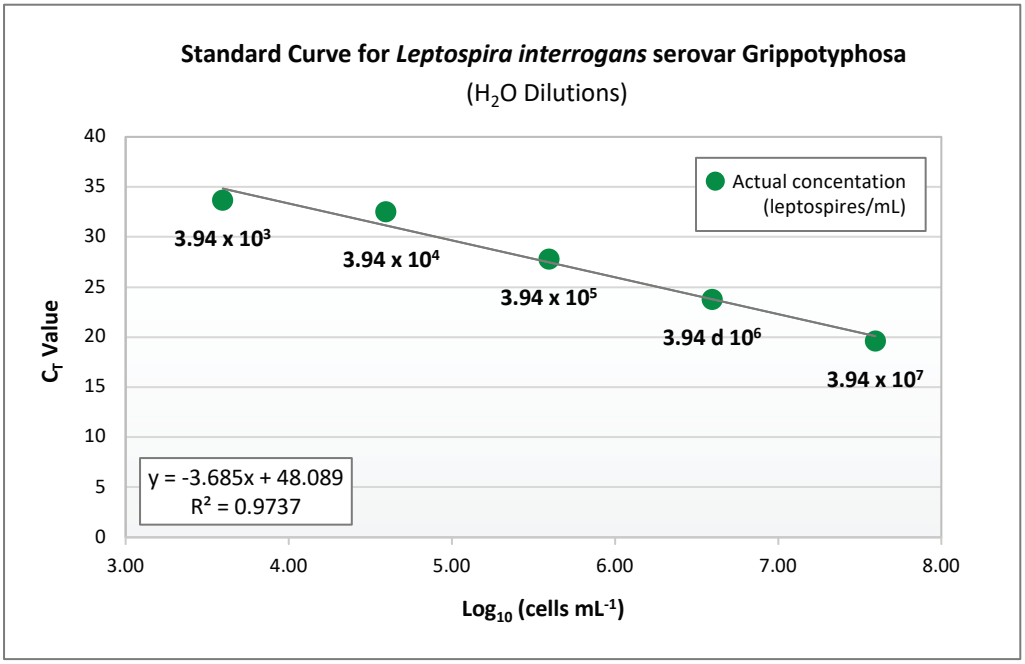

Figure B.2. Cycle threshold value per log concentration for ten-fold serial dilutions of *Leptospira interrogans* serovar Grippotyphosa.

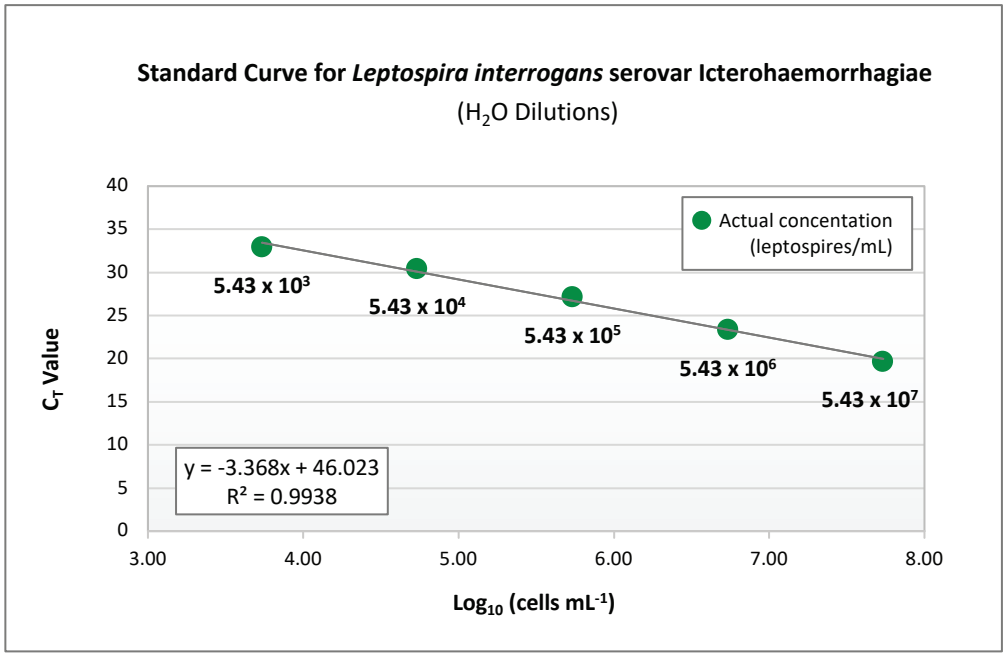

Figure B.3. Cycle threshold value per log concentration for ten-fold serial dilutions of *Leptospira interrogans* serovar Icterohaemorrhagiae.

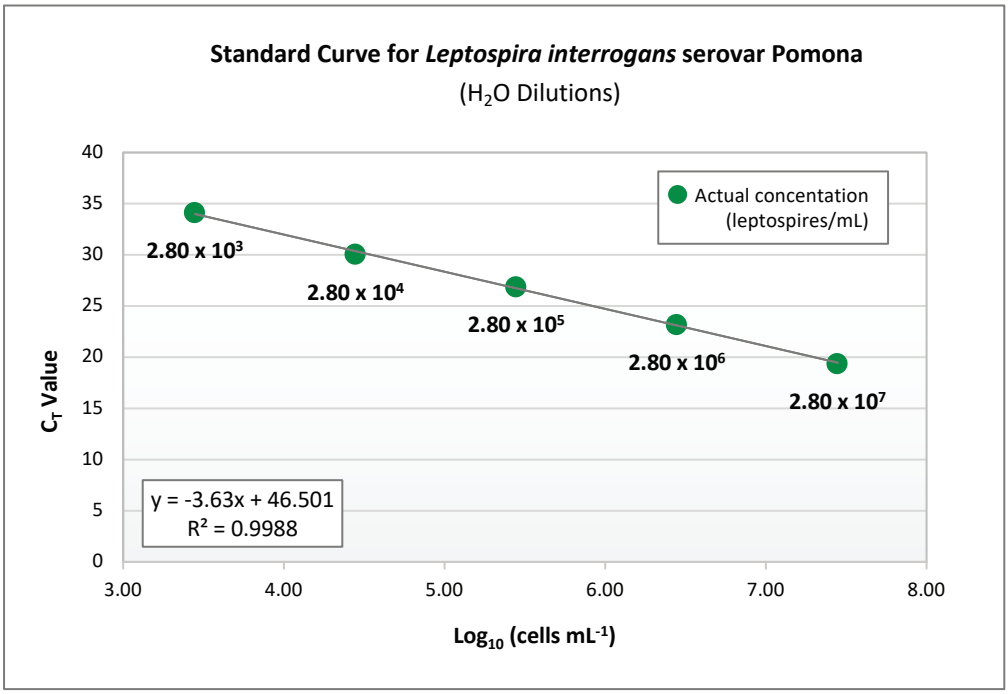

Figure B.4. Cycle threshold value per log concentration for ten-fold serial dilutions of *Leptospira interrogans* serovar Pomona.

APPENDIX C: Individual Standard Curves for *Leptospira interrogans* serovars Grippotyphosa and Pomona

(Urine Dilutions)

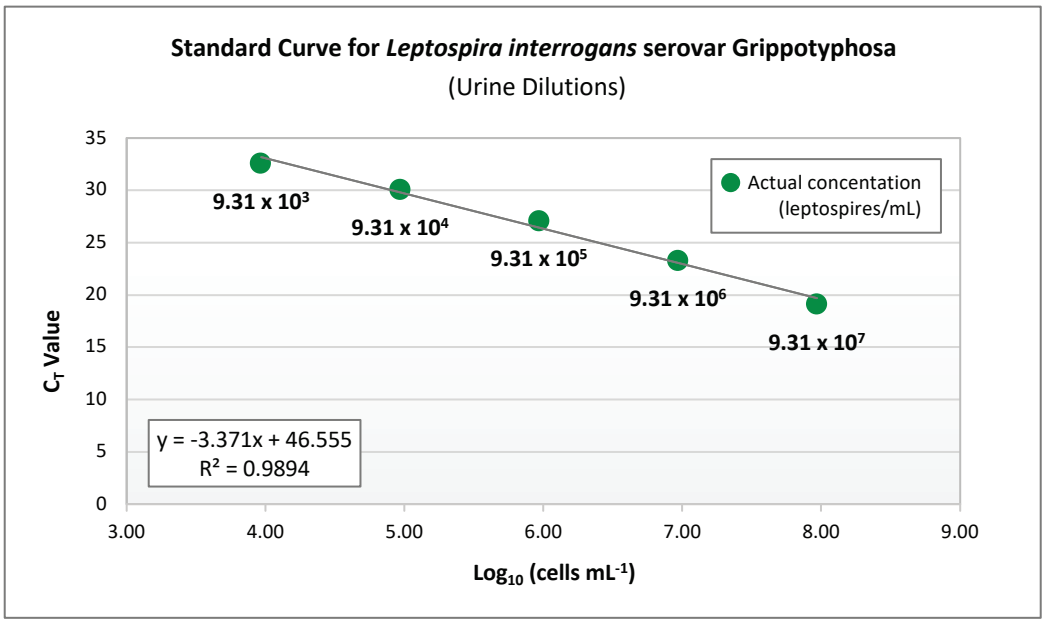

Figure C.1. Cycle threshold value per log concentration for ten-fold serial dilutions of *Leptospira interrogans* serovar Grippotyphosa.

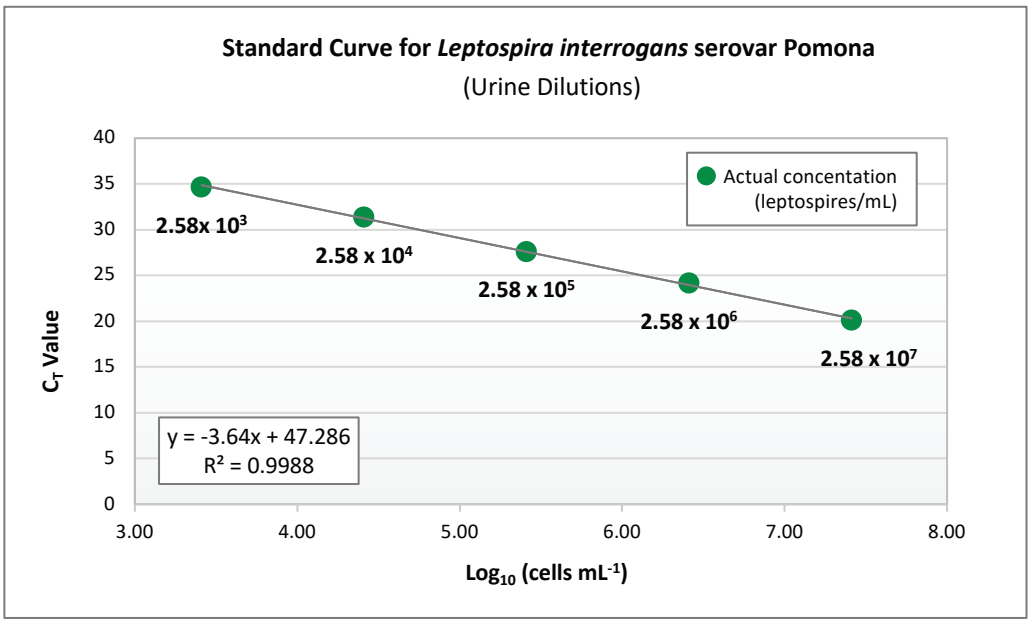

Figure C.2. Cycle threshold value per log concentration for ten-fold serial dilutions of *Leptospira interrogans* serovar Pomona.