2015

Primer efficacy in the DNA barcoding of spiders

Rhennetta Jo Bork
University of Northern Iowa

PRIMER EFFICACY IN THE DNA BARCODING OF SPIDERS

A Thesis Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors

Rhennetta Jo Bork

University of Northern Iowa

May 2015

PRIMER EFFICACY IN THE DNA BARCODING OF SPIDERS

This Study By: Rhennetta Jo Bork

Entitled: Primer Efficacy in the DNA Barcoding of Spiders

has been approved as meeting the thesis project or requirement for the Designation University Honors

_____ _____
Date Dr. Peter Berendzen, Biology

_____ _____
Date Dr. Jessica Moon, Director, University Honors Program

Abstract

DNA barcoding is the process of amplifying a 650 base pair segment of the sequence of the mitochondrial gene cytochrome c oxidase (COI), and amplifying this gene with a polymerase chain reaction (PCR). It is used to help identify and distinguish animal species and also to help determine genetic differences in species. DNA barcoding can be especially useful when working with spiders since they tend to be very small and hard to distinguish. However, achieving a DNA barcode can be difficult and thus techniques to improve the method of DNA barcoding can be helpful. This research looked into the different primer combinations previously used to amplify the barcode region in spiders. These primers were used in nine different combinations and tested on 43 spiders from four different counties in South Dakota. Spiders that were amplified successfully by the universal primer pair (LCO-1490 and HCO-700ME) were tested with the other eight primer sets. Primer sets that contained Chelicerate Reverse 1, Chelicerate Reverse 2, Lepidoptera Forward or LCO-1490 worked the best by successfully amplifying the most number of species. Overall, nine different species of spiders were barcoded successfully and will be entered into GenBank.

TABLE OF CONTENTS

PRIMER EFFICACY IN THE DNA BARCODING OF SPIDERS

Introduction

One of the most complex fields in biology is that of taxonomy. Establishing accurate evolutionary relationships is important to help create taxonomy. Taxonomy is the system in which species are classified and identified. Species classification allows scientists to determine the relationships between species and species identification allows for determining what specimens have been collected. Taxonomy is important in helping scientists understand the complex biological relationships and the diversity of life because this diversity underpins all biological studies (Hebert, Cywinska, & Ball, 2003). However, taxonomy identifications are often misdiagnosed due to phenotypic and genetic variability, morphologically cryptic data, limited identification during certain life stages, and a need for expertise in a field to confidently identify species (Hebert et al., 2003). To resolve these misdiagnoses, a new and emerging field in taxonomy is DNA barcoding.

DNA barcoding provides an efficient method for species level identifications (Sun, Li, Kong, & Zhang, 2012). The DNA barcode is achieved by amplifying a 650 base pair segment of the sequence of the mitochondrial gene cytochrome c oxidase (COI) by using polymerase chain reaction (PCR). This amplification is typically done with universal primers that anneal or attach to the COI gene on many different species types. The 650 base pair segment obtained from amplification can then be sequenced and utilized to identify species based on the sequence variation between species (Hajibabaei et al., 2006; Lin et al., 2014). A barcode is successfully made for a species when genetic diversity occurs between species rather than within a species, which occurs when interspecific distances are much higher than intraspecific distances (Mankga, Yessoufou, Moteetee, Daru, & van der Bank, 2013). These distances can be determined using statistical analyses and phylogenetic trees to analyze the barcode sequences obtained. The

current study tested the nine different primer sets using seven of the various primers used in the DNA barcoding of spiders to determine which primers most efficiently amplified the region for sequencing so that a barcode could be established.

The purpose of this research was to obtain a DNA barcode from the COI region of spiders to determine genetic variances within species and between species. DNA barcoding and the input of barcode sequences into databases allows for easier species identification in scientific research, commercial markets and forensic testing. A DNA barcode of spiders is achieved by extracting DNA from the leg of a spider using a kit and then running the genetic material through a polymerase chain reaction (PCR) to amplify the genetic material specified in the barcode. The genetic material will then be sent to an outside lab for sequencing, so the sequences can eventually be analyzed using computer techniques. Another component of the current research involved the use of nine different DNA primers to determine which primer works best for the barcoding of spiders. This will help researchers around the world to more easily identify the species of hard to identify spiders.

This research could greatly improve the ability and efficiency of barcoding and data basing spiders with DNA barcodes by helping to determine what primer set most efficiently amplifies the barcode region of Midwestern spider species. These barcodes could be entered into international databases and help make the identification and species resolution of spiders much easier. This research will also be able to specifically add new species of South Dakotan spiders to the GenBank database and help to determine the various spider species residing in four counties in South Dakota.

Literature Review

Barcoding Benefits

Due to the difficulty taxonomists face in determining evolutionary relationships to identify species, there must be a simple and easy method to identify specimens. DNA barcoding works well for this due to the large amount of barcode divergence within species (Hebert & Gregory, 2005). There are many other additional factors that make the use of barcoding effective. For one, DNA barcoding is a rapid form of molecular data collection and, along with morphology based identification techniques, can be considered extremely accurate in species identification and resolution (Hebert & Gregory, 2005; Prendini, 2005; Sun et al., 2012). Furthermore, DNA barcoding is beneficial because DNA remains constant throughout an animal's life, is present in all cells, has thermostability, and DNA contains molecular markers, such as COI, that can be easily obtained from animal tissue. (Greenstone et al., 2005; Lin et al., 2014). Meanwhile, the mitochondrial COI region is useful in barcoding because all animals possess this gene, COI barcodes diverge enough to allow species resolution between very closely related species, and insertions and deletions in the processing of this gene rarely occur so the gene can still be amplified and identified by primers and is fairly uniform in length (Barrett & Hebert, 2005). Moreover, DNA primers are robust and easy to use and DNA barcoding is very cheap (only about 2 dollars per specimen), which adds to the utility of barcoding (Hebert & Gregory, 2005; Sun et al., 2012).

Barcode Libraries

Once these barcodes are obtained, the goal is to assemble a reference library of barcode sequences for all species so that species identification can be obtained quickly and during any stage of life. This is called the Barcode of Life Initiative and species are being inputted by

scientists around the world (Barrett & Hebert, 2005; Greenstone et al., 2005; Bergsten et al., 2012; Sun et al., 2012). This will help aid ecologists, conservationists, and workers involved in the control of pests, invasive species, and food safety by allowing these specialists to determine the identification of species in an area and how these species interact (Hebert & Gregory, 2005). In fact, this database currently consists of over 500,000 COI barcodes and has already been established for over 38,000 species. This means that not only have 38,000 species been added to this international database, named GenBank, but also that multiple specimens from each species have been barcoded as well. The utility of such databases resides in its ability to assign individuals to known species and to establish new species from diverging taxa (Moulton, Hojun, & Whiting, 2010) (Hebert & Gregory, 2005). These libraries also have the potential to become a tool that helps assess intraspecific variation across geographic regions (Miller, Beentjes, Helsdingen, & IJland, 2013). Overall, the purpose of barcode libraries is to allow DNA taxonomists to create a global inventory of life's diversity (Hebert & Gregory, 2005).

Cryptic Species

In addition, barcoding libraries accelerate the process of identifying cryptic species, which occur widely in museum specimen collections and voucher species collections (Sun et al., 2012). Cryptic species are species that are morphologically identical and can only be identified by molecular means. On the molecular level, a cryptic species is a species that has two % or more variation in its sequences from another organism identified as a member of the same species. DNA barcoding can be helpful in resolving cryptic species by matching organisms on the molecular level with either the holotype, the species it was originally thought to belong to, or the new species when morphological data is not able to determine the difference between the two (Hajibabbaei et al., 2006; Blagoev, Nikolova, Sobel, Hebert, & Adamowicz, 2013; Hernandez-

Triana, Prosser, Rodriguez-Perez, Chaverri, & Hebert, 2014). The resolution of cryptic species is especially important when barcoding is used on older specimens, such as voucher specimens or museum specimens, because these organisms are often old and degraded. Nonetheless, it is important to obtain barcodes from museum species libraries in order to make connections between collections of different ages. This allows for the observation of species distribution, integration of morphological and molecular data, the inclusion of rare species in barcode libraries, and the creation of libraries with broad geographical taxa data in a cost effective manner (Prendini, 2005; Hajibabaei et al., 2006; Miller et al., 2013; Hernandez-Triana et al., 2014).

Previous Barcoding Studies

While barcoding has already accomplished the identification of species of snails, butterflies, fish, amphibians and birds, the hope of barcode libraries in the future is that they will contain all of species of animals and be able to be easily accessed by scientific researchers and lay people alike (Sun et al., 2012). In addition, GenBank, an international barcode database, has already been used in research to analyze barcode data and this research has indicated that species boundaries can also be successfully deduced by COI gene divergence (Hebert, Penton, Burns, Janzen, & Hallwachs, 2004). Barcoding has also been used to successfully identify degraded organisms in feces in predator/prey studies, to match eggs and larval stage carabids and spiders with adult parents, to identify instances of food fraud in consumer meat markets, to determine the identification of medicinal plants in South African marketplaces where morphology based identification can be dangerous and misleading, and in forensics to identify black fly species when defining the time or location of death of a human victim (Wells & Sperling, 2001;

Greenstone et al., 2005; Zeale, Butlin, Barker, Lees, & Jones, 2011; Mankga et al., 2013; Lin et al., 2014).

Problems with DNA Barcoding

While barcoding is an innovative field, it can also be a problematic field and there are some problems that still must be worked out in order to effectively database all animals. Many scientists are skeptical as to whether or not barcoding will be able to efficiently identify all animal species, as barcoding has already failed in some species of cnidarians (jellyfish) and the universal primers used in barcoding are not compatible with all species (Sun et al., 2012). Furthermore, universal primer binding sites for the COI gene are not present in all species, such as nematodes, and the DNA markers used in barcoding are not conserved enough to be amplified in all domains of life (Prendini, 2005).

Critics also find fault in barcoding due to the variation that naturally occurs in all animals. DNA barcoding is based on sequence divergence percentages even though these percentages are mostly unknown and likely vary from species to species, meaning that there is no universal divergence percentage that would indicate if specimens are a part of the same species or not (Prendini, 2005). DNA barcoding efficiency also decreases if a species has more overlap in variation within a species and less divergence separating sister species (Sun et al., 2012). Opponents also criticize the use of mitochondrial DNA fragments for species based identifications since mitochondrial DNA is maternally inherited and may fail to identify certain species due to male biased gene flow (Prendini, 2005). Lastly, DNA barcoding is highly criticized due to the fact that it is very difficult to barcode specimens older than 10 years, such as museum organisms or voucher specimens, because at this point DNA may begin to degrade owing to drying out and uncertainty in storage techniques (Hajibabaei et al., 2006; Mankga et al.,

2013; Hernandez-Triana et al., 2014). However, the problem of old specimen degradation cannot always be remedied by using fresh specimens alone, since fresh specimen collection may not yield complete species variation and weather conditions, time of year, and collecting methods can influence what types of species and how many total species are actually caught (Miller et al., 2013).

Numts

Numts provide another barrier in the accuracy and efficiency of DNA barcoding. Numts are nuclear mitochondrial pseudogenes, mitochondrial genes that have been inserted into the nuclear genome and no longer code for proteins. They can also be amplified by the universal primers used in barcoding. (Moulton et al., 2010; Lobo et al., 2013). Numts cause incorrect species identification and overestimation of species because when amplified in barcode sequences, numts provide higher percentages of sequence divergence than actually exists (Moulton et al., 2010). Furthermore, most numts do not contain stop codons, which makes them difficult to distinguish from mitochondrial gene material. This decreases the accuracy of DNA barcoding because the numts are then not able to be sorted out of the DNA sequence (Moulton et al., 2010).

Primer Design

It has been suggested that careful primer design is important in achieving the correct barcode fragment in PCR. It has also been suggested that increasing the primer specificity when barcoding can help eliminate numt co-amplification and increase the recognition of hard to distinguish stop codons (Moulton et al., 2010). The reason that universal primers do not work well for all species is because primer length, degeneracy, G to C ratios, melting temperatures, and the degree to which a primer matches the complimentary fragment sequence varies in

primers and can affect the ability of a primer to anneal to the extracted DNA. Although more testing is needed to verify the ability of primer specificity to eliminate numts, it has been suggested that taxonspecific primers that are exact complimentary matches of the target sequences has proven useful for isolating small mitochondrial DNA fragments (Moulton et al., 2010). Additional experiments are also needed to determine what primers work best for what species, as a tenant of barcoding is that primers must be accurate and cost effective (Lobo et al., 2013). For this reason, creating primer cocktails has been suggested as a solution to barcode species precisely and in a cost effective manner. In fact, these cocktails have already successfully barcoded birds, fish, butterflies, and some species of mammals and look to be a promising technique in the barcoding of many other species (Lobo et al., 2013).

Barcoding in Spiders

Increased barcoding efficiency and barcoding databases will prove especially useful in the identification and species resolution of spiders as they are small and can be difficult to identify based on morphology alone (Barrett & Hebert, 2005). Spiders can also be difficult to identify as there are more than 44,000 species of spiders and spiders can vary in appearance depending on life stage and gender (Barrett & Hebert, 2005; Greenstone et al., 2005; Prendini et al., 2005; Blagoev et al., 2013). However, it is important to properly identify spiders as they are predators and their behavior and appearances can help us to understand cases of sexual dimorphism, mating systems, and studies of social interactions in many biological situations. Furthermore, the 44,000 species of spiders already known are thought to be only fraction of the species of spiders in existence due to the difficulty in identifying and resolving spider species (Barrett & Hebert, 2005). While it has been proven that spiders and other arachnids can be successfully identified using COI barcoding, it is still important to improve and create accurate

spider species identification techniques so that spider identification and species resolution can be both accessible and efficient (Barrett & Hebert, 2005).

Materials and Methods

The purpose of this research was to see what primers work best to amplify the COI gene region in spiders from South Dakota. Increased primer specificity has been suggested to improve the barcode region amplification in animals, so the current research was done to determine if improving the primer specificity in spiders helps to improve the amplification process.

2014 Collections

Fresh specimens were caught by hand or with a vacuum aspirator in Davison County, South Dakota during the months of May, June and July 2014. The spiders were then preserved in 70 % ethanol and stored at room temperature. A total of 18 fresh specimens were used for this research.

2010 Collections

Spiders were selected from a group of voucher specimens obtained from the lab of Dr. Brian Patrick at Dakota Wesleyan University. Specimens were caught in the Fort Pierre National Grasslands in Stanley Country, Jones County, and Lyman County South Dakota in the summer of 2010. Samples were either caught by hand or using a ramp trap. Some 2010 species were tested twice due to a high abundance of voucher specimens in that species. Gender of 2010 spiders was not considered in this study, but it was noted. Specimens were kept in glass vials with rubber stoppers and stored in 70 % ethanol at room temperature. A total of 21 specimens from 2010 that had been stored in ethanol were used in this research. A map of the locations of fresh and voucher specimens can be seen in Figure 1.

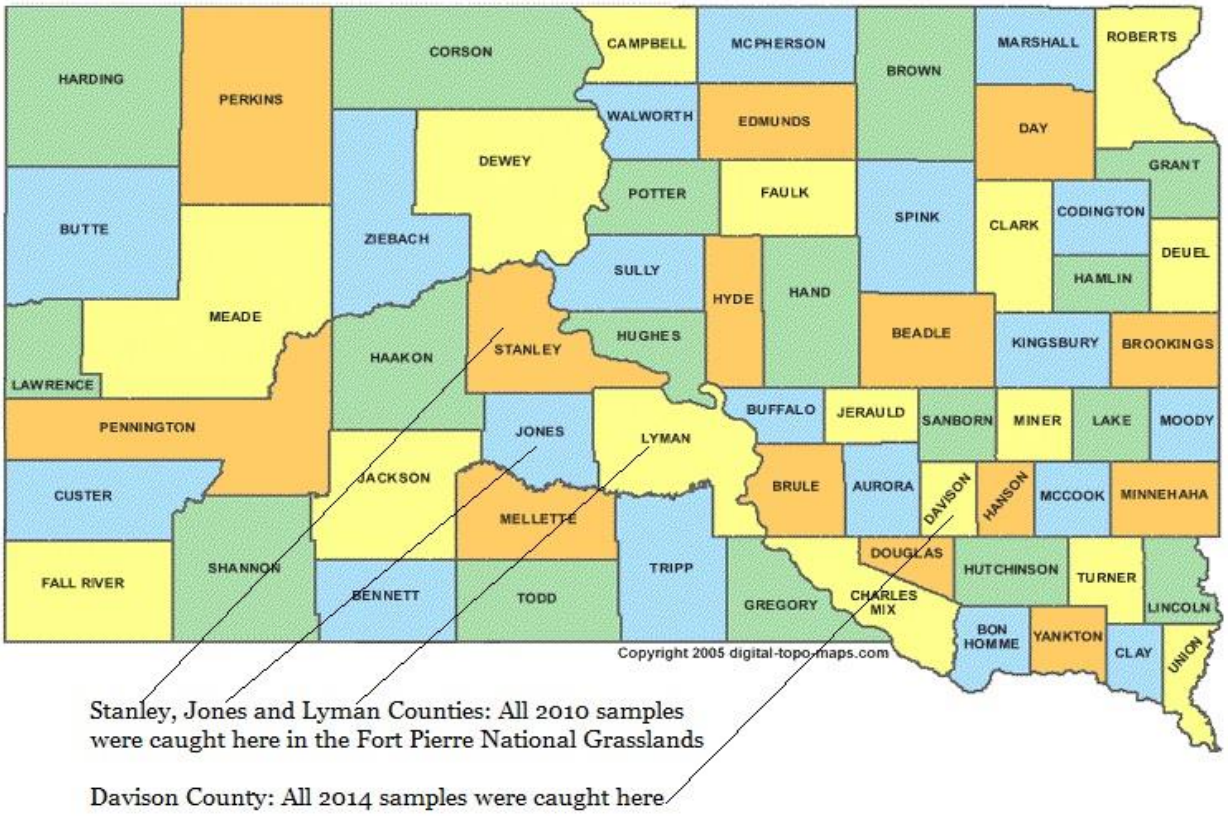

Figure 1: A map of South Dakota with Davison, Jones, Lyman and Stanley Counties highlighted.

PCR Optimization

Source tissue was obtained from large specimens by removing a leg with tweezers. Small specimens (1 to 4 mm) were placed in entirety into the micro-centrifuge tube. Genomic DNA was extracted using a Qiagen DNeasy® quit following the manufacturer's recommendations (Qiagen Inc. https://www.qiagen.com/us/). Both specimen types were rinsed in reverse osmosis water prior to being placed in the micro-centrifuge tubes as ethanol inhibits proteinase K. Samples were incubated for four hours at 56 degrees Celsius, rather than the standard 3three hour time period to enhance lysis and DNA extraction. Small specimens subjected to whole body extraction were not vortexed to mix to ensure that the bodies were not destroyed. These samples were instead pipetted to mix and then returned to their vial filled with ethanol after the

incubation step. This ensured that any small spiders used could continue to be used for extraction or identification purposes in the future.

To obtain the mitochondrial COI gene, several primers were used based on what was effective in prior spider research. Various primers were paired up from previous research to determine which combination best amplified the COI gene in spiders. A list of the primer "cocktails" used and their sequences are given in Table 1. Since DNA quantities were not available early in the research, samples were amplified with other primer combinations after being effectively amplified by universal primers to ensure that DNA extraction occurred. PCR reactions contained 2.5 µL buffer solution, 2 µL nucleotides (dNTPs), 0.5 µL of primer, 0.125 µL Takara Taq polymerase, 2.5 µL template DNA and 16.875 µL water for a reaction total of 25 µL. The initial denaturation step of PCR occurred at 94°C for 1 min, followed by a denaturing step of 94°C for 45 seconds and an annealing step of 48°C for 45 seconds. Extension occurred at 72°C for 30 seconds, the final extension occurred at 72°C for 5 minutes, and the samples were then refrigerated at 4°C until PCR terminated. Samples were stored at -20°C until they were sequenced. Samples were packed in vials with lids on dry ice to prevent evaporation during the mailing process. Samples were sent to the CORE Sequencing Facility at Black Hills State University where gel pictures were obtained. Gel pictures were analyzed see which primer set works best at amplification of DNA by looking for the presence of two bands next to the sample number. This can be seen in Figure 2. DNA quantities were also found using a Cubit Fluorometer.

Table 1- Primers used and their sequences

Primer Name	Sequence
LCO 1490 (Miller et. al., 2013)	5'-GGTCAACAAATCATAAAGATATTGG-3'
HCO-700ME (Breton et. al., 2006)	5′-TCAGGGTGACCAAAAAATCA-3′
Chelicerate Forward 1 (Briscoe et. al., 2013)	5'- TACTCTACTAATCATAAAGACATTGG-3'
Chelicerate Reverse 1 (Barret & Hebert, 2005)	5'-CCTCCTCCTGAAGGGTCAAAAAATGA-3'
Chelicerate Reverse 2 (Miller et. al., 2013)	5'-GGATGGCCAAAAAATCAAAATAAATG-3'
Lepidoptera Forward (Hebert et. al., 2004)	5′-ATTCAACCAATCATAAAGATAT-3′
Lepidoptera Reverse (Hebert et. al., 2004)	5′-TAAACTTCTGGATGTCCAAAAA-3′

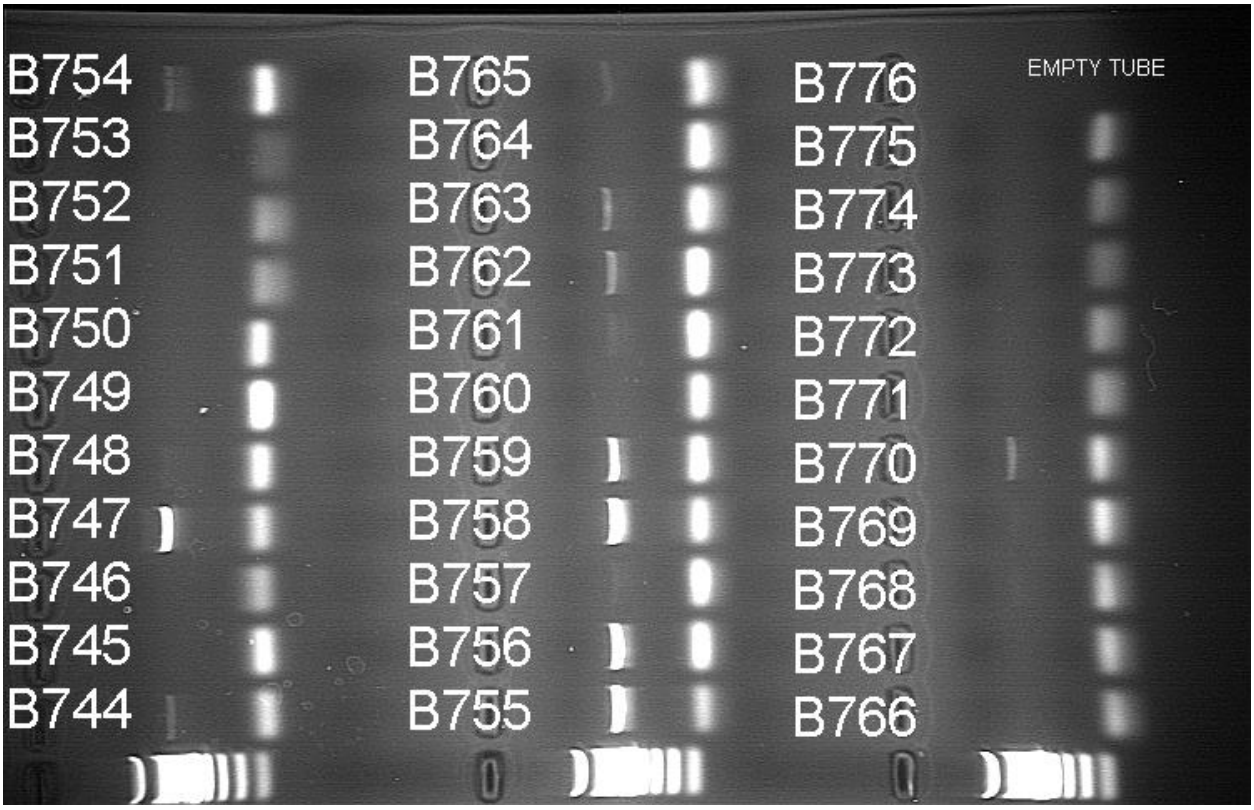

Figure 2: Agarose gel image of positive amplifications of DNA seen where there are multiple bands next to the sample number.

Results

Primers LCO 1490 and HCO-700ME successfully amplified all 43 samples used based on the appearance of two bands on their gel pictures, and thus had 100% amplification. DNA concentration values were tested for with a cubit fluorometer at a later date. It was found that four samples did not contain any measurable amounts of DNA and perhaps only produced two bands due to contamination. Percentages of samples amplified were then adjusted after the four samples that were found not to contain any DNA were removed from the sample list. These samples were removed from the rest of the data tables, so any percentages seen do not contain these values. The percent of amplification achieved by each primer set can be seen in Table 2. LCO 1490 and Chelicerate Reverse 1 amplified on average 67.4 % of the samples. Lepidoptera forward and Chelicerate Reverse 1 amplified 51.2 % of the samples, while LCO 1490 and Chelicerate Reverse 2 amplified 44.2 % of samples. Chelicerate Forward 1 with Chelicerate Reverse 1, Chelicerate Forward 1 with Chelicerate Reverse 2, and Lepidoptera Forward with Chelicerate Reverse 2 had an average amplification of 24.0 %. LCO 1490 with Lepidoptera Reverse, Chelicerate Forward 1 with Lepidoptera Reverse, and Lepidoptera Forward with Lepidoptera Reverse only amplified an average of 0.76 % of the samples.

Table 2: Percent of samples amplified for each primer set, with and without the four samples that contained no DNA

Primer Sets	# of Samples Amplified	% of Samples Amplified	# of Samples Amplified (With 4 removed with no DNA)	% of Samples Amplified
LCO 1490 & HCO 700ME	43/43	100%	39/39	100%
LCO 1490 & Chelicerate Reverse 2	19/43	44.2%	19/39	48.7%
LCO 1490 & Chelicerate Reverse 1	29/43	67.4%	29/39	74.4%
LCO 1490 and Lepidoptera Reverse	0/43	0%	0/39	0%
Chelicerate Forward 1 & Chelicerate Reverse 1	7/43	16.3%	7/39	17.9%
Chelicerate Forward 1 & Chelicerate Reverse 2	9/43	20.9%	9/39	23.1%
Chelicerate Forward 1 & Lepidoptera Reverse	1/43	2.3%	1/39	2.56%
Lepidoptera Forward & Chelicerate Reverse 1	22/43	51.2%	22/39	56.4%
Lepidoptera Forward & Chelicerate Reverse 2	15/43	34.9%	15/39	38.5%
Lepidoptera Forward & Lepidoptera Reverse	0/43	0%	0/39	0%

Additionally, in order to see if any primer sets could be deemed more efficacious for 2010 or fresh specimens the number and percentage of amplifications for 2010 versus fresh specimens were calculated (Table 3).

Table 3: The number and % of amplifications for 2010 vs. 2014 samples

Primer Sets	Number of 2010 Samples Amplified	Number of 2014 Samples Amplified	% of 2010 Samples Amplified Vs. % of 2014 Samples Amplified
LCO 1490 & HCO 700ME	21/21	18/18	2010= 53.8 2014= 46.2
LCO 1490 & Chelicerate Reverse 2	7/21	12/18	2010= 36.8 2014= 63.2
LCO 1490 & Chelicerate Reverse 1	15/21	14/18	2010= 51.7 2014= 48.3
LCO 1490 and Lepidoptera Reverse	0/21	0/18	2010= 0 2014= 0
Chelicerate Forward 1 & Chelicerate Reverse 1	4/21	3/18	2010= 57.1 2014= 42.9
Chelicerate Forward 1 & Chelicerate Reverse 2	5/21	4/18	2010= 55.6 2014= 44.4
Chelicerate Forward 1 & Lepidoptera Reverse	0/21	1/18	2010= 0 2014= 100
Lepidoptera Forward & Chelicerate Reverse 1	11/21	11/18	2010= 50 2014= 50
Lepidoptera Forward & Chelicerate Reverse 2	7/21	8/18	2010= 46.7 2014= 53.3
Lepidoptera Forward & Lepidoptera Reverse	0/21	0/18	2010= 0 2014= 0

The samples that were found to lack any measurable DNA were analyzed to determine if they were amplified by any other primer sets than the initial LCO 1490 with HCO-700me combination. From the analysis, it was determined that two samples with no measurable DNA were able to be amplified by more than one primer set (Table 4). These samples were left in all data sets, while the other four samples containing no DNA were removed.

Table 4: Samples with no measurable DNA that were able to be amplified

Samples with no Measurable DNA	Sample 11	Sample 17	Sample 26	Sample 47	Sample 48	Sample 65
Number of Primers Able to Amplify with PCR	6	5	1	1	1	1

Finally, the amplified samples were analyzed to determine what species were successfully amplified and how often each species was successfully amplified with each of the 10 primer sets. In total, nine species of spiders were able to be successfully amplified from the old (2010) tissues (Table 5).

Table 5: A list of species (2010) successfully amplified and also the percentage of times the species was able to be amplified

Species Amplified	Number of Samples Amplified	Number of Samples Tested	% of Samples Amplified
Schizocosa crassipalpata	44	140	31.4
Xysticus bicuspis	4	10	40.0
Xysticus luctans	13	40	32.5
Eridantes erigonoides	1	10	10.0
Schizocosa mccooki	1	10	10.0
Ceratinops littloralis	1	10	10.0
Islandiana flaveola	1	10	10.0
Hogna frondicula	5	10	50.0
Cheiraconthiam mildri	4	10	40.0

Finally, the average number of samples amplified by each primer "cocktail" were determined in order to determine what primers were more robust than others. Primer sets using Lepidoptera Reverse had an average amplification of 0.85 %, while the primer sets that utilized Chelicerate Reverse 1 and Chelicerate Reverse 2 successful amplified an average of 49.6 % of samples and 36.7 % of samples respectively. HCO-700ME successfully amplified 100 % of its samples, but was only used in one primer set. When looking at the forward primers, LCO-1490 amplified an average of 55.8 % of samples successfully, Lepidoptera Forward successfully amplified an average of 31.6 % of samples and Chelicerate Forward effectively amplified 14.5 % of samples.

Discussion

LCO 1490 with HCO-700ME were the most effective primer set, followed by LCO 1490 with Chelicerate Reverse 1 and Lepidoptera Forward with Chelicerate Reverse 1. In all, universal primers worked very well in the amplification of spiders, but it should also be noted that any primer set using Chelicerate Reverse 1 as its reverse primer worked very well too. Finally, primer sets using Lepidoptera Reverse amplified less than one % of samples, and should not be used on spiders to obtain DNA barcode sequences.

LCO 1490 and HCO-700ME appear to be 100 % effective in this research because they were used as a standard to determine which samples would contain DNA and could likely be amplified by other primers. DNA values in nanogram/microliter were measured after the cessation of this research and have been noted, but they were not initially used to determine which samples could likely be amplified again. While this primer combination likely works well with spiders, since it is considered to be a universal primer pair, further research should be done where DNA concentrations are found first and no primer set is used preferentially to determine which samples will likely give results.

LCO 1490 and Chelicerate Reverse 1 was the second best primer pair. This pair successfully amplified 74.4 % of samples (when the four samples that were deemed to have no DNA in them were removed). Additionally, this primer pair successfully amplified 15 voucher specimens and 14 fresh specimens. From the data, it can be determined that this primer set would be fairly effective at amplifying the barcode region of spider tissue samples, old or new.

LCO 1490 with Chelicerate Reverse 2 and Lepidoptera Forward with Chelicerate Reverse 1 were both about equal in primer efficacy. Each of these sets successfully amplified 48.7 % and 56.4 % of the samples respectively (when the four DNA free samples were

removed). While neither of these primer sets were particularly robust, researchers may consider using these sets if they do not have other primers needed and would still like to perform their PCR steps.

The pairs of Chelicerate Forward 1 with Chelicerate Reverse 1, Chelicerate Forward 1 with Chelicerate Reverse 2, and Lepidoptera Forward with Chelicerate Reverse 2 displayed only 17.9 %, 23.1 %, and 38.5 % successful amplifications respectively. Although these primers are more specific towards spiders, they did not give great amplification results and thus should not be used in the PCR based identification of spiders.

The pairs of Chelicerate Forward 1 with Lepidoptera Reverse, LCO 1490 with Lepidoptera Reverse, and Lepidoptera Forward with Lepidoptera Reverse gave 2.56 %, 0 %, and 0 % amplification of samples respectively. None of these combinations should be utilized for the amplification of the DNA barcode in spiders.

From looking at the primers used to successfully amplify spiders, inferences about the success of certain primers can be made. For instance, the primer "cocktails" using Lepidoptera Reverse had an average amplification of 0.85 % which is a very poor result. From this data, it can be assumed that Lepidoptera Reverse should not be used as a reverse primer in the barcoding of spiders. The primer sets that utilized Chelicerate Reverse 1 and Chelicerate Reverse 2 successful amplified an average of 49.6 % of samples and 36.7 % of samples meaning they both had a fairly robust amplification rate and thus would work quite well in the amplification of spiders in other settings. HCO-700ME was used in only one primer set, and thus had an average amplification of 100 %. It works well with spiders from these results and could also be utilized in the future. When looking at the forward primers, LCO-1490 and Lepidoptera Forward had fairly high success rates and would both work very well in the future, especially if used with any of the

reverse primers besides Lepidoptera Reverse. Chelicerate Forward had a moderate amplification rate and should not be used in cases were a more robust primer set is needed for amplification.

When looking at the difference in the ability of these primer sets to amplify old and new tissues, there did not appear to be a great difference in amplification, meaning that the age of the tissue was not a determining factor in the success of the amplification. All primer sets that were able to amplify samples, with the exception of Chelicerate Forward 1 with Lepidoptera Reverse, were able to amplify both old and new specimens, despite the fact that the old specimens have been stored in alcohol for four years and likely had some degradation of DNA. Furthermore, if amplification was achieved, a near equal number of old and new specimens were able to be barcoded each time. On average, primer sets that achieved amplification successfully amplified 44.0 % old specimens and 56.0 % new specimens, which is relatively equal. Overall, there was no obvious pattern found and amplification seemed to be an issue of primer specificity rather than tissue age and preservation.

Of the samples reported to have no measureable DNA, two samples were able to have their COI region amplified by more than one primer set. Sample 11, one of the *Schizocosa crassipalpata* samples, and Sample 17, of the species *Xysticus luctans*, were able to be amplified by six and five primer sets respectively. Sample 11 was amplified by primer sets with an average amplification of 53.1 %. Sample 17 was amplified by primer sets with an average amplification of 59.5 %. Since these samples were able to be amplified by primer sets that were fairly efficacious, it may be assumed that even a small amount of DNA (small enough to not be measured) can be amplified if the primer set is robust enough. For the other four samples that were amplified by one primer set but contained no measurable DNA, it can either be inferred that

the universal primers are most efficacious with little DNA quantities or that the initial amplification with the universal primers occurred due to contamination.

Of the older 2010 spider samples, nine species stored in ethanol for four years were able to be amplified by the various primer cocktails. Of these nine species, no species was able to be amplified 100 % of the time. Successful amplification ranged from 10.0 % of samples to 50.0 % of samples. DNA from both *Xysitcus bicuspis* and *Xysticus luctans* was able to be amplified by at least four primer sets and was amplified by primers that had an amplification average of 65.7 %. DNA from *Shizocosa crassipalpata* was able to be amplified by seven primer sets with an average amplification of 47.8 %. However, these results may be a bit skewed because this species had the most tissue samples included in this research due to the high number of voucher specimens. In addition, DNA from *Hogna frondicula* and *Cheiraconthiam mildri* was able to be amplified by multiple primer sets, but more research needs to be done all around to determine if certain species work better with certain primer sets.

Conclusions

The purpose of this research was to determine which primer set best amplified the DNA barcode region of spiders and to determine genetic variances within species and between species of spiders. This will help researchers around the world to more easily identify what spider species they are looking at and will allow more spiders to be entered into GenBank for public access of the diversity of siders in South Dakota.

There are a few limitations to this research. For one, fresh spiders in this research were caught exclusively in Davison County, South Dakota, while the 2010 spiders were caught only in Lyman County, Jones County, and Stanley County, South Dakota. In future research, fresh and old spiders should be caught in the same counties to make sure there is not a difference in spider populations between counties. Additionally, all spiders in this research were caught between the months of May and August. This is a problem because it may not represent the total diversity of spiders in these counties. Future research should include spiders caught at all times of the year, especially in the early spring and late fall months. Finally, a limitation to this research was that DNA values were not available at first, so the universal primers were used to preferentially determine what samples should be amplified with the other primer sets. The thought behind this strategy was that if a sample was able to be amplified by universal primers, DNA must have been extracted out of it and could be amplified by other primer sets in the future. In future research, DNA values should be determined before any amplifications are performed, and samples with high DNA values should be used with all primer sets. This will eliminate one primer set being preferentially used over the others.

Overall, it did not appear that specificity greatly helped the amplification process in this research. From this research, universal primers provided the most efficacy in amplifying the

barcode region of these samples and a pair combining a universal primer with a more specific spider primer was the next most efficacious. Additionally, this research was successfully able to amplify the barcode regions of 43 spiders total, including 24 old specimens stored in ethanol since 2010 and 19 fresh specimens from 2014. This allowed for nine different species of old specimens to be sequenced so that they may be entered into GenBank and help to determine the genetic diversity and identifications of spider populations in spiders from four South Dakota counties.

References

Barrett, R. H., & Hebert, P. N. 2005. Identifying spiders through DNA barcodes. Canadian Journal Of Zoology 83(3): 481-491.

Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., & ... Vogler, A. P. 2012. The effect of geographical scale of sampling on DNA barcoding. Systematic Biology 61(5): 851-869.

Blagoev, G., Nikolova, N., Sobel, C., Hebert, P., & Adamowicz, S. 2013. Spiders (Araneae) of Churchill, Manitoba: DNA barcodes and morphology reveal high species diversity and new Canadian records. BMC Ecology, 1344.

Breton, S., Burger, G., Stewart, D.T., Blier, P.U., 2006. Comparative analysis of gender-associated complete mitochondrial genomes in marine mussels (Mytilus spp.). Genetics 172, 1107–1119.

Greenstone, M., Rowley, D., Heimbach, U., Lundgren, J., Pfannenstiel, R., & Rehner, S. 2005. Barcoding generalist predators by polymerase chain reaction: carabids and spiders. Molecular Ecology 14(10): 3247-3266.

Hajibabaei, M., Smith, M., Janzen, D. H., Rodriguez, J. J., Whitfield, J. B., & Hebert, P. N. 2006. A minimalist barcode can identify a specimen whose DNA is degraded. Molecular Ecology Notes 6(4): 959-964.

Hamilton, C. A., Formanowicz, D. R., & Bond, J. E. 2011. Species delimitation and phylogeography of Aphonopelma hentzi (Araneae, Mygalomorphae, Theraphosidae): cryptic diversity in North American tarantulas. *PloS one* 6(10): e26207.

Hebert, P. D., Cywinska, A., & Ball, S. L. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(1512): 313-321.

Hebert, P. D., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14812-14817.

Hebert, P. N., & Gregory, T. (2005). The Promise of DNA Barcoding for Taxonomy. Systematic Biology, 54(5), 852-859.

Hernandez-Triana, L. M., Prosser, S. W., Rodriguez-Perez, M. A., Chaverri, L. G., & Hebert, P. N. 2014. Recovery of DNA barcodes from blackfly museum specimens. *Molecular Ecology Resources* 14: 508-518.

Ivanova, N., & Grainger, C. 2006. Protocols: Primer Sets . In *Canadian Centre for DNA Barcoding* .

Lin, C., Fung, L., Chan, P., Lee, C., Chow, K., & Cheng, S. 2014. A rapid low-cost high-density DNA-based multi-detection test for routine inspection of meat species. Meat Science 96(Part A): 922-929.

Lobo, J., Costa, P., Teixeira, M., Ferreira, M., Costa, M., & Costa, F. 2013. Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. BMC Ecology, 1334.

Mankga, L. T., Yessoufou, K., Moteetee, A. M., Daru, B. H., & van der Bank, M. 2013. Efficacy of the core DNA barcodes in identifying processed and poorly conserved plant materials commonly used in South African traditional medicine. Zookeys (365): 215-233.

Moulton, M. J., Hojun, S., & Whiting, M. F. 2010. Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from Orthoptera (Arthropoda: Insecta). Molecular Ecology Resources 10(4): 615-627.

Miller, J. A., Beentjes, K. K., van Helsdingen, P., & IJland, S. 2013. Which specimens from a museum collection will yield DNA barcodes? A time series study of spiders in alcohol. Zookeys (365): 245-261.

Prendini, L. 2005. Comment on "Identifying spiders through DNA barcodes". Canadian Journal of Zoology. pp. 498-504.

Sun, Y., Li, Q., Kong, L., & Zheng, X. 2012. DNA barcoding of Caenogastropoda along coast of China based on the COI gene. Molecular Ecology Resources 12(2): 209-218.

Wells, J. D., & Sperling, F. A. 2001. DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). *Forensic Science International* 120(1): 110-115.

Zeale, M. K., Butlin, R. K., Barker, G. A., Lees, D. C., & Jones, G. 2011. Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. Molecular Ecology Resources 11(2): 236-244.