FECAL CONTAMINANT SOURCE TRACKING USING MITOCHONDRIAL DNA DETECTION OF TARGET ORGANISM BY PCR AND QPCR

By

DAVID KYLE TAYLOR

Bachelor of Science in Environmental, Health, and Safety Management

Northeastern State University

Tahlequah, Oklahoma

2016

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2018

FECAL CONTAMINANT SOURCE TRACKING

USING MITOCHONDRIAL DNA DETECTION OF

TARGET ORGANISM BY PCR AND QPCR

Thesis Approved:

Mark Krzmarzick, Ph.D., P.E.

Thesis Adviser

Babu Fathepure, Ph.D.

Greg Wilber, Ph.D., P.E.

ACKNOWLEDGEMENTS

This project was funded with the support of Accurate Environmental LLC. I would like to thank Dr. Ali Fazel, CEO and Director of Environmental Services for Accurate Environmental, for his support, which was paramount to the success of this research and my master's degree.

I started as a part-time environmental science graduate student working a full-time position at an environmental services company in Tulsa when approaching Dr. Mark Krzmarzick about a position in his lab under his advisement. After a semester of engineering-focused courses, it was a clear decision for me to transfer into the Environmental Engineering program to pursue research and coursework full-time. It did not take me long to feel welcomed by my fellow lab mates, Xuewen, Xiang, and Matt; to which I owe sincere gratitude for their friendships. I especially owe thanks to Dr. Krzmarzick for his unflinching support from day one, and in-depth training to carry out quality research.

I would also like to thank my committee members Dr. Babu Fathepure and Dr. Gregory Wilber for participating in my defense, and providing sage counsel to improve and refine my thesis document.

Finally, I would like to thank my family for their constant encouragement and support. My partner, Cassandra Sturycz, was my rock throughout this degree, and she continues to support me in the decision to pursue a Ph.D. degree.

Name: DAVID KYLE TAYLOR

Date of Degree: DECEMBER, 2018

Title of Study: FECAL CONTAMINANT SOURCE TRACKING USING MITOCHONDRIAL DNA DETECTION OF TARGET ORGANISM BY PCR AND QPCR

Major Field: ENVIRONMENTAL ENGINEERING

Abstract: Water quality impairment by fecal waste pollution of surface water and groundwater is a public health issue by introducing pathogenic microorganisms and/or excessive nutrients leading to eutrophication of surface water. Current molecular methods in fecal source tracking commonly incorporate the detection of fecal bacteria that are unique to the polluting species gut microorganism community found in fecal matter. This method has been thoroughly researched without conclusive evidence of a sure and fast method for multiple species detection. In this study, we investigate and develop a standardized PCR and qPCR method for the detection and quantification of host species mitochondrial DNA (mtDNA) in affected water sources. Mitochondria are a promising candidate for fecal polluter detection, as a huge number of intestine epithelial cells are exfoliated during feces evacuation, and numerous mitochondria are present in each cell which contain multiple copies of the mitochondrial genome. We developed a comprehensive detection method of mtDNA by designing novel primers for bison, cattle, duck, geese, human, and swine. Clone libraries were developed for a standardized DNA template of the PCR product genes inserted into a commercial vector plasmid. A dilution gradient of the standards was performed for the quantification of unknown samples. Water samples with unknown quantities of mtDNA were collected from 10 locations along the Illinois River and analyzed alongside the standardized mtDNA dilution series. Our results showed that each primer is specific to the target organism and did not produce false positives, mtDNA has a low detection limit in environmental samples, and clone libraries are an effective approach to long term storage of mtDNA standards. This approach is a viable method for rapid detection of fecal waste polluters with direct specificity to the contributing species.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Surface water and groundwater are susceptible to fecal waste contamination from a number of sources, such as livestock feedlot runoff, untreated waste water, land applied manure fertilization, faulty septic tank systems, and wildlife excretions (Ritter et al., 2002; Simpson et al., 2002). Exposure to these wastes put humans and animals at risk of infections from waterborne pathogens or eutrophication of water sources by excessive nutrient inputs. Waters affected by these harmful contributions will contaminate drinking water sources, wildlife well-being, agricultural crop irrigation, marine harvesting, and human recreation activities. The major contributors of these inputs are non-point sources (NPS) which are difficult to accurately characterize as pollution sources, as many NPSs with diffuse host origins may exist along a single water source. Additionally, rainfall events will exacerbate the contamination of NPS polluters by carrying contaminants away from land to flowing water sources. These events will lead to increases of harmful pollutants into water sources that will require rapid detection methods in order to promptly notify affected users and facilitate swift corrective actions. Current detection methods of pathogenic bacteria in water sources are primarily culture-dependent methods, which detect fecal indicator bacteria (FIB) in water samples that are filtered then grown on

selective media for microorganisms that mimic pathogens and satisfy the indicator organism criteria stated by the EPA (EPA, 2006). Culture-independent methods have more recently been developed to eliminate the culture growing step and potentially decrease the amount of time necessary to report public health risk of valued waters (Field and Samadpour, 2007; Wade et al., 2006).

Locally, water source contamination is a concern due to the amount of livestock operations existing along water sources that span multiple states. Agricultural operations are often located near water sources; such as rivers, tributaries, and creeks for the ease of irrigation and other water intensive operations. This proximity may pose a risk to the water quality downstream of an agricultural facility as a result of runoff contaminants from livestock fecal waste (Crane et al, 1983). Livestock waste is not the sole culprit of fecal waste in water sources. Human fecal matter is a known contaminant to water sources from faulty septic systems, wastewater contamination via infrastructure distribution failures, and/or mishandling of waste (Sauer et al., 2011; Newton et al., 2011; Sidhu et al., 2013). Additionally, wildlife are also sources of fecal contamination; however these organisms are rarely the cause of pathogenic outbreaks (Cox et al., 2005). A 2017 Public Water Supply Annual Compliance Report by the Oklahoma Department of Environmental Quality (ODEQ) indicated that there were a total of seven acute maximum contaminant level (MCL) violations from six different systems for coliform detection, resulting in mandatory boil orders for the associated public water systems (ODEQ, 2017). Failure to detect and treat pathogens in affected water can lead to a multitude of side effects from illnesses caused by various microorganisms. Bacterial pathogens include *Escherichia coli* (strain 0157:H7), *Shigella*, *enterococcus*, *Salmonella typhi*,

Campylobacter jejuni, *Vibrio cholerae*, etc., which are likely to cause side effects including diarrhea, vomiting, fever, and cramps. When water sources that supply irrigation water for crops becomes contaminated, it is likely that the food will become inoculated with the pathogen, further spreading probability of illness and contamination. Aquatic wildlife is also affected by fecal waste contaminants; shellfish are particularly susceptible to infection of fecal coliforms.

Excessive fecal waste, especially from agricultural runoff, may also play a part in the eutrophication of water sources from the additional nutrients present in runoff (Blann et al., 2009). Eutrophication will lead to fish kills and compromised water quality as the dissolved oxygen is primarily absorbed by aquatic photosynthetic organisms such as algae (Smith and Schindler, 2009).

Current detection methods of fecal contaminants in surface waters are primarily carried out by the presence of FIB. These organisms show morphological and physiological characteristics resembling bacteria present in warm blooded animals' intestinal tracts. These methods allow regulators to monitor the presence of pathogenic microorganisms in sensitive surface waters with relative ease. There are well practiced culture-dependent methods that have been implemented and recommended by the EPA since 1976: Fecal coliform, total coliform, *Escherichia coli* (*E. coli*), fecal streptococci, and enterococci detection (EPA, 1976). These methods range in specificity and are implemented based on the necessity of characterizing specific microorganisms under various growth conditions.

Fecal coliform tests are widely implemented as the standard for coliform detection due to the relative specificity it has for organisms that primarily exist in

intestinal environments. Fecal coliform tests are carried out by taking water samples at the location of interest and vacuum filtering the samples, either on site or in the laboratory, through 0.45-micron pore size filter paper. Then, as inoculum, the filter paper is placed on a selective MFC media agar containing (1.5% agar, 1.5 g/L bile salts, 12.0 g/L lactose, 0.1 g/L methyl blue, 5.0 g/L proteose peptone, 5.0 g/L sodium chloride, 10.0 g/L tryptose, and 3.0 g/L yeast extract). After incubation of the sample at 44.5 °C for 24 hours, colonies form on the plate, which will indicate the presence of fecal indicator bacteria by the growth of blue colonies. Total coliform tests involve the same sample collection and preparation steps, but the incubation temperature is 35 °C and the nutrient agar selects for organisms in the presence of bile salts and production of acid and gas during fermentation of lactose. *E. coli* tests specifically measure for the *E. coli* genera by performing a total coliform test with the additional step for beta-glucuronidase activity, which is mostly specific for *E. coli* amongst all other enteric coliforms. Fecal streptococci counts require the preparation of a complex organ media that includes chemicals (sodium azide, cycloheximide, and 2,3,5-triphenyltetrazolium chloride) that are typically toxic to the majority of enteric microorganisms, but fecal streptococci are resistant to them Enterococci counts are performed by inoculating a standardized media notably constituted of 6.5% NaCl and hydrolyse pyyrolidonyl-beta-naphthylamide then incubated at 10 °C or 45 °C.

These culture based methods are effective at determining the presence of indicator organisms; however, they are unable to determine the source of the pathogenic organisms contributing to the pollution. In complex systems, there are many inputs to a water source that are potential pollution contributors, which cannot be characterized by culture based

tests. Identification methods of fecal contaminant sources can be classified as culture-dependent and culture-independent, which are either library-dependent and library-independent (Field & Samadpour, 2007). Molecular and non-molecular laboratory methodologies can be used as tools to achieve specificity and reproducibility. Conventional approaches use culture-based non-molecular methods that rely on microbiological techniques. Research is progressing to more molecular-based techniques as bioinformatic technologies are more user-friendly and genome data for a wide range of organisms are available.

Molecular based methods include polymerase chain reaction (PCR), denaturing gradient gel electrophoresis (DGGE), host specific 16S rRNA, length heterogeneity (LH) PCR, phage analysis, pulse field gel electrophoresis (PFGE), quantitative real-time PCR (qPCR), repetitive DNA sequences, ribotyping, terminal restriction fragment length polymorphism (T-RFLP), and toxin biomarkers. Non-molecular approaches include antibiotic resistance analysis (ARA), caffeine presence, and optical brighteners. The molecular method of amplifying regions of mitochondrial DNA (mtDNA) from eukaryotic host cells is an emerging approach to detect fecal pollution sources. Mitochondria are organelles responsible for aerobic respiration in most eukaryotic cells. These organelles also contain their own genomes and ribosomes, which is hypothesized to be a remnant of an endosymbiotic relationship between a host eukaryotic cell and endosymbiont respiratory bacterial cell (mitochondria) giving rise to eukaryotic cells. Mitochondrial genomes are convenient for species detection as there are varying regions of genetic conservation and variation between species. Conserved regions of the mitochondrial genome allow for the development of universal primers in order to detect

multiple species with a single primer set, which could serve as a pass/fail step for the bulk presence of mtDNA within an environmental sample. Variable regions of the mitochondrial genome allow for selectivity of one species over another. Assuming there is enough mtDNA in abundance for detection in the environment, this method would prove effective in characterizing the organism contributing to the bulk mtDNA. Conveniently, warm blooded animals excrete large numbers of epithelial cells during defecation, and these cells contain many numbers of mitochondria (Iyengar et al., 1991; Andreasson et al., 2002; Gerber et al., 2001). An individual mitochondrion contains approximately 2-10 mtDNA copies (Wiesner et al., 1992). Once in the environment, the persistence of mtDNA and its molecular stability remains intact and within detection range for up to 14 days (Baker-Austin et al., 2010). Considering the characteristics of mitochondrial genomes and the abundance of mitochondria excreted from exfoliated epithelial cells during defecation, detection of pathogenic microorganism contamination sources may be linked by pairing mtDNA based detection methods with coliform presence assays. By implementing the principles of PCR and the characteristics of mtDNA, I expect to find that the developed methods in this study will:

1. Provide enough specificity to detect mtDNA for target animals instead of non-target DNA sequences in environmental samples
2. Clone library standards will be an effective tool for quantification of target animal mtDNA in unknown samples
3. qPCR will provide greater sensitivity over standard PCR for detecting and quantifying mtDNA

CHAPTER II

LITERATURE REVIEW

2.1 Culture-Dependent, Library-Independent

As mentioned in the previous chapter, current standards for fecal contaminant detection relies on culture methods from water sample inoculum. Such approaches do not need library databases for verification of organism detection. Methods in this section involve growing bacteria or viruses that are specific to the source organism.

2.1.1 Bacterial

Previous researchers began developing methods to determine fecal contaminant sources in surface waters using culture based methods that involved growing microorganisms obtained from environmental samples on selective media and growth conditions for characterization of microbes present. Principles of temperature variation for microbial differentiation were first studied by researchers in 1904 (Eijkman, 1904). The EPA recommended fecal coliforms as indicator bacteria in 1976 (U.S. EPA, 1976). Approaches in determining the sources of fecal pollutants emerged by analyzing fecal coliform (FC) to fecal streptococci (FS) ratios (Feachem, 1975). Gelreich 1976 came up with standard ranges that defined the general source of fecal pollution based on the ratio of FC/FS ranging from: > 4 suggesting human, between 0.1 & 0.6 indicating domestic animals, and < 0.1 indicating wildlife. The aforementioned approach was dropped when

the application to agricultural settings became inconsistent and difficult to utilize, as streptococci and coliforms have different survival rates (Sinton et al., 1998). Another bacterial approach in differentiating between human and agricultural fecal inputs is possible by comparing the quantities of atypical colonies (AC) to total coliform (TC) colonies after a membrane filtration and incubation on selective media. This ratio is indicative of values similar to raw wastewater values when high flows are present in water sources (Nieman & Brion, 2003). These ratios are affected by the age of contaminants in the water source, compromising the usefulness of the approach. No comparative tests are known between this approach and other identification methods.

Numerous methods exist to culture bacteria specific to the host of interest. Main efforts have been to isolate and culture strains that are unique to humans (*Bifidobacterium adolescentis* and sorbitol-fermenting Bifidobacterium) (Resnick & Levin, 1981; Mara and Oragui, 1983, 1985; Lynch et al., 2002). *Rhodococcus coprophilus* has also been isolated as an indicator bacterium from grazing animal species (Rowbotham & Cross, 1977; Mara and Oragui, 1983, 1985). Savill et al. (2004) applied PCR techniques to amplify regions of 16S rRNA genes of *Rhodococcus coprophilu*s in addition to culture techniques. These methods show a strong specificity for determining between human and grazing animals (Blanch et al., 2006), but the ability for gut microorganisms to exist in the environment is short-lived and their persistence is decreased as seasonal variations occur.

2.1.2 Bacteriophage

It has been found that strains of *Bacteroides fragilis* in certain regions of the world are capable of growing bacteriophages only from human waste water, while other

strains support phage growth from human and various animal waste (Tartera et al., 1989; Tartera & Jofre, 1987; Puig at al., 1999). Research with Bacteroides fragilis has been done in Europe and South Africa (Grabow et al., 1995; Payan et al., 2005), and a comparative study indicated that the method was effective for determination of source organisms (Blanch et al., 2006). Isolation of bacteriophage infected *Bacteroides thetaiotamicron* and *Bacteroides ovatus* has been developed to broaden the possibilities of detection using this method, which might be applicable for North American use. In 2015, the EPA suggested coliphages as potential fecal contaminant indicators in ambient water quality due to their occurrence in the environment, epidemiological correlates, and characteristics (EPA, 2015).

2.2 Culture-dependent, Library-Dependent

The approaches to fecal source tracking described below require cultured microorganisms and libraries that are databases with host origin isolates or patterns for a chosen method. Methods that identify patterns for a specific organism are typically known as DNA fingerprinting.

2.2.1 DNA Fingerprinting Methods

Multiple researchers investigated ribotyping methods applied to bacterial source tracking by means of restriction enzyme digested genomic sequences cut from Southern blot gels and the fragments were then probed with 16S ribosomal RNA sequences in order to determine discriminate species (Tynkkynen et al., 1999; Parveen et al., 1999; Carson et al., 2001; Farag et al., 2001; Hager, 2001a; Samadpour, 2002; Hartel et al., 2002). Scott et al. (2003) concluded that *E. coli* isolated from humans differed significantly from animals in a study that cultured bacteria from feedlot runoff waters and

wastewater effluents. Ribotyping methods provide easily reproducible results, but the labor intensive work flow and database requirements make this method less appealing for rapid detection of contaminants in a water source.

Pulse field gel electrophoresis (PFGE) is another fingerprinting method used by researchers to determine species using restriction enzymes that cut and separate large fragments of genes within the entire genome of an organism. Fragments are separated on a gel by an alternating pulsed electrical field (Tynkkynen et al., 1999; Simmons et al., 2000; Hager, 2001b). A comparative study between single restriction enzyme ribotyping method, PFGE, and randomly amplified polymorphic DNA (RAPD) concluded that PFGE was the most discriminant in determining two *Lactobacillus* strains (Tynkkynen et al., 1999). Later, research indicated that accuracy of ribotyoping experiments are greatly improved when using more than just one restriction enzyme (Samadpour, 2002). In 2004, a comparative method study of blinded control samples found that a PFGE method was only able to detect 1 out of 37 known bacterial isolates (Mathes et al., 2004).

Denaturing Gradient Gel Electrophoresis (DGGE) methods were investigated to detect single nucleotide differences in sequences of the same length by separation. The bands will separate based on the melting properties of the bases in sequence, and each band will define a specific gene sequence. Bands can be excised and sequenced, then compared to sequences of known organisms from a database. Farnleitner et al. (2000) conducted a study coupling DGGE and PCR to detect variations of the functional gene, uidA Beta-D-Glucuronidase, found in *E. coli* strains. This method allowed for the detection and differentiation of *E. coli* strains from environmental water samples without the need for isolation and pure culture identification.

Another DNA fingerprint and PCR based method, repetitive DNA sequences PCR (rep-PCR) was evaluated by amplifying *E. coli* DNA between adjacent repetitive extragenic elements and analyzing the fingerprint patterns against a database through recognition software. The *E. coli* were cultured directly from humans and animals, not environmental samples (Dombek et al. 2000, Carson et al. 2003). Holloway (2001) conducted a study applying the same methods as Dombek et al. (2000) but with the addition of *Enterococcus faecalis* analysis. This study did not provide any confirmatory results of species-specific isolates from their datasets; thus concluding that the method is not reliable for fecal source tracking applications.

2.2.2 Antibiotic Resistance Analysis (ARA)

A method to discriminate between human and animal fecal pollution is to culture the host fecal microorganism and expose it to antibiotics to test for resistance. The assumption is based on the premise that humans and animals, agricultural and wild, are exposed to antibiotics with different constituents, at different intervals, and varying concentrations. Antibiotics of the same class are commonly used amongst humans and livestock animals, resulting in the same antibiotic resistance mechanisms in the animal. When wild animals exist within close range to livestock animals, they are exposed to the antibiotics through eating the livestock feed or ingestion of the drugs through waste runoff. This was observed in a study examining intestinal microorganisms of wild animals from natural populations and subsequent antibiotic resistance of isolated bacteria from the hosts (Souza et al., 1999). Bacteria are known to show antibiotic resistance traits very rapidly and not to remain resistant if unnecessary. It has been observed in a single study that individual isolate resistance can change within the timeframe of a single study

(Samadpour et al., 2005). Fecal source tracking comparative methods studies have shown unfavorable outcomes for ARA performance when compared with other identification methods on blind sample experiments (Griffith et al., 2003; Moore et al., 2005; Samadpour et al., 2005).

2.2 Culture-Independent, Library-Inependent

Initial efforts to detect host-specific microorganisms were carried out by (Hagedorn et al., 2003; Hagedorn et al., 1999; Sinton et al., 1998). Detection was performed by polymerase chain reaction (PCR) assays in order to amplify genetic markers specific to the microorganisms known to reside in different species gut flora. This proved to be a difficult task, as the great majority of microorganisms in warm-blooded animals' intestines are quite similar, and the sheer quantity of microbes residing in intestinal environments ranges from approximately 10^{11}–10^{12} microbial cells/g feces (Guarner and Malagelada, 2003). Within this highly concentrated community of gut microbes, there are an approximate 300-500 bacterial species represented (Guarner and Malagelada, 2003). Stoeckel and Harwood, 2007 have listed the known animal-specific bacterial genetic markers that are specific to the host organism.

2.2.1 Caffeine Detection

Caffeine presence in surface and ground waters has been thought of as a suitable indicator of human waste detection in combined waste water overflows and direct discharge; however, sensitivity of detection was shown to be dependent upon regional conditions and WWTP elimination efficiencies (Buerge et al., 2003). Standley et al. (2000) found that a combination of caffeine and fragrance levels were indicative of human waste water; whilst determining that agricultural waste was characterized by

unique steroids and wildlife waste could be characterized by different, unique steroid ratio (Standley et al., 2000). This method would prove as a quick assay to determine a human waste containing sample versus a non-human waste sample, as caffeine is mostly anthropogenic, however the specificity to source organisms is lacking.

2.2.2 Fecal Stanols

Sterols are the by-products of cholesterol and other steroid metabolism, which are metabolized to stanols in the gut of animals. When used as an indicator for animal activity, these compounds are known as biochemical tracers. The end-products and concentrations of by-products vary between animals due to diets and intestinal microbial diversity (Leeming et al., 1996). Formation of coprostanol, a by-product of cholesterol, is higher in human guts than in other animals' intestines, thus lending the ability to distinguish between human and non-human samples (Blanch et al., 2006).

2.2.3 Optical Brighteners

Optical brighteners (OBs) are compounds commonly found in household detergents and clothing whitening products; they are also known as fluorescent whitening agents. These compounds absorb UV light around 365 nm and emit a range of ~ 415-435 nm, resulting in the affected fabric to lose the natural yellow appearance and look bright white. Detection methods for OB presence in water are relatively simple; cotton fabric filters or pads are placed in a water source for a duration of 2-3 days, where binding of OBs to the fibers in the cotton material will occur, resulting in a fluorescence around 415-435 nm (Dixon et al., 2005). Disadvantages to this method include the inability to quantify contaminants and false positives from a multitude of other chemicals that can be found in the environment (Dickerson et al., 2007).

2.2.4 Mitochondrial

Martellini et al. (2005) were the first researchers to detect fecal contaminants based on abundance of host exfoliated epithelial cells excreted in tandem with fecal matter. This study focused on the detection of human fecal pollution by PCR amplification of sequences unique to human, sheep, cow, and pig. The researchers performed standard single-, multiplex-, and nested PCR of samples followed by amplicon analysis via gel electrophoresis. The samples were sourced from animal tissue, animal feces, and wastewater effluents of varying treatment conditions. The researchers concluded that the general method of mitochondrial DNA detection is a rapid and sensitive approach with detection limits of conventional and multiplex PCR varying between 10–500 pg of genomic DNA, corresponding to 3–150 genome equivalents, and for nested PCR (0.1–1 pg) to less than one genome equivalent. They also concluded that further research should be done to increase the sensitivity of multiplex PCR, as this method would be a time saving and economical approach to source tracking applications.

Caldwell et al. (2007) followed up the pioneer study by investigating real-time PCR amplification of human, cow, and pig mitochondrial DNA found in feces, WWTP influents and effluents, and various farm influents and effluents. Novel primers and probes were designed for single and multiplex real-time PCR. The detection limit they found was 2.0×10^6 mitochondrial copies per 100mL sample water and 11×10^6 mtDNA copies/gram human feces (0.2g per 100 mL effluent/ wastewater). Multiplex experiments resulted in inconsistent detections of species from replicate tests that the probes were used singly. The authors mention continued work to develop dual-labeled probes and primers for more target host animals, and surface water testing for mitochondrial DNA in

order to determine detection limits and baseline determinations to compare with other source tracking methods. Schill and Mathes (2008) conducted a validation study using their mtDNA real-time PCR assays against blinded fecal suspensions from nine vertebrate species (dog, cow, chicken, sheep, horse, pig, goose, deer, and human). The authors additionally added non target DNA to fecal suspensions for potential false amplifications. All assays were reported as successful and specific amplification of targeted species was achieved. Spurious amplification was observed when concentrations of the target exceeded 10^2 copies/μL and the non-target DNA was present in very large quantities; therefore, the authors recommend a "real-world" threshold of 10^2 copies/μL for mtDNA detection in environmental samples.

Single reaction PCR assays were then investigated by Kortbaoui et al. (2009) using universal PCR primers that are selective for five species: human, bovine, porcine, ovine, and chicken mtDNA; dot blot hybridization protocols were also developed and reported for species differentiation from the universal PCR products, providing an additional step of specificity. Authors of this study reported that there were no observed differences of detection between PCR and dot-blot assays, suggesting equal detection sensitivities and potential development of microarray experiments for future studies.

A study four years later developed a mitochondria-based microarray for the pre-screening of environmental samples in order to detect 28 animals that potentially contribute to fecal pollution (Vuong et al., 2013). Researchers of this study also included a clamping PCR mechanisms that inhibited the amplification of fish mtDNA with the universal primer set in order to improve amplification sensitivity of the target region. An additional nested PCR step with two universal primers common to the target animals was

performed to improve the microarray sensitivity was conducted prior to the microarray. Authors state that a microarray approach would be most effective as a pre-screening method to qPCR analysis, as the environmental sample has the potential to contain many different contributing species to a given water source. It should be noted that in order for the microarray probes to detect targets adequately, a preliminary PCR is required in order to bring the copies to a detection level for microarray. Following the microarray, qPCR analysis is carried out for a quantification step. Due to the price of microarray assays and the relative ease and availability of thermocycling equipment, it appears that a quality control step consisting of PCR with universal primers that have selectivity for different animals in varying groups would be a more economical choice. While microarray assays have the advantage of detecting large ranges of sequences at once, the issues with lower sensitivity to targets and the cost currently make standard PCR and qPCR a better option for the majority of users.

CHAPTER III

METHODOLOGY

In the described experiments below, multiple laboratory methods were used from commercial kits manufactured with defined end-products. Each kit will be listed with manufacturer information and the kit model for easy reference if further information is required. The methods described in this study are standardized methods used in various fields for molecular research. None of the experiments in this study were novel to the field.

3.1 Sample Collection

Muscle tissue samples were collected from a local food market. Samples were selected based on date of processing by the market and whether or not a history of freezing was known. The saliva sample was collected and stored according to the instructions provided with the Oragene DISCOVER OGR-500 collection kit (DNA Genotek, CA).

Environmental river samples were collected along the Illinois River under late summer/early fall conditions in the state of Oklahoma. A total of ten 500 mL samples were collected starting at the southern region of the river and collected along the coastline until reaching the Oklahoma-Arkansas Stateline (Figure 1). Longitude and latitude of sampling locations are indicated in Appendix A. Water collection and equipment preparation strategies were modeled after the U.S. Geological Survey (USGS) Fecal

Indicator Bacteria section (7.1) of the National Field Manual for the Collection of Water Quality Data (Myers et al., 2014). Samples were collected at the midline of the river when possible or until chest depth. Sample depths ranged from the water surface to three inches below the surface, and the water was collected upstream before it could come in contact with the sample collector to prevent contamination. The sterile Nalgene 500 mL HDPE collection bottles (mfr#: 2189-0016) were rinsed three times with the site river water and discarded downstream before collecting the final sample water. Samples were immediately stored on ice and vacuum filtered using 0.45-micron pore size filter paper within 24 hours. Following filtration, the filter paper with DNA attached were stored in autoclaved 50 mL Falcon tubes at -80 °C for approximately 48 hours before DNA extractions.

3.2 Muscle Tissue & Saliva DNA Extraction

The target species DNA were extracted from muscle tissue samples according to the standard protocol provided by the QIAGEN DNeasy® Blood& Tissue Kit (Qiagen, USA). DNA from saliva samples were extracted by a user-defined protocol with QIAGEN DNeasy® Blood & Tissue Kit (Qiagen, USA) materials, sourced from the manufacturer's website. DNA was extracted from each sample in duplicates and stored at-20 °C. Frozen muscle tissue stocks were kept at -20 °C for redundant tissue sample sources. Working muscle tissue samples were stored as (less than or equal to) 25 mg masses in 1.5 mL microcentrifuge tubes at -20 °C. Quantification of extracted DNA was carried out using the Promega QuantiFluor® dsDNA kit and Promega Quantus™

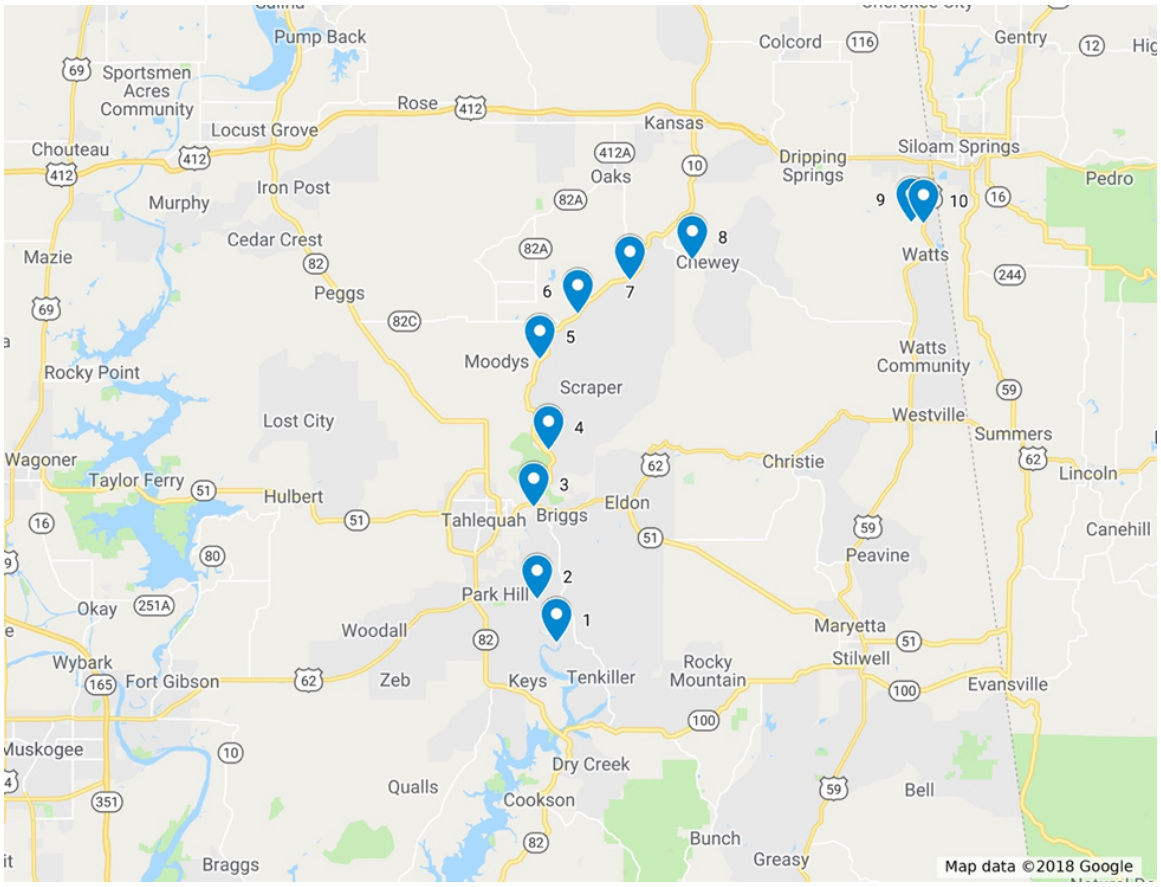

Figure 1. Sample collection map along the Illinois River.

fluorometer (Promega, USA). All muscle tissue DNA samples (excluding human) were diluted to the lowest occurring concentration amongst the highest yielding samples from the duplicates. This led to a working DNA template concentration of 2 ng/uL, except the human sample that was extracted from saliva.

3.3 River Sample DNA Extraction

Environmental DNA (eDNA) was extracted from each 500 mL river water samples with the Qiagen DNeasy® PowerWater® Kit (Qiagen, USA) after vacuum filtration with a 0.45-micron filter. The -80 °C stored filter papers were transferred to individual Qiagen BeadTubes with 1.0 mL of Solution 1, then the tubes were horizontally fastened to a modified vortex plate. The plate was modified by drilling holes equidistant around the middle radius in pairs in order for a twist-tie to wrap around the bottom of the tube to prevent excessive movement (Appendix G). Additionally, the lids were taped down to the plate with Fisher Scientific Blue Tape. The remaining steps of the extraction process were carried out as described in the manufacturer's protocol manual. Quantification of extracted DNA was carried out using the Promega QuantiFluor® dsDNA kit and Promega Quantus™ fluorometer (Promega, USA).

3.4 Primer Design

Primers were designed or incorporated from previous research (Vuong et al., 2013) for the specificity of target organisms (bison, cow, chicken, duck, goose, human, and sheep) and universal specificity in order to detect multiple organisms in one standard PCR assay. When designing the primer sets, certain parameters were set as guidelines for cross-compatibility between standard PCR and qPCR. In order to achieve this, amplicon sizes were to be kept under 200 basepairs; self-complementarity between primers at

minimum values; and GC-rich primer sequences, as well as GC-rich amplicon composition, when possible.

Accession numbers for target animals were based on the subspecies or breed likely to be involved in fecal contamination of water sources. After designating suitable accession numbers for each target organism from NCBI BLASTn suite, the accession number of interest was entered into the PCR template textbox on the query page NCBI's Primer-BLAST online software. The following parameters were changed from the default settings: 200 maximum PCR product size, user guided search mode, refseq representative genomes database, exclude uncultured/environmental sample sequences, and clear the organism specificity text box. Primer sets were compared in MEGA 7.0 software against ClustalW aligned full mitochondrial genomes of target animals of interest, as well as domestic and wildlife animals in order to rigorously test specificity. Additionally, the PCR binding sites and subsequent PCR product sequences were compared between the animal breed/subspecies of interest and other breeds/subspecies closely related to the animal of interest in order to determine the variability between these animals' genomes at sites of interest. Once suitable primers sets were decided, stocks were ordered from Invitrogen in 100 mM concentrations and subsequently diluted to a working stock of 25 mM in 100 µL aliquots.

Initial verification of primer sets was performed on the extracted muscle and saliva tissue of the following positive controls: bison, cow, chicken, duck, goose, human, pig, and sheep; negative controls: mussels, salmon, shrimp, trout, and turkey. Gel agarose electrophoresis of the polymerase chain reaction (PCR) products indicated specificity for

the positive controls. If a primer set was unsuccessful in amplifying the intended target(s), a redesign of the primer set was required.

3.5 PCR

PCR was used as a qualitative method to amplify regions of the target species mitochondrial DNA. Universal and specific primers were used and developed based on the species of interest and the most relevant accession number available from NCBI BLASTn database. Each PCR reaction mixture volume totaled to 50 µL using a chemical mastermix consisting of 25 mM $MgCl_2$ (Promega, USA); 5X Colorless GoTaq® Flexi Buffer (Promega, USA); 20 mg/mL BSA; 10 mM dNTP Mix (Promega, USA); 5u/µL GoTaq® G2 Flexi DNA Polymerase (Promega, USA); and molecular grade water to final volume. Primers concentrations were 25 mM. Thermocycling conditions were as follows: 95 °C for 5 minutes followed by 34 cycles of 95 °C for 30 seconds, 58 °C for 30 seconds, and 72 °C for 1 minute, then one cycle at 72 °C for 5 minutes. PCR products were loaded into 1.5% agarose gel electrophoresis, ran on a Mupid®-exU electrophoresis system (Helixx, CA), and stained with Invitrogen 1X SYBR Safe DNA Gel Stain (Fisher Scientific, USA). Gels were analyzed using Bio-Rad Image Lab™ Software in a Bio-Rad Gel Doc™ XR+ Gel Documentation System (BIO-RAD, USA).

3.6 Clone Libraries

PCR cloning methods were carried out using the Promega pGEM®-T Easy Vector System (Promega, USA) to the manufacturer's specifications. Following verification of primer set specificities, standards were prepared from electrocompetent cell transformation with the PCR amplicons as the plasmid insert. The cloning process began with same-day PCR products that were purified with a MoBio UltraClean® PCR Clean-Up Kit (MoBio, USA). Verification of the amplification and purification of the PCR

products was carried out by loading the samples into 1.5% agarose gel electrophoresis, ran on a Mupid®-exU electrophoresis system (Helixx, CA), and stained with Invitrogen 1X SYBR Safe DNA Gel Stain (Fisher Scientific, USA). Gels were analyzed using Bio-Rad Image Lab™ Software in a Bio-Rad Gel Doc™ XR+ Gel Documentation System (BIO-RAD, USA). Upon confirmation of the PCR products, the Promega pGEM®-T Easy Vector System (Promega, USA) to the manufacturer's specifications. ChromoMax™ IPTG/X-Gal (Fisher BioReagents, USA) and FisherSci Ampicillin (Fisher Scientific, USA) were used alongside the Promega cloning kit.

Analysis of transformants via plasmid extractions were performed after harvesting the electrocompetent JM109 *E. coli* cells by aliquoting 1.5 mL of liquid cultures and centrifuging the cells at 3,000 x g for 5 minutes and carefully discarding the supernatant by pipetting. This aliquot/centrifuge step was repeated two more times, for a total of 4.5 mL of harvested liquid cultures. Plasmid extractions were performed following the manufacturer's protocol for the Qiagen QIAprep Spin Miniprep Kit. DNA concentrations of the plasmids were quantified using the Promega QuantiFluor® dsDNA kit and Promega Quantus™ fluorometer (Promega, USA). Insertion of the amplicon into the vector plasmid was verified by standard PCR with the associated primer set to the gene sequence of interest. Gel analysis was then performed on the plasmid template PCR product in a 1.5% agarose gel stained with Invitrogen 1X SYBR Safe DNA Gel Stain at 100V for 20 minutes. Gels were analyzed using Bio-Rad Image Lab™ Software in a Bio-Rad Gel Doc™ XR+ Gel Documentation System (BIO-RAD, USA). Long term storage of transformants were prepared from the liquid LB cultures at the plasmid extraction step, which 1.0 mL of the liquid cultures were aliquoted into a 2.0 mL Corning™ cryogenic

storage tube with 1.0 mL of 30% glycerol. After the cell/glycerol mixture were mixed gently by pipetting, the tubes were stored at -80 °C.

3.7 Sequencing

The clone library plasmid DNA was amplified using M13 primers, and the subsequent PCR products were purified with a MoBio UltraClean® PCR Clean-Up Kit (MoBio, USA) to meet OSU's DNA and Protein Core facility. PCR products were Sanger sequenced bi-directionally using M13 primers. 10μL of the cleaned M13 products were used as the template for sequencing. The M13 PCR products were loaded into 1.5% agarose gel electrophoresis, ran on a Mupid®-exU electrophoresis system (Helixx, CA), and stained with Invitrogen 1X SYBR Safe DNA Gel Stain (Fisher Scientific, USA). Gels were analyzed using Bio-Rad Image Lab™ Software in a Bio-Rad Gel Doc™ XR+ Gel Documentation System (BIO-RAD, USA). Sequences obtained from OSU's DNA and Protein Core facility can be found under Appendix F in FASTA format.

3.8 Quantitative PCR

Quantitative PCR (qPCR) of the clone library plasmid DNA was performed to develop a standard curve for detection limits. Standards for each clone library were prepared from known concentrations of clone library plasmid extracts containing the mitochondrial gene of interest. Serial dilution stocks starting with the initial plasmid DNA concentration were constituted as a 1:10 dilution ratio with EDTA for a total of 9 dilutions. Each clone library dilution set was analyzed with qPCR in duplicate, the duplicates were log 10 transformed and averaged. The clone library dilution samples were amplified as the DNA template with the same species-relevant primers as standard

PCR, a chemical mastermix consisting of BIO-RAD iTaq™ Universal SYBR® Green Supermix, 20 mg/mL BSA, and PCR-grade water to a reaction volume of 10 µL. The unknown river samples were ran with the standard plasmid DNA dilutions to quantify the detection limits of mitochondrial DNA present in the water samples. Thermocycling conditions and mastermix constituents were followed as outlined for the standards protocol.

CHAPTER IV

FINDINGS

4.1 Template DNA Preparation

Muscle tissue DNA was extracted using QIAGEN DNeasy® Blood& Tissue Kit (Qiagen, USA). The resultant DNA concentrations (ng/µL) of the tissue samples for bison, chicken, cow, duck, goose, and sheep were: 13E+3, 5.4E+3, 7.1E+3, 18E+3, 9.3E+3, and 13E+3, respectively. Saliva tissue DNA was extracted using a user-defined protocol with QIAGEN DNeasy® Blood & Tissue Kit materials. The resultant DNA concentration (ng/µL) of the saliva sample was 1.7E-2. Table 1 shows the comprehensive list of extracted DNA concentrations of muscle and saliva samples, including negative control samples and redundant muscle tissue samples. The extraction concentrations indicate a consistent DNA extraction product for the muscle tissue samples.

4.2 PCR

Standard PCR was performed on muscle tissue DNA extractions, saliva DNA extractions, and clone library plasmid DNA extractions. Initial verification of primer sets was performed on the extracted muscle and saliva tissue of the following positive controls: bison, chicken, cow, duck, goose, human, pig, and sheep; negative controls: mussels, salmon, shrimp, trout, and turkey. Each sample was amplified with all primer sets to test for amplification specificity. Gel agarose electrophoresis of the PCR products

Table 1. Target animal, source of DNA, significance to the study, and the concentrations of each sample extraction.

Target Animal	DNA Origin	Purpose for Study	Concentration (ng/µL)
Bison	Muscle	(+) Control	1.3E+04
Chicken	Muscle	(+) Control	5.4E+03
Cow	Muscle	(+) Control	7.1E+03
Duck	Muscle	(+) Control	1.8E+04
Goose	Muscle	(+) Control	9.3E+03
Human	Saliva	(+) Control	1.7E-02
Pig	Muscle	(+) Control	2.9E+03
Sheep	Muscle	(+) Control	1.3E+04
Mussels	Muscle	(-) Control	4.2E+03
Salmon	Muscle	(-) Control	5.7E+03
Shrimp	Muscle	(-) Control	3.2E+03
Trout	Muscle	(-) Control	2.1E+03
Turkey	Muscle	(-) Control	3.6E+03
Bison	Muscle	Duplicate Extraction	1.2E+04
Bovine	Muscle	Duplicate Extraction	3.6E+03
Chicken	Muscle	Duplicate Extraction	5.1E+03
Duck	Muscle	Duplicate Extraction	1.6E+04
Goose	Muscle	Duplicate Extraction	5.7E+03
Mussels	Muscle	Duplicate Extraction	4.0E+03
Pig	Muscle	Duplicate Extraction	2.1E+03
Salmon	Muscle	Duplicate Extraction	4.0E+03
Sheep	Muscle	Duplicate Extraction	1.3E+04
Shrimp	Muscle	Duplicate Extraction	1.7E+03
Trout	Muscle	Duplicate Extraction	1.2E+03
Turkey	Muscle	Duplicate Extraction	3.5E+03

provided visual verification of successful target amplification. A comprehensive set of gel images are located in Appendix B, displaying muscle tissue DNA and saliva DNA amplifications indicating primer specificities.

The following primers were designed to replace the first round of primers that either detected the target organism and a non-target organism, only detected non-target organisms, or did not detect any organisms in the presence of template DNA (target in parenthesis): Bis15914F/Bis1611R, Bov1482F/Bov16116R, Duck14644F/Duck14788R, Goose4728F/Goose4861R, and Hum15425F/Hum15594R. Table D3 shows the final primer sets used for the PCR amplification of targeted animals (Appendix D). Appendix C contains figures indicating the target gene region for each primer set designed in this study specific to the animal of interest.

Clone library plasmid DNA was PCR amplified using M13 sequencing primers for quality control measures in order to test for successful transformations. The M13 products were quantified to meet OSU's DNA and Protein Core facility template concentration requirements. Quantification data for the M13 PCR products are found in Table 2. PCR products from this experiment are the expected amplicon size of the target-specific primer set plus the plasmid sequence basepairs before and after M13 binding sites. The additional basepair lengths are dependent upon the vector used in the ligation process, which for this study, the Promega pGEM® -T Easy Vector Plasmid was used. This plasmid would add 264 bp to the specific primer amplicon size. Table 2 shows the expected amplicon size of the M13 amplified PCR products and concentrations of the cleaned PCR products of the M13 amplified vector with inserts. Visual analysis of the M13 PCR products was carried out as a quality control step before sending the PCR

Table 2. Target animals, primer sets used for clone libraries, primer pair amplicon size, amplicon size with sequencing primers, and concentration of cleaned M13 PCR products.

Target Animal	Primer	Amplicon Size (bp)	M13 Amplicon Size (bp)	Concentration (ng/µL)
Bison	Bis15914F Bis16116R	202	466	20
Chicken	Ckmitol-G Ckmitol-D	565	829	50
Cow	Bov1482F Bov1571R	90	354	19
Duck	Duck14644F Duck14788R	121	385	14
Goose	Goose4728F Goose4861R	134	398	4.6
Human	Hum15425F Hum15594R	170	434	19
Pig	Pomito4-G Pomito4-D	723	987	6.8
Sheep	OvmitoN2-G OvmitoN2-D	370	634	40

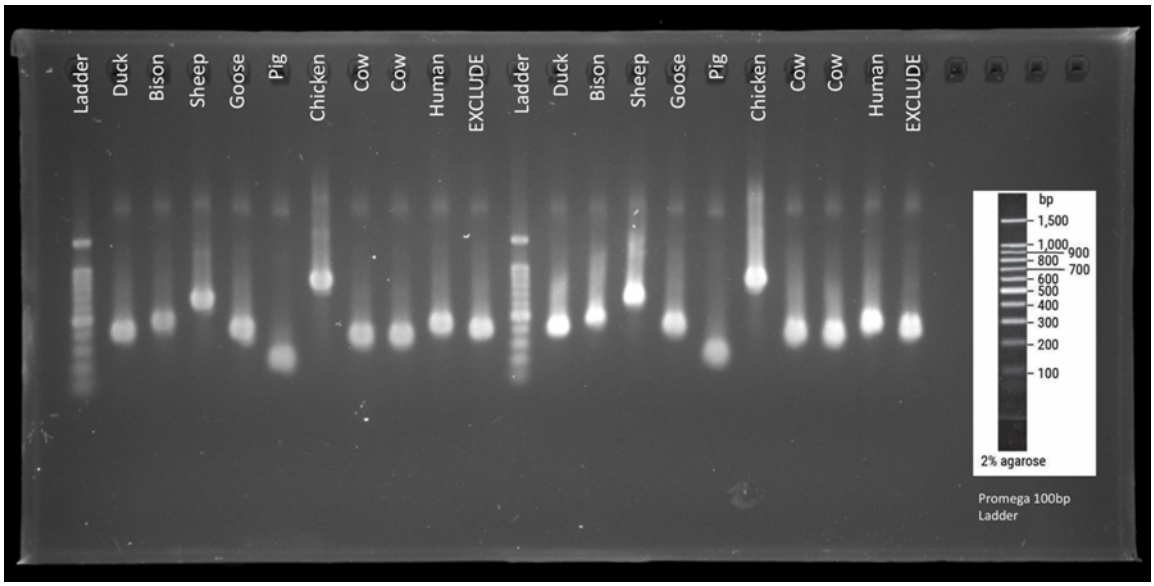

Figure 2. Agarose gel electrophoresis of M13 amplified clone library plasmid DNA.

products out for sequencing. Figure 1 shows the M13 PCR products before the clean-up step

Figure 1 indicates all clone library M13 PCR products are of an expected length, with the exception of the pig dataset, which should be ~987 bp but is present at ~250 bp. It is thought that the failure was due to the initial primer set selection. New primers have been designed for the pig sample set and remain to be tested.

4.3 Clone Library Standards

Dilution standards of the clone library plasmid DNA were made from starting concentrations of plasmid DNA extracted from the clone library transformants. A 1:10 dilution ratio was used to create the clone library standards used for PCR and qPCR assays. Concentrations of the plasmid DNA are listed in Table 3.

Standard PCR experiments were ran with the diluted plasmid DNA samples and their respective primer sets. PCR products were subsequently analyzed on agarose gel electrophoresis. Bison, chicken, cow, duck, goose, human, pig, and sheep all had limits of detection (LoD) corresponding to the following standard dilution sets: -4, -4, 0, -4, -6, -6, 0 and -4. Table 4 shows the lowest concentration for each target animal, indicating the specificity when analyzing data using standard PCR and agarose gel electrophoresis qualitative methods. Agarose gels that provided qualitative verification of the lowest detection limits are found in Appendix B (Figures B18-B21).

4.4 Sequencing

The clone library plasmid DNA that was transformed with a target animal's mitochondrial DNA was sequenced and analyzed at the OSU DNA and Protein Core facility. Sequencing results indicated that the transformations were successful for seven

Table 3. Concentration of Transformant Plasmid DNA from Quantification of Clone Library

Concentration of Transformant Plasmid DNA (copies/µL)										
Animal	**0**	**-1**	**-2**	**-3**	**-4**	**-5**	**-6**	**-7**	**-8**	**-9**
Bison	5.0E+10	5.0E+09	5.0E+08	5.0E+07	5.0E+06	5.0E+05	5.0E+04	5.0E+03	5.0E+02	5.0E+01
Chicken	6.4E+10	6.4E+09	6.4E+08	6.4E+07	6.4E+06	6.4E+05	6.4E+04	6.4E+03	6.4E+02	6.4E+01
Cow	5.7E+10	5.7E+09	5.7E+08	5.7E+07	5.7E+06	5.7E+05	5.7E+04	5.7E+03	5.7E+02	5.7E+01
Duck	2.9E+10	2.9E+09	2.9E+08	2.9E+07	2.9E+06	2.9E+05	2.9E+04	2.9E+03	2.9E+02	2.9E+01
Goose	5.4E+10	5.4E+09	5.4E+08	5.4E+07	5.4E+06	5.4E+05	5.4E+04	5.4E+03	5.4E+02	5.4E+01
Human	4.4E+10	4.4E+09	4.4E+08	4.4E+07	4.4E+06	4.4E+05	4.4E+04	4.4E+03	4.4E+02	4.4E+01
Pig	5.2E+10	5.2E+09	5.2E+08	5.2E+07	5.2E+06	5.2E+05	5.2E+04	5.2E+03	5.2E+02	5.2E+01
Sheep	3.3E+10	3.3E+09	3.3E+08	3.3E+07	3.3E+06	3.3E+05	3.3E+04	3.3E+03	3.3E+02	3.3E+01

Table 4. Comparison of Clone Library Limits of Detection via Gel Analysis & qPCR

Target	Limits of Detection (Copies/µL)	
	PCR	**qPCR**
Bison	5,000,000	-
Chicken	6,439,500	644
Cow	-	-
Duck	2,941,626	29,416
Goose	535,255	5,353
Human	43,743	437
Pork	-	-
Sheep (Nested)	3,256,007	3,256

of the eight sequences. When the sequencing results were aligned using ClustalW in MEGA 7.0, all of the sequences were the length expected, except for the pig sample. The identity of the sequences was high, averaging 98.7%. The data indicating the identities of related subspecies/breeds of animals of interest to the sequence data from the clone library standards can be found in Appendix I. Sequencing data is shown in Table 5 providing data for bison, cow, chicken, duck, goose, human, and sheep; data for the pig standard is not provided due to insufficient results. It is thought that the failure was due to the initial primer set selection. New primers have been designed for the pig sample set and remain to be tested.

4.5 qPCR

Quantitative PCR (qPCR) was used to quantify mitochondrial DNA target regions in standards prepared from clone library plasmid DNA and unknown samples collected from the Illinois River. Primers used for qPCR analysis were the same sets used for standard PCR assays. Table D3 in Appendix D shows the final primer sets used for the qPCR amplification of standards and unknown DNA samples.

Development of standard curves from the qPCR data were unsuccessful for the bison, cow, and pig samples. This is thought to be due to the primers not being optimized for qPCR assays, or the chain of custody for the clone library plasmid DNA was mishandled. Mismanagement of the clone library plasmid DNA is indicated by the sequenced PCR product of the pig sample resulting in an insufficient sequence, but the agarose gel (Figure 1) indicating an insert in one of the clone libraries, albeit an incorrect insert. The mishandling of the clone library samples is primarily thought to have occurred with the pig sample, as the sequencing results indicated the pig data set was not sequenced with

Table 5. Clone Library Sequencing Performance Data

Target	Primer Name	Designed Amplicon Size (bp)	Actual Amplicon Size (bp)	Identity (%)
Bison	Bis15914F Bis16116R	202	202	99
Chicken	Ckmitol-G Ckmitol-D	565	565	100
Cow	Bov1482F Bov1571R	90	90	100
Duck	Duck14644F Duck14788R	121	121	100
Goose	Goose4728F Goose4861R	134	134	93
Human	Hum15425F Hum15594R	170	170	100
Pork	Pomito4-G Pomito4-D	723	-	-
Sheep	OvmitoN2-G OvmitoN2-D	370	371	99

the designed sequence profile, as well as the clone library plasmid DNA concentration gradient agarose gel indicating no amplification (Figure B20 in Appendix B). The bison and cow qPCR results suggested that there were overwhelming primer-dimer formations occurring during the reaction. It is likely that the primers exhibited excessive self-complementarity, resulting in repetitive amplifications of incorrect sequences.

Standard curve datasets for chicken, duck, goose, human, and sheep were successfully developed with the clone library plasmid DNA and their respective primer sets. The chicken standard curve data showed detection limits as low as 6.4E+03 copies/µL with the -9 dilution standards below detection limit. Amplification efficiencies were 76.6% for the dilution sets -1 through -8. Initial qPCR assay of the duck sample standard DNA resulted in large amounts of primer-dimer formations in the lower dilution range (-6 through -9 dilutions). Following this, a procedure to optimize the duck standards was carried out by amplifying the clone library template on a primer concentration gradient of 25 mM, 18.75 mM, and 12.5 mM at an increased extension temperature of 61 °C from 59 °C. This process did not prove successful to reduce primer-dimer formations or amplification efficiencies. Duck standard curve data showed detection limits as low as 290,000 copies/µL with the -7 dilution standards below detection limit. Amplification efficiencies were 84% for the dilution sets -1 through -6. Goose standard curve data showed detection limits as low as 54,000 copies/µL with the -8 dilution standards below detection limit. Amplification efficiencies were 78% for the dilution sets -1 through -7. Human standard curve data showed detection limits as low as 4,400 copies/µL with the -9 dilution standards below detection limit. Amplification efficiencies were 78% for the dilution sets -1 through -7. Human standard curve data showed detection limits as low as

4,400 copies/µL with the -9 dilution standards below detection limit. Amplification efficiencies were 84% for the dilution sets -1 through -8. Sheep standard curve data showed detection limits as low as 33,000 copies/µL with the -8 dilution standards below detection limit. Amplification efficiencies were 60.8% for the dilution sets -1 through-7. Standard curve plots can be found in Appendix E.

4.6 Detection & Quantification of Mitochondrial DNA in the Illinois River

qPCR was used to quantify several mitochondrial genes from animals of interest that might be known contributors to fecal contamination in the Illinois River. The target animal, primers, and sequences for qPCR approaches are listed in Table 6. For example, primers Hum15425F//Hum15594R are specific for Homo sapiens and primers Ckmitol-G//Ckmitol-D are specific for Gallus gallus (chicken).

River samples were collected downstream (southern region) of the Illinois river and collected progressively travelling northbound. Chest-high waders were worn in order to collect samples at the river midline, or at chest depth where the river depth was too great. Sampling capture depths were a composite of three inches below the surface and water at the surface. Three rinses with the river water were done before collecting the final sample (500 mL) and stored on ice until same-day vacuum filtration. Once extracted, the DNA recovered from the samples were qPCR amplified alongside the clone library dilution standards, in order to measure detection limits and quantities of target DNA in the water at a given sampling location. Each sample location was tested for the presence of mitochondrial DNA from bison, chicken, cow, duck, goose, human, pig, and sheep. As indicated in the previous section, qPCR data for bison, cow, and pig were unavailable, as the standards from the clone library were unsuccessfully

Table 6. Target animal, primer set, and corresponding sequences used in qPCR amplification of unknown river samples

Target	Primer Name	Primer Sequence
Bison	Bis15914F	GTCACTCACCCCCAAAATGC
	Bis16116R	TCTACCCTTGGCAACATGCA
Chicken	Ckmitol-G	ACCCTATTTGACTCCCTCAA
	Ckmitol-D	ATGTCGACCAGGGGTTTATG
Cow	Bov1482F	CGACTAAACAACCAAGATAG
	Bov1571R	TTCTCTATAGCGCCGTACTT
Duck	Duck14644F	AACCTACTTAGGATCTTTCGCCC
	Duck14788R	TATTAGAGGGTGCGGGAAGGT
Goose	Goose4728F	AGTACTAAATGCCACCCTGA
	Goose4861R	CAATTGCTATGGCTGCTGG
Human	Hum15425F	GACGCCCTCGGCTTACTTC
	Hum15594R	ACGGATCGGAGAATTGTGTAGG
Pig	Pomito4-G	CCCATTATCTACACTACCCTTATC
	Pomito4-D	TTAGGCTTGTGATGACGGGTAT
Sheep	OvmitoN2-G	TACACTGTTACAGGCATCAG
	OvmitoN2-D	CGTGAAGTTAGTTAGGAGAGTA

transformed at the ligation step or the primers were not optimized for qPCR. Chicken was detected at sample location #1 at an approximate concentration of 6,497 copies/µL, which is slight above the threshold of detection for the chicken. Duck was not detected at concentrations in the river that exceeded the lowest detection limit of 290,000 copies/µL. Goose was detected at collection sites #6 and #7 at concentrations of 60,115 and 10,642 copies/µL. The latter detection was quantified below detection limits, however amplification was still recorded and occurred at the correct melting temperature, indicating the correct amplification of target sequences.

Human mtDNA was detected at multiple sample locations: #1-4 and #6-10. The average concentration of human mtDNA found was 129,871 copies/µL. The collection sites with the greatest concentrations, found at an order of magnitude greater than all other sample sites, were locations #2 & #8. Location #2 was directly downstream from a housing development that appears to have approximately thirty single family homes. Due to the rural location of this housing development, septic systems are the likely methods for wastewater treatment. Soil conditions in the area are likely rocky and have much variability with expansion and contraction events with seasonal changes, causing septic tanks that are embedded in the soil to develop fractures, thus introducing human waste into the water table. Location #8 concentrations also appear to be driven by faulty septic systems, as there are residential plots all along the coastline of the river and no apparent wastewater treatment plants upstream of the collection site or downstream from locations #9 and #10. Sheep mtDNA was detected at sample site #10 at 28,452 copies/µL. Table H1 in Appendix H provides an overview of the mtDNA concentrations detected at each sampling location and the respective target animal.

CHAPTER V

CONCLUSION

Discussion

Water affected by fecal contamination is likely to result in diminished water quality by introducing pathogens and/or excessive nutrients into previously unaffected waterways. Development of a variety of methods is important in order to effectively detect and characterize the source at which the pollution is migrating from. This study is a progression of previously established methods (Martellini et al., 2005; Caldwell et al., 2007; Kortbaoui et al., 2009; Vuong et al., 2013). The aim of this work was to develop a comprehensive method from primer development to detection and quantification of contaminant sources in unknown environmental samples.

Results of this study can be concluded with mixed success, as 6 of the 8 clone library standards datasets indicated amplification of the intended DNA sequences in standard PCR assays; 5 of the 8 clone library standards datasets indicated amplification and quantification of the intended DNA sequences in qPCR assays. Sequencing results indicated that one of the clone library standards sets (pig) did not successfully undergo a proper transformation, likely to have occurred during the ligation step. Insufficient primer specificity is thought to be the other factor in the unsuccessful results of the remaining clone library standards (bison & cow). Further refinement of these primer sets and clone

library standards are currently being worked on.

Detection of target species was identified at 9 out of 10 sample collection sites. Chicken mtDNA was quantified above detection limits at site #1. Goose mtDNA was quantified above detection limits at collection sites #6 and #7. Human mtDNA was quantified above detection limits at all collection sites, with the exception of site #5. The abundance of human mtDNA is attributed to the amount of residential development near the river shoreline and the likelihood of septic systems as the treatment technology for wastewater, as well as the proximity of a city with an approximate population of 16,000 people within 10 miles (Northwest) of the sampling location and within the hydrological basin.

Our methods concluded:

- PCR and qPCR methods with specific primer development will accurately select for target animal mtDNA and not produce false positives by amplifying other environmental mtDNA
- Clone library dilution standards are an effective method for determining the presence and relative quantity of mtDNA in an unknown sample
- qPCR provided greater detection sensitivity of standard and unknown than PCR methods
- The mtDNA methods suggest that the Illinois river watershed in eastern Oklahoma is impacted by fecal contamination by human wastewater through most

of the river system, and is impacted by chicken waste near the upper reaches of Tenkiller Ferry Lake.

REFERENCES

1. Andréasson, H., Gyllensten, U., & Allen, M. (2002). Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. Biotechniques, 33(2), 402-411.
2. Baker-Austin, C., Rangdale, R., Lowther, J., & Lees, D. N. (2010). Application of mitochondrial DNA analysis for microbial source tracking purposes in shellfish harvesting waters. Water Science and Technology, 61(1), 1.
3. Blanch, A. R., Belanche-Muñoz, L., Bonjoch, X., Ebdon, J., Gantzer, C., Lucena, F., ... & Mocé, L. (2006). Integrated analysis of established and novel microbial and chemical methods for microbial source tracking. Applied and environmental microbiology, 72(9), 5915-5926.
4. Blann, K. L., Anderson, J. L., Sands, G. R., & Vondracek, B. (2009). Effects of agricultural drainage on aquatic ecosystems: a review. Critical reviews in environmental science and technology, 39(11), 909-1001.
5. Bonjoch, X., Balleste, E., & Blanch, A. R. (2004). Multiplex PCR with 16S rRNA gene-targeted primers of Bifidobacterium spp. to identify sources of fecal pollution. Applied and Environmental Microbiology, 70(5), 3171-3175.
6. Buerge, I. J., Poiger, T., Müller, M. D., & Buser, H. R. (2003). Caffeine, an anthropogenic marker for wastewater contamination of surface waters. Environmental science & technology, 37(4), 691-700.

7. Buerge, I. J., Poiger, T., Müller, M. D., & Buser, H. R. (2006). Combined sewer overflows to surface waters detected by the anthropogenic marker caffeine. Environmental science & technology, 40(13), 4096-4102.
8. Caldwell, J. M., Raley, M. E., & Levine, J. F. (2007). Mitochondrial multiplex real-time PCR as a source tracking method in fecal-contaminated effluents. Environmental science & technology, 41(9), 3277-3283.
9. Carson, C.A., Shear, B.L., Ellersieck, M.R., Asfaw,, A., 2001. Identification offecal Escherichia coli from humans and animals by ribotyping. Applied and Environmental Microbiology 67, 1503–1507.
10. Cox, P., Griffith, M., Angles, M., Deere, D., & Ferguson, C. (2005). Concentrations of pathogens and indicators in animal feces in the Sydney watershed. Applied and environmental microbiology, 71(10), 5929-5934.
11. Crane, S. R., Moore, J. A., Grismer, M. E., & Miner, J. R. (1983). Bacterial pollution from agricultural sources: A review. Transactions of the ASAE, 26(3), 858-0866.
12. Dickerson Jr, J. W., Hagedorn, C., & Hassall, A. (2007). Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods. Water Research, 41(16), 3758-3770.
13. Dixon, L. K., Taylor, H. M., Staugler, E., & Scudera, J. O. (2005). Development of a fluorescence method to detect optical brightners in the presence of varying concentrations of fluorescent humic substances: Identifying regions influenced by OSTDS in the estuarine waters of Charlotte Harbor. Mote Marine Laboratory, Sarasota, FL

14. Dombek, P. E., Johnson, L. K., Zimmerley, S. T., & Sadowsky, M. J. (2000). Use of repetitive dna sequences and the pcr to differentiateescherichia coli isolates from human and animal sources. Applied and Environmental Microbiology, 66(6), 2572-2577.
15. Eijkman, C. (1904). The fermentation test at 46 C as means of testing potable water. Zentr. Bakteriol. Parasitenk., I, Orig.(Ger.), 37, 742.
16. Farag, A.M., Goldstein, J.N., Woodward, D.F., Samadpour, M., 2001. Water quality in three creeks in the backcountry of Grand Teton National Park, USA. Journal of Freshwater Ecology 16, 135–143.
17. Farnleitner, A. H., Kreuzinger, N., Kavka, G. G., Grillenberger, S., Rath, J., & Mach, R. L. (2000). Comparative analysis of denaturing gradient gel electrophoresis and temporal temperature gradient gel electrophoresis in separating Escherichia coli uidA amplicons differing in single base substitutions. Letters in applied microbiology, 30(6), 427-431.
18. Feachem, R. (1975). An improved role for faecal coliform to faecal streptococci ratios in the differentiation between human and non-human pollution sources. Water research, 9(7), 689-690.
19. Field, K. G., & Samadpour, M. (2007). Fecal source tracking, the indicator paradigm, and managing water quality. Water research, 41(16), 3517-3538.
20. Geldreich, E. E., & Litsky, W. (1976). Fecal coliform and fecal streptococcus density relationships in waste discharges and receiving waters. Critical reviews in environmental science and technology, 6(4), 349-369.
21. Gerber, A. S., Loggins, R., Kumar, S., & Dowling, T. E. (2001). Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes?. Annual review of genetics, 35(1), 539-566.

22. Griffith, J. F., Weisberg, S. B., & McGee, C. D. (2003). Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. Journal of Water and Health, 1(4), 141-151.

23. Guarner, F., & Malagelada, J. R. (2003). Gut flora in health and disease. The Lancet, 361(9356), 512-519.

24. Hagedorn, C., Robinson, S. L., Filtz, J. R., Grubbs, S. M., Angier, T. A., & Reneau, R. B. (1999). Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Applied and Environmental Microbiology, 65(12), 5522-5531.

25. Hagedorn, C., Crozier, J. B., Mentz, K. A., Booth, A. M., Graves, A. K., Nelson, N. J., & Reneau Jr, R. B. (2003). Carbon source utilization profiles as a method to identify sources of faecal pollution in water. Journal of applied microbiology, 94(5), 792-799.

26. Hager, M.C., 2001a. Detecting bacteria in coastal waters. Part 1. The Journal for Surface Water Quality Professionals 2 (3), 16–25.

27. Hager, M.C., 2001b. Detecting bacteria in coastal waters. Part 2. The Journal for Surface Water Quality Professionals 2 (4).

28. Hartel, P.G., Summer, J.D., Hill, J.L., Collins, J.V., Entry, J.A., Segars,, W.I., 2002. Geographic variability of Escherichia coli ribotypes from animals in Idaho and Georgia. Journal of Environmental Quality 31 (4), 1273–1278.

29. Iyengar, V., Albaugh, G. P., Lohani, A., & Nair, P. P. (1991). Human stools as a source of viable colonic epithelial cells. FASEB journal: official publication of the Federation of American Societies for Experimental Biology, 5(13), 2856.

30. Jagals, P. W. O. K., Grabow, W. O. K., & De Villiers, J. C. (1995). Evaluation of indicators for assessment of human and animal faecal pollution of surface run-off. Water Science and Technology, 31(5-6), 235-241.
31. Kortbaoui, R., Locas, A., Imbeau, M., Payment, P., & Villemur, R. (2009). Universal mitochondrial PCR combined with species-specific dot-blot assay as a source-tracking method of human, bovine, chicken, ovine, and porcine in fecal-contaminated surface water. Water research, 43(7), 2002-2010.
32. Leeming, R., Ball, A., Ashbolt, N., & Nichols, P. (1996). Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. Water research, 30(12), 2893-2900.
33. Lynch, P. A., Gilpin, B. J., Sinton, L. W., & Savill, M. G. (2002). The detection of Bifidobacterium adolescentis by colony hybridization as an indicator of human faecal pollution. Journal of applied microbiology, 92(3), 526-533.
34. Mara, D. D., & Oragui, J. I. (1983). Sorbitol-fermenting bifidobacteria as specific indicators of human faecal pollution. Journal of applied bacteriology, 55(2), 349-357.
35. Mara, D. D., & Oragui, J. (1985). Bacteriological methods for distinguishing between human and animal faecal pollution of water: results of fieldwork in Nigeria and Zimbabwe. Bulletin of the World Health Organization, 63(4), 773.
36. Martellini, A., Payment, P., & Villemur, R. (2005). Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Research, 39(4), 541-548.

37. Mathes, M. V., O'Brien, T. L., Strickler, K. M., Hardy, J. J., Schill, W. B., Lukasik, J., ... & Fenger, T. L. (2007). Presumptive sources of fecal contamination in four tributaries to the New River Gorge National River, West Virginia, 2004. U. S. Geological Survey.

38. Myers, D.N., Stoeckel, D.M., Bushon, R.N., Francy, D.S., and Brady, A.M.G., 2014, Fecal indicator bacteria (ver. 2.1): U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A7, section 7.1, May 2014, accessed __ *October 2018*__ from http://pubs.water.usgs.gov/twri9A/.

39. Moore, D.F., Harwood, V.J., Ferguson, D.M., Lukasik, J., Hannah, P., Getrich, M., Brownell, M., 2005. Evaluation of antibiotic resistance analysis and ribotyping for identification of fecal pollution sources in an urban watershed. J. Appl. Microbiol. 99, 618–628.

40. Newton, R. J., VandeWalle, J., Borchardt, M. A., Gorelick, M. H., & McLellan, S. L. (2011). Lachnospiraceae and Bacteroidales alternative fecal indicators reveal chronic human sewage contamination in an urban harbor. Applied and environmental microbiology, AEM-05480.

41. Nieman, J., & Brion, G. M. (2003). Novel bacterial ratio for predicting faecal age. Water science and technology, 47(3), 45-49.

42. Oklahoma Department of Environmental Quality. (2017). *Public Water Supply Program Annual Compliance Report.* Oklahoma City, OK: (n.p.).

43. Parveen, S., Portier, K. M., Robinson, K., Edmiston, L., & Tamplin, M. L. (1999). Discriminant analysis of ribotype profiles of Escherichia coli for differentiating human and nonhuman sources of fecal pollution. Applied and Environmental Microbiology, 65(7), 3142-3147.

44. Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G. T., ... & Muniesa, M. (2005). Method for isolation of Bacteroides bacteriophage host strains suitable for tracking sources of fecal pollution in water. Applied and Environmental Microbiology, 71(9), 5659-5662.

45. Puig, A., Queralt, N., Jofre, J., & Araujo, R. (1999). Diversity of Bacteroides fragilis Strains in Their Capacity To Recover Phages from Human and Animal Wastes and from Fecally Polluted Wastewater. Applied and Environmental Microbiology, 65(4), 1772.

46. Resnick, I. G., & Levin, M. A. (1981). Assessment of bifidobacteria as indicators of human fecal pollution. Applied and environmental microbiology, 42(3), 433-438.

47. Ritter, Keith Solomon, Paul Sibley, Ken Hall, Patricia Keen, Gevan Mattu, Beth Linton, L. (2002). Sources, pathways, and relative risks of contaminants in surface water and groundwater: a perspective prepared for the Walkerton inquiry. *Journal of Toxicology and Environmental Health Part A*, *65*(1), 1-142.

48. Rowbotham, T. J., & Cross, T. (1977). Ecology of Rhodococcus coprophilus and associated actinomycetes in fresh water and agricultural habitats. Microbiology, 100(2), 231-240.

49. Samadpour, M., 2002. Microbial source tracking: principles and practice, US EPA Workshop on Microbial Source Tracking, 5–9 http://www.sccwrp.org/tools/workshops/source_tracking_agenda.html.

50. Samadpour, M., Roberts, M. C., Kitts, C., Mulugeta, W., & Alfi, D. (2005). The use of ribotyping and antibiotic resistance patterns for identification of host sources of Escherichia coli strains. Letters in applied microbiology, 40(1), 63-68.

51. Sauer, E. P., VandeWalle, J. L., Bootsma, M. J., & McLellan, S. L. (2011). Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water research, 45(14), 4081-4091.

52. Savill, M. G., Murray, S. R., Scholes, P., Maas, E. W., McCormick, R. E., Moore, E. B., & Gilpin, B. J. (2001). Application of polymerase chain reaction (PCR) and TaqMan™ PCR techniques to the detection and identification of Rhodococcus coprophilus in faecal samples. Journal of microbiological methods, 47(3), 355-368.

53. Schill, W. B., & Mathes, M. V. (2008). Real-time PCR detection and quantification of nine potential sources of fecal contamination by analysis of mitochondrial cytochrome b targets. Environmental science & technology, 42(14), 5229-5234.

54. Scott, T.M., Parveen, S., Portier, K.M., Rose, J.B., Tamplin, M.L., Farrah, S.R., Koo, A., Lukasik, J., 2003. Geographical variation in ribotype profiles of Escherichia coli isolates from humans, Swine, Poultry, Beef, and Dairy Cattle in Florida. Applied and Environmental Microbiology 69 (2), 1089–1092.

55. Sidhu, J. P. S., Ahmed, W., Gernjak, W., Aryal, R., McCarthy, D., Palmer, A., ... & Toze, S. (2013). Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. Science of the Total Environment, 463, 488-496.

56. Simmons, G.M., Waye, D.F., Herbein, S., Myers, S., Walker, E., 2000. Estimating nonpoint fecal coliform sources in Northern Virginia's four mile run watershed, in: Younos, T., Poff, J. (Eds.), Proceedings of the Virginia Water Research Symposium, 248–267 VWRRC Special Report SR-19-2000.

57. Simpson, J. M., Santo Domingo, J. W., & Reasoner, D. J. (2002). Microbial source tracking: state of the science. Environmental science & technology, 36(24), 5279-5288.

58. Sinton, L. W., Finlay, R. K., & Hannah, D. J. (1998). Distinguishing human from animal faecal contamination in water: a review. New Zealand journal of marine and freshwater research, 32(2), 323-348.

59. Smith, V. H., & Schindler, D. W. (2009). Eutrophication science: where do we go from here?. Trends in ecology & evolution, 24(4), 201-207.

60. Staley, C., Gordon, K. V., Schoen, M. E., & Harwood, V. J. (2012). Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. Applied and environmental microbiology, 78(20), 7317-7326.

61. Stoeckel, D. M., & Harwood, V. J. (2007). Performance, design, and analysis in microbial source tracking studies. Applied and Environmental Microbiology, 73(8), 2405-2415.

62. Souza, V., Rocha, M., Valera, A., & Eguiarte, L. E. (1999). Genetic structure of natural populations of Escherichia coli in wild hosts on different continents. Applied and Environmental Microbiology, 65(8), 3373-3385.

63. Standley, L. J., Kaplan, L. A., & Smith, D. (2000). Molecular tracers of organic matter sources to surface water resources. Environmental science & technology, 34(15), 3124-3130.

64. Tartera, C., & Jofre, J. (1987). Bacteriophages active against Bacteroides fragilis in sewage-polluted waters. Applied and environmental microbiology, 53(7), 1632-1637.

65. Tartera, C., Lucena, F., & Jofre, J. (1989). Human origin of Bacteroides fragilis bacteriophages present in the environment. Applied and environmental microbiology, 55(10), 2696-2701.

66. Tynkkynen, S., Satokari, R., Saarela, M., Mattila-Sandholm, T., & Saxelin, M. (1999). Comparison of Ribotyping, Randomly Amplified Polymorphic DNA Analysis, and Pulsed-Field Gel Electrophoresis in Typing of Lactobacillus rhamnosus and L. caseiStrains. Applied and Environmental Microbiology, 65(9), 3908-3914.

67. U.S. EPA (1976). Quality Criteria for Water. U.S. Environmental Protection Agency: Washington, DC.

68. U.S. EPA (2015). Review of Coliphages as Possible Indicators of Fecal Contamination for Ambient Water Quality. U.S. Environmental Protection Agency Office of Water: Washington D.C.

69. Vuong, N. M., Villemur, R., Payment, P., Brousseau, R., Topp, E., & Masson, L. (2013). Fecal source tracking in water using a mitochondrial DNA microarray. Water research, 47(1), 16-30.

70. Wade, T. J., Calderon, R. L., Sams, E., Beach, M., Brenner, K. P., Williams, A. H., & Dufour, A. P. (2005). Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. Environmental health perspectives, 114(1), 24-28.

71. Wiesner, R. J., Rüegg, J. C., & Morano, I. (1992). Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues. Biochemical and biophysical research communications, 183(2), 553-559.

APPENDICES

Appendix A: Latitude and longitude of sampling location for the lakes sampled in this study.

Table A1. Latitude and Longitude coordinates for sample Locations 1-10.

Sample Number	Latitude	Longitude
1	35° 49' 16.9"	-94° 54' 11.9"
2	35° 51' 22.6"	-94° 55' 16.7"
3	35° 55' 24.3"	-94° 55' 27.0"
4	35° 57' 54.7"	-94° 54' 38.9"
5	36° 01' 52.9"	-94° 55' 07.3"
6	36° 03' 52.9"	-94° 53' 03.8"
7	36° 05' 20.6"	-94° 50' 14.7"
8	36° 06' 14.1"	-94° 46' 52.3"
9	36° 07' 52.3"	-94° 35' 00.1"
10	36° 07' 49.9"	-94° 34' 20.6"

Appendix B: PCR Gels

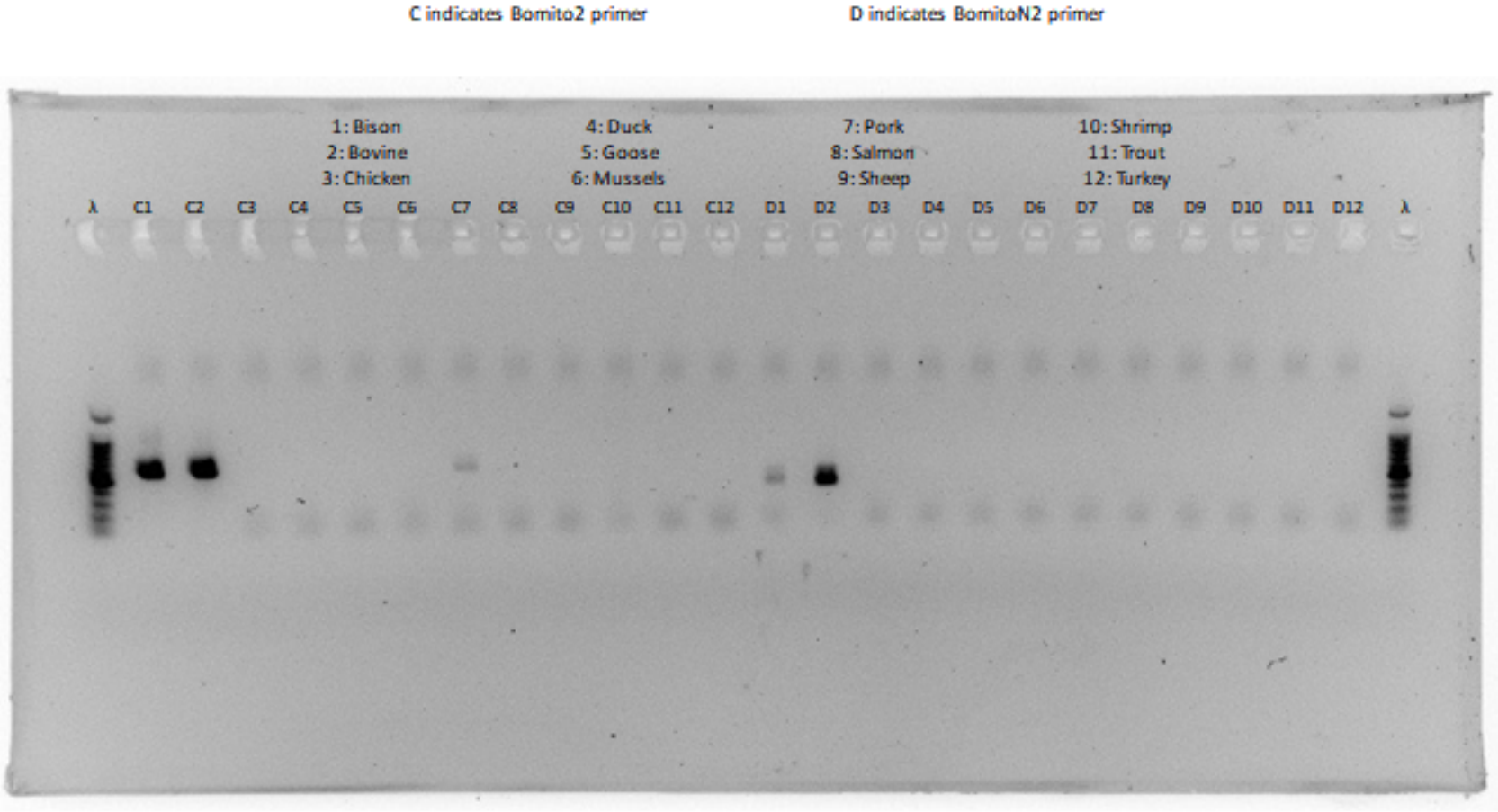

Figure B1. Agarose gel electrophoresis analysis of PCR products from Bomito2 (cow) & BomitoN2 (nested cow) primer sets on muscle tissue DNA.

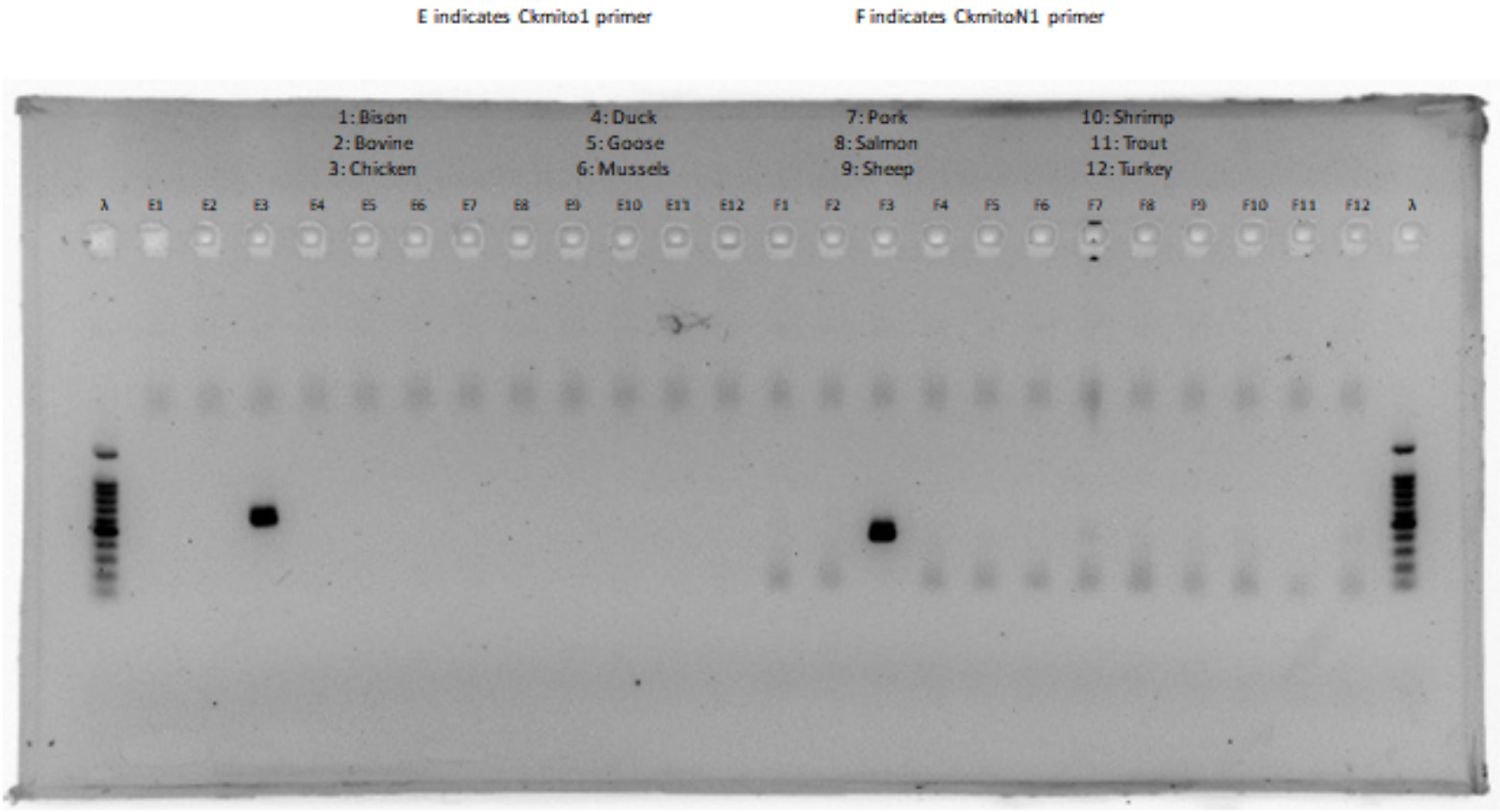

Figure B2. Agarose gel electrophoresis analysis of PCR products from Ckmito1 (chicken) & CkmitoN1 (nested chicken) primer sets on muscle tissue DNA.

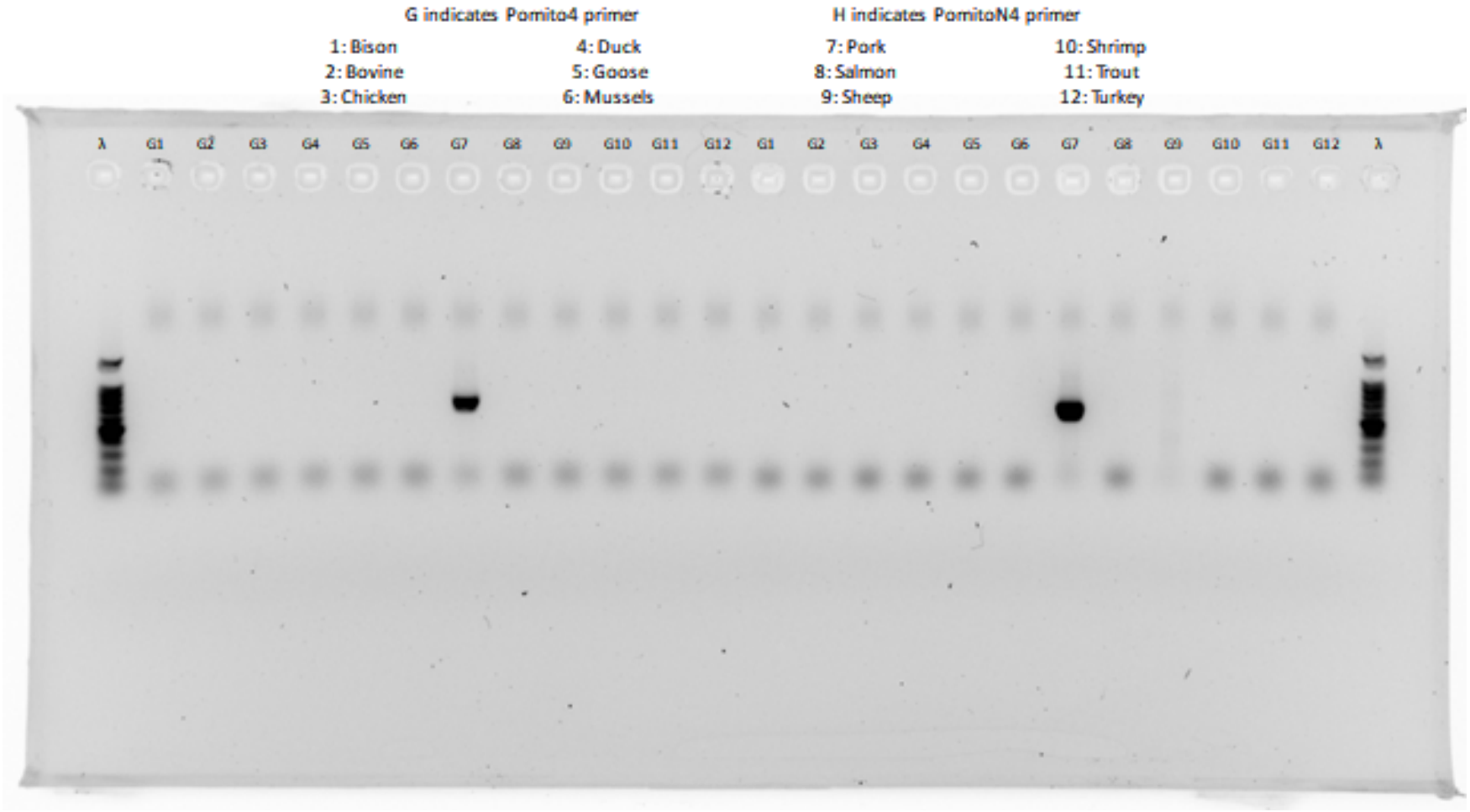

Figure B3. Agarose gel electrophoresis analysis of PCR products from Pomito4 (pig) & PomitoN4 (nested pig) primer sets on muscle tissue DNA.

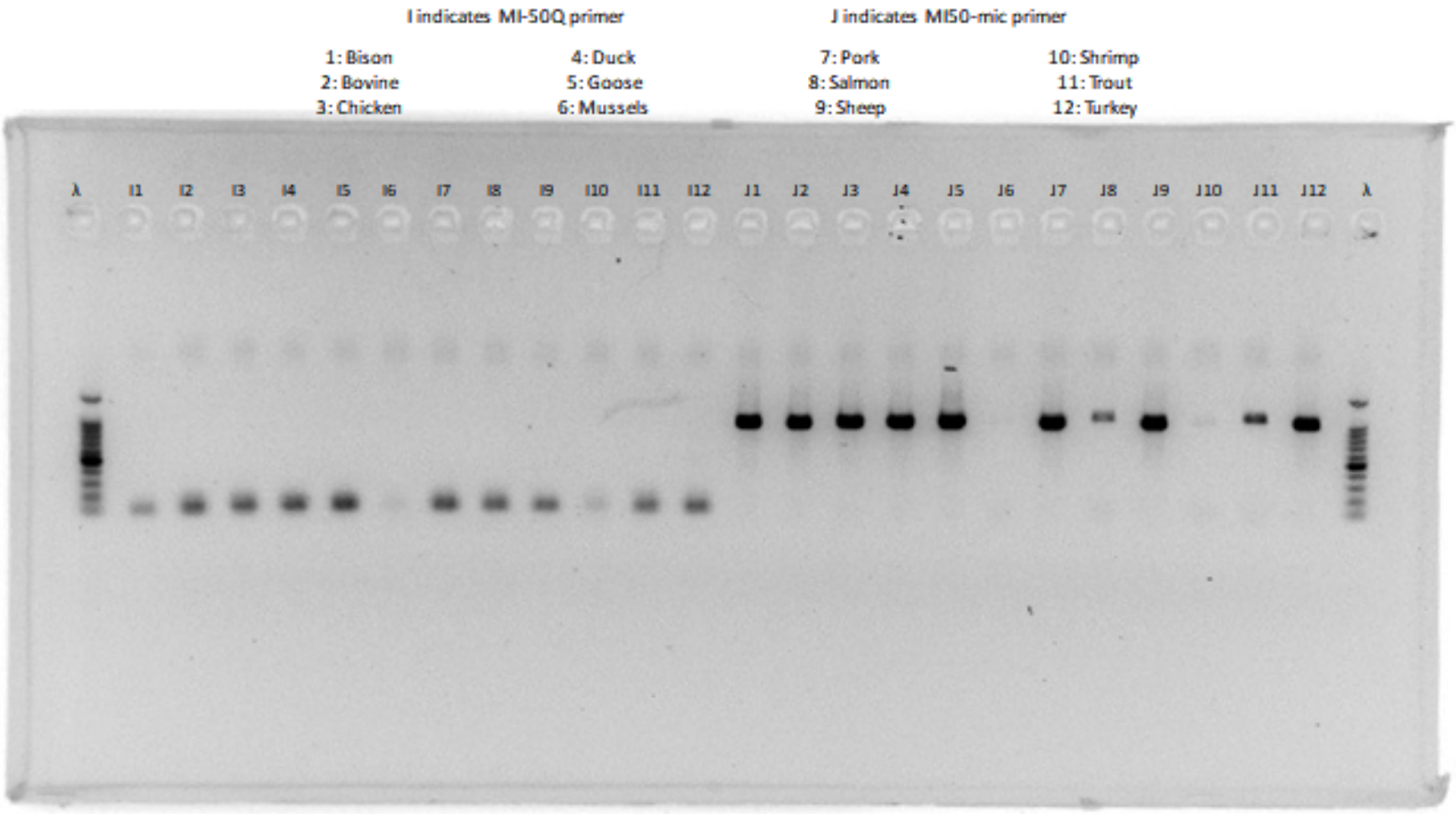

Figure B4. Agarose gel electrophoresis analysis of PCR products from MI-50Q (universal qPCR) & MI50-mic (universal) primer sets on muscle tissue DNA.

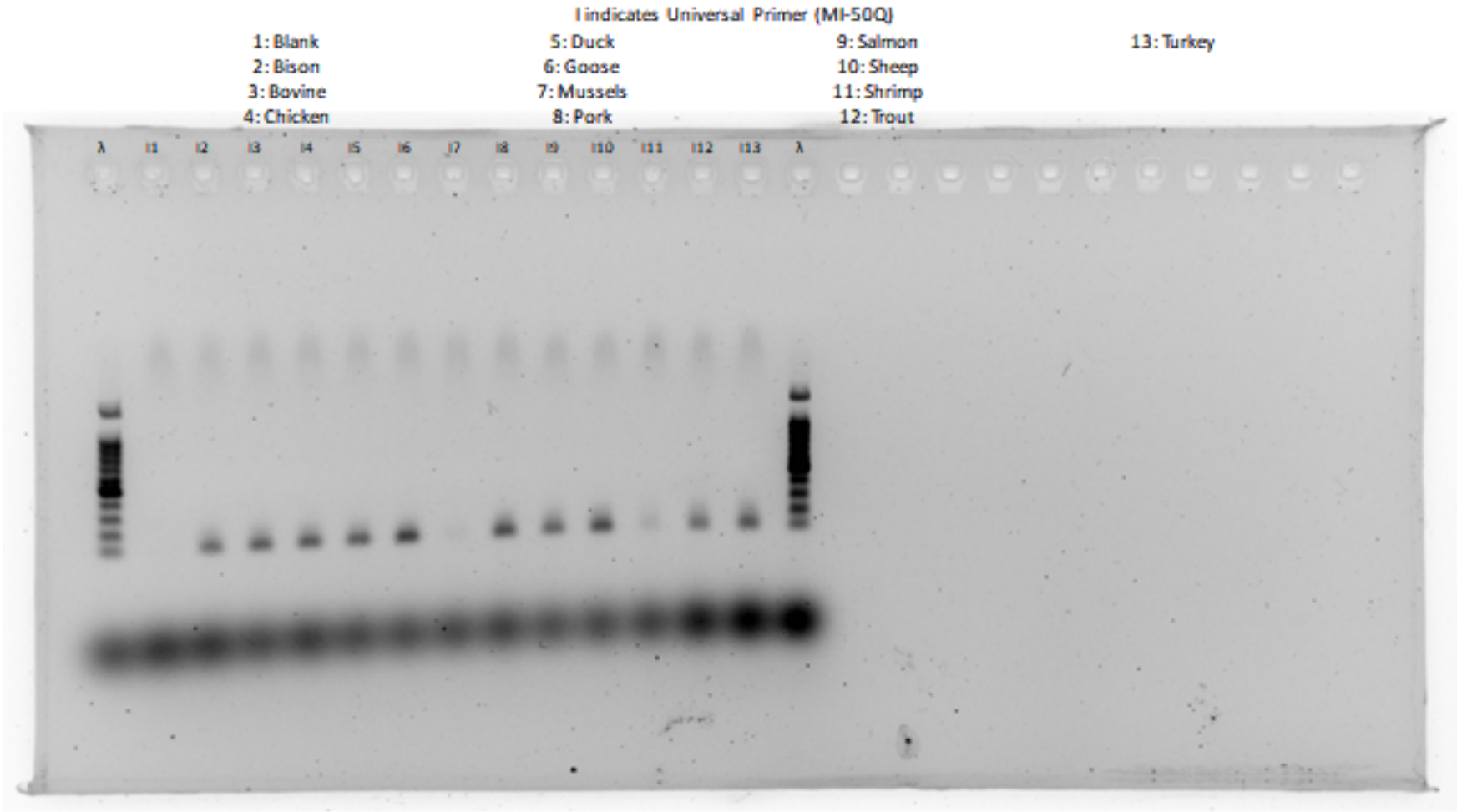

Figure B5. Agarose gel electrophoresis analysis of PCR products from the MI-50Q (universal qPCR) primer set on muscle tissue DNA.

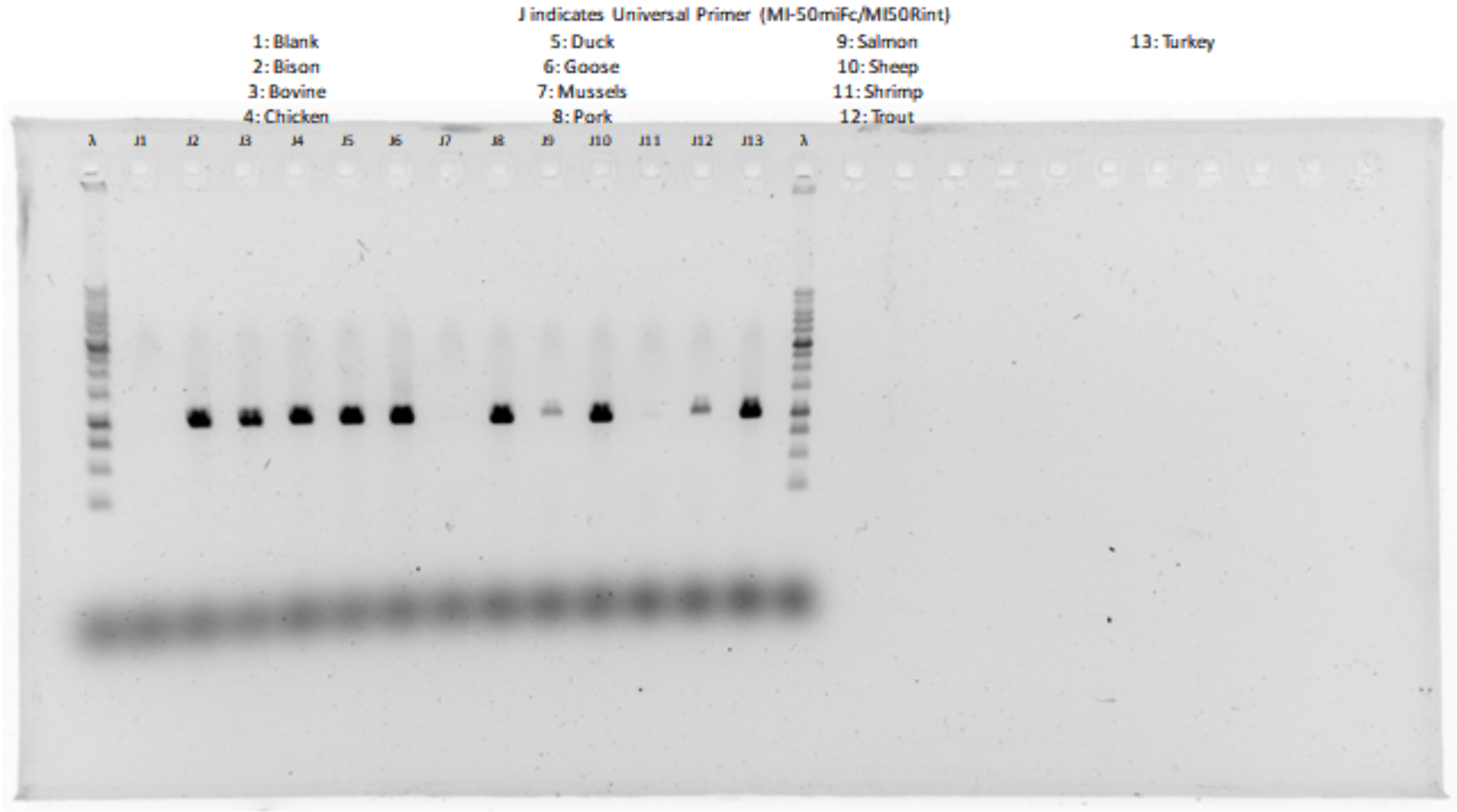

Figure B6. Agarose gel electrophoresis analysis of PCR products from the MI50-mic (universal) primer set on muscle tissue DNA.

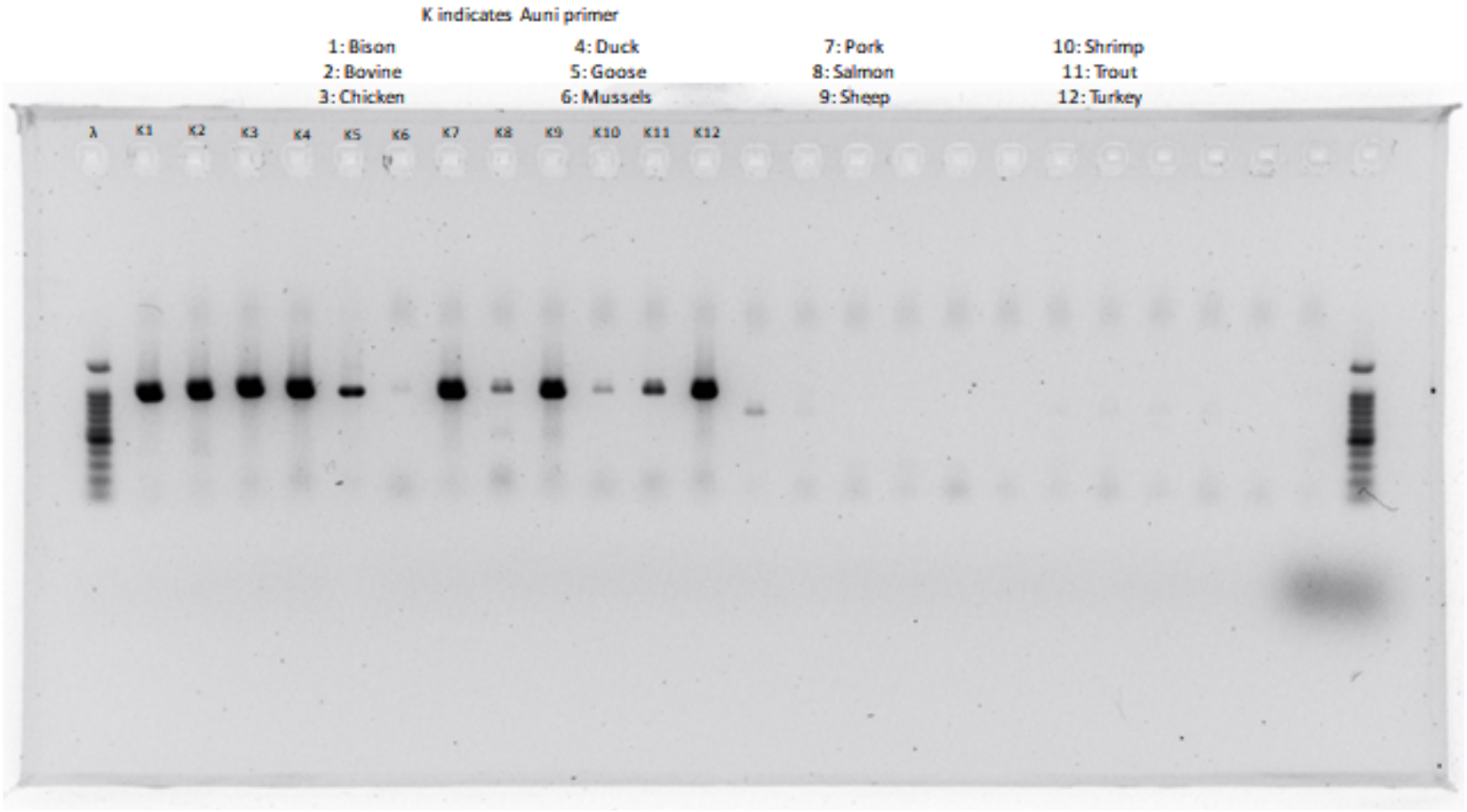

Figure B7. Agarose gel electrophoresis analysis of PCR products from the Auni (universal) primer set on muscle tissue DNA.

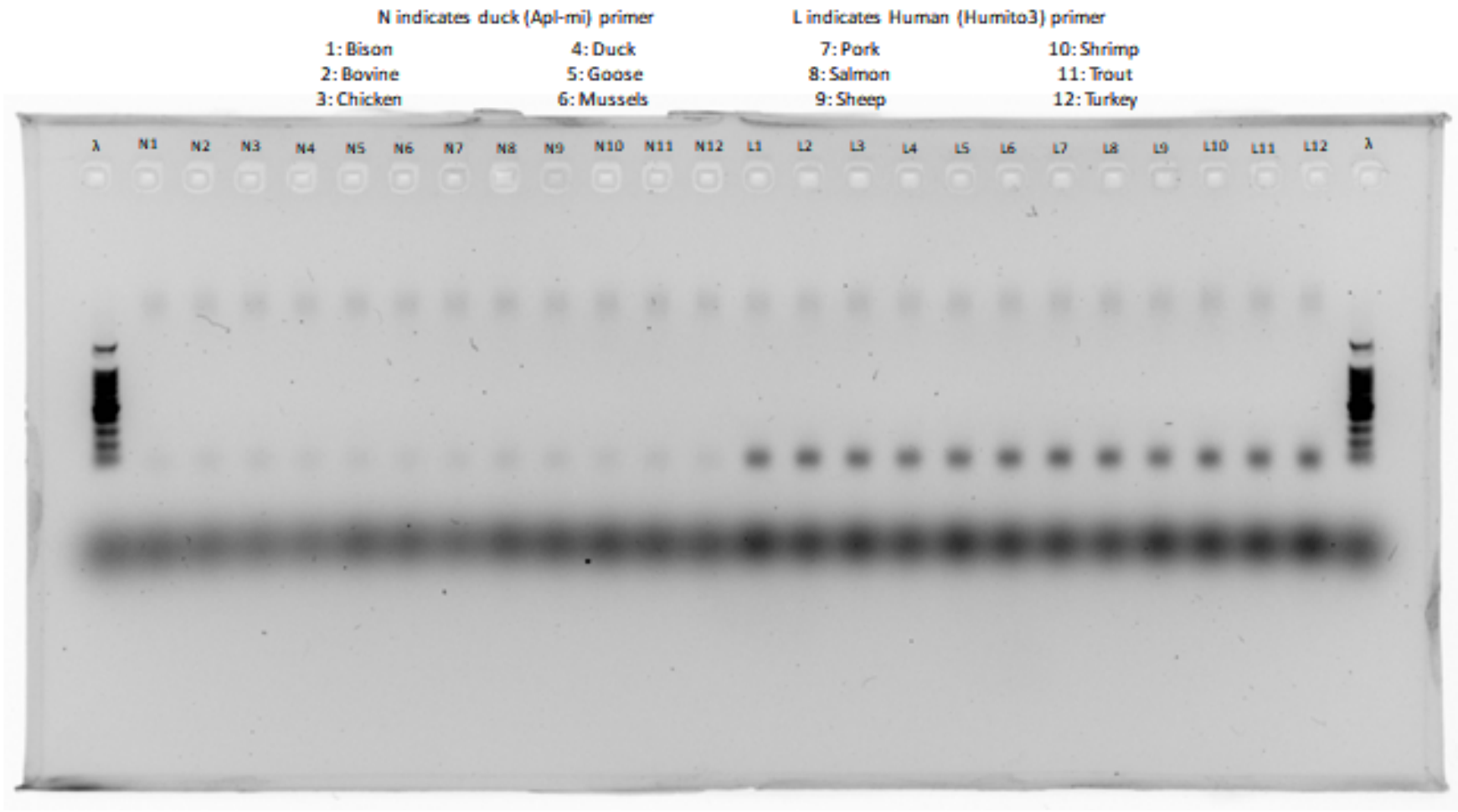

Figure B8. Agarose gel electrophoresis analysis of PCR products from Apl-mi (duck) & Humito3 (human) primer sets on muscle tissue DNA.

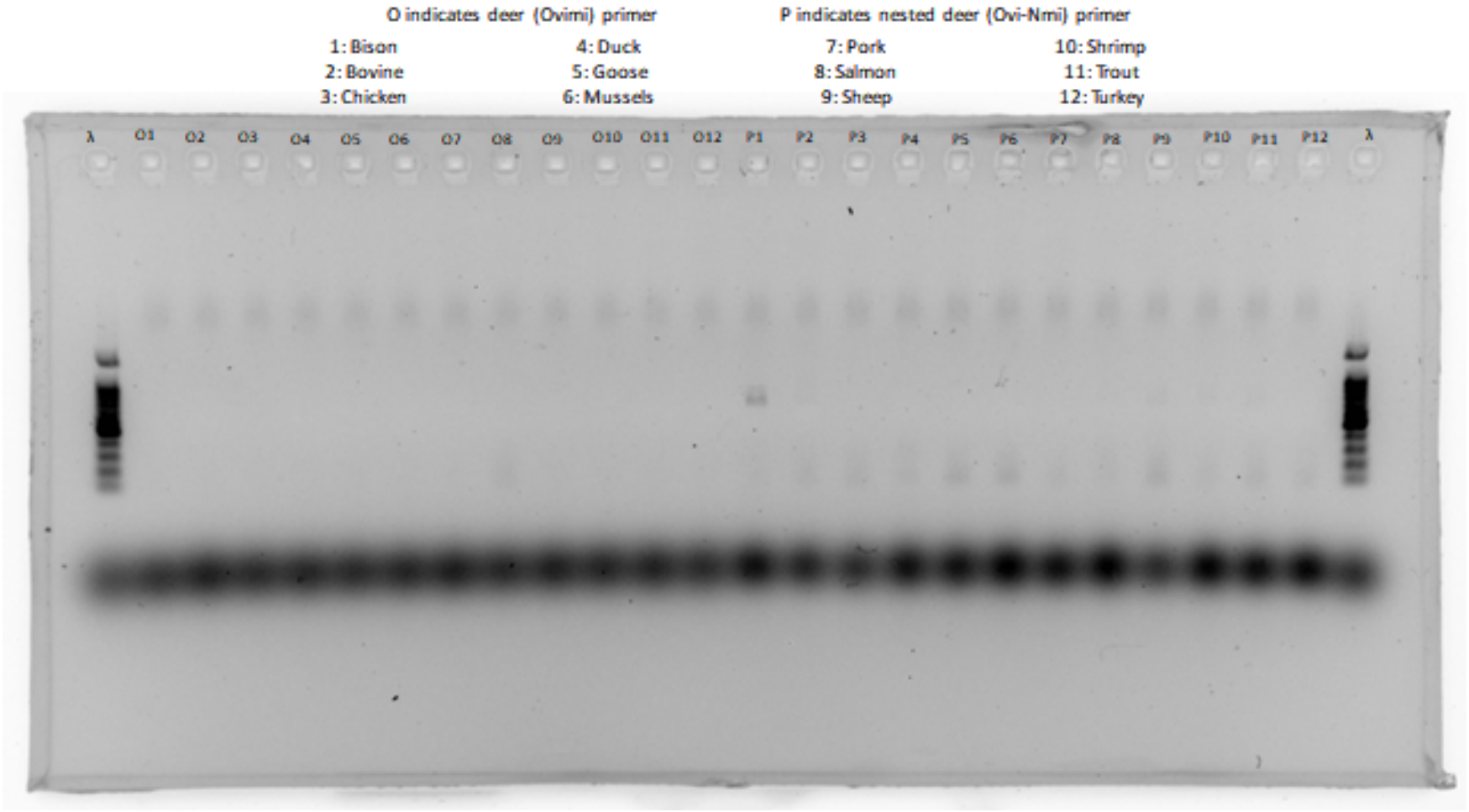

Figure B9. Agarose gel electrophoresis analysis of PCR products from Ovimi (deer) & Ovi-Nmi (nested deer) primer sets on muscle tissue DNA.

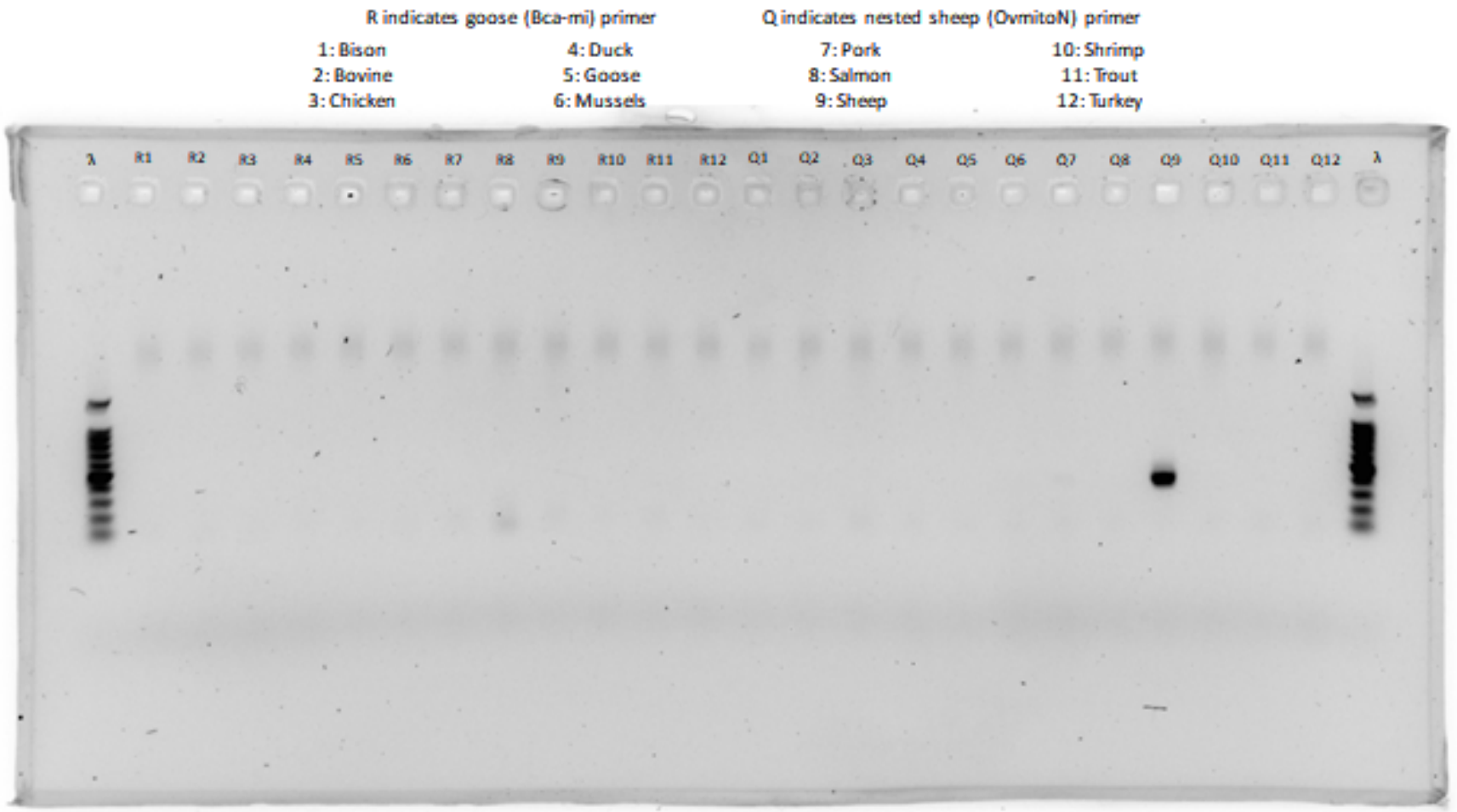

Figure B10. Agarose gel electrophoresis analysis of PCR products from Bca-mi (goose) & OvmitoN (nested sheep) primer sets on muscle tissue DNA.

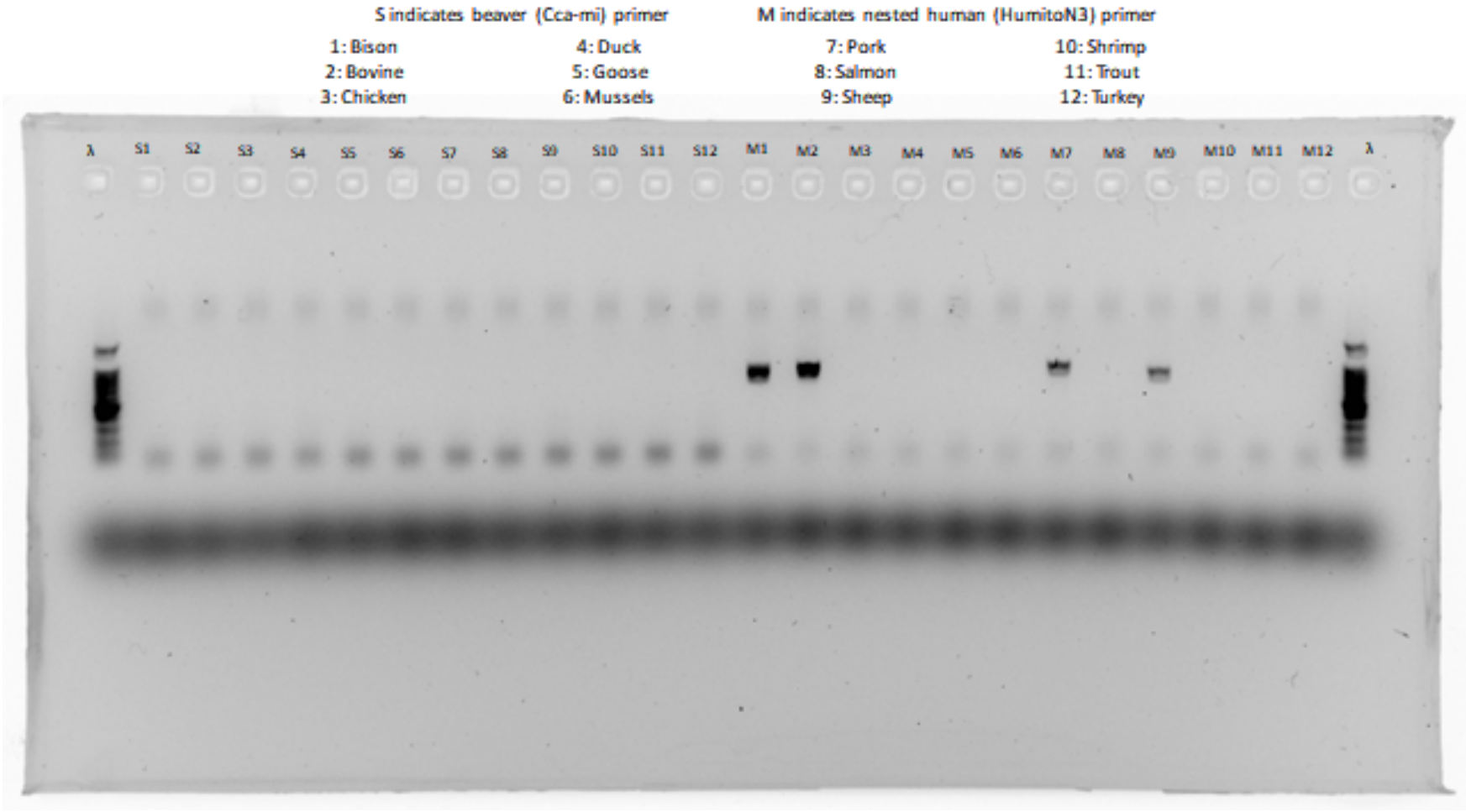

Figure B11. Agarose gel electrophoresis analysis of PCR products from Cca-mi (beaver) & HumitoN3 (nested human) primer sets on muscle tissue DNA.

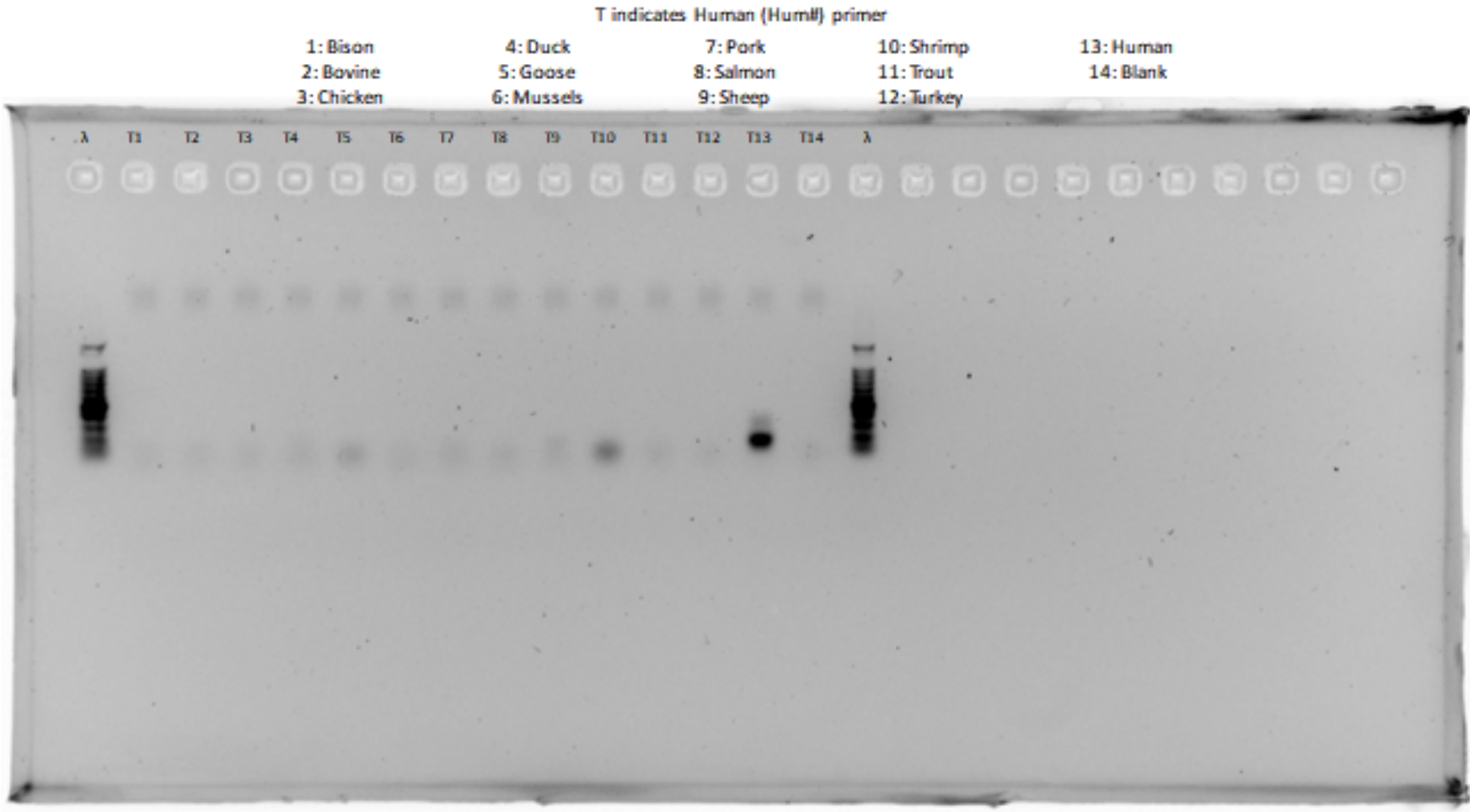

Figure B12. Agarose gel electrophoresis analysis of PCR products from the Hum15425F// Hum15594R (human) primer set on muscle tissue DNA.

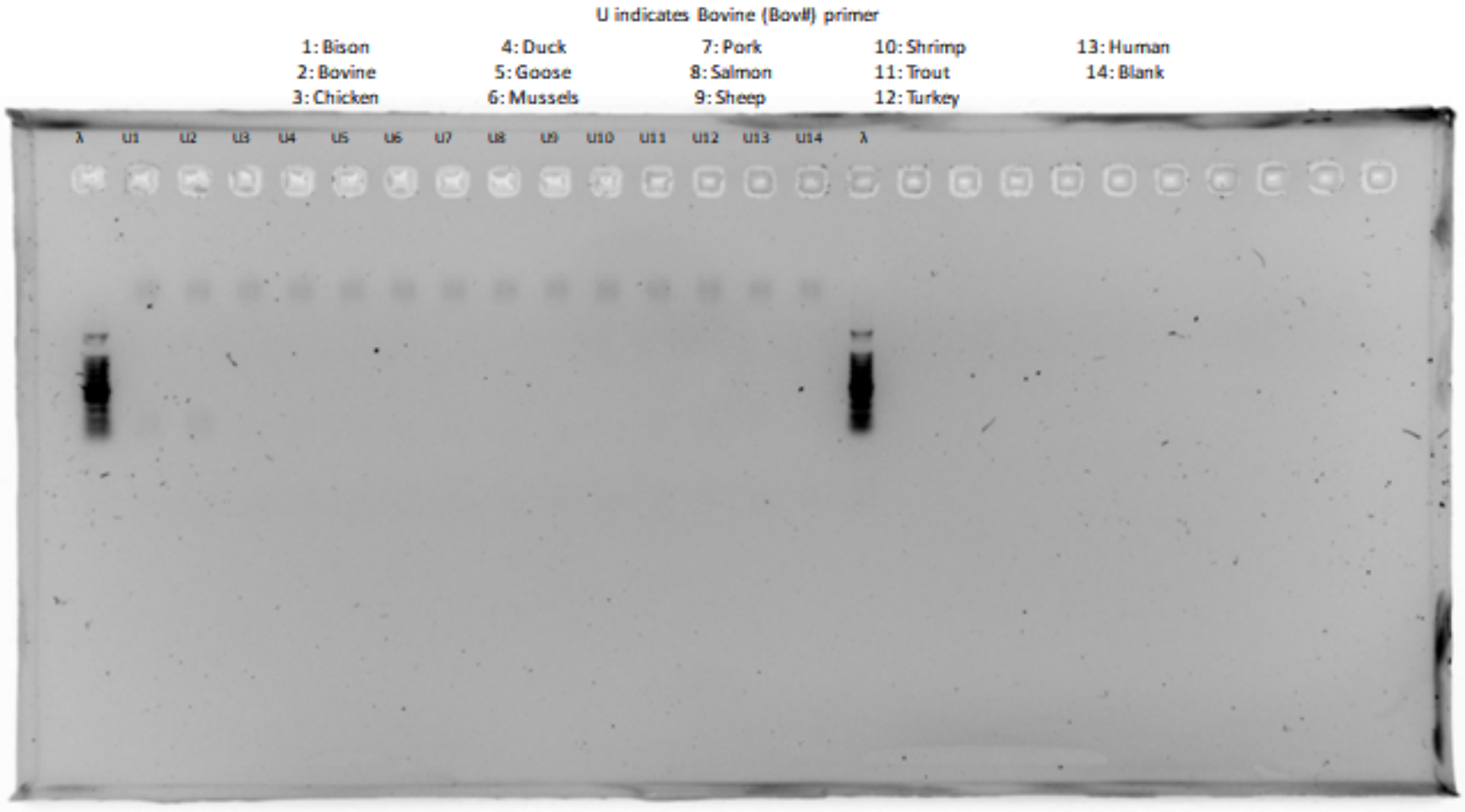

Figure B13. Agarose gel electrophoresis analysis of PCR products from the Bov1482F// Bov1571R (cow) primer set on muscle tissue DNA.

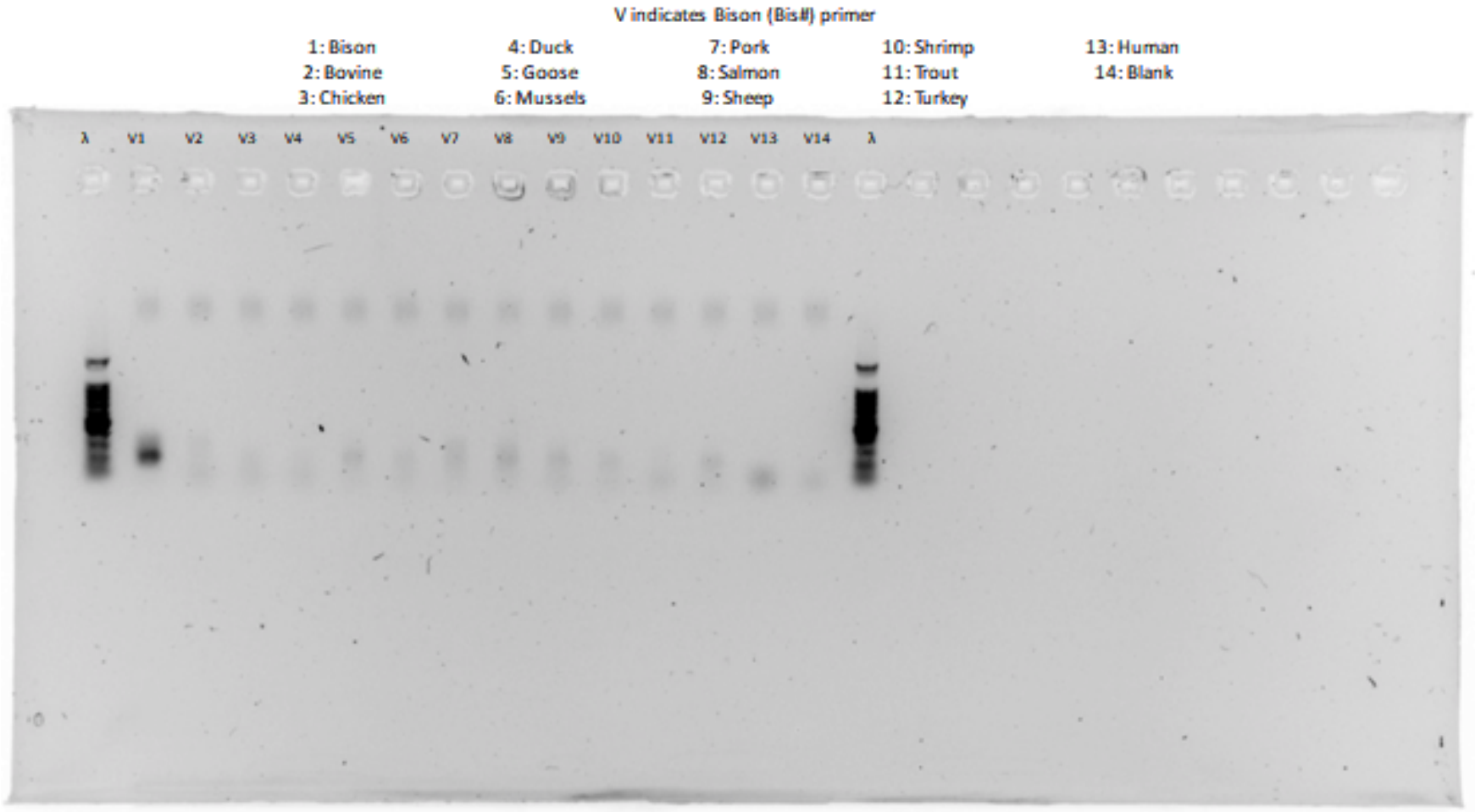

Figure B14. Agarose gel electrophoresis analysis of PCR products from the Bis15914F// Bis16116R (bison) primer set on muscle tissue DNA.

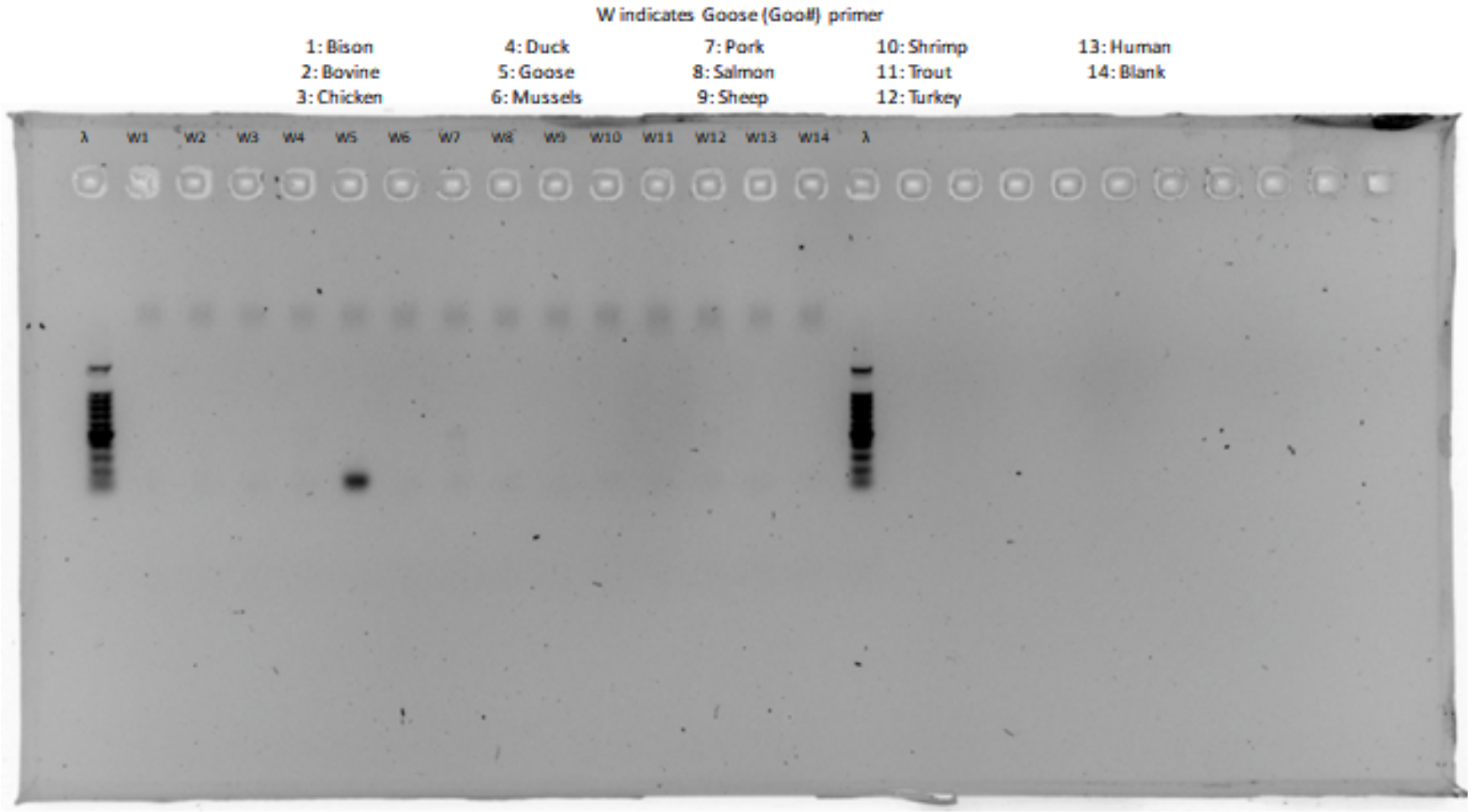

Figure B15. Agarose gel electrophoresis analysis of PCR products from the Goose4728F// Goose4861R (goose) primer set on muscle tissue DNA.

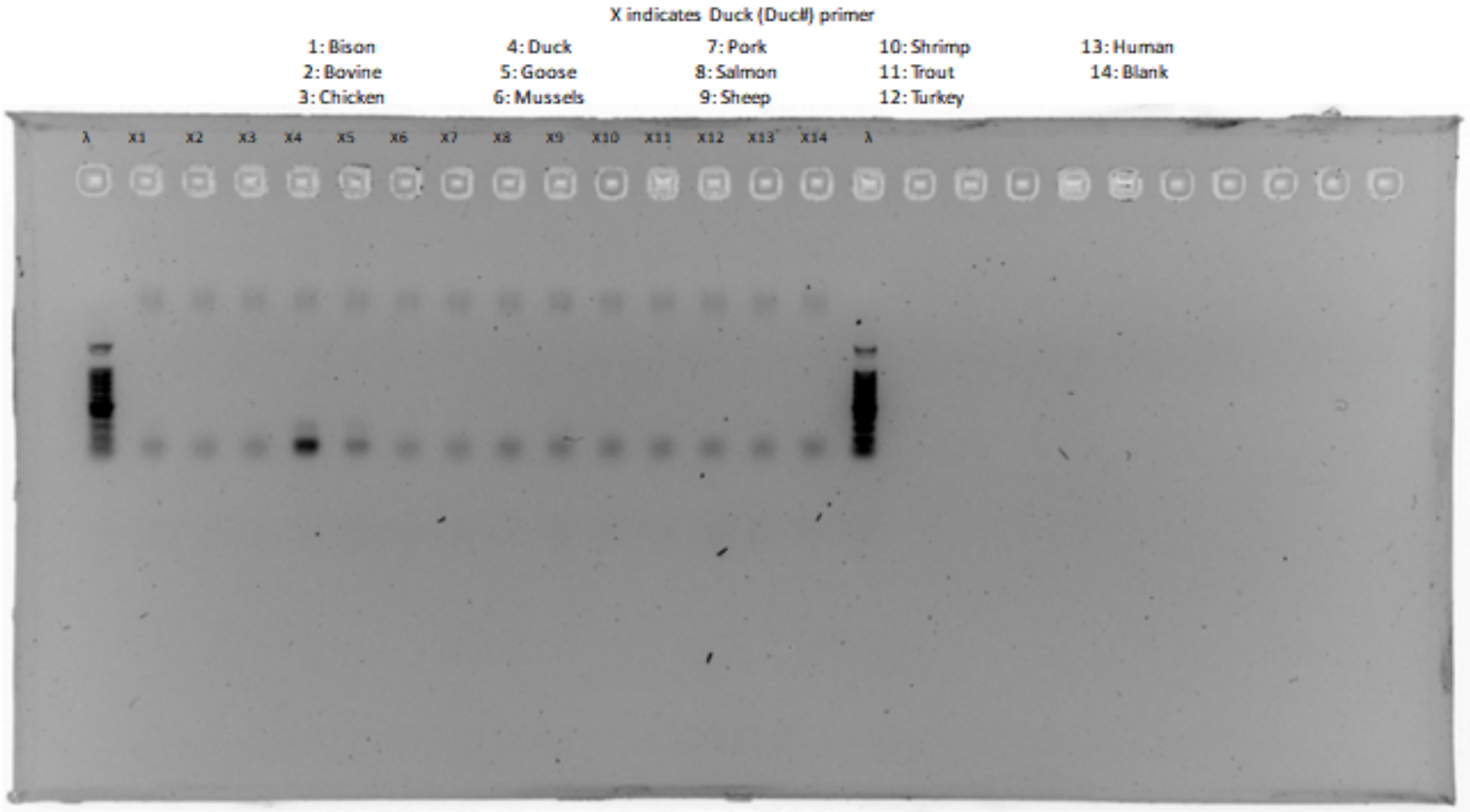

Figure B16. Agarose gel electrophoresis analysis of PCR products from the Duck14644F// Duck14788R (duck) primer set on muscle tissue DNA.

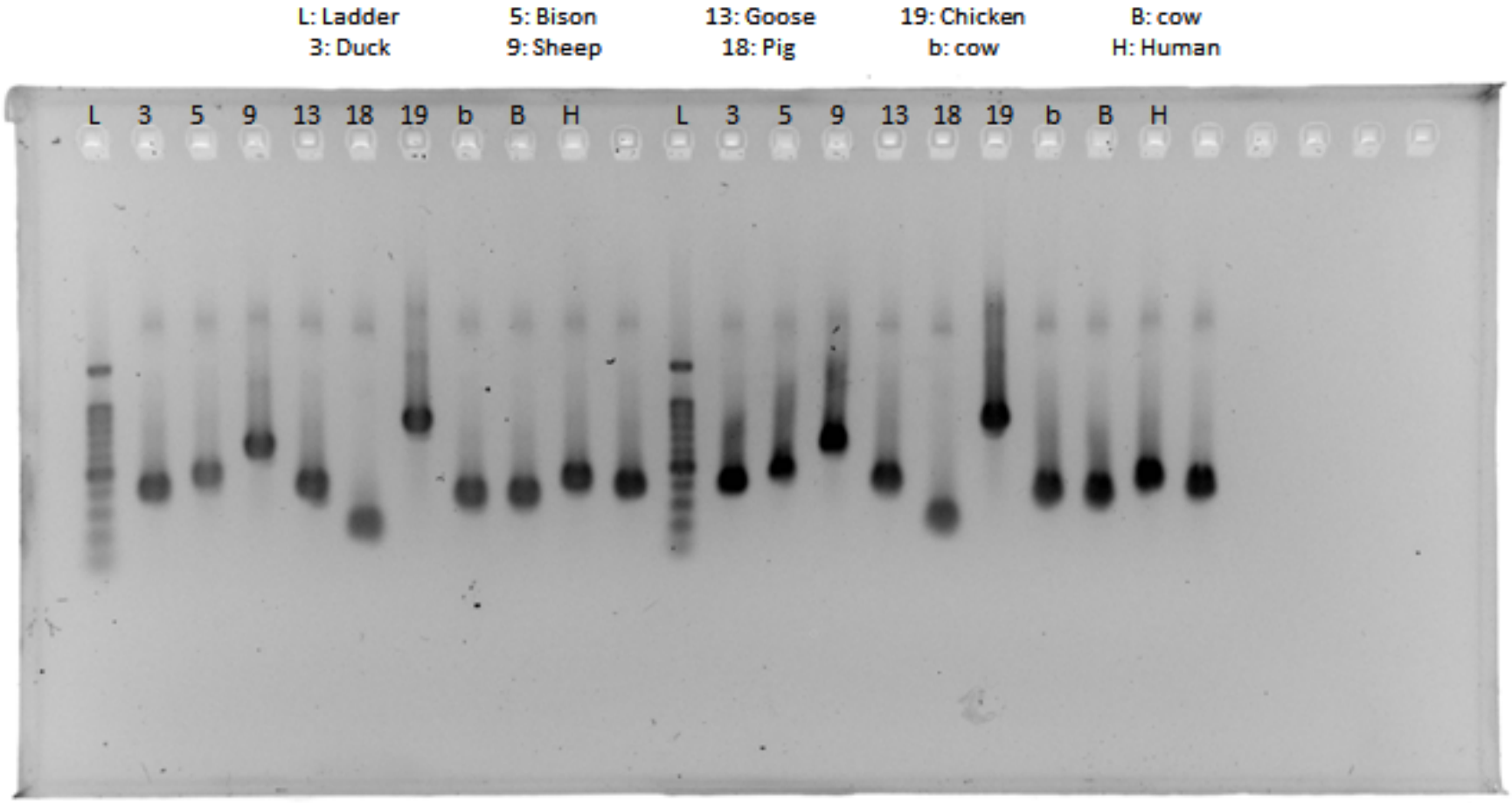

Figure B17. Agarose gel electrophoresis analysis of PCR products from the M13 primer set on clone library plasmid DNA for insert verification.

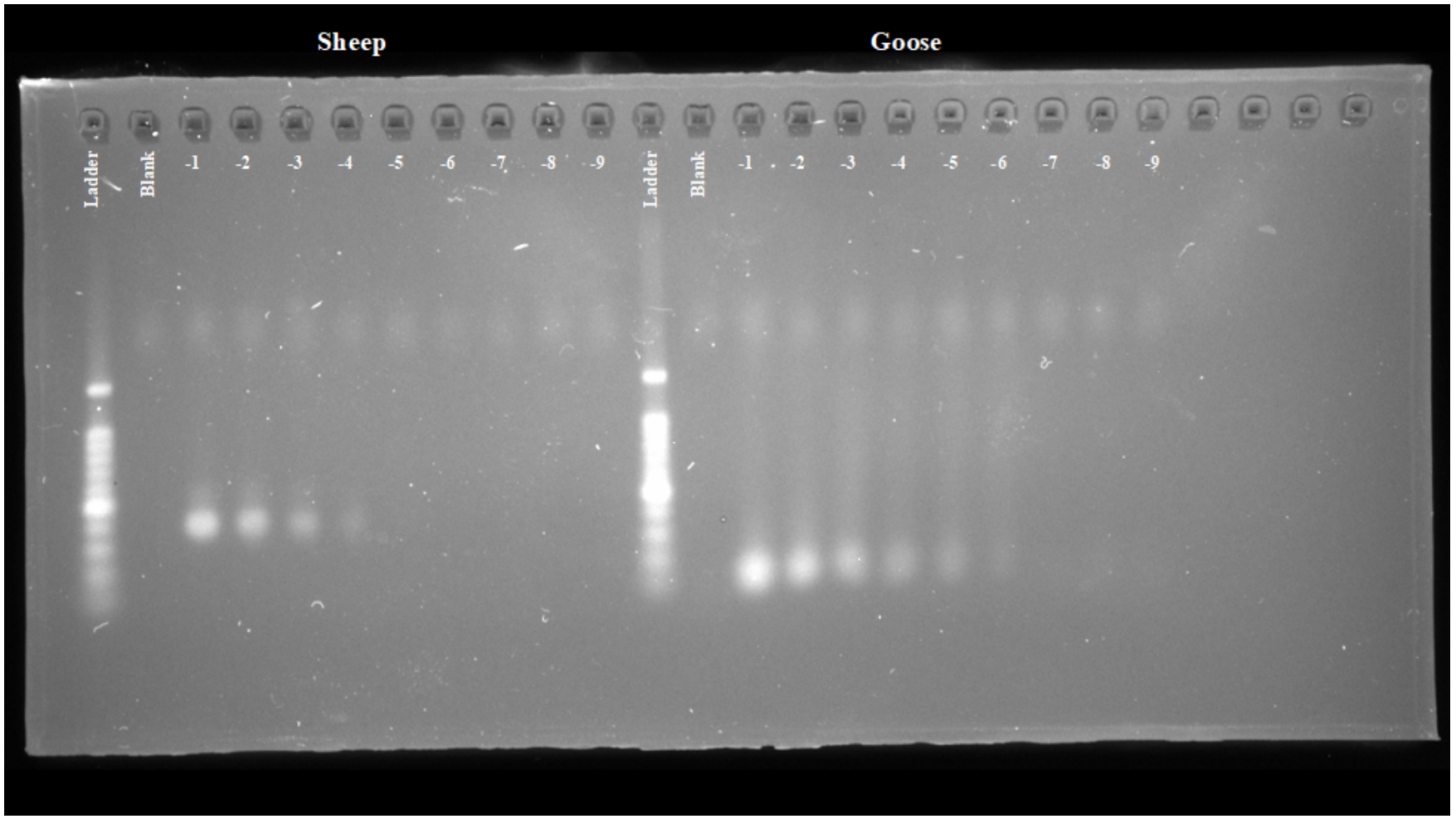

Figure B18. Agarose gel electrophoresis analysis of PCR products from OvmitoN2-G//OvmitoN2-D and Goose4728F//Goose4861R on clone library plasmid DNA dilution standards.

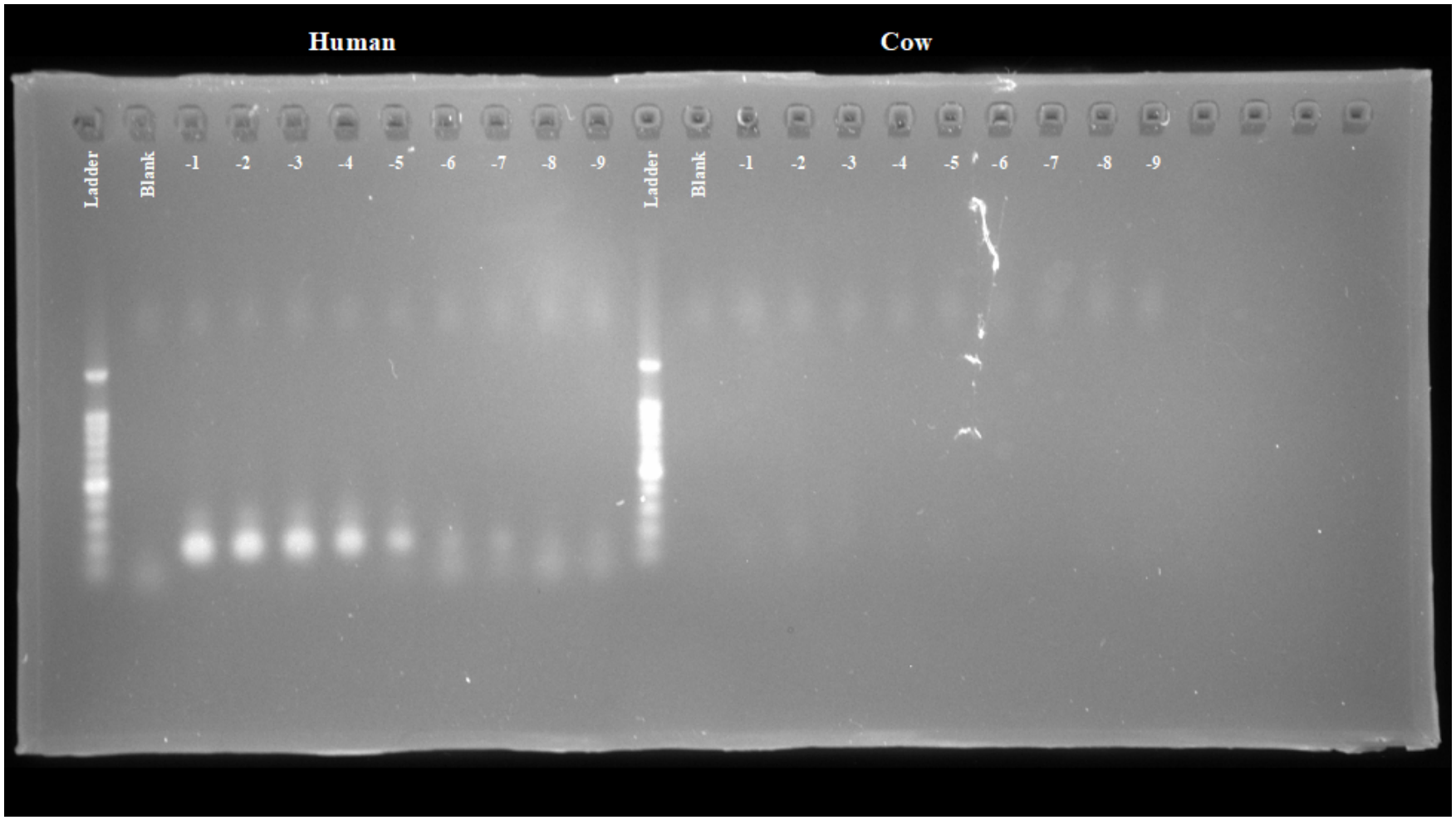

Figure B19. Agarose gel electrophoresis analysis of PCR products from Hum15425F// Hum15594R and Bov1482F// Bov1571R on clone library plasmid DNA dilution standards.

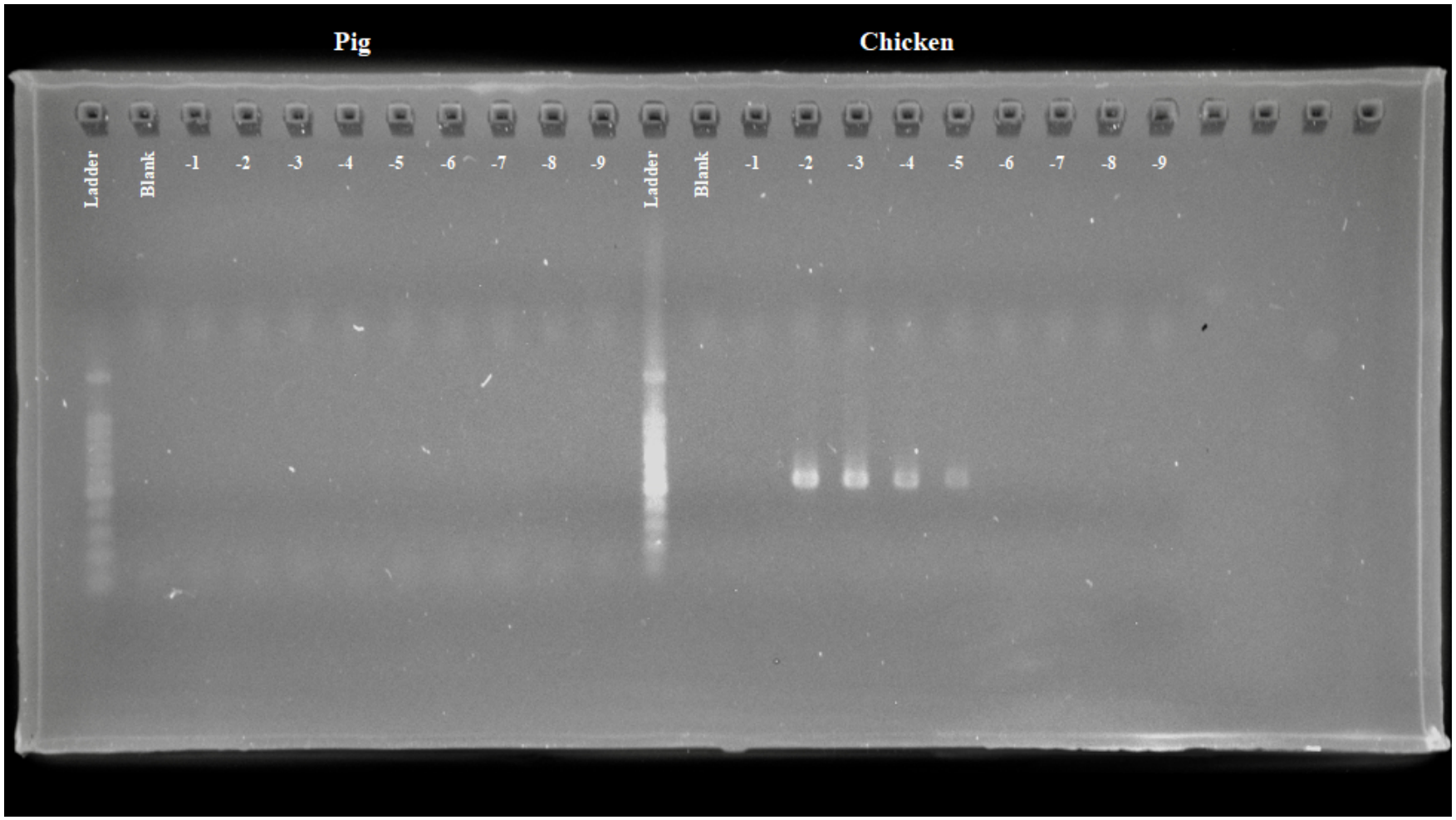

Figure B20. Agarose gel electrophoresis analysis of PCR products from Hum15425F// Hum15594R and Bov1482F// Bov1571R on clone library plasmid DNA dilution standards.

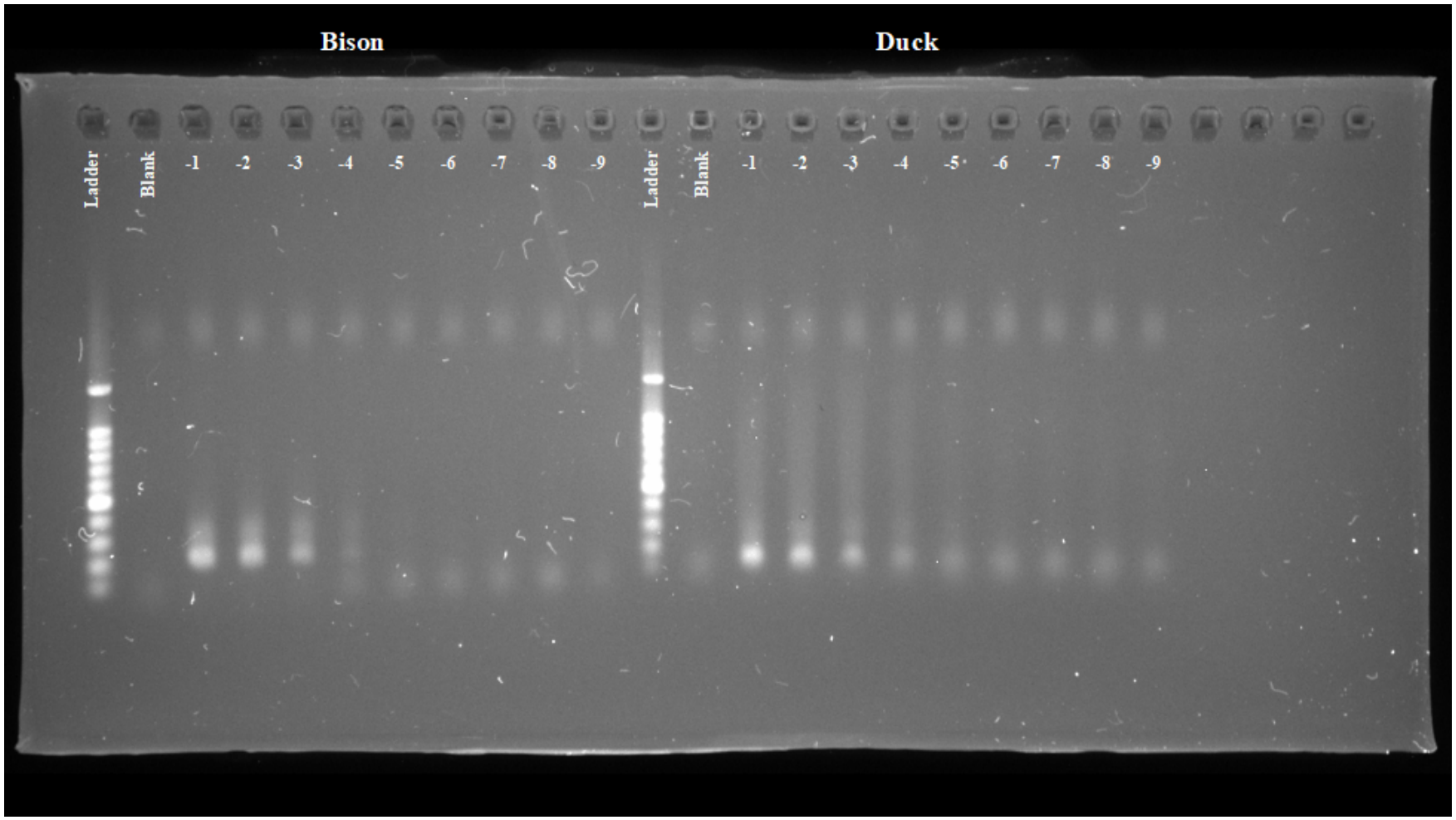

Figure B21. Agarose gel electrophoresis analysis of PCR products from Bis15914F// Bis16116R and Duck14644F// Duck14788R on clone library plasmid DNA dilution standards.

Appendix C: Gene Target Regions

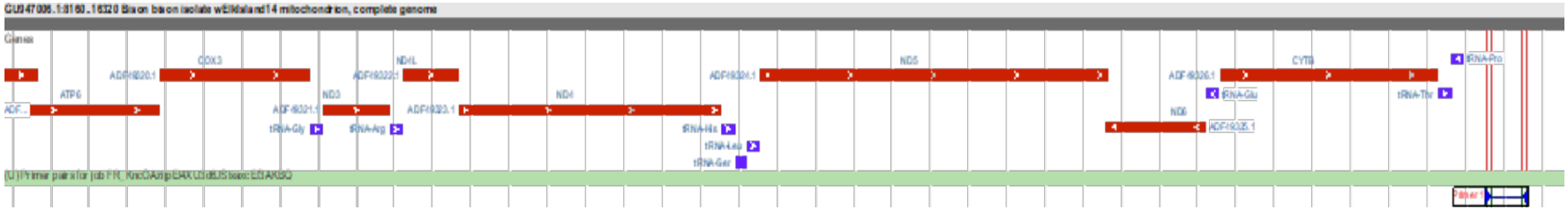

Figure C1. Gene target region for *Bis15914F// Bis16116R*

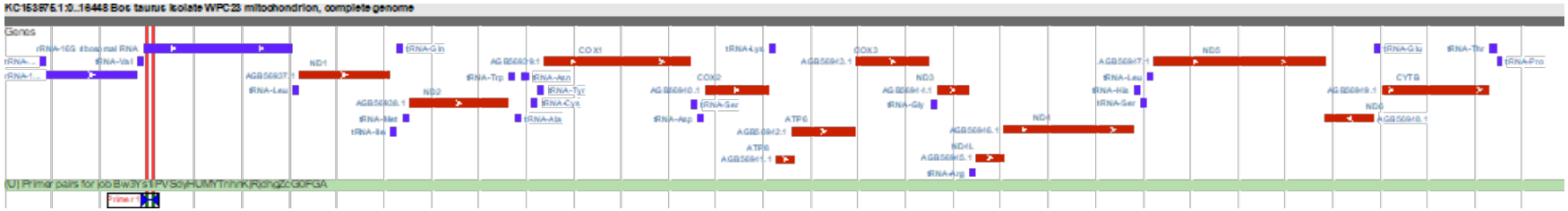

Figure C2. Gene target region for *Bov1482F // Bov1571R*

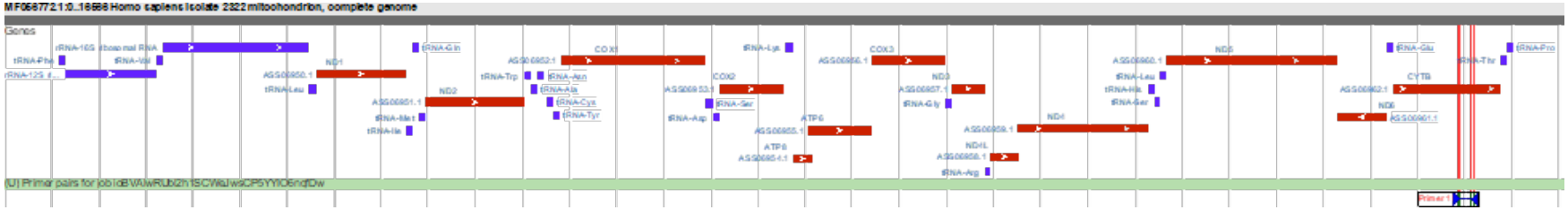

Figure C3. Gene target region for *Hum15425F // Hum15594R*

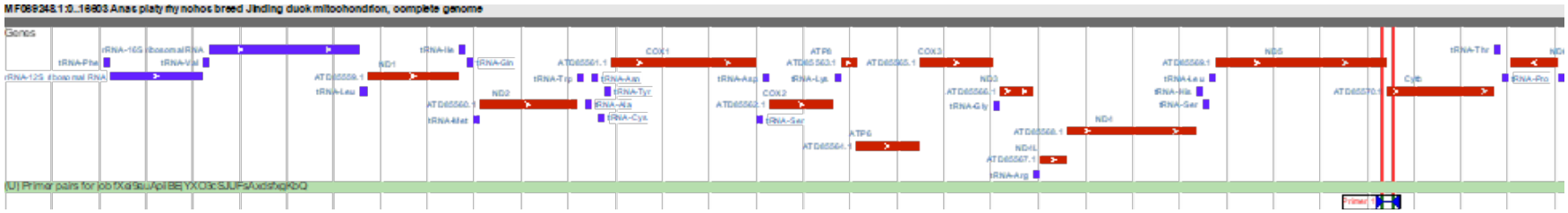

Figure C4. Gene target region for *Duck14644F // Duck14788R*

Appendix D: Primer Set Data Throughout Experiment

Table D1. Initial primer sets used to test specificity amongst positive and negative control samples.

Target	Accession #	Primer Name	Primer Sequence	Target Gene	Amplicon Size (bp)
Beaver	NC_015108	Cca-miF	CACAAAACTACATCACGTCATTTAT	ND2	1123
		Cca-miR	TCTAAACACAGGGGTCAAGTC	tRNA-Tyr	
Chicken	AP003580.1	Ckmitol-G	ACCCTATTTGACTCCCTCAA	16S rRNA	565
		Ckmitol-D	ATGTCGACCAGGGGTTTATG	16S rRNA	
Chicken (nested)	AP003580.1	CkmitoNl-G	CCCCCACACTAACAAGCAAT	16S rRNA	381
		CkmitoNl-D	GGTTGTAAGGTGGTCGTGAT	16S rRNA	
Cow	GU947006.1	Bomito2-G	CATAGCAATTGCCATAGTCC	Cytochrome b	554
		Bomito2-D	TTTTCGATTGTGCCGGCCGTT	Cytochrome b	
Cow (Nested)	KC153975.1	BomitoN2-G	CCCTCTTACTAATTCTAGCTC	Cytochrome b	401
		BomitoN2-D	TTAGCACTAGGATGAGGAGA	Cytochrome b	
Deer	NC015247.1	Ovi-miF	ATTTATAGTATCTCTCGCAGGACTA	ND4L	1591
		Ovi-miR	GTAGAGGTAGAATGTGCAATGATAT	ND4	
Deer (Nested)	NC015247.1	Ovi-NmiF	TAAACACGCACTTCACTTTAGCAAG	ND4L	1246
		Ovi-NmiR	CAGGTTGGTCAAGCTTGCTAG	ND4	
Duck	MF069248	Apl-miF	AACAATATGGTCTATCGAGAGCCAA	ND4L	123
		Apl-miR	TGCGTGTAGAGGCTACTAGGAT	ND4L	
Goose	NC007011.1	Bca-miF	GCTTCTACTCAGCCTTCATCTTCAG	ND4L	124
		Bca-miR	GATCATATAGACAGGCCGACGAAT	ND4L	
Human (Nested)	MF056772.1	HumitoN3-G	CTACTCTACCATCTTTGCAGG	ND2	651
		HumitoN3-D	CGTGGTGCTGGAGTTTAAGTTG	ND2	
Human2	MF056772.1	Humito3-G	CCCAACCCGTCATCTACTCT	ND2	151
		Humito3-D	GCTTCTGTGGAACGAGGGTT	ND2	

All primer set in Table D1 were developed in Vuong et al., (2013)

Table D1. Continued

Target	Accession #	Primer Name	Primer Sequence	Target Gene	Amplicon Size (bp)
Human	MF056772.1	AuniF	CACGAGGGTTCAGCTGTCTCTTAC	16S rRNA	1002
		AuniR	GGCTAGGCTAGAGGTGGCTAGAAT	ND1	
Pig	AJ002189	Pomito4-G	CCCATTATCTACACTACCCTTATC	ND2	723
		Pomito4-D	TTAGGCTTGTGATGACGGGTAT	ND2	
Pig (Nested)	AJ002189	PomitoN4-G	CAGTAATGTCCGGAACCATACTAG	ND2	643
		PomitoN4-D	TGTGGTTGCTGAGCTGTGGATT	ND2	
Sheep (Nested)	AY858379.1	OvmitoN2-G	TACACTGTTACAGGCATCAG	COX3	370
		OvmitoN2-D	CGTGAAGTTAGTTAGGAGAGTA	ND3	
Universal	-	MI-50F	ACTGGGATTAGATACCCCACTATG	12S rRNA	2011
		MI-50R	CGGTCTGAACTCAGATCACGTA	16S rRNA	
Universal (Nested)	-	MI-50Fint	GAGGAGCCTGTTCYRYAAYCGA	12S rRNA	1804
		MI-50Rint	TGATCCAACATCGAGGTCGTAAA	16S rRNA	
Universal (qPCR)	-	MI-50Q-F	TTTACGACCTCGATGTTGGATCA	16S rRNA	102
		MI-50R	CGGTCTGAACTCAGATCACGTA	16S rRNA	
Universal2	-	MI-50miF	GAGCYKGGTGATAGCTGGTT	12S rRNA	999
		MI-50Rint	TGATCCAACATCGAGGTCGTAAA	16S rRNA	

All primer set in Table D1 were developed in Vuong et al., (2013)

Table D2. Comprehensive list of final primer sets used for experiment.

Target	Accession #	Primer Name	Primer Sequence	Target Gene	Amplicon Size (bp)
Beaver	NC_015108	Cca-miF*	CACAAAACTACATCACGTCATTTAT	ND2	1123
		Cca-miR*	TCTAAACACAGGGGTCAAGTC	tRNA-Tyr	
Bison	GU947006	Bis15914F	GTCACTCACCCCCAAAATGC	-	202
		Bis16116R	GTCACTCACCCCCAAAATGC	-	
Chicken	AP003580.1	Ckmitol-G*	ACCCTATTTGACTCCCTCAA	16S rRNA	565
		Ckmitol-D*	ATGTCGACCAGGGGTTTATG	16S rRNA	
Cow	KC153975	Bov1482F	CGACTAAACAACCAAGATAG	16S rRNA	90
		Bov1571R	TTCTCTATAGCGCCGTACTT	16S rRNA	
Deer	NC_015247.1	Ovi-miF*	ATTTATAGTATCTCTCGCAGGACTA	ND4L	1591
		Ovi-miR*	GTAGAGGTAGAATGTGCAATGATAT	ND4	
Duck	MF069248	Duck14644F	AACCTACTTAGGATCTTTCGCCC	ND5	121
		Duck14788R	TATTAGAGGGTGCGGGAAGGT	CYTB	
Goose	NC_007011	Goose4728F	AGTACTAAATGCCACCCTGA	ND2	134
		Goose4861R	CAATTGCTATGGCTGCTGG	ND2	
Human	MF056772	Hum15425F	GACGCCCTCGGCTTACTTC	CYTB	170
		Hum15594R	ACGGATCGGAGAATTGTGTAGG	CYTB	
Pig	AJ002189	Pomito4-G*	CCCATTATCTACACTACCCTTATC	ND2	723
		Pomito4-D*	TTAGGCTTGTGATGACGGGTAT	ND2	
Sheep (Nested)	AY858379.1	OvmitoN2-G*	TACACTGTTACAGGCATCAG	COX3	370
		OvmitoN2-D*	CGTGAAGTTAGTTAGGAGAGTA	ND3	
Universal	-	MI-50F*	ACTGGGATTAGATACCCCACTATG	12S rRNA	2011
		MI-50R*	CGGTCTGAACTCAGATCACGTA	16S rRNA	
Universal (Nested)	-	MI-50Fint*	GAGGAGCCTGTTCYRYAAYCGA	12S rRNA	1804

*** Indicates primers developed in Vuong et al., (2013)**

Table D2. Continued

Target	Accession #	Primer Name	Primer Sequence	Target Gene	Amplicon Size (bp)
Universal (qPCR)	-	MI-50Q-F*	TTTACGACCTCGATGTTGGATCA	16S rRNA	102
		MI-50R*	CGGTCTGAACTCAGATCACGTA	16S rRNA	
Universal2	-	MI-50miF*	GAGCYKGGTGATAGCTGGTT	12S rRNA	999
		MI-50Rint*	TGATCCAACATCGAGGTCGTAAA	16S rRNA	

*** Indicates primers developed in Vuong et al., (2013)**

Table D3. Final primers used for standards detection and quantification via PCR and qPCR analysis.

Target	Accession #	Primer Name	Primer Sequence	Target Gene	Amplicon Size (bp)
Bison	GU947006	Bis15914F	GTCACTCACCCCCAAAATGC	-	202
		Bis16116R	GTCACTCACCCCCAAAATGC	-	
Chicken	AP003580.1	Ckmitol-G*	ACCCTATTTGACTCCCTCAA	16S rRNA	565
		Ckmitol-D*	ATGTCGACCAGGGGTTTATG	16S rRNA	
Cow	KC153975	Bov1482F	CGACTAAACAACCAAGATAG	16S rRNA	90
		Bov1571R	TTCTCTATAGCGCCGTACTT	16S rRNA	
Duck	MF069248	Duck14644F	AACCTACTTAGGATCTTTCGCCC	ND5	121
		Duck14788R	TATTAGAGGGTGCGGGAAGGT	CYTB	
Goose	NC_007011	Goose4728F	AGTACTAAATGCCACCCTGA	ND2	134
		Goose4861R	CAATTGCTATGGCTGCTGG	ND2	
Human	MF056772	Hum15425F	GACGCCCTCGGCTTACTTC	CYTB	170
		Hum15594R	ACGGATCGGAGAATTGTGTAGG	CYTB	
Pig	AJ002189	Pomito4-G*	CCCATTATCTACACTACCCTTATC	ND2	723
		Pomito4-D*	TTAGGCTTGTGATGACGGGTAT	ND2	
Sheep (Nested)	AY858379.1	OvmitoN2-G*	TACACTGTTACAGGCATCAG	COX3	370
		OvmitoN2-D*	CGTGAAGTTAGTTAGGAGAGTA	ND3	

*** Indicates primers developed in Vuong et al., (2013)**

Appendix E: qPCR Data

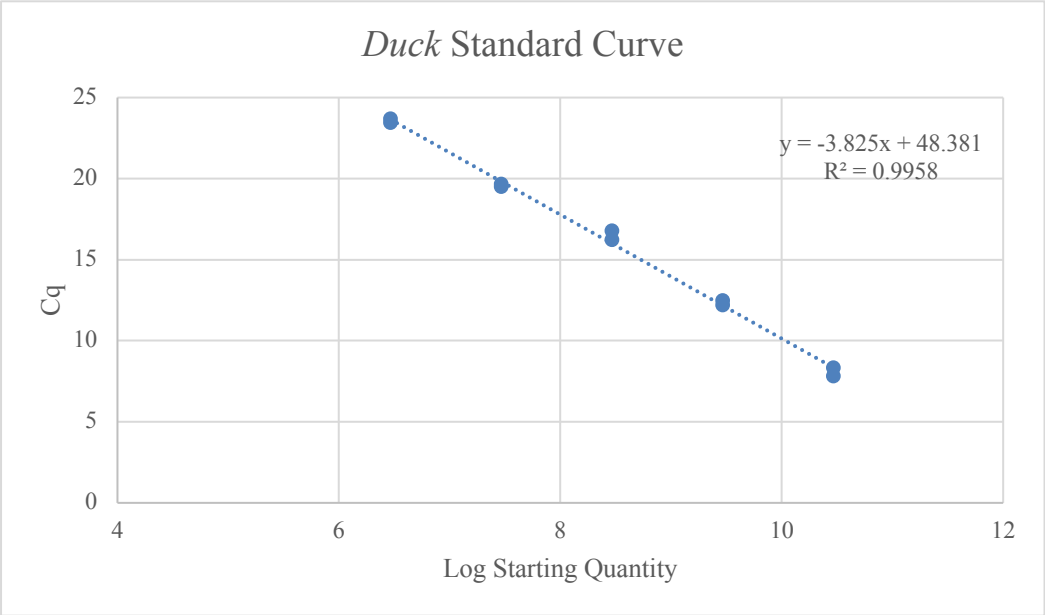

Figure E1. *Duck* standard curve plotted with measured Cq values from qPCR against the log of the relative concentration. R^2 is 0.9958, indicating log starting quantity values are closely related to Cq values.

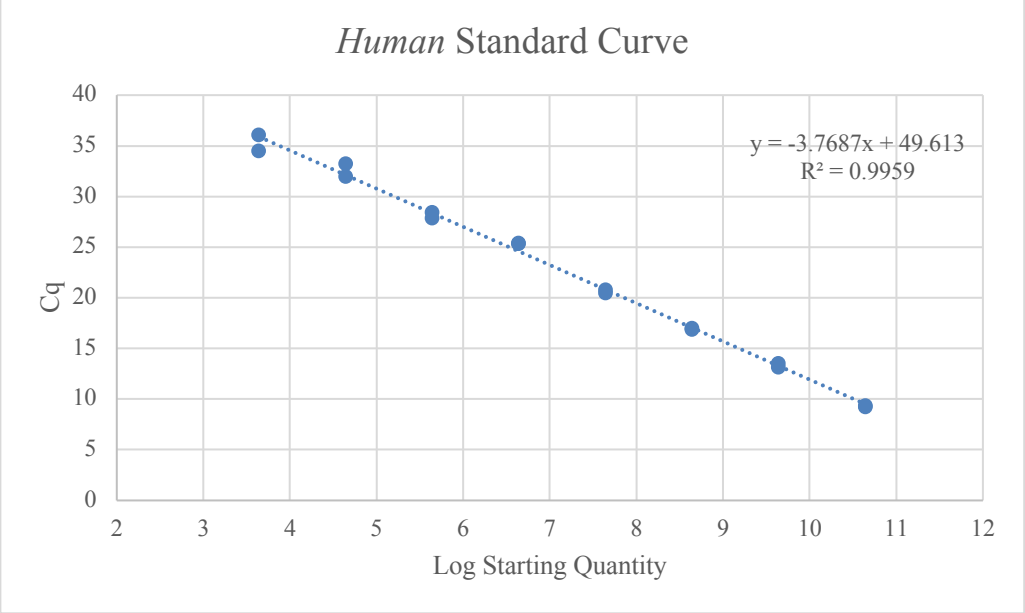

Figure E2. *Human* standard curve plotted with measured Cq values from qPCR against the log of the relative concentration. R^2 is 0.9959, indicating log starting quantity values are closely related to Cq values.

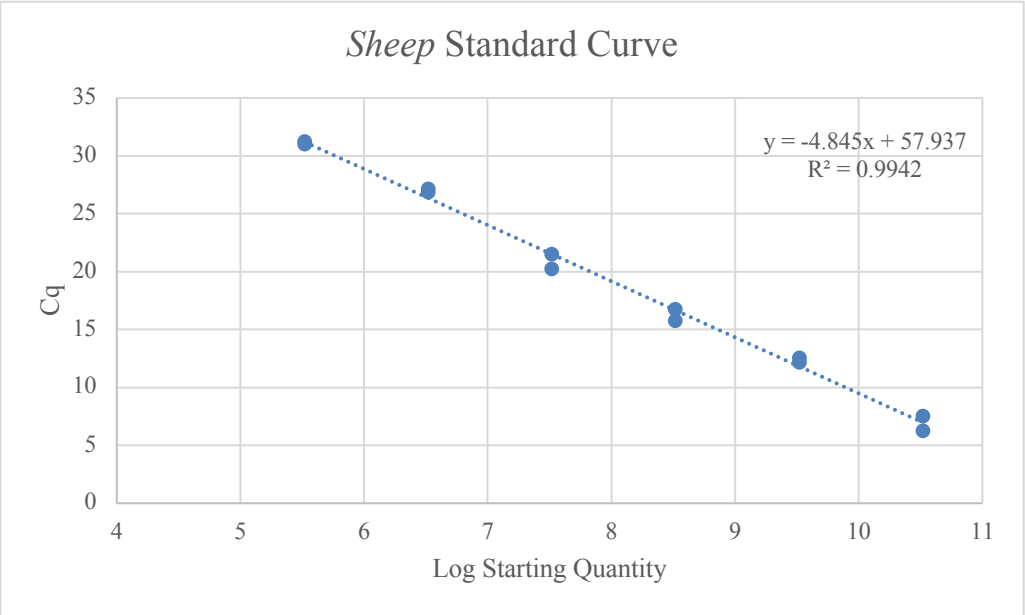

Figure E3. Sheep standard curve plotted with measured Cq values from qPCR against the log of the relative concentration. R^2 is 0.9942, indicating log starting quantity values are closely related to Cq values.

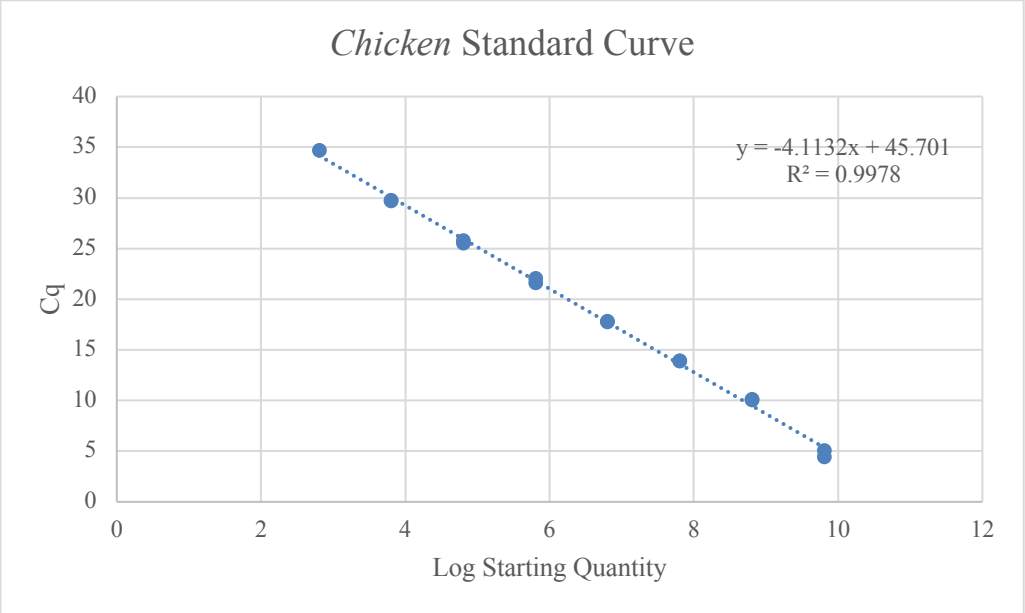

Figure E4. Chicken standard curve plotted with measured Cq values from qPCR against the log of the relative concentration. R^2 is 0.9978, indicating log starting quantity values are closely related to Cq values.

Appendix F: Sequences Obtained from Clone Library Plasmid DNA

> Bison
GTCACTCACCCCCAAAATGCATTACCCAAACGGGGGAATATACATAACATTA
ATGTAATAAAAACATATTATGTATATAGTACATTAAATTATATGCCCCATGCA
TATAAGCAAGTACTTATCCTCTATTGACAGTACATAGTACATAAAGTTATTAA
TTGTACATAGCACATTATGTCAAATCTACCCTTGGCAACATGCA

> Chicken
ACCCTATTTGACTCCCTCAACCAAAGCAGGTTAACCTATGACAATAGAAGAA
TCAATGCTAAAATGAGTAATCTGGAACCTATCCTCCTACGGCGTAAACTTACA
TTAATACATTATTAACAGAACTCAACTTATACCCCCACACTAACAAGCAATAC
GTATTCCTCAATCTGTTAAGCCAACCCAGGAGCGCCCACAGGATGATTAAAA
CCTACAGAAGGAACTCGGCAAACCAAAGACCCGACTGTTTCCCAAAAACATA
GCCTTCAGCTAACAACAAGTATTGAAGGTGATGCCTGCCCAGTGACCCCCAA
AGTTCAACGGCCGCGGTATCCTAACCGTGCGAAGGTAGCGCAATCAATTGTC
CCGTAAATTGAGACTTGTATGAATGGCTAAACGAGGTCTTAACTGTCTCCTGT
AGGTAATCTATGAAATTAGTATTCCCGTGCAAAAACGAGAATGTGAACATAA
GACGAGAAGACCCTGTGGAACTTTAAAATCACGACCACCTTACAACCTTACA
CAGCCCCACTGGGTCCACCCACACATAAACCCCTGGTCGACAT

> Cow
CGACTAAACAACCAAGATAGAATAAAACAAAACATTTAATCCCAATTTAAAG
TATAGGAGATAGAAATCTAAGTACGGCGCTATAGAGAA

> Duck
AACCTACTTAGGATCTTTCGCCCTATCCATCCTAGTAATAATCCTGACCACAC
AGACCTTCTAATGGCCCCAAACATCCGCAAATCCCACCCCCTACTAAAAATA
ATCAACAACTCCCTAATCGACCTTCCCGCACCCTCTAATA

> Goose
AGTACTAAATGCCACCCTGATACTAGTACTTTTATCCCTAGCAGGCCTCCCCC
CATTAACAGGCTTTATACCAAAGTGACTTATTATCCAAGAACTAACTAAACA
GGAAATAACACCAGCAGCCATAGCAATTG
> Human
GACGCCCTCGGCTTACTTCTCTTCCTTCTCTCCTTAATGACATTAACACTATTC
TCACCAGACCTCCTAGGCGACCCAGACAATTATACCCTAGCCAACCCCTTAA
ACACCCCTCCCCACATCAAGCCCGAATGATATTTCCTATTCGCCTACACAATT
CTCCGATCCGT

> Sheep
TACACTGTTACAGGCATCAGAGTATTATGAAGCACCCTTTACAATCTCAGACG
GGGTTTACGGTTCAACTTTCTTCGTAGCTACAGGATTTCACGGCCTCCATGTC
ATCATCGGATCCACCTTCCTAATTGTCTGCTTCTTCCGCCAATTGAAATTTCAT
TTCACCTCTAGTCACCATTTCGGTTTCGAAGCCGCTGCCTGATACTGACACTT
CGTAGATGTAGTATGACTTTTCCTCTATGTATCCATCTACTGATGAGGCTCAT

GTCCTTTTAGTATTAATTAGTACAACTGACTTCCAATCAGTTAGTTTCGGTCTA
ATCCGAAAAAGAACAATAAACCTTATAATTACTCTCCTAACTAACTTCACG

Appendix G: Modified Plate for River Sample DNA Extraction

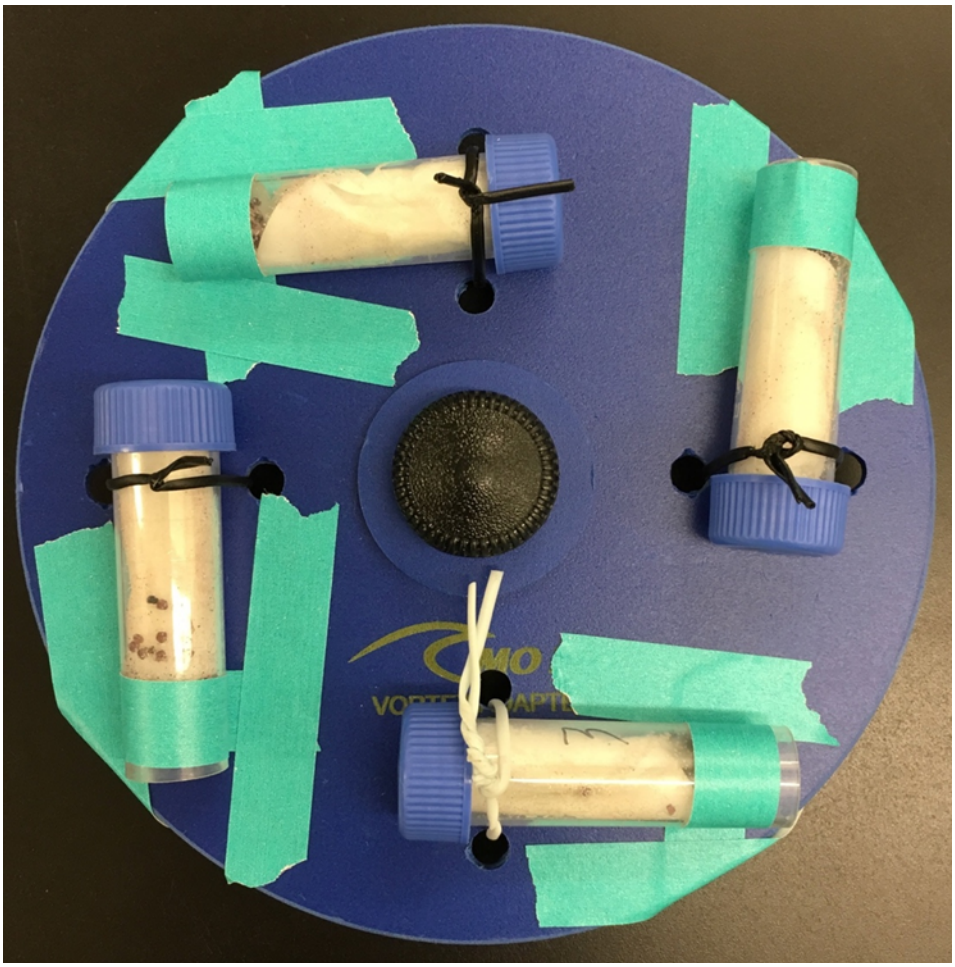

Figure G1. Modified vortex plate to accommodate additional Qiagen BeadTubes.

Appendix H: Illinois River Sample Data

Table H1. mtDNA concentrations detected at sampling locations along the Illinois River.

Target animal	Sampling Location Concentrations (copies/µL)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Bison	-	-	-	-	-	-	-	-	-	-
Chicken	6,497.1	-	-	-	-	-	-	-	-	-
Cow	-	-	-	-	-	-	-	-	-	-
Duck	-	-	-	-	-	-	-	-	-	-
Goose	-	-	-	-	-	60,114.8	10,641.6	-	-	-
Human	56,511.2	467,746.6	66,209.9	58,798.3	-	17,718.2	10,915.3	378,377.0	44,744.9	5,770.9
Pig	-	-	-	-	-	-	-	-	-	-
Sheep	-	-	-	-	-	-	-	-	-	28,451.6

Appendix I: Comparison of Target Animal Subspecies/Breeds to Clone Library Standard Sequences

Table I1. Comparison of bison subspecies/breed to bison sequence data

Bison Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_033873.1	Bison schoetensacki	80%
NC_027233.1	Bison priscus	95%
NC_014044.1	Bison bonasus	86%
NC_012346.1	Bison bison	100%

Table I2. Comparison of chicken subspecies/breed to chicken sequence data

Chicken Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
CM008858.1	Gallus gallus (Yeonsan ogye)	100%
NC_007239.1	Gallus lafayetti	97%
NC_007240.1	Gallus sonnerati	95%
NC_007238.1	Gallus varius	99%

Table I3. Comparison of cow subspecies/breed to cow sequence data

Cow Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_013996.1	Bos primigenius	100%
NC_025563.1	Bos mutus	92%
NC_012706.1	Bos javanicus	100%
AY126697.1	Bos indicus (Nellore breed)	99%
NC_006380.3	Bos grunniens	92%
NC_024818.1	Bos gaurus (Mondulkiri breed)	89%
NC_036020.1	Bos frontalis	89%
NC_006853.1	Bos taurus	100%

Table I4. Comparison of duck subspecies/breed to duck sequence data

Duck Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_028346.1	Anas clypeata	97%
NC_024631.1	Anas acuta	98%
NC_023832.1	Anser cygnoides	89%
NC_022452.1	Anas crecca	99%
NC_022418.1	Anas poecilorhyncha	100%
NC_009684.1	Anas platyrhynchos	100%

Table I5. Comparison of goose subspecies/breed to goose sequence data

Goose Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_039888.1	Anser albifrons	97%
NC_025654.1	Anser indicus	96%
NC_016922.1	Anser fabalis	97%
NC_007011.1	Branta canadensis	93%
NC_011196.1	Anser anser	96%
NC_039888.1	Anser albifrons frontalis	97%

Table I6. Comparison of human subspecies/breed to human sequence data

Human Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_012920.1	Homo_sapiens	100%
CM003501.1	Homo sapiens isolate CHM1	100%
CM003309.1	Homo sapiens isolate CHM13	100%

Table I7. Comparison of sheep subspecies/breed to sheep sequence data

Sheep Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_001941.1	Ovis aries	99%
AY858379.1	Ovis aries	99%
NC_026064.1	Ovis vignei	99%
NC_026063.1	Ovis orientalis breed Asian mouflon	99%
NC_020656.1	Ovis ammon isolate h77	99%
NC_015889.1	Ovis canadensis	96%

Table I8. Comparison of pig subspecies/breed to pig primer design data

Pig Accession Comparison		
Accession Number	**Breed**	**Identity to Clone Library Sequence**
NC_023536.1	Sus verrucosus	99%
NC_014692.1	Sus scrofa taiwanensis	100%
NC_000845.1	Sus scrofa	100%
NC_012095.1	Sus scrofa domesticus	100%
NC_039090.1	Sus scrofa cristatus	100%
NC_024860.1	Sus celebensis	99%
NC_023541.1	Sus cebifrons	98%
NC_026992.1	Sus barbatus	100%

VITA

David Kyle Taylor

Candidate for the Degree of

Master of Science

Thesis: FECAL CONTAMINANT SOURCE TRACKING USING MITOCHONDRIAL DNA DETECTION OF TARGET ORGANISM BY PCR AND QPCR

Major Field: Environmental Engineering

Biographical:

Education:

Completed the requirements for the Master of Science in Environmental Engineering at Oklahoma State University, Stillwater, Oklahoma in December, 2018.

Completed the requirements for the Bachelor of Science in Environmental, Health, and Safety Management at Northeastern State University, Tahlequah, Oklahoma in 2016.

Experience:

Teaching Assistant Position, advised by Dr. Brad Rowland

Graduate Research Assistant, dissertation focused on fecal contaminant source tracking project advised by Dr. Mark J. Krzmarzick

Project Supervisor at Refluoresce Inc., Tulsa, Oklahoma; Supervisor role overseeing 10 employees as the team lead for projects such as hazardous waste management, combustible dust remediation, mold abatement, and plant turnaround services.