University of Northern Iowa [UNI ScholarWorks](https://scholarworks.uni.edu/)

[Honors Program Theses](https://scholarworks.uni.edu/hpt) **[Student Work](https://scholarworks.uni.edu/sw_uhp) Student Work**

2014

A preliminary systematic analysis of the species of genus Physaria (Brassicaceae) of Wyoming

Kayla Rethwisch University of Northern Iowa

[Let us know how access to this document benefits you](https://scholarworks.uni.edu/feedback_form.html)

Copyright © 2014 Kayla Rethwisch

Follow this and additional works at: [https://scholarworks.uni.edu/hpt](https://scholarworks.uni.edu/hpt?utm_source=scholarworks.uni.edu%2Fhpt%2F105&utm_medium=PDF&utm_campaign=PDFCoverPages)

Part of the [Botany Commons](https://network.bepress.com/hgg/discipline/104?utm_source=scholarworks.uni.edu%2Fhpt%2F105&utm_medium=PDF&utm_campaign=PDFCoverPages)

Recommended Citation

Rethwisch, Kayla, "A preliminary systematic analysis of the species of genus Physaria (Brassicaceae) of Wyoming" (2014). Honors Program Theses. 105. [https://scholarworks.uni.edu/hpt/105](https://scholarworks.uni.edu/hpt/105?utm_source=scholarworks.uni.edu%2Fhpt%2F105&utm_medium=PDF&utm_campaign=PDFCoverPages)

This Open Access Honors Program Thesis is brought to you for free and open access by the Student Work at UNI ScholarWorks. It has been accepted for inclusion in Honors Program Theses by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

[Offensive Materials Statement:](https://scholarworks.uni.edu/offensivematerials.html) Materials located in UNI ScholarWorks come from a broad range of sources and time periods. Some of these materials may contain offensive stereotypes, ideas, visuals, or language.

A PRELIMARY SYSTEMATIC ANALYSIS OF THE SPECIES OF THE GENUS *PHYSARIA* **(BRASSICACEAE) OF WYOMING**

A Thesis Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors with Distinction

Kayla Rethwisch

University of Northern Iowa

May 2014

This study by: Kayla Rethwisch

Entitled: A preliminary systematic analysis of the species of genus *Physaria* (Brassicaceae) of Wyoming

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

 \mathcal{L}_max , and the contract of the contract of

Date Dr. Steve L. O'Kane, Jr., Honors Thesis Advisor, Biology

 \mathcal{L}_max , and the contract of Date Dr. Jessica Moon, Director, University Honors Program

TABLE OF CONTENTS

FIGURES

Introduction

Physaria is a recent and rapidly evolving genus of plants in the Brassicaceae family. Particularly problematic are the taxa in the "old" *Physaria* sensu stricto found in Wyoming, characterized by highly inflated, didymous fruits. Due to the plants in this group being of recent origin, the relationships among them are unknown. This study examined several samples from each taxon in a preliminary study of the taxonomy of these species. The resulting phylogenies were analyzed to determine relatedness among the species. By looking at gene regions, sections on the genomic sequence, it is possible to determine how these various plants are related.

The hopes of this study were to solve remaining taxonomic problems among the *Physaria* group, specifically what makes up 'old' *Physaria.* For example, are the subspecies of *Physaria saximontana* and *Physaria didymocarpa* really distinct?; do specimens with elongating caudices represent new taxa; does *Physaria vitulifera* grow in Wyoming? The relatedness among the species of 'old' *Physaria* were to be determined, as well as whether the species in this group are where they belong and what the monophyletic groups within 'old' *Physaria* are. It is important to understand the relationships among these species due to their recent and rapid origin.

History and Identification

Brassicaceae, the mustard family, is composed of 350 to 381 genera and 3,000 species mostly distributed in the northern temperate zones (Walters and Keil, 1996). Members of the Brassicaceae family are characterized by petals arranged in the form of a cross, alternative leaves with branched hairs, tetradynamous stamens, indicating six male fertilizing organs, a membranous septum, or replum that divides the ovary into two locules, and fruits called siliques, a long narrow seedpod, or silicles, a short silique (Glisson, 2004). Brassicaceae was originally

1

systematically analyzed based on morphological variation of the silique and silicle (Al-Shehbaz and O'Kane, 2002). Recently, molecular systematics has revealed new insights into their evolutionary relationships. Analyses of ribosomal DNA and chloroplast genomes combined with analyses of gene sequences and morphological data place Brassicaceae in a monophyletic order, Brassicales (Al-Shehbaz and O'Kane, 2002).

Originally, *Lesquerella* and *Physaria* were both included in the Old World genus *Vesicaria,*

with *Physaria* later becoming a separate genus (Heidel and O'Kane, 2010).The genus *Lesquerella* was later created for species of *Vesicaria* with uninflated, nondouble fruits (Heidel and O'Kane, 2010). Species within *Lesquerella* are mainly perennials and some form woody tissue in their caudices, the base of the stem (Payson,

Figure 1: Physaria acutifolia Photo courtesy of Steve O'Kane

1921). The most common classification of assessing evolutionary trends within this group is seen by looking at fruit morphology and seed anatomy (Beilstein, Al-Shehbaz, and Kellog, 2006).Because the structures of the flowers are relatively stable, the few differences between them offer value in determining relationships between the taxa (Payson, 1921). Evolutionary trends of the group can also be seen by looking at the shapes of leaves which range in shape from entire to pinnate to lobed (Payson, 1921). Evolutionary trends can be seen by also looking at the pedicels, a stalk bearing an individual flower, which are an important indicator for determining species because of a relatively constant position assumed (Payson, 1921). In some primitive species the pedicels are straight or curved upwards, other species are recurved, and yet a third specialized type is S-shaped (Payson, 1921). The species studied in this thesis have the

Figure 2: Twinpods of *Physaria vitulifera* **Photo courtesy of Steve O'Kane**

primitive form of pedicels, being straight or slightly curved. Figure 1 shows the slightly curved pedicels of *Physaria acutifolia* found in 'old' *Physaria.* Recently, most of the species of the genus *Lesquerella,* characterized as bladderpods, has been united with an enlarged *Physaria,* characterized as twinpods (Al-Shehbaz and O'Kane, 2002). Twinpods are distinguished

as having a fruit with a pair of enlarged valves, which can be seen in Figure 2*.* All Wyoming and most of the North American species of *Lesquerella* were transferred to *Physaria* when this occurred (Heidel and O'Kane, 2010). It has been found that *Physaria* is nested within and has evolved more than once from *Lesquerella* (Heidel and O'Kane, 2010). *Physaria* is of very recent origin (Rollins, 1939). *Physaria* and *Lesquerella* are very similar morphologically. They are almost indistinguishable in leaf morphology, trichome type, inflorescence, flower color, fruiting pedicels, and seed-coat sculpture and embryo type (Al-Shehbaz and O'Kane, 2002). The geographic ranges of the *Physaria* falls within the dry western United States and northern Mexico range of *Lesquerella* (Al-Shehbaz and O'Kane, 2002)*.* The differing features are nondidymous fruits with no distal sinuses of *Lesquerella,* and strongly didymous fruits with deep sinuses between the valves distally and proximally of *Physaria* (Al-Shehbaz and O'Kane, 2002)*.* The main distinguishing feature between the two is the doubly inflated fruits of 'old' *Physaria* (Al-Shehbaz and O'Kane, 2002). The newly expanded *Physaria,* that includes *Lesquerella,* is

characterized by stem leaves that are attenuate (reduced) or cuneate (wedge-shaped), to a petiolelike base (leafstalk); a perennial life form; chromosome numbers based on $n = 4, 5, 6, 7$; stellate to stellate-peltrate trichomes (star shaped to shield-shaped); and no documented interspecific hybridization (Al-Shebaz and O'Kane, 2002). The distribution of *Physaria* extends from the southwestern United States to the Canadian border, and from the Great Plains to the Sierra Nevada and Cascade mountain ranges, and can be found in dry barren regions where sunlight is intense and competition is minimal (Payson 1921). These plants have developed an adaptation for these conditions seen in the heavy pubescence on their gray green leaves (Waite, 1973). Descriptions of the species of 'old' *Physaria* found in the Wyoming area are given below (O'Kane, 2000, and Heidel and O'Kane, 2012).

Physaria acutifolia grows in high elevations, open soil patches, hillsides, roadcuts, sagebrush, pinyon-juniper, Gambel oak, and ponderosa pine communities. It is found in Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, South Dakota, Utah, and Wyoming. It has a branched caudex, the axis of the plant, and long styles, the reproductive part of the plant. The basal leaves, those that grow from the lowest part of the stem, are slender and often narrowly winged. The replum is oblong and constricted in the center. The fruiting pedicels are 6-12mm and nearly straight, and petals are yellow. The fruits are didymous and inflated, 8- 20mm. The trichomes, outgrowths from the epidermis of the plant, are appressed in the fruits.

Physaria brassicoides grows on bare hillsides, dry gravel and clay soil, badlands, clay knolls, and banks. It is found in Colorado, Montana, Nebraska, North Dakota, South Dakota, and Wyoming. The caudex is branched and stems several from the base. The basal leaves are somewhat winged. The fruiting pedicels are 5-12mm and straight to somewhat curved, and petals are yellow. The fruits are didymous, moderately inflated, and trichomes are spreading on

the fruits.

Physaria condensata grows on calcareous knolls and ridges, clay banks, limey slopes, shaley hills, and clay patches. It is found in Wyoming. The caudex is usually simple, with branching seen rarely and stems several from the base. The racemes, or flower cluster, are congested and the basal leaves are mostly horizontal. The fruiting pedicels are straight, 5-10mm, and the flower sepals are yellowish-green and fruits are didymous and inflated. The trichomes are loosely spreading on the fruits.

Physaria didymocarpa subspecies *didymocarpa* grows on gravel bars, steep shale outcrops, rocky flats, gravelly prairies, talus slopes, dry hillsides, road cuts, mountains and foothills. It is found in Idaho, Montana, Washington, and Wyoming. Surfaces appear somewhat fuzzy. Petals are yellow and fruits are moderately inflated. There are spreading trichomes on the fruits, some more appressed. There are 4-6-8 ovules per ovary.

Physaria didymocarpa subspecies *lanata* grows on steep limey banks, rock and sand road cuts. It is found in Montana and Wyoming. Trichomes are spreading throughout and surfaces appear woolly. Petals are yellow and fruits are inflated. There are 2-4-8 ovules per ovary.

Physaria dornii grows on calcareous shale, slopes, and ridges. It is found in Wyoming. Caudex is simple and stems are simple from the base. Basal leaves are ascending or erect, and racemes are compact. Fruiting pedicels slightly curved, 7-18mm, and petals are yellow. Fruits are didymous and highly inflated. There are 4-8-12 ovules per ovary.

Physaria eburnifolia grows on limestone hills, red soil, rocky calcareous slopes, clay depressions, granite and marble detritus. It is found in Wyoming. The caudex is simple and stems simple from the base, arising to a rosette. Fruiting pedicels are nearly straight, 6-10mm, the flower sepals are purplish to greenish, and the petals are white. Fruits are strongly didymous, irregular in shape and size, and strongly to somewhat inflated. There are 4-8 ovules per ovary.

Physaria integrifolia grows in calcareous hills and slopes, shale-limestone cliffs, bare steep slopes, red clay banks, and shale. It is found in Idaho, Montana, and Wyoming. Caudex is usually branched, and basal leaves form a strong rosette. Racemes are congested. Fruiting pedicels are spreading, straight or slightly curved, 7-11mm. Fruits are didymous and highly inflated. Trichomes are appressed in the fruit. There are 8 ovules per ovary.

Physaria saximontana subspecies *saximontana* grows on red shaley banks and ledges, heavy clay, and hillsides. It is found in Wyoming. Basal leaves are apex rounded to angled. Styles 3- 5mm. 4 ovules per ovary.

Physaria saximontana subspecies *dentata* grows on open gravelly slopes, scree, and mountain fellfields on limestone. It is found in Montana. The basal leaves are apex angled. The styles are 4-7mm. There are 4-8 ovules per ovary.

Physaria vitulifera grows on rocky hillsides, dry banks, gravel and sand, granitic slopes, soil scree, and red shale. It is found in Colorado. The caudex is simple or branched and stems several from the base. Racemes are congested. Fruit pedicels are usually curving upward, 6- 10mm. Fruits are didymous, inflated, and irregular in shape. Trichomes are appressed or spreading on the fruits. There are 4 ovules per ovary.

The purpose of this study was to solve the remaining taxonomic problems with the 'old' *Physaria* group found in Wyoming. This study examined several samples from each taxon in a preliminary study of the taxonomy of these species. By looking at gene regions, sections on the genomic sequence, it is possible to determine how these various plants are related. The relatedness among the species of 'old' *Physaria* were to be determined, as well as whether the species in this group are where they belong and what the monophyletic groups within 'old'

Physaria are. Several programs can be used in order to determine the relatedness among these species.

Programs

Several programs can be used to prepare phylogenetic trees. These programs include MEGA5, CLUSTAL, MAFFT, and MRBAYES. Through various studies, these programs have been found to be very effective and accurate.

MEGA5 offers two methods for aligning nucleotide and protein sequences: MUSCLE and Clustal (Hall, 2011). MEGA5 uses maximum likelihood, evolutionary distance, and maximum parsimony methods to provide tools for statistical analyses of DNA (Tamura et al, 2011). The maximum likelihood analyses allow users to infer evolutionary trees, select the best-fit substitution models, and infer ancestral state and sequences (Tamura et al, 2011). Neighborjoining algorithms are used to optimize the model that best fits the data, and the bootstrap method used allows users to use data that contains sites with insertion-deletions and missing data (Tamura 2011). Ancestral states and sequences can be found by estimating branch lengths and using Bayesian posterior probabilities to generate the possible ancestral state of each node (Tamura 2011). Users are able to explore ancestral sequences inferred using maximum parsimony methods (Tamura et al, 2011).

CLUSTAL X provides an integrated system for performing multiple sequence and profile alignments, realigning sequences, and analyzing the results (Thompson, 1997). The sequences are automatically colored to highlight certain regions on the alignment (Thompson, 1997). Figure 3 displays a set of data in multiple alignment mode in the CLUSTAL program that has been realigned. Low scoring segments are highlighted using a white character on a black background and exceptional residues are shown as a white character on a grey background (Thompson,

 \bf{B}

EFTU ANANI VTGVEMFOKTLDEGLAGDNVGLLLRGIOKADIERGMVLAKPGSITPHTKFESEVYVLKKEEGGRHTPFFPGYRPOFYVR											345
EFTU SPIPL VTCAEMFOKTLEEGMAGDNVGLLLRGIOKNDVORGMVIAKPKSITPHTKFEAEVYILKKEEGGRHTPFFKGYRPOFYVRJ											346
EFTU ODOSI	TTGTEMFOKTLDEGTAGDNVGTLLRGVAREDTEAGMVLSEPGTTTPHTMFESEVYVLTKDEGGRHTPFFTGVRPOFYVRT										345
EFTU CHACO	VIGLEMFRRLLEOG AGENIGVLLRGIEKKDIERGMVIAGPGTIEPHTRFEAOVYILRKEEGGRHSPFF GYRPOFFVRT										346
EFTU CHLRE	VTGLEMFKKTLDE/VAGDNVGVLLRGVOKKDIERGMVIAKPGTITPHTKFEAOVYVLTKEEGGRHS/FMLGYOPOFYVRT										345
EFTU ARATH	VIGVEMFOKILDEAFAGDNVGLLLRGIOKADIORGMVLAKPGSIDPHTKFEATIYVLKKEEGGRHSPFFAGYRPOFYMRI										345
EFTU MYCGE	VTGIEMFKKELDSAMAGDNAGVLLRGVERKEVERGOVLAKPGSIKPHKKFKAEIYALKKEEGGRHTCFLNGYRPDFYFRI										335
EFTU MYCPN	VTGIEMFKKVLDSAMAGDNAGVLLRGVDRKEVERGOVLAKPGSIKPHKKFKAEIYALKKEEGGRHTCFLNGYRPOFYFRT										335
EFTU CHLVI	VIGIEMFOKTLDEGOAGDMAGLLLRGVDKEALERGMVIAKPGSITPHTKFKAEVYILKKEEGGRHTPFFNGYRPOFYFRI										334
EFTU BACSU	VTGVEMFRKLLDVADAGDNIGALLRGVSREFIGRGOVLAKPGTITPHSKFKAEVYVLSKEEGGRHTPFFSNYRPOFYFRT										337
EFTU ECOLI	CTGVEMFRKLLDEGRAGENVGVLLRGIKREEIERGOVLAKPGTIKPHTKFESEVYILSKDEGGRHTPFFRGYRPOFYFRT										334
EFTU RICPR	CTGVEMFRKLLDEGOSGDNVGILLRGAKREEVERGOVLAKPGSIKPHDKFEAEVYVLSKEEGGRHTPFANDVRPOFYFRT										335
EFTU BACFR	VIGVEMFRKLLDOGEAGDNVGLLLRGVDKNEIKRGMVLCKPGOIKPHSKFKAEVYILKKEEGGRHIPFHNKYRPÖFYLRI										335
EFTU CHLTR	VTGVEMFRKELPEGRAGENVGLLLRGIGKNDVERGMVVCLPNSVKPHTOFKCRVYVLOKEEGGRHKPFFTGYRPOFFFRT										334
EFTU BORBU	VTGVEMFOKTLEOGOAGDNVGLLLRGVDKKDIERGQVLSAPGTITPHKKFKASIYGLTKEEGGRHKPFFPGYRPQFFFR										336
	ruler 280290300310320330340350										
		M MWWWWWWWW									

Figure 3: Realigned sequence in Clustal Photo from Thompson, 1997

MAFFT covers various types of Multiple Sequence Alignment problems ranging from small alignments consisting of distantly related sequences to large-scale alignments (Katoh and Toh 2008). MAFFT uses sequences that are assumed to be homologous; when there is an insertion or deletion in the alignment, the position is displayed as a gap (Katoh and Toh 2008). MAFFT uses a two-cycle progressive method to balance speed and accuracy and two techniques, group to group alignment algorithm and 6mer method, to reduce CPU time of progressive methods (Katoh and Toh, 2008). Iterative refinement method used in MAFFT uses the weighted sum-of-pairs which are refined using an approximate group-to-group alignment algorithm and repeated until no improvements are made (Katoh and Toh, 2008). MAFFT has been improved by adding two techniques to improve scalability of progressive alignment and accuracy of ncRNA alignment: the PartTree algorithm and the Four-way consistency objective function (Katoh and Toh, 2008).

 Bayesian inference has many advantages including easy interpretation of results and the ability to incorporate prior information. In a bayesian analysis, inferences of phylogeny are

based upon posterior probabilities of phylogenetic trees (Huelsenbeck and Ronquist, 2001). MRBAYES uses Markov chain Monte Carlo (MCMC) which uses samples of posterior probabilities of phylogenetic trees to determine summation and integrals required for the bayesian analysis (Huelsenbeck and Ronquist, 2001). MRBAYES has a command-line interface that reads in an aligned matrix of DNA sequences in the NEXUS format (Huelsenbeck and Ronquist, 2001). The user can change assumptions of the substitution model, details of the analysis, and can delete and restore taxa in the analysis (Huelsenbeck and Ronquist, 2001). The program can infer ancestral states while accommodating uncertainty about the phylogenetic tree and model parameters (Huelsenbeck and Ronquist, 2001).

Methods

This project entailed a series of steps: (1) DNA extraction, (2) PCR amplification, (3) DNA sequencing, (4) DNA alignment, and (5) phylogenetic analysis. Previous work in O'Kane's lab has identified several non-coding DNA regions that show promise. The chloroplast *rps* intron and the *ndh*C*-trn*V intergenic spacer were studied. The nuclear internal transcribed spacer of ribosomal DNA (rITS) and two introns of *Pgi*C were examined. Resulting phylogenies were examined for how well resolved they are and for the underlying statistical support.

DNA Extraction

Dneasy Plant Mini Kit and Dneasy Plant Maxi Kit Handbook were used to complete DNA Extraction (Qiagen, 2001). Roughly 20 mg of plant tissue was ground to a fine powder using a mortar and pestle. Using a micropipette, 400µL Buffer API and 4µL Rnase A were added to the plant tissue. The mixture was incubated for 30 minutes at 65° C. The mixture was inverted 2 to 3 times during incubation. Next, 130µL of Buffer P3 was added to the lysate mix and incubated for 5 minutes on ice. After 5 minutes, the mixture was centrifuged using a 16M Microcentrifuge for 5 minutes at full speed. The lysate mixture was then pipetted to a QIA Shredder Spin column sitting in a 2mL collection tube. The contents were centrifuged for 2 minutes at full speed. The flow-through fraction was transferred to a new tube without disturbance of the pellet. Next, 1.5 volumes of Buffer P3 was added to the lysate and mixed by pipetting. 650uL of the mixture was applied to Dneasy mini spin column sitting in a 2mL collection tube. The mixture was centrifuged for 1 minute at 9000rpm and the flow-through was discarded. The remaining mixture not added by the 650μ L addition was added to the Dneasy column mini spin column, centrifuged and discarded as before. The Dneasy column was placed in a new 2mL collection

tube, 500µL Buffer AW was added to Dneasy column and centrifuged for 1 minute at 10,000rpm. The flow-through was discarded. 500µL Buffer AW was added to the Dneasy column and centrifuged for 2 minutes at full speed to dry the membrane. The Dneasy column was then transferred to a 2mL microcentrifuge tube and 75μ L of Buffer AE preheated at 65^oC was pipetted directly onto the Dneasy membrane. The mixture was allowed to incubate at room temperature for 10 minutes and then centrifuged for 1 minute at 9000rpm to elute. Lastly, the elution steps were repeated once.

PCR Amplification

The non-coding gene regions amplified by polymerase chain reaction (PCR) were chloroplast *rps*16 intron and *ndhC*-*trn*V intergenic spacer and nuclear ribosomal internal transcribed spacers (rITS) and *pgi*C 6F-9R introns. Cocktails for each region were made in 1.5mL tubes. 24µL of the cocktail was then distributed among the 0.2mL PCR centrifuge tubes already containing 1μ L of extracted DNA. The total volume in each PCR centrifuge tube was 25µL. They were then amplified using a Biometra thermocycler. The two Biometra thermocycler used throughout this thesis are Tprofessional Basic Gradient Thermocycler and Tgradient Thermocycler.

Rps16 intron had a Phusion (ThermoScientific) cocktail consisting of 1µL DNA, 5µL 5X Phusion HF Buffer, 2.5µL dNTP, 2.5µL rpsR₂, 2.5µL rpsf, 0.25µL Phusion II enzyme, and 11.25 μ L distilled water. The mixtures were denatured for 1 minute at 98 \degree C. Then they were ran through 30 cycles of 10 seconds at 98° C, 25 seconds at 66.8° C, and 30 seconds at 72° C, followed by 7 minutes at 72. A final soak was done at 16° C.

Ndhc-trnv intergenic spacer had a Phusion cocktail consisting of 1µL DNA, 5µL 5X Phusion

HF Buffer, 2.5µL dNTP, 2.5µL ndhc-f, 2.5µL trnv-r, 0.25µL Phusion II enzyme, and 11.25µL distilled water. The mixtures were denatured for 1 minute at 98° C. Then they were ran through 30 cycles of 10 seconds at 98 $^{\circ}$ C, 25 seconds at 63 $^{\circ}$ C, and 1 minute and 30 seconds at 63 $^{\circ}$ C. A final soak was done at 16° C.

rITS had a taq cocktail consisting of 1µL DNA, 2.5µL Buffer, 2.5µL rITS-f, 2.5µL rITS-r, 2.5µL dNTP, 0.15µL taq, and 13.85µL distilled water. The mixtures were denatured for 2 minutes at 94 $^{\circ}$ C. Then they were ran through 30 cycles of 45 seconds at 94 $^{\circ}$ C, 1 minute at 60 $^{\circ}$ C, and 1 minute at 72° C, followed by 5 minutes at 72° C. A final soak was done at 16° C.

PgiC6F-9R had a Phusion cocktail consisting of 2µL DNA, 5µL 5X Phusion HF Buffer, 2.5µL PgiC6F primer, 3.75µL PgiC9R primer, 2.5µL dNTP, 0.25µL Phusion II enzyme, and 9µL distilled water. The mixtures were denatured for 1 minute at 98 °C. Then they were ran through 32 cycles of 98°C for 10 seconds, 54°C for 25 seconds, and 72°C for 7 minutes. A final soak was done at 15^oC.

Running a Gel. 15mL of 20X SB gel and sterile water were combined to make 300mL buffer solution. 15µL ethidium bromide was added to the buffer solution and mixed using parafilm. 35mL of the buffer solution was added to 0.385g agarose. The agarose mixture was boiled 3 times in the microwave. The mixture was swirled gently between each heating. Once cool, the gel was poured slowly into a tray, bubbles were popped that were on the surface, and the gel was cooled in the refrigerator for 15-20 minutes. The gel was placed in the electrophoresis apparatus and buffer was added until it slightly covered the wells. Next, 2.5µL SB loading dye, 2.0μ L of DNA, and 7.5μ L distilled water were added to each well to total 12μ L. The pipet tip was rinsed in the buffer solution between applications. The remaining buffer was

added to cover the gel. The gel was run negative to positive at 120V. Once the gel had run, it was placed in a distilled water bath on a stir plate at speed 4 for 5 minutes. A Kodak EDAS 290 camera was used to take digital images using a UV illuminator. The PCR products were then run through a clean-up procedure using 0.5 units exonuclease I and 0.25 units shrimp alkaline phosphatase for each PCR product. The mixtures then were ran through the Biometra thermocycler with a lid temperature of 104° C for 15 minutes at 37 $^{\circ}$ C, 15 minutes at 80 $^{\circ}$ C, and a soak at 10^oC. This process removed leftover dNTP's and primers not used in the original PCR product. The resulting PCR products were then sent to the DNA Facility at the Iowa State University Office of Biotechnology in Ames, Iowa. PCR products and primers were sent in plates. 7µL of primer was sent for each PCR product and 10µL of each PCR product was placed in the wells.

DNA Sequencing

The sequences returned from the DNA Facility at Iowa State University were then examined using Chromas, a program that allows the user to view the forward and reverse regions for each sequence. The sequences were analyzed closely to ensure accuracy of the read for each base. Ambiguity codes were used when there was uncertainty in the identification of the base. Ambiguity codes are shown in Figure 4. An ambiguity can be seen at base 321 and 322 in the *rps* intron forward sequence of *Physaria eburnifolia* 3121, shown in Figure 5. The forward and reverse sequences were compared against each other using ClustalX which aligns the sequences and displays any differences between the sequences.

Possible base identification Ambiguity Code	
G or T	K
A or C	М
A or G	R
C or G	S
C or T	V
A or T	

Figure 4: Ambiguity Codes

Figure 5: Ambiguity in *Physaria eburnifolia 3121*

Once the sequences were thoroughly analyzed, they were uploaded to an online program called MAFFT that aligns and analyzes a large amount of data rather quickly. The parameters were set to BLOSUM62 for the scoring matrix for amino acid sequences, $1PAM/\kappa=2$ for scoring matrix for nucleotide sequences, and the strategy was set to G-INS-i. Once the results were ready to be viewed, they were sent to the email the user provided. The results for each region were then saved as a FASTA file. Each region was opened in BioEdit and analyzed and edited to ensure that all sequences were correctly aligned in MAFFT.

The next step was finding the best model to use in MrBayes using MEGA6 (Tamura, et al, 2013). Each region was opened and changed to MEGA format. "Distance" was chosen and within "Compute pair-wise distances" was chosen. Each region was uploaded. The next window allowed for the user to change preferences. These changes are shown below in Figure 6. Matrices were then created to determine which species had the same sequences. Those sequences were combined in the master file.

Options Summary									
Option	Selection								
Analysis	Distance Estimation								
Scope	Pairs of taxa								
Estimate Variance									
Variance Estimation Method	None								
No. of Bootstrap Replications	Not Applicable								
Substitution Model									
Substitutions Type	Nucleotide								
Genetic Code Table	Not Applicable								
Model/Method	Maximum Composite Likelihood								
Fixed Transition/Transversion Ratio	Not Applicable								
Substitutions to Include	d: Transitions + Transversions								
Rates and Patterns									
Rates among Sites	Uniform rates								
Garnma Parameter	Not Applicable								
Pattern among Lineages	Same (Homogeneous)								
Data Subset to Use									
Gaps/Missing Data Treatment	Partial deletion								
Site Coverage Cutoff (%)	95								
Select Codon Positions	V 2nd V 3rd V Noncoding Sites J 1st								

Figure 6: Pair-wise distances settings

Still within MEGA6, "Models" was chosen and within models, "Find Best DNA/Protein models (ML)" was selected. Each region was uploaded to find the best maximum likelihood tree. The next window displayed allowed for changes in settings. The settings used are displayed in Figure 7 shown below. The results displayed models from best fit to least. The models for best fit were T92+G for *rps* intron, T92 for *ndhC*-*trn*V intergenic spacer, T92+G for Chloroplast, and K2 for rITS.

Figure 7: Model setting changes

Gaps were coded for each region in a program called FastGap 1.2. These files were saved in Fasta format. Using these files, a network tree was made for each region using SplitsTrees. Each region was uploaded into the program and a network tree was made using the best maximum likelihood tree determined using MEGA in the previous step.

The next step was to use MrBayes (Huelsenbeck and Ronquist, 2001), which interprets the results and creates phylogenies. A block was made for each gene region that contained information regarding each data set. The rITS block is seen in Figure 8. The block was then uploaded to CIPRES online in order to run the block for a large number of generations. The temp and stopval values were changed in increments of 0.001 until the nchain values were between 0.1 and 0.7, and the PSRF+ values were all between 1.00 and 1.02 indicating

```
divergence.
```

```
begin mrbayes;
 log start filename=bayes output;
 set autoclose = yes nowarn=no;
 [begin sets;]
 charset ITS = 1-664;
 charset ITS gaps = 665-683;
partition nuclear = 2: ITS, ITS_gaps;
set partition = nuclear;
lset applyt=(1) nst=2 rates=equal;
lset applyto=(2) nst=1;
 unlink shape=(all) statefreq=(all) revmat=(all) pinvar=(all);
prset applyto=(1) ratepr=variable statefreqpr=fixed(empirical);
showmodel;
 mcmc ngen=15000000 printfreq=100 samplefreq=1000 starttree=random
       nchains=4 temp=0.134 [checkfreq=500] diagnfreq=1000 stopval=0.005 stoprule=yes;
 sumt relburnin=yes burninfrac = 0.33 contype = halfcompat;
 sump relburnin=yes burninfrac=0.33;
 log stop;
end;
```
Figure 8: rITS MRBayes block

Partitionfinder (Lanfear 2012) was used to determine whether the trees could be concatenated or not. For partition finder, the sequences for each taxa were combined. They were concatenated in the order of rITS, *rps* intron, *ndhC*-*trn*V intergenic spacer, and chloroplast. Although the order of combination does not matter, the same order needed to be maintained for

each taxa. Figure 9 displays the block used to run the program. Figure 10 displays the summary of the run. It indicated that rITS should be one partition and *rps* intron, *ndhC*-*trn*V intergenic

```
spacer, \frac{H}{\text{alignment}} = \text{kayla}, \text{phy}; and
           ## BRANCHLENGTHS: linked | unlinked ##<br>branchlengths = linked;
           \frac{m}{2} models = all;
           # MODEL SELECCTION: AIC | AICC | BIC #<br>model_selection = BIC;
           ## DATA BLOCKS: see manual for how to define ##<br>[data_blocks]
           Lata_Diocks;<br>ITS = 1-664;<br>RPS = 665-1656;<br>NDHC = 1657-2295;<br>Chloroplast = 2296-3937;
           ## SCHEMES, search: all | greedy | rcluster | hcluster | user ##
           [schemes]<br>search = all;
           #user schemes go here if search=user. See manual for how to define.#
```
chloroplast should be in another partition.

Figure 9: Partitionfinder block

Best partitioning scheme Scheme Name: 10 Scheme lnL: -6944.49299 Scheme BIC: [14344.285566](tel:14344.285566) Number of params: 55 Number of sites: 3937 Number of subsets : 2

Subset | Best Model | Subset Partitions| Subset Sites| Alignment 1| TrNef+G| ITS| 1- 664| .\analysis\phylofiles\fcdb76644228e946ce4cbed84d141a38.phy 2| K81uf+G| Chloroplast, NDHC, RPS| 665-1656, 1657-2295, 2296-3937 | .\analysis\phylofiles\1d06c082bec1665ff3114d620d2debeb.phy

Scheme Description in PartitionFinder format Scheme $10 = (ITS)$ (Chloroplast, NDHC, RPS);

RaxML-style partition definitions DNA, $p1 = 1-664$ DNA, p2 = 665-1656, 1657-2295, 2296-3937

Figure 10: Partitionfinder summary

There were a few changes made to each tree in the FigTree application. 'Node labels' was checked and under the display drop down within node labels, 'prob' was selected. Font size was increased as needed. Under the 'appearance' drop down, line weight was increased to three. Under the 'trees' dropdown, 'root tree' was checked. The branch leading to outgroups, ARGY7548 and FEND4355, was selected. The tree was rerooted using the 'Reroot' button on the top of the display. The file was exported as a graphic and saved as an Encapsulated PostScript.

Each file was opened in Adobe Illustrator CC. The image opened was ungrouped and unwanted boxes were removed. The text was edited for size, font, and format. Once complete, the file was saved in a pdf format.

Results

Looking at pairwise distances shows how close the sequences are to each other. Each pairwise distance is close to zero, representing sequences that are closely related. These values can be seen below in tables 11-14 of pairwise distances calculated in MEGA.

Figure 11: rITS Pairwise Distances

Figure 12: *rps* **intron Pairwise**

Figure 13: *ndhC***-***trn***V intergenic spacer Pairwise Distances**

Figure 14: Chloroplast Pairwise Distances

The network trees created in Splits Trees are shown below in figures 15-18. The outgroups used were *Lesquerella fendleri* 4355 and *Lesquerella argyraea* 7548.

Figure 15: rITS network tree

For the outgroups in Figure 15, FEND4355 and ARGY7548, referred to in Appendix A, the double banded line reaches a long distance indicating much variation from the other taxa, as expected. PINTE3149 and PSAXID89 are most similar to the outgroup as they are placed near the outgroup. PDORN4376, PEBUR3121, and PBRAS1811, PDIDYL3136, and PDIDYL3138 have double banded lines leading to them. The length of the double banded line indicates how apart the taxa are, whereas a long double banded line indicates more distance between the taxa and a short double banded line indicates a closer distance between the taxa (Bryant and Moulton 2002). PDIDYLY3855 and PDIDYLY2689 share a double banded line indicating a conflicted relationship between the two. PDORN4376 and PEBUR3121 also share a double banded line

and the remaining lines are short, indicating a close relationship between the two taxa. The other taxa are represented with straight lines indicating that there is no 'noise' in the data or conflicting evolutionary history (Bryant and Moulton 2002). PDIDYLY3855 and PDIDYLY2689 are next to each other in the network indicating a close relationship. There is a trifurcation event for PSAXID79, PBRAS3130, and PDIDYD3794, indicating a close relationship between the three taxa.

The outgroups for Figure 16 again have a long double banded line. PDIDYLY3855 comes out of the cluster with a long straight line indicating that it is not as closely related to the other taxa. On the other hand, there are short lines coming out of the cluster for PCOND3788, PDORN4376, PVITU0489, PVITU0500, and PACUT12821, PBRAS1811, PBRAS3130, PCOND3787, PDIDY551, PSAXI74315, indicating close relationships to each other.

Figure 17: *ndhC***-***trn***V intergenic spacer network tree**

For Figure 17, the outgroup has a long thick double line extending from the cluster, indicating a distant relationship from the other taxa. The remaining taxa extend from a cluster of parallel lines in the middle. PVITU0489 and PVITU0500 share a parallel line indicating a close relationship to each other.

For Figure 18, PDIDYLY3855 has a long line extending from the center indicating a distant relationship from the other taxa. PDIDY2689 and PDIDLY3855 share the same parallel line indicating a close relationship between the two. A close relationship is indicated for PVITU0489 and PVITU0500 as they either share a parallel line in the tree or are short lines coming out of the cluster.

Figure 19: rITS phylogenetic tree. Numbers at nodes are posterior probabilities.

Figure 20: *rps* **intron phylogenetic tree.** Numbers at nodes are posterior probabilities.

Figure 21: *ndhC***-***trn***V intergenic spacer phylogenetic tree.** Numbers at nodes are posterior

probabilities.

Figure 22: Chloroplast phylogenetic tree. Numbers at nodes are posterior probabilities.

Discussion and Conclusion

There were several similarities between the network trees seen in Figures 19-22. For all four network trees, the outgroup extends from the center with a double banded line, indicating a distant relationship from the rest of the taxa. PDIDYLY3855 extends from the center with a long straight line indicating some distance from the other taxa.

There were several similarities between the phylogenetic trees. For Figures 19-22, ARGY7548 and FEND4355 are the outgroups which enable the reader to distinguish the differences between the resulting data. When comparing the *rps* intron phylogenetic tree and the ndhC-trnV intergenic spacer phylogenetic tree, there were several similarities. PDIDYLY2689 and PDIDYLY3855 are sister taxa in both trees indicating a close relationship between the two taxa. In the *rps* intron tree PVITU0489 and PVITU0500 were the same sequence but in the ndhC-trnV intergenic spacer tree, they were sister taxa. Both relationships indicate the two species are very similar. PACUT3721, PDIDYD3794, PDIDYL3136, PDIDYL3138, and PSAXID89, are all the sequence sequence in both the ndhC-trnV intergenic spacer and Chloroplast trees.

The *rps* intron and ndhC-trnV intergenic spacer phylogenetic trees also shared similarities with the Chloroplast tree. One being that PACUT12821, PBRAS3130, PCOND3787, PDIDYD551, and PSAXI74315 are the same sequences for all three. The final similarity was that PVITU0489 and PVITU0500 are sister taxa in the *rps* intron, ndhC-trnV intergenic spacer, and also the chloroplast phylogenetic trees.

There were a couple of similarities between the ndhC-trnV intergenic spacer and Chloroplast phylogenetic trees. PACUT12821, PBRAS3130, PCOND3787, PDIDYD551, and PSAXI74315 are sister taxa with PBRAS1811 in both phylogenetic trees. Lastly, PACUT3721, PDIDYD3794, PDIDYL3136, PDIDYL3138, and PSAXID89, are all the same sequence in both.

There were only a couple similarities between the rITS and ndhC-trnV intergenic spacer phylogenetic trees. The first being that PACUT3721, PDORN17503, and PEBUR3120 were all the same sequence. The second is that PACUT12821 and PSAXI74315 are also the same sequence, which is also seen in the Chloroplast tree as well.

There was one similiarity between the *rps* intron, rITS, and Chloroplast phylogenetic trees. PDIDYLY2689 and PDIDYLY3855 are sister taxa in these trees. There were several similarities between the *rps* intron and Chloroplast phylogenetic trees. PDIDYD3794, PDIDYL3136, PDIDYL3138, PSAXID79, and PSAXID89 are all in the same position. PCOND3788, PDORN4376, PINTE3149, and PSAXI3153 are all in the same position. PDORN17503 as well as PEBUR3121 are also in the same position.

This study examined several taxa from the *Physaria* group, found in Wyoming. The hopes of this study was to solve the remaining taxonomic problems within the group, in order to understand the phylogeny of the group in better detail. While there were several similarities between the phylogenetic trees from this study, there were also many differences, as well as trifurcations which indicate there was not enough data to resolve the phylogeny of the group. While the results were inconclusive, they provided a foundation for future studies in that the gene regions used in this study did not provide enough information to resolve the phylogenies. In the following study done by a grad student, one or more additional chloroplast gene regions will need to be examined in order to resolve the phylogeny of the group.

Literature Cited

Al-Shehbaz, I.A., and O'Kane Jr, S.L. 2002. *Lesquerella* is united with *Physaria (*Brassicaceae*).* Novon, 12: 319-329.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the compositae. Molecular Phylogenetics and Evolution, 1(1): 3-16.

Beilstein, M.A., Al-Shehbaz, I.A., and Kellogg, E.A. 2006. Brassicaceae phylogeny and trichome evolution. American Journal of Botany, 93 (4): 607-619.

Bryant, D., and Moulton, V. 2002. NeighborNet: An agglomerative method for the construction of planar phylogenetic networks*.* Algorithms in Bioinformatics Lecture Notes in Computer Science, 2453: 375-391.

Dress, A., Huson, D., and Moulton, V. 1996. Analyzing and visualizing sequence and distance data using SplitsTree. Discrete Applied Mathematics, 71: 95-109.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5): 1792-1797. doi: 10.1093/nar/gkh340.

Glisson, B. 2004. *Physaria saximontana* Rollins var. *saximontana* (Fremont County twinpod): Technical Conservation Assessment. USDA Forest Service, Rocky Mountain Region.

Hall, B.G. 2011. Phylogenetic trees made easy. Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts, USA.

Heidel, B., and O'Kane Jr, S.L. 2010. *Physaria* species of Wyoming: Part 1. Plant Taxonomy in Action. Castilleja. Publication of the Wyoming Native Plant Society 29 (3). [www.uwyo.edu/wyndd/wnps/wnps_home.htm.](http://www.uwyo.edu/wyndd/wnps/wnps_home.htm)

Heidel, B., and O'Kane Jr., S.L. 2012. *Physaria* Species of Wyoming. Part 2. *Physaria* Key. Castilleja. Publication of the Wyoming Native Plant Society, 31 (3). [www.wynps.org.](http://www.wynps.org/)

Huelsenbeck, J.P. and Ronquist, R. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17 (8): 754-755.

Huson, D.H. 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics, 14 (1):68-73.

Huson, D.H., and Bryant, D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Molecular Biology and Evolution, 23 (2): 254-267.

Katoh, K and Toh, H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinformatics, 9 (4): 286-298. doi: 10.1093/bib/bbn013.

Lanfear, R., Calcott, B., Ho, S.Y.W., and Guindon, S. 2012. PartitionFinder: Combined Section of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Molecular Biology and Evolution, 29 (6): 1695-1701.

Mensen, T. The utility of phosphoglucose isomerase (PgiC) intron sequences in the phylogenetics of *Physaria* (Brassicacae). 2013. Unpublished Senior Thesis, Honors Program, University of Northern Iowa.

Mulligen, G.A. 1966. Two new species of *Physaria* (Cruciferae) in Colorado. Can. J. Bot, 44: 1661-1665.

Missouri Botanical Garden Press. 2002. *Novon,* 12 (3): 319-319. [http:www.jstor.org/stable/3393073.](http://www.jstor.org/stable/3393073)

ThermoScientific Molecular Biology. 2003. *Phusion High-Fidelity DNA Polymerase.*

O'Kane, S.L., Jr. 2000. *Physaria.* In Flora of North America, 7: 226-241, 612-617.

Payson, E.B. 1921. A Monograph of the genus *Lesquerella.* Annals of the Missouri Botanical Garden, 8 (2): 103-236.

Qiagen. 2001. Dneasy Plant Mini Kit and Dneasy Plant Maxi Kit Handbook 1: 1-23.

Rollins, R.C. 1939. The cruciferous genus *Physaria.* Rhodora, 41: 392-415.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., and Small, R.L. 2004. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. American Journal of Botany, 92 (1): 142-166.

Shaw, J., Lickey, E.B., Schilling, E.E., and Small, R.L. 2006. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. American Journal of Botany, 94 (3): 275-288.

Tamura, K, Peterson, D, Stecher, G, Nei, M, and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution, 28 (10): 2731-2739.

Tamura, K, Stecher, G, Peterson, D, Filipski, A, and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. [Molecular Biology and Evolution,](http://mbe.oxfordjournals.org/content/30/12/2725.abstract.html?etoc) 30: 2725- 2729.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The ClUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research, 14 (24): 4876-4882.

Waite, S.B. 1973. A taxonomic revision of *Physaria (*Cruciferae) in Utah, 33 (1): 31-36.

Walters, D.R. And D.J. Keil. 1996. Vascular Plant Taxonomy, 4th edition. Kendall/Hunt Publishing Company. Dubuque, IA.

Appendix A

This study by: Kayla Rethwisch

Entitled: A preliminary systematic analysis of the species of genus *Physaria* (Brassicaceae) of

Wyoming

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

 $\frac{11/3c}{2c/3c/4}$
Date

Dr. Steve L. O'Kane, Jr., Honors Thesis Advisor, Biology

Dr. Jessica Moon, Director, University Honors Program