A taxonomic revision of the spider genus *Neodietrichia* (Araneae: Linyphiidae) through supplementary COI gene sequence analysis

By

James Bertaux

Honors Project Thesis Submitted in Partial Fulfillment

Of the Requirements for a Baccalaureate Degree “With Distinction”

In the Ron and Laura Strain Honors College of

THE UNIVERSITY OF INDIANAPOLIS

December 2023

Faculty Advisor: Marc Milne, Ph.D.

Executive Director of Honors: James B. Williams, Ph.D.

Abstract

The linyphiid spider, *Neodietrichia hesperia*, was described by Crosby and Bishop in 1933 and the genus has held this lone species (is monotypic) for decades. In 2016, Dr. Milne collected spiders at Blossom Hollow Nature Preserve in southern Indiana in a similar genus, *Lophomma*, that looked similar to *N. hesperia*. Upon examination, he concluded that morphological evidence supported that this *Lophomma* belonged in *Neodietrichia* as a second species in that genus. I used molecular evidence in the form of COI barcoding to determine if these morphological conclusions were backed up by the genetics among closely related species. I extracted, amplified, confirmed via gel electrophoresis, and purified DNA from multiple *N. hesperia* and *N. depressum* specimens from various locations around the United States. Additionally, I helped create a phylogeny of *Neodietrichia* specimens (both from us and from genetic databases) to help confirm their relationship. We confirmed that the collected specimens from Blossom Hollow belonged to the sister species of *N. hesperia*, just as the morphological evidence had suggested. These data helped us revise the genus and publish our results in the animal taxonomic journal, *Zootaxa*, in 2023.

Table of Contents

List of Tables

List of Figures

Statement of Purpose

Marusik et al. (2006) revised *Lophomma* and found that one species, *Lophomma depressum*, did not seem to belong, but they were unsure of its true taxonomic placement. They therefore dubbed it "*Lophomma*" *depressum* and left it at that. In 2016, Dr. Milne collected spiders at Blossom Hollow Nature Preserve in southern Indiana that fit the description of "*Lophomma*" *depressum*. However, these specimens were initially thought to belong to a different genus, *Neodietrichia*, due to their similar morphology. Upon closer inspection and an examination of the literature, Dr. Milne hypothesized that morphological evidence supported that this "*Lophomma*" was actually a *Neodietrichia.* The purpose of this project was to confirm or refute this morphological hypothesis using COI data.

Introduction

Spiders are a very successful group of organisms due to their abundance, diversity, and widespread global distribution. They are found on every continent except for Antarctica (Ubick et al. 2017). Accordingly, spiders play important roles within almost every terrestrial ecosystem. Spiders serve as a crucial intermediate link in terrestrial food webs since they largely prey on small insects and are commonly preyed upon by larger vertebrates. Spiders also serve as excellent environmental bioindicators due to their sensitivity to habit change, rapid response following environmental disturbances, and ease and low cost of sampling (Schwerdt 2018). Spiders also act as efficient biological control agents. Spiders help control mosquitos and other insect populations which act as vectors for various infectious diseases. Therefore, spiders help reduce the spread of these diseases (Ndava et al. 2018). Additionally, medical products derived from spider venom as well as the engineering applications of spider silk are other reasons why spiders are important (Ubick et al. 2017).

Even though spiders play such a critical role in most terrestrial ecosystems, they are vastly understudied. Spider taxonomy is woefully incomplete and new distribution records and species are discovered every year (Dubois 2003). Spiders are of the Araneae taxonomic order which contains numerous families, genera, and over 51,000 species. One of the largest and most taxonomically complex families of spiders in the world is Linyphiidae Blackwall, 1859 and it contains over six hundred genera and almost five thousand species (World Spider Catalog 2023). Linyphiid spiders are also known as sheet web spiders. They are small spiders ranging from 1 to 7 mm in size, and are found globally with great diversity in northern temperate regions and abundance in North America. Relative to their large family size, little is known regarding

Linyphiidae and taxonomic research concerning them is difficult due to their taxonomic and morphological complexity (Ubick et al. 2017).

One genus within Linyphiidae that merits more taxonomic attention is *Neodietrichia* (Crosby and Bishop 1933). The genus *Neoditrichia* was first described by Crosby and Bishop in 1933 when they discovered a new species of spider that did not match the descriptions of any pre-existing genera. They named their newly described species *Dietrichia hesperia,* and the genus remained monotypic prior to this project. The genus and species name was later changed to *Neodietrichia hesperia* because the term *Dietrichia* was already being used to describe a fossil bivalve (Ozdikmen 2008). There has been little work on this genus since its discovery in 1933. One exception is that in the mid-1950's, Hackman (1954) found a specimen similar to, but slightly different from, *N. hesperia*. He suggested that it was a novel subspecies, but he did not formally describe it because the morphological differences were small, and he only possessed one specimen - which he noted was typically insufficient evidence to describe a new species. It would be 52 years before further work was done on this genus.

In 2006, Marusik et al. (2006) conducted a revision of the linyphiid genus, *Lophomma* (Emerton, 1882), a group of spiders similar to *Neodietrichia*. One species, *Lophomma depressum*, stumped the authors. They determined that it was not part of the *Lophomma* genus, but were unsure of its true placement, thereby referring to it as "*Lophomma*" *depressum*. This species was originally described by Emerton (1882) in the late 1800's. In December 2016, Dr. Marc Milne of the University of Indianapolis collected 21 specimens of "*Lophomma*" *depressum* in central Indiana and then another 20 in February and March of 2020 (M. Milne, personal communication). Milne examined the morphology of these specimens and proposed that there was morphological evidence to suggest that there are at least two species within *Neodietrichia*:

N. hesperia and Marusik et al. (2006)'s "*Lophomma*" *depressum* (M. Milne, personal communication). Since that time, Milne has gathered and examined "*Lophomma*" *depressum* from private and public collections collected from multiple Canadian provinces and US states (M. Milne, personal communication).

The study of grouping organisms based on their similarities and differences in relation to each other, such as with *Neodietrichia*, is known as taxonomy. Taxonomic research is very important in the field of biology. Perhaps most noteworthy is its importance in conservation biology. Without taxonomic completeness and accuracy, it is hard to protect organisms from possible threats. Therefore, taxonomists need to properly categorize these organisms. Yet, our current taxonomic understanding of the planet is far from complete and far from accurate. For these reasons, taxonomists and conservation biologists should work closely together (Dubois 2003). Taxonomists can help provide foundational work for making conservation biology efforts more efficient while conservation biologists could provide taxonomists with relevant information and access to specimens (Dubois 2003). It is important that organisms are properly classified by taxonomists; however, it is equally important that scientists in other fields, such as conservation biologists, use this information to further the understanding of these organisms within their respective fields.

Morphology is essential to taxonomic research. Although supplementary methods such as DNA barcoding have proven useful, they cannot serve as a replacement for morphology (Hajibabaei et al. 2007). Morphology is the study of the structure and form of organisms. In regard to spiders, there are certain morphological features that are of increased interest since they allow for better taxonomic comparison. These features include an examination of the intricate complexities of the genitalia, total spider length, carapace length and width, femur length of the

first leg, and measurements of the a/b ratio of the first leg. The a/b ratio represents the position where the trichobothrium (a prominent, chemosensory hair) is located on the metatarsus where “a" is the distance that the trichobothrium is from the tibia and “b” is the length of the metatarsus (Ubick et al. 2017). Morphologically, “*Lophomma*” *depressum* seems to belong in *Neodietrichia*, sister to *N. hesperia*, due to both species sharing: 1) a similarly raised carapace (dorsal portion of cephalothorax) in males, 2) a pit behind the posterior lateral eyes with an excavated portion in males, 3) a similar size and coloration, 4) male genitalia that both possesses similar features such as the shape of certain sclerites in the pedipalp, and a long, pointed, and sometimes toothed tibial apophysis (a protrusion off of the tibia of the leg) (Milne et al. 2023). However, there is variation among specimens in some of these characteristics and in order to confirm that the multiple morphologically-varying specimens we possessed all fell under the same *L. depressum* species, we sought the use of molecular tools to aid the taxonomic placement of these specimens.

We could provide additional evidence towards the hypothesis that there are two species, and not just variations of one species by sequencing the COI genes of available specimens via DNA barcoding. DNA barcoding has become an efficient molecular technique to determine species identity and to aid in species delimitation (Rach et al. 2017). The COI gene has been established as a particularly useful mitochondrial gene for its use in DNA barcoding in metazoans (a broad range of multicellular organisms with cells differentiated into tissues and organs) for several reasons (Rach et al. 2017). The first reason is the conserved size and structure of the COI gene (Rach et al. 2017). It is ideal when selecting a gene for barcoding analysis to choose one that will be present in as many organisms as possible. However, a gene that has too much variation (especially concerning overall size and structure) from organism to organism would not serve as an optimal reference point for genetic comparisons. In other words, the ideal

barcoding gene would be one that appears in as many organisms as possible without risking excessive variation within the gene itself. A second reason that COI is an effective barcoding gene is due to the heterogeneous substitution patterns that are present in its variety of functional domains (Rach et al. 2017). While the overall size and structure of the gene needs to be conserved, the gene must also provide sufficient genetic differences in order to barcode species. The less genetic differences there are in a gene the more likely that the two species could end up with very similar barcodes. The fact that COI contains various functional domains means that there are more opportunities for different sequence patterns to occur within those domains. The ample number of heterogeneous substitution patterns that occur within these functional domains will allow for greater genetic differences to arise. The more that these genetic differences are maximized, without compromising the conserved nature of the gene, the better the gene will be for genetic barcoding; this is something that COI is thought to do well. Another reason why COI is an effective barcoding gene is because COI is a mitochondrial gene. The advantages of mitochondrial genes are that: they are haploid, they have multiple copies per cell, and that they evolve faster than nuclear genes since they lack a proofreading mechanism (Rach et al. 2017). The haploid nature and high copy number of mitochondrial genes makes them easy to amplify (Rach et al. 2017). Additionally, the haploid nature of COI means that there will only be one barcode for the COI gene instead of two; a single barcode allows for simplicity. The lack of a proofreading mechanism in mitochondria leads to an increased mutation chance. An increased mutation chance results in more genetic differences and therefore points of comparison. Despite the benefits of using COI, there is some disagreement among the scientific community when it comes to the use of mitochondrial genes in DNA barcoding; however, mitochondrial genes still seem to be a great option (Rach et al. 2017).

These COI sequence data (our molecular data) may then be combined with existing morphological data to make the most informed taxonomic decision possible. The combination of using multiple taxonomic methodologies to arrive at the best taxonomic decision possible is known as integrative taxonomy. Integrative taxonomy is increasingly being considered an essential tool for species delimitation (accurate taxonomic placement of species), especially since species are constantly changing (Montes De Oca et al. 2015). The act of using morphology alone in the process of delimiting species may not always be sufficient, especially when similar features across a species make morphological delimitation difficult (Montes De Oca et al. 2015). Therefore, molecular techniques are often used in conjunction with morphology. This is often seen in the taxonomic study of spiders (Montes De Oca et al. 2015). In addition, evolutionary biologists generally agree that species result from evolving lineages; however, they disagree on which of these lineages result in separate species (Padial et al. 2010). Since disagreement mostly surrounds what is necessary for a lineage to constitute a new species, it makes sense to use integrative taxonomy. Integrative taxonomy comprises multiple research techniques, so there will be fewer opinions regarding the delimitation of a species when using integrative taxonomy as opposed to a single research method such as morphology.

Preliminary Experiments

For this project, we started with a template protocol that was developed by Dr. Milne and previous UIndy students. However, we needed to modify this protocol for COI primers since this was our gene of interest. The following preliminary experiments were conducted to modify the template protocol into the final protocol. The final protocol is described in Methods.

First, we needed to find out what COI primer pair(s) would be successful in amplifying COI for *Neodietrichia*. Primers work in pairs and bind to a gene to prepare the formation of complementary DNA strands. We had several options at our disposal, but we found through experimentation that primer pair L-COI-1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and H-CO1-2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') worked best – referred to as 1490/2198 for simplicity. This was determined by testing several primer pairs with a freshly caught, non-*Neodietrichia* spider (to conserve our *Neodietrichia* specimens until we were confident in our protocol). We also had some success with primer pair 1490/2776. However, we decided to move forward with 1490/2198 (despite its shorter sequence of ~710 bp compared to ~1250 bp of 1490/2776) since it was more successful in amplifying preserved specimens. We obtained thermocycler recipes for gene amplification from colleagues also attempting gene amplification on linyphiid spiders.

Second, we found that we had significant primer dimers (a band in a gel that occurs when primers bind together rather than the template DNA strand) when running our gels with 1490/2198. This led us to believe that our primer concentrations were too high. A high primer concentration could lead to unwanted primer interactions (such as primer dimers) instead of the desired interactions between the primers and DNA in the amplification process. Therefore, we

performed a serial dilution experiment to determine the primer concentration at which we could still achieve clear bands while minimizing the presence of primer dimers. Concentrations of 10 (original protocol), 5, 1, 0.5, 0.1, 0.05, and 0.01μM were tested on non-target specimens in the family Lycosidae. A primer concentration of 5μM per primer was determined ideal since the gel had clear bands and minimal primer dimers.

Even after these initial preliminary experiments, our DNA extractions and amplifications continued to have a high rate of failure. We therefore sought to improve our extraction technique. Initially, the first step of the extraction protocol consisted of removing a sample from its ethanol filled vial, soaking it in water for 30 minutes with occasional mixing, and letting it dry. However, upon requesting advice from an expert in molecular techniques on spider tissue, Dr. Shahan Derkarabetian (San Diego Natural History Museum), we began removing the specimen from ethanol and fully submerging it in water and refrigerating it overnight. This seemed to improve our success rate for DNA extractions. We speculated this improvement to be from allowing adequate time for the ethanol to be removed from the specimen and/or allowing the tissue of the specimen to rehydrate. Ethanol is known to greatly interfere with the DNA extraction and amplification process. Sometimes, however, an unsuccessful extraction was unavoidable if a sample contained contaminated or degraded DNA.

Additionally, we modified our purification process to increase efficiency and success rate. Initially, our purification process was as follows:

> We started with 20μl PCR product and added 100μl of buffer PB and vortexed. Then pH was checked. If pH was not less than or equal to 7.5 then we added increments of 10μl sodium acetate until desired pH was achieved and vortexed the solution. The solution was

then added to a spin column with a collection tube and centrifuged for 1 min at 13,000 rpm. The spin column was then placed on a new collection tube, 750µL buffer PE was added, and column was centrifuged for 1 min at 13,000rpm and again for 2 min at 13,000rpm. The spin column was then placed on a 1.7ml vial. 30µl of buffer EB was added to the spin column and allowed to sit for 5 min with lid open. The spin column was then centrifuged for 1 min at 13,000rpm and collected purified product was stored in the freezer until sent off for sequencing. 10µl of purified product and 5µl of 5µM forward primer 1490 were combined before sending off to sequence.

Our sequence results using this method, or lack thereof, revealed that our DNA concentrations were greater than ideal. This was surprising because we initially thought that failed PCR attempts were due to low DNA concentrations; however, other factors, such as DNA quality, could have caused failed attempts. To increase the purity of our DNA products, we began using a purifying solution called ExoSap-It (Applied Biosystems, ThermoFisher Scientific, Cat. No. 78205.10.ML) instead of the column purification system we were using previously. We switched to ExoSap-It because the purification process was much easier, and it had a higher success rate for obtaining sequences. Even when we diluted our purified DNA product with the original method (1µL purified product and 9µL ultra pure H_2O), the ExoSap-It procedure still out-performed our original purification protocol due to its ease of use and higher sequence success rate.

We also tried using devices to check DNA concentration and/or quality. This would allow us to save time by not using solutions with no or poor DNA present. The first device we tried was a Qubit Fluorometer (Qubit 3.0 Fluorometer, Life Technologies, Cat. No. Q33216). A Qubit Fluorometer measures the fluorescence of a fluorescently labeled sample. If the target substance

is present, the Qubit Fluorometer measures the changes of light emitted from the fluorescently labeled sample to determine DNA concentration. Unfortunately, we had issues with calibrating and operating the Qubit Fluorometer at first due to user error. Once we were successful at operating it, the Qubit Fluorometer provided insight into DNA concentration. However, we found that samples with lower concentrations could provide successful results when samples with higher concentrations would not. This led us to believe that DNA quality was more important than DNA quantity. To address this, we tried using a NanoDrop to provide insight into DNA purity. A NanoDrop is a machine that measures a target substance with ultraviolet and visible spectrum electromagnetic radiation. The target substance, DNA, should absorb at a certain wavelength ($\approx 260nm$) while common contaminants would absorb at a different wavelength ($\approx 280nm$). The DNA absorbance should provide insight into DNA concentration, and the ratio between DNA and contaminant absorbances should provide insight into DNA purity. However, at the time of the experiment, we failed to get the NanoDrop to work. Due to these reasons, we didn't receive much utility from trying to implement a device to check for DNA concentration and/or quality during this project.

Methods

This project consisted of three phases: identification and morphological analysis, gene sequencing, and taxonomic analysis. The first phase consisted of spider identification and morphological analysis. Collected specimens were identified with the help of Dr. Milne and other student researchers according to their appropriate taxa by using dissecting microscopes and research articles that describe *Neodietrichia* such as Crosby and Bishop (1933), Hackman (1954), and Marusik et al. (2006). Dr. Milne guided and assisted me with identification when needed. In the identification process, morphological features were examined and logged. Quantitative data included body length, carapace length and width, femur length of the first leg, and the (a/b) ratio. In addition, qualitative data, such as visual inspection, were used in conjunction with quantitative data to differentiate *Neodietrichia* specimens from each other and other genera (Figure 1). Once specimens were assigned to the *Neodietrichia* genus, they were compared within the genus by their genitalia alone. Overall, morphological data was very important since it provided the initial framework for molecular data comparison.

The second phase, gene sequencing, comprised the core of this honors project. In this phase we attempted to sequence COI genes from various *Neodietrichia* samples obtained from museums (CNC = Canadian National Collection of Insects, Arachnids, and Nematodes, MCZ = Museum of Comparative Zoology at Harvard, and OSU = Ohio State University Biodiversity Collection), private collections (DB = the private collection of Don Buckle, GB = the private collection of Gergin Blagoev, MD = the private collection of Dr. Michael Draney, and PP = the private collection of Dr. Pierre Paquin), and those collected by Dr. Marc Milne (MM). Specimens from museums and private collections were obtained by loan requests made to those institutions and individuals by Dr. Milne. After morphological examination (see above), requests

to destroy specimens had to be made to museums and private collectors since the COI extraction process would damage or destroy the spiders. Once permission was obtained, this step began.

The various samples were referred to by their sampling location. Samples were from the United States and Canada. Specifically, in the United States specimens were examined from Indiana (MM), Maine (MCZ), North Carolina (CNC), Ohio (OSU), and South Carolina (MD) and from Canada specimens were examined from Alberta (DB and GB), British Columbia (CNC), Manitoba (CNC), New Brunswick (CNC), Nova Scotia (CNC), Ontario (GB), Quebec (PP), and Saskatchewan (DB). The second phase consisted of five subphases: DNA extraction, polymerase chain reaction (PCR), gel electrophoresis, purification, and sequencing.

For extraction, we used spider legs - or often the whole spider - to extract raw spider DNA. Ideally, we would only use legs to preserve the rest of the specimen - especially for spiders on loan from museums or private collections. However, the whole spider was often used, with permission of the owner, to provide sufficient biological material for successful DNA extraction. Spider DNA was extracted from the specimen using the DNeasy Blood & Tissue Kit with columns that contained various buffers and enzymes (Qiagen, Cat. No. 69504). The first step of the initial extraction protocol consisted of removing a sample from its ethanol-filled vial and fully submerging it in water and refrigerating it overnight. Once the spider was removed from the overnight soak, it was then allowed to dry and was put into a vial with buffer ATL and proteinase K. Buffer ATL was used to break open cellular components to free the DNA stored within cells. Proteinase K also assisted in the liberation of stored DNA by breaking down proteins. It also helped to remove potential contamination. The vial was then soaked in a hot water bath at 56°C overnight with occasional vortexing (machine aided mixing) every few hours or when possible. Next, the contents of the vial were transferred into a spin column (a vial with a filter that can separate a solution in a centrifuge). Buffer AL and ethanol were added and the spin column was centrifuged. The spin column then went through a series of buffer washes with buffer AW1 and AW2. This first series of washes served to remove non-DNA substance from the spin column while leaving the DNA molecules embedded in the filter. The spin column was then washed twice with 75µl of buffer AE (we changed this to 50µl because a smaller final volume meant a higher DNA concentration). Buffer AE dissolved the DNA and allowed it to pass through the spin column filter. Then DNA was then left in the buffer AE and stored in a freezer.

PCR utilizes DNA primers to amplify a single desired gene. For this project we used the primer pair consisting of forward primer LCO1490 and reverse primer HCOI2198 as determined

in the preliminary experiments. These primers were designed to amplify approximately half of the COI gene. Primers were used to initiate cycles of PCR; however, QIAGEN Master Mix Kit (Qiagen, AllTaq Master Mix Kit (500), Cat. No. 203144) provided the necessary reagents for building the complementary DNA strands. The Master Mix included: DNA taq polymerase (an enzyme responsible for adding nucleotide bases to primers to build complementary DNA strands), QIAGEN PCR buffer (promoted ideal PCR conditions), $MgCl_2$ (enhanced activity of taq polymerase), and dNTPs (the nucleotide building blocks necessary to build DNA strands). Our PCR solution was prepared by combining 25µl MasterMix, 15µl ultra pure H_20, 2.5µl forward primer at 5µM concentration, 2.5µl reverse primer at 5µM concentration, and 5µl DNA. Our PCR solution was then put into a thermocycler (machine used for PCR) to facilitate amplification for the COI gene. Each cycle in the thermocycler results in a set of new complementary DNA strands. After many cycles, there are many copies of the DNA strand for the desired gene. There are 3 phases in one PCR cycle - denaturation, annealing, and extension. Denaturation is when the DNA molecule separates into two strands; annealing is when the primers bind to the separated strands; and extension is when nitrogenous bases are added to form a new complementary DNA strand. Each phase of the thermocycle has corresponding temperature and time characteristics, and these parameters are unique to the primer pair being used. Therefore, we utilized a specific thermocycler protocol (2 minutes at 94°C; 34 cycles of 20 seconds at 94°C, 35 seconds at 49°C, and 30 seconds at 65°C; and 3 minutes at 72°C) which we coded into the PCR machine for our primer pair.

Gel electrophoresis determined whether PCR was successful in copying our COI gene. An agarose-based gel dyed with ethidium bromide was used to run our PCR product. We mixed 0.5g agarose with 50ml 0.5x TBE buffer and heated until a very light boil. We cooled the

mixture, added 1µl ethidium bromide, and poured it into a mold. Once cooled, we were left with an electrically active gel with well slots at one end. Our gel was then placed in an electrophoresis chamber and submerged in 0.5x TBE buffer. Our dyed PCR product is then placed in the well slots (one specimen per slot) along with a DNA ladder and/or any positive or negative controls. DNA is negatively charged, so the gel is positioned with the well slots towards the cathode in the chamber. Therefore, when we added electrical current to the chamber, the negatively charged DNA would migrate towards the positively charged anode. The rate at which DNA moves through the gel is inversely proportional to its size, so smaller pieces of DNA travel farther than larger ones. The DNA ladder contained DNA fragments of various sizes which allowed us to estimate the size of our bands (dyed DNA) to confirm they were the section of the COI gene we were amplifying. Sometimes, clear bands would not form or they were too far advanced along the gel (too small of fragments) to be COI. This was usually the result of primers binding to each other (primer dimers) instead of binding to DNA. Once it appeared that our bands were amplified COI, we moved on to purify our PCR product.

Purification prepared our PCR product to be sent off to a sequencing facility. We utilized a purification reagent called ExoSap-It (see preliminary experiments). We combined 5µl of PCR product and 2µl of ExoSap-It and incubated them in the thermocycler for 30 min following Exo Sap It protocol guidelines. The solution was then stored in the freezer until we sent it off for sequencing to Genewiz, a DNA sequencing facility. We priority-mailed our samples since they were temperature sensitive. Genewiz processed our samples and determined their sequences (if possible). The results were then relayed back to us electronically.

Once we completed the second phase, we conducted the third and final phase - taxonomic analysis. Taxonomic analysis took into account both morphological and gene sequencing data.

By looking at those two sets of data we were able to make insightful decisions regarding taxonomic placement of the specimens analyzed. Additionally, we were able to provide more insight into how many *Neodietrichia* species there may be. Analysis of the sequenced COI genes was done primarily through a multistep process. First, sequences were imported into MEGA X (Molecular Evolutionary Genetics Analysis version 10; a nucleotide editing software program). Next, we checked for successful sequences and eliminated failed ones. Failed sequences did not provide a sequence, lacked a complete sequence, or were contaminated with DNA (often human or from other spiders). The successful sequences were then trimmed of erroneous or failed nucleotide bases. Next, the sequences were aligned in MEGA X. We subsequently performed Neighbor Joining (NJ) analysis. NJ is a mathematical technique that groups taxa based on their similarities via computer software (such as MEGA X). This allowed us to create phylogenetic trees for our COI sequence data. The NJ analyses were manually rooted with a designated outgroup (organism unrelated to and outside our taxa - *Neodietrichia* - of interest). Our outgroup was *Mermessus maculatus* (Linyphiidae), a species in a closely-related genus. The more closely related that the genetic samples are, the more closely related they will appear in the phylogenetic trees. Our constructed phylogenetic tree for the COI gene of *Neodietrichia* supplemented our morphological data to provide evidence for species delimitation within *Neodietrichia*.

Results

We were successful in obtaining the COI sequences for three spiders: Indiana (Johnson Co., Blossom Hollow Nature Preserve), Ohio (Vinton Co., Vinton Furnace State Forest), and New Brunswick (Kent Co., Kouchibouguac National Park) (Table 1). The remaining specimens

Table 1: Successful COI DNA sequences of *Neodietrichia.*

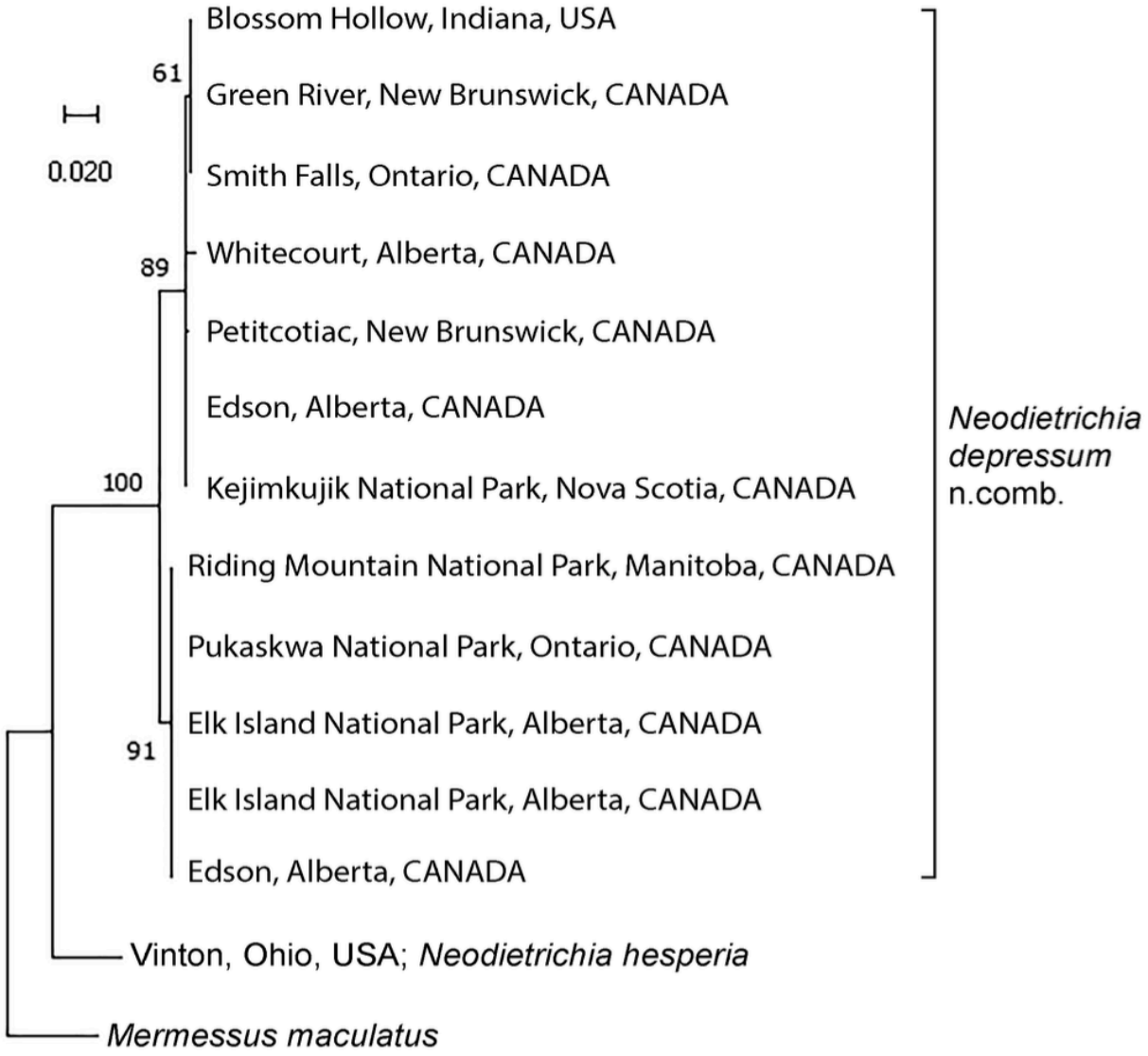
N. depressum (IN: Blossom Hollow Nature Preserve)	NNNNNACTCNTGATNNNNNGGNNNNNNNNNGNNNNNNGNAGTTTAACTGGCCCCC TTGTTCGAAACNAACCGCGNTTTNCTGGGAAGCTTCTAGGGNANNATGNTTCGACT AGGGTAACGGCACNNGCNNNGNNTCCGCAATANTCTTCTTCATNATAATACCCNTCA NANTCCGAGGNNNNGGCNACTGNCTACTTCCTCNATGAAATTNNGCGCCCCTGATG NNGTTTCCCCTCATAAACAAAATAANCTACTGANACTTACCTCCCTTTCTCCTACTCC TGNACGTATCTGCTAAANNGAAGGNAGGAGCTTGAACAGGGTGAGAACNGTACCC TCCCCTANTGGCAACTGCTCCGACCTGGAGCCNNCNTNGANCNAACCATCATCTCC TTACACCTAGTANGTGGCGCCTCTATCTTAGGGGGCATCATTTACTTCATTTNNCTTAT TNNNNNAAANCCCCTGNNGGAATAACTACCAAACAACCNNCTTCGTCTGATCCGTC TCACTCACTTCANTCCTACTTCTTCTATCTCTCCCTTCCNTANAACGCTGGNGNNNC AATACTACACTCAGACCGGANCCTNAACATCACCTTCNTNGACTCNNCCGGANGAN GAGANCCNATTCNATAANNANNCCNATTCTTACTCTGTGGTCACNGTGANCTTGAN NNNAANNNNNACTGTNTATANCTCNNTTCNGGAANNGCNATCNGGACTNNNTNAN TGNTTTNCNNTACNNNNCTNNNCTNNNNNNNNNN
N. hesperia (OH: Vinton Furnace State Forest)	NNNNNNTTTNNTTTTTGGGGCTTGAGCTGCTATAGTAGGGACAGCAATAAGAGTATT AATTCGAGTGGAGTTAGGACAAATTGGTAGATTGTTGGGGGACGATCAGTTATATAA TGTTATTGTTACTGCTCATGCTTTTGTAATAATTTTTTTTATAGTTATACCTATTTTAATT GGGGGATTTGGAAATTGGTTAGTCCCTTTAATGTTAGGAGCACCTGATATAGCTTTTC CTCGTATAAATAATTTAAGATTTTGATTATTACCCCCCTCATTGTTTTTATTGTTTATTT CAAGAATAGATGAAATAGGGGTGGGAGCAGGCTGGACTGTGTACCCTCCTCTTGCA TCTTTGGAGGGGCATTCTGGAAGTGCTGTAGATTTTGCTATTTTCTCTTTACATTTGG CTGGGGCGTCATCTATTATGGGGGCTATTAATTTTATTTCTACTATTTTAAATATACGCG GATATGGGATGTCTATAGAGAGGGTTCCATTATTTGTTTGGTCTGTTTTGATCACGGC TGTTTTATTACTTTTATCTTTACCTGTTTTAGCAGGTGCGATTACAATACTTTTAACTG ATCGAAATTTTAATACATCATTTTTTGACCCTTCTGGGGGAGGGGATCCAGTGTTGTT TCAACATCTGTTTTGATTTTTTGGTCANNNNNAAAANTTNAANNNNNCNNNTCNNN NNNNNNNNNTNNNNNNNNTNNNTTTANNNNNNNNNNNNNNNNN
N. depressum (NB: Green River)	NNNNNTTTATTTTTGGGGCTTGGGCTGCATAGTGGGGACAGCAATGAGAGTGTTAAT TCGAATTGAGTTAGGACAAACTGGTAGTTTGTTGGGTGATGATCAGTTGTATAATGT CATTGTTACTGCCCATGCTTTTGTGATAATTTTTTTTATAGTAATACCTATTTTAATTGG GGGATTTGGAAATTGGTTAGTGCCTTTGATATTAGGGGCTCCTGATATGGCTTTTCCT CGGATAAATAATTTAAGATTTTGATTATTACCTCCTTCTTTATTTTTATTATTTATTTCAA GAATAGATGAGATAGGGGTTGGGGCAGGATGAACTGTTTATCCCCCTCTTGCTTCCT TAGAAGGACACTCTGGAAATTCGGTGGATTTTGCTATTTTTTCCTTGCACTTAGCTGG

	GGCGTCTTCTATTATGGGAGCTATTAATTTTATTTCTACTATTTTAAATATACGAGGGTA TGGGATAACAATAGAGAAAGTGCCATTGTTTGTATGATCTGTTTTGATTACAGCTGTT TTATTACTTTTGTCTCTTCCTGTATTGGCCGGGGCGATTACAATACTCTTAACAGATCG AAATTTTAATACGTCATTTTTTGACCCGTCTGGGGGTGGGGATCCAGTATTGTTTCAG CATTTGTTTTGATTTTTTGGTCACNTNNAAGTTNAANNNNNTCNGGCCTGNNTCNCN CAGNNNCCNNNNNNNGNTCGNTCGNCCCTCGNTNNNNCNANNNNNCNNNNNNCN NANNNATCCCNNNNNNNN

either failed to provide a sequence or were contaminated with foreign DNA. By combining our three successful sequences with other known COI sequences from GenBank we were able to form a phylogenetic tree of *Neodietrichia* specimens through neighbor-joining analysis (Figure 2).

The COI p-distances between specimens within species were 0.000-0.010, and the COI p distances between *N. hesperia* and *N. depressum* were 0.040-0.044. *Neodietrichia* and the outgroup *M. maculatus* were separated by p-distances of 0.044-0.057. The bootstrap value for all *N. depressum* specimens was 100% which means that every time the computer recreated the phylogenetic tree, all of these specimens were grouped within *N. depressum*. This is very strong evidence for the *N. depressum* branch to exist. A bootstrap value for the *N. hesperia* branch could not be determined because there is only one specimen for the computer program to sample from.

Conclusion



The p-distances obtained from the neighbor-joining phylogenetic tree support the morphological evidence that *N. hesperia* and *N. depressum* are indeed two different species within the *Neodietrichia* genus. The relatively low p-distance of 0.000-0.010 between specimens within species means that the specimens are in fact not that different and do belong in the same species. The relatively large p-distance of 0.040-0.044 supports that *N. hesperia* and *N. depressum* are different enough to constitute different species, especially in conjunction with the morphological differences between the two species. These morphological differences include similar positioning of sulci on the carapace of males, unique large, scimitar-shaped terminal

apophysis in males, longer tibial apophysis in males, epigyna with sclerotized hood that is near the midline and is most anterior to other epigynum parts in females, median plate of epigyna is wider posteriorly and narrows anteriorly in females. The combination of these morphological characteristics (and those discussed in the introduction) along with our COI sequence data is enough evidence to conclude that *Neodietrichia* contains two species: *N. hesperia* and *N. depressum*.

The investigation of *Neodietrichia* via molecular means - this project - was conducted over the course of a few years with great attention to detail. Most of the effort directed into this project, such as the many preliminary experiments and failed trials, did not directly translate into our final results. Instead, most of the effort spent on this project was to make sure that we had our procedures optimized to ensure a successful, high-quality result - which was thankfully the outcome of this project. It was years of preparation, trial and error, testing on other spider genera or spare *Neodietrichia* specimens, and procedure refinement before we obtained successful COI sequences for *Neodietrichia* specimens. The meticulous execution of this project over the course of years with the help of Dr. Milne and other UIndy students, along with our results, is testament to the merit of this honors project.

Reflection

We were successful in obtaining three COI sequences for *Neodietrichia*. While more sequences would have been nice, these sequences were adequate for analyzing the taxonomic make up of *Neodietrichia* since one of the three was *N. hesperia*. While we had several *N. depressum* sequences (our two sequences and those from GenBank) we were only able to sequence one *N. hesperia* specimen. Since *N. hesperia* has not been sequenced prior to this project, obtaining this one *N. hesperia* sequence was critical because it allowed for the taxonomic comparison of *N. hesperia* and *N. depressum*. The combination of our sequences, and those from GenBank, allowed us to complete a phylogenetic analysis and an integrative analysis (morphological and phylogenetic) for *Neodietrichia*. The successful sequencing and analysis of *Neodietrichia* allowed us to meet our project goals of determining the taxonomic make up of this genus.

There were many difficulties surrounding this project which can be grouped into three main categories: 1) obtaining *Neodietrichia* samples, 2) determining a viable COI PCR procedure, and 3) successfully extracting and amplifying *Neodietrichia* from a very limited specimen supply. The first category, obtaining *Neodietrichia* samples, was difficult for a few reasons. All specimens were obtained from museums or private collections with their permission or were obtained via collection by Dr. Milne. A *Neodietrichia* community collection project initiated by Dr. Milne failed to produce any *Neodietrichia* specimens. *Neodietrichia* spiders are difficult to collect and this is thought to be due to their tree-dwelling nature. Difficulty in obtaining *Neodietrichia* spiders meant that we had a limited supply of *Neodietrichia* to work with. Therefore, we had little to no room for error because we often only had one copy of a *Neodietrichia* specimen to work with. This usually translated into only having one or a few

extraction attempts - especially if a specimen was on loan from a museum or a private collection and the number of legs we could remove was limited by the specimen loaner.

The second difficulty, determining a viable COI procedure for preserved *Neodietrichia* specimens, was challenging due to its trial-and-error nature. We were unable to use a device such as a Qubit Fluorometer or Nanodrop to check purities or concentrations to save time. Instead, if a procedure did not work, we would hypothesize what the problem was and attempt to solve it or we would just redo the procedure altogether. Since extraction to sequencing is a multistep process we could not pinpoint where a mistake occurred. We only knew a mistake occurred when a specimen failed to give bands. This was quite frustrating because anything could have gone wrong prior to running a gel such as: PCR, extraction, or starting with contaminated or degraded DNA. This forced us to be very problem-solving oriented. If something failed, we would hypothesize the most likely reason for failure and attempt to correct it. If the correction was unsuccessful, we would systematically move on to the next most likely reason for failure and attempt to correct that and so on until results were achieved. Although this process is highlighted by larger problems in the preliminary experiments section, these iterations of failure and correction were happening constantly in the lab for even the most minute details.

After obtaining a viable protocol, the last main difficulty was successfully sequencing *Neodietrichia* from the limited specimens we had. We only had one copy of most of our specimens. We also could only perform very few extraction attempts because the small size of *Neodietrichia* required multiple legs for extraction. Additionally, specimens on loan from museums or private collections often wanted minimal material removed from the specimens to preserve the specimen for future reference. For these reasons, we had to be very meticulous, especially during the extraction process when working with the *Neodietrichia* of interest since

we had one or few attempts for success. Luckily, we were able to obtain a *N. hesperia* and two *N. depressum* sequences which were enough to conduct a taxonomic analysis of *Neodietrichia*.

Having the opportunity to work with and learn from great people was why this project was important to me. I was able to learn from and work with Dr. Milne and other UIndy students who have done research with Dr. Milne in the past. I remember starting research in the molecular lab with Dr. Milne after briefly doing some field research collecting spiders. I was unfamiliar with the molecular protocols, and I was eager to learn. I was constantly asking the other research students questions when working with them to improve my lab technique and knowledge. In time, I went from the curious, new recruit to the confident, sole-remaining lab researcher after the other research students had graduated. It is satisfying to look back on this journey of improving my confidence and skill in the lab – although the curiosity remains. Having the opportunity to work with people of great skill and personality was, and still is, very important to me. That was one of the main reasons I decided to undertake this research product. While I do enjoy molecular research, I wasn't as interested in spiders at the start of this project as Dr. Milne. However, through hard work and dedication, he tricked me into becoming a true spider enthusiast.

The undertaking of this project honed some of my professional skills including: problem solving, communication, cooperation, patience, and goal setting. The trial and error nature of developing a working protocol for COI for *Neodietrichia* constantly required problem solving. Every time a procedure failed (which was often) we had to hypothesize what went wrong, adjust the procedure, retest the procedure, and repeat until reaching the desired result. Without problem solving, the execution of this project was simply not possible. Communication and cooperation were also integral to this project since it was a group effort. I was always asking questions,

scheduling meetings with other research students in the lab, and with Dr. Milne, and bouncing ideas off of the other research students and Dr. Milne. Even when the other research students graduated, I was constantly discussing and meeting with Dr. Milne to improve the project's procedure to maximize our results. Communication and cooperation with the other research students in the beginning of the project was critical because I knew that I would soon be the only molecular research student on this project. By communicating and cooperating with the other research students in the beginning, I developed the confidence and capability necessary to proceed without them once they were gone. Even though this project was a success in the end, the day-to-day operations consisted mostly of failure. In other words, we learned how to do something right by doing it wrong many times first. This greatly tested my patience. I rarely get upset or impatient easily, but it is quite frustrating to see failure over and over again without results. It is also frustrating to have to redo the same processes many times, with slight modifications, knowing that it will most likely result in another failed attempt. On the bright side, this made success so much more satisfying. Seeing COI bands for *Neodietrichia* for the first time after several months of failure was deeply satisfying, and it rewarded my patience. The multiyear nature of this project also presented a unique goal setting challenge. I needed to take a long term, ambitious goal and break it down into short term, achievable goals. Without properly outlining our project into smaller goals at the start, we would have lacked direction and would have failed to achieve the same results – if any.

Overall, I thoroughly enjoyed working on this project. I improved my molecular research techniques, learned about spiders, honed my professional skills, and had fun working with wonderful people. Due to my efforts, I also became a coauthor on the peer-reviewed paper, "A taxonomic revision of *Neodietrichia* (Araneae: Linyphiidae), a rarely encountered but

widespread spider taxon.” which is an impressive achievement especially for an undergraduate student.

Acknowledgements

I thank fellow UIndy students Spencer Burton, Alex Flickinger, Carly Mabie, and Kyle Humerckhouse for creating the initial template protocols and training me in proper lab techniques. I also thank fellow UIndy students Caylie Wimmersberger, Casey Venable, and Lauren Cochran for their contributions towards Neodietrichia research and support towards this project. I would also like to thank Gergin Blagoev for making our phylogenetic analyses more complete by providing COI sequences and Shahan Derkarabetian for providing insight into optimal extraction techniques. I am also grateful to Spencer Burton, Leah Milne, Ted Milne, Esther Mullins, Caitlin Stahl, Casey Venable, and the many participants in the Search for *Neodietrichia* community science project in efforts to collect *Neodietrichia* specimens. I would like to thank the Indiana Department of Natural resources (for Blossom Hollow and Glacier's End Nature Preserves and Ginn Woods Nature Preserve), the US Forest Service, Keith Kanoti and the University of Maine's School of Forest resources (for Penobscot Experimental Forest), the Parks, Outreach, & Service Delivery Division of the Department of Lands & Forestry of Nova Scotia (for Coldbrook Provincial Park), Cliff Chapman, Joanna Woodruff, and the Central Indiana Land Trust (for Blossom Hollow and Glacier's End Nature Preserves), John taylor and Ball State University, and Caitlin Stahl and Southeastway Park for allowing us to collect on their properties. I thank the University of Indianapolis for providing equipment, resources, and faculty to guide me. I thank Dr. James Williams and the Honors College of the University of Indianapolis for giving me the opportunity to complete such an amazing life-changing project. Finally, I thank my advisor Dr. Marc Milne, the best advisor in the entire world, for making this project possible and an absolute blast!

References

Blackwall J. 1859. Descriptions of newly discovered spiders captured by James Yate Johnson Esq., in the island of Madeira. Annals and Magazine of Natural History (3) 4(22):255-267. doi:10.1080/00222935908697122

Crosby CR, Bishop SC. 1933. American spiders: Erigoninae, males with cephalic pits. Ann Entomol Soc Am. 26:105-172.

Dubois A. 2003. The relationships between taxonomy and conservation biology in the century of extinctions. C R Biol. 326(1):9–21. doi:10.1016/s1631-0691(03)00022-2

Emerton JH. 1882. New England spiders of the family Theridiidae. Trans Conn Acad Arts Sci. 6:1–86. https://doi.org/10.5962/bhl.part.7410

Hackman W. 1954. The spiders of Newfoundland. Acta Zool Fenn. 79:1-99.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA. 2007. DNA barcoding: How it complements taxonomy, Molecular Phylogenetics and population genetics. Trends Genet. 23(4):167–172. doi:10.1016/j.tig.2007.02.001

Marusik Y, Gnelitsa V, Koponen S. 2006. A survey of Holarctic Linyphiidae (Aranei). 4. A review of the erigonine genus *Lophomma* Menge, 1868. Arthropoda Sel. 15(2):153-171.

Milne MA, O'Neil C, Bertaux J. 2023. A revision of *Neodietrichia* (Araneae: Linyphiidae), a rarely encountered but widespread taxon. Zootaxa 5296(1): 31-44.

Montes de Oca L, D'Elía G, Pérez-Miles F. 2015. An integrative approach for species delimitation in the spider genus *Grammostola* (Theraphosidae, Mygalomorphae). Zool Scr. 45(3):322–333. doi:10.1111/zsc.12152

Ndava J, Llera SD, Manyanga P. 2018. The future of mosquito control: The role of spiders as biological control agents: A review. Int J Mosq Res. 5(1), 6-11.

Özdikmen H. 2008. *Neoditetrichia* nom. nov., a replacement name for the genus *Dietrchia* Crosby & Bishop, 1933 (Arenea: Linyphiidae) non Reck, 1921. Mun Ent and Zool. 3(1): 537-538.

Padial JM, Miralles A, De la Riva I, Vences L. 2010. The integrative future of taxonomy. Front Zool. 7(16):1-14. doi:10.1186/1742-9994-7-16

Rach J, Bergmann T, Paknia O, DeSalle R, Schierwater B, Hadrys H. 2017. The marker choice: Unexpected resolving power of an unexplored COI region for layered DNA barcoding approaches. PLoS One. 12(4):e0174842. doi:10.1371/journal.pone.0174842

Schwerdt L, Elena de Villalobos A, Miles FP. (2018). Spiders as potential bioindicators of Mountain Grasslands Health: The Argentine tarantula *Grammostola* Vachoni (Araneae,

Theraphosidae). Wildlife Res. 45(1):64. doi:10.1071/wr17071

Ubick D, Paquin P, Cushing PE, Roth V, editors. 2017. Spiders of North America: an identification manual, 2nd Edition. American Arachnological Society. New Hampshire, USA.

World Spider Catalog (2023) World Spider Catalog. Version 23.5. Natural History Museum Bern. Available from: http://wsc.nmbe.ch (accessed 6 November 2023) https://doi.org/10.24436/2