Primer Development Methodology for RT-qPCR Genetic Studies of Lodgepole Pine Dwarf Mistletoe (*Arceuthobium americanum*)

Cameron Boyda 2021

B.Sc., Thompson Rivers University, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE IN ENVIRONMENTAL SCIENCE

Thesis Examination Committee:

Mila Kwiatkowska (PhD) Associate Professor in the Department of Computing Science and Thesis Co-Supervisor, Thompson Rivers University

Joanna Urban (PhD) Associate Teaching Professor in the Department of Biological Sciences and Thesis Co-Supervisor, Thompson Rivers University

Kingsley Donkor (PhD) Professor in the Department of Chemistry and Committee member, Thompson Rivers University

March 2021

Thompson Rivers University

Cameron Boyda

ABSTRACT

The genus *Arceuthobium* (Dwarf Mistletoe) of the family Santalaceae, is a group of dioecious parasitic flowering plants with a unique, explosive, seed dispersal system. *Arceuthobium americanum* infects Lodgepole Pine (*Pinus contorta* var. *latifolia*) in North Western America and is considered as a pest in the forestry industry. Mistletoe is also an ecologically important organism, promoting wildlife habitat, and biodiversity. Current literature on *Arceuthobium* shows incomplete understanding of molecular mechanisms underlying explosive seed dispersal, and critical mechanisms linking parasitic infection and mortality of the Lodgepole Pine host. Experimentation using Real-time quantitative polymerase chain reaction (RT-qPCR) measures the expression of genes, giving insight into the molecular mechanisms. Determination of stably expressed genes, known as reference genes, is an important step for conducting reproducible experiments and to obtain reliable RT-qPCR expression data. The objective of this thesis is to create an *in silico* methodology to develop primer pairs to target reference genes as a basis to study Dwarf Mistletoe using RT-qPCR. Highly conserved genes, such as Ubiquitin, Actin, EF-a, and GAPDH, are targeted using sequences from same family Santalaceae plant *Viscum album*, model organism *Arabidopsis thaliana*, as well as *Solanum lycopersicum*, *Coffea canephora*, *Paeonia suffruticosa*, *Zea mays*, *Kalanchoe dalgremontiana*, *Triticum aestiyum*, and *Nicotiana benthamiana*. Algorithms designed to develop primer pairs, Primer-3, Beacon Designer 8, and Primer-BLAST, were used to determine primers in published sequences of the *Arceuthobium oxycedri* aquaporin gene, a gene of interest in seed dispersal in *A. americanum*. In experimental validation, the primer pair candidates taken from literature produced no significant results in RT-qPCR. A primer pair produced from primer development algorithm Primer-BLAST using the *Arceuthobium oxycedri* aquaporin gene sequence provided consistent and reliable RT-qPCR

data. Although a reliable reference gene was not developed, the research presented in this thesis establishes methodological steps for *in silico* RT-qPCR primer design, it demonstrates that genetic sequences from organisms within the same genus can be used for RT-qPCR primer design, and it provides a primer pair for future studies of the *A. americanum* aquaporin gene using RT-qPCR.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors Mila Kwiatkowska and Joanna Urban, as well as the late Cynthia Ross Friedman for the opportunity to do this study and the

support. Their support, patience, and guidance helped me greatly during the research and writing of this thesis. I would also like to thank my committee member Kingsley Donker, for his continued encouragement throughout the project. My sincere thanks also goes to Diana Bizecki Robson, Matt Reudink, and Wendy Gardner, for their insightful comments and encouragement. I thank Rolena DeBruyn, Rachel Lobay, and Erika Koeck for assistance in the laboratory. I would especially like to thank my friends and family for their personal and moral support.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

Arceuthobium americanum (Lodgepole Pine Dwarf Mistletoe)

Arceuthobium americanum is dioecious hemi-parasitic flowering plant, having male and female reproductive organs on separate individuals. *A. americanum* has a life cycle lasting six years; female flowers appear two years following shoot development and persist up to two years after that. Each growing season, flowers appear on a shoots in summer, dispersing their seeds in the late summer. It has a multitude of hosts, with its main host being *Pinus contorta* Dougl. Ex Loud. var. *latifolia* Engelm. (Lodgepole Pine) (Broshot et al. 1986, Ip 1992). *A. americanum* is able to re-infect host trees, and gradually accumulate in the same host over time, as a systemic or localized infection. Trees that are affected see a reduction of overall growth, and heavy infestation can cause mortality (Hawksworth and Wiens 1996). The infection of the host tree causes a tightening of branches in the area called “Witches’ Brooms”.

The infection of the Lodgepole Pine can also warp the wood, making it unusable for harvesting. In addition, the infection of taller and older trees can pose a significant risk to younger, understory trees by infecting them, hindering replanting efforts *A. americanum* is of particular interest to silviculture because of its economic impact to forestry, and with a potential loss of 37 million m^3 of coniferous forest per year (Unger 1992). Current management options for forestry includes harvest by clearcutting, reduction of edge ratio in a clear cut, pruning, prescribed burning, thinning, and overstory removal (Broshot 1986, Ip 1992).

Geographical Distribution

The genus *Arceuthobium* was thought to comprise about 40 different species found across the globe, but more recently the species number was reduced to 26 by comparing nuclear and chloroplast DNA sequences (Hawksworth and Wiens 1996, Nickrent et al. 2004). *A. americanum* is found in North Western America (see Figure 1.1).

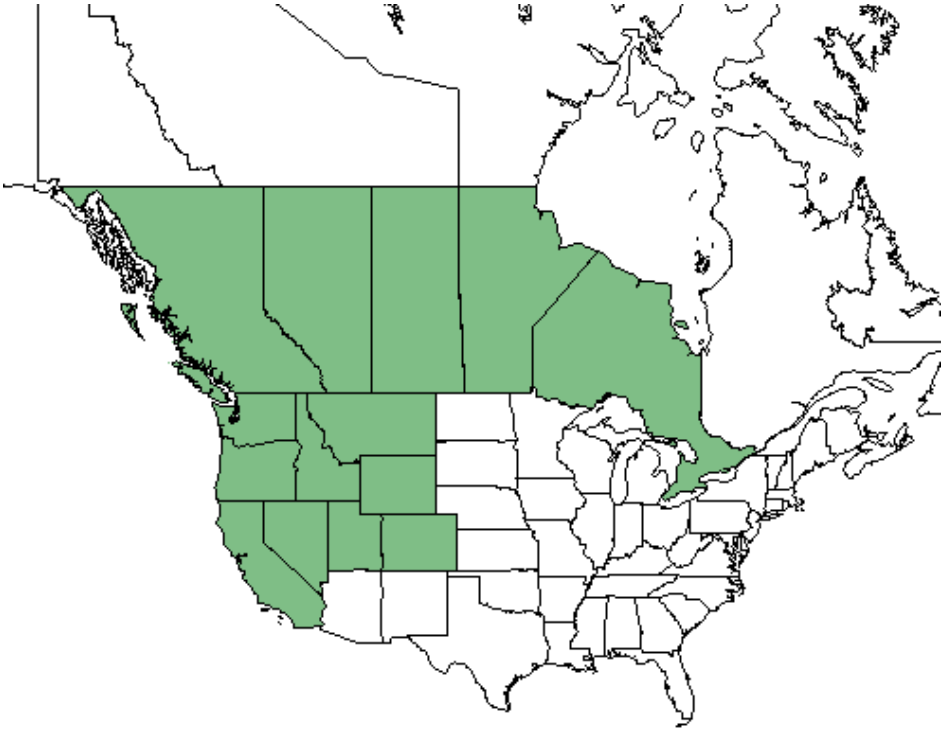

Figure 1.1. Data from the Natural Resources Conservation Services USA, showing the distribution of *Arceuthobium americanum* based on multiple studies and sample collections. (USDA, NRCS 2021)

Ecological Importance

Ecologically, *Arceuthobium americanum* provides food and shelter for forest wildlife. The fruits and stems are used as food by mammals, birds, and other animals. Witches' Brooms are used as

natural shelters by some bird species because of the dense arrangement of branches and needles, which provide protection and support (Shaw 2004).

Seed Dispersal

Dwarf Mistletoe does not rely on birds or mammals to disperse its seeds; it employs an interesting and rare mechanism to do so. Hygroscopic forces build up in the fruits over the summer, and the seeds eventually explode from the ripened fruit, ejecting them onto a new host plant in late August and early September (Hawksworth and Wiens 1996). The seed is discharged at velocities reaching 25 m/s and thrown as far as 20 m from the parent plant (Hawksworth et al. 1977).

Hygroscopic forces needed to expel the seed are thought to originate in the viscin tissue, which is also responsible for adhering to its potential host (Ross et al. 2006). This mucilaginous tissue is formed around the seed in the viscin cells as it grows during the summer months (see Figure 1.2). When the seed is ripe, a thermogenesis trigger causes the dispersal event, although the molecular mechanics are not yet known (deBruyn et al. 2015).

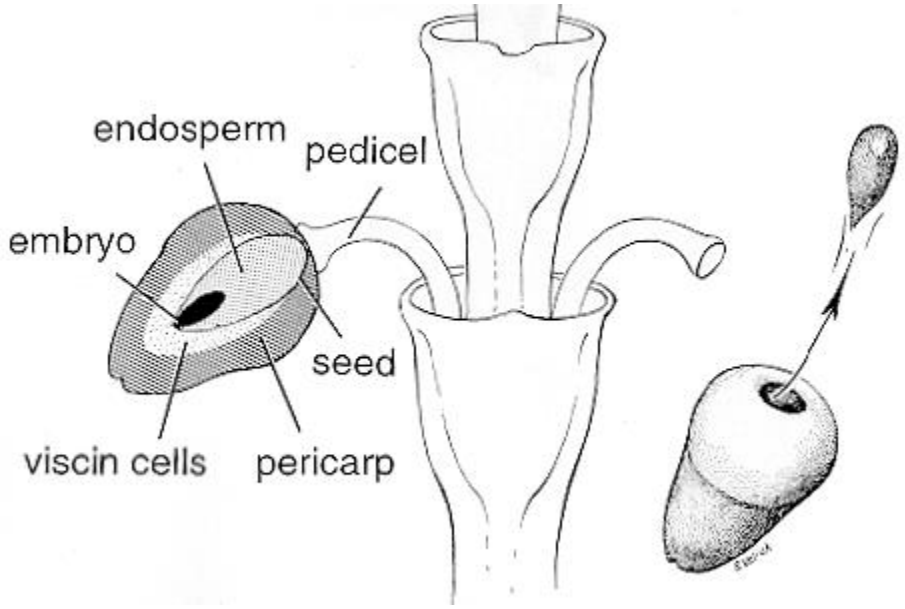

Figure 1.2. A cross section and breakdown of the Dwarf Mistletoe (left). An artist's depiction of the seed dispersal (right). (Hawksworth and Wiens 1996).

Molecular Interests

Currently it is estimated that many genes are involved in the seed dispersal process, such as aquaporin proteins (PIP;2.1), which assist the movement of water across membranes and cyanide-resistant alternative oxidase (AOX), associated with thermogenesis (Maurel et al. 2002, deBruyn et al. 2015, Ross Friedman et al. 2009). A similar water-driven dispersal process is employed in tobacco anther dehiscence, however, as of yet, there is no known similar thermogenesis-triggered dispersal in the plant world (deBruyn et al. 2015, Bots et al. 2005).

Quantitative Polymerase Chain Reaction

In molecular genetics, a common method to quantify expressive genes is reverse transcription quantitative polymerase chain reaction (RT-qPCR). It can accurately quantify target sequences of nucleic acids in a given sample (Saiki et al. 1988). The expression of a particular gene can then

tell a lot about the biology of an organism. It can be used for tracking changes in the expression of genes while tissues develop overtime in a particular organism, as in the case of this thesis.

Relationship Between Genes and RNA

Genomic deoxyribonucleic acids (DNA) are made up of base-pair sequences of Arginine, Thymine, Cytosine and Guanine which are grouped into genes such as PIP;2.1. For a cell to express the gene in the form of a functional protein, the sequence is reverse transcribed into a complementary messenger ribonucleic acid (mRNA, referred to as simply RNA in this thesis). The RNA is a short-lived nucleic acid that lasts only a few seconds to hours; therefore, the cell produces RNA only when the gene is being actively expressed (Caixeiro et al. 2016, Fleige and Pfaffl 2006).

Successful quality extraction of RNA is essential for the construction of cDNA and downstream genetic applications. A quality extraction and subsequent application of RNA is dependent on the quality and purity of the extraction from the tissue sample, i.e., the sequences of the RNA are intact and without a significant number of contaminants (Imbeaud et al. 2005). The intactness of the RNA sequences, or its integrity, can be compromised because of the volatility of the RNA strands and also presence of RNA-degrading enzymes known as RNAses. Degrading-enzyme RNAses break down RNA into smaller pieces of ribonucleic acids. These RNA identifying proteins are ubiquitous; the presence of even small amounts of these proteins can quickly degrade the RNA sequences, causing their analysis to be unreliable (Breslow and Chapman 1996). When reverse transcribing RNA into cDNA, compromised RNA will result in an incomplete cDNA library. Thus, the amount of cDNA is not a representation of the actual expression of RNA in the given sample. This is caused by incomplete or complete lack of

transcription of RNA when reverse transcribing the RNA strands into cDNA. Incomplete cDNA might not have complementary sequences available for RT-qPCR primers to bind to, giving results that falsely negative expression in expression analysis (Fleige and Pflaff 2006). The integrity of the RNA can be measured using a cheap and relatively quick method called gel electrophoresis. This method uses a portion of the RNA sample from the extraction protocol, and separates the strands based on sizes (Stellwagen 2009). Although mRNA is in relatively low amounts in the sample, there is still a large amount of ribosomal RNA (rRNA). When ran in a bleach-agarose gel, 18S and 28S rRNA molecules are dissociated and clearly visible with an expected ratio of 1:2 (Ishikawa 1977). A dye, such as GelRed™ Nucleic Acid Stain (Biotium), or Ethidium Bromide interacts with the RNA, allowing visualization of it under UV light; the intensity of the photon emission can be measured to relate the 18S and 28S ratios (Becker 2010). The resulting RNA extraction must also be free of polymerase inhibitors for the RT-qPCR reaction to take place. Proteins, phenols, and carbohydrates are all known to interfere with both the reverse translation to cDNA and the resulting RT-qPCR reaction (Tichopad 2004). This is especially important when extracting from plant material, which contains these molecules in great abundance (Loomis 1974). It is important to use a protocol that efficiently isolates RNA; this protocol may differ from plant to plant depending on the biochemical composition of the tissues (Salzman 1999). To ensure the purity of the RNA, it can be measured using a spectrophotometer at two particular wavelengths: 260 nm and 280 nm. The ratio between the two wavelength readings indicates the presence of most unwanted molecules with expected wavelength ratios of 1:2. An offset of this ratio indicates substance contamination and could result in polymerase inhibition for cDNA and RT-qPCR reactions (Fleige and Pfaffl 2006). Once the RNA has been determined to meet quality standards for downstream application, quantity of

the RNA should be measured so that RT-qPCR reactions remain consistent across all samples. The measurement of RNA samples can be made with a fluorometric device and dye with high accuracy (Aranda et al. 2009).

Immediately following protocols to identify the properties of the RNA reverse transcription, methods should be followed to create a complementary cDNA library of RNA samples. Using manufactured kits, RNA is easily reverse transcribed using specialized reverse transcriptase and randomized primers. The random primers bind to all RNA in the sample, allowing for the reverse transcriptase to construct the complementary DNA strand (Bustin et al. 2005). The resulting cDNA can then be used in RT-qPCR reactions for expression analysis.

Thermal Cycler

The amount of mRNA is relatively proportional to how much protein is being produced, and so gene expression can be quantified by measuring its corresponding mRNA sequences. This is done by first reverse transcribing isolated RNA into cDNA, which is more stable and can be read by RT-qPCR machines. Reverse transcription into cDNA is done with reverse transcriptase, a protein derived from viral replicating systems (Fleige and Pfaffl 2006).

The core function of the RT-qPCR machine is to use the polymerase chain reaction to replicate the target cDNA. Using reagents in the sample, cDNA, DNA polymerase, free nucleotides, and primers to target the cDNA in question, it is able to double the target cDNA each cycle. During a cycle the RT-qPCR machine heats to 90°C to dissociate the cDNA into single strands, allowing the primers to bind to the target cDNA after cooling to ~60°C (Saiki et al. 1988). While primers are bound to the target cDNA, the DNA polymerase can then replicate the cDNA sequence by

adding the free nucleotides, ending the cycle. Each cycle in the machine then effectively doubles the amount of target cDNA (see Figure 1.3).

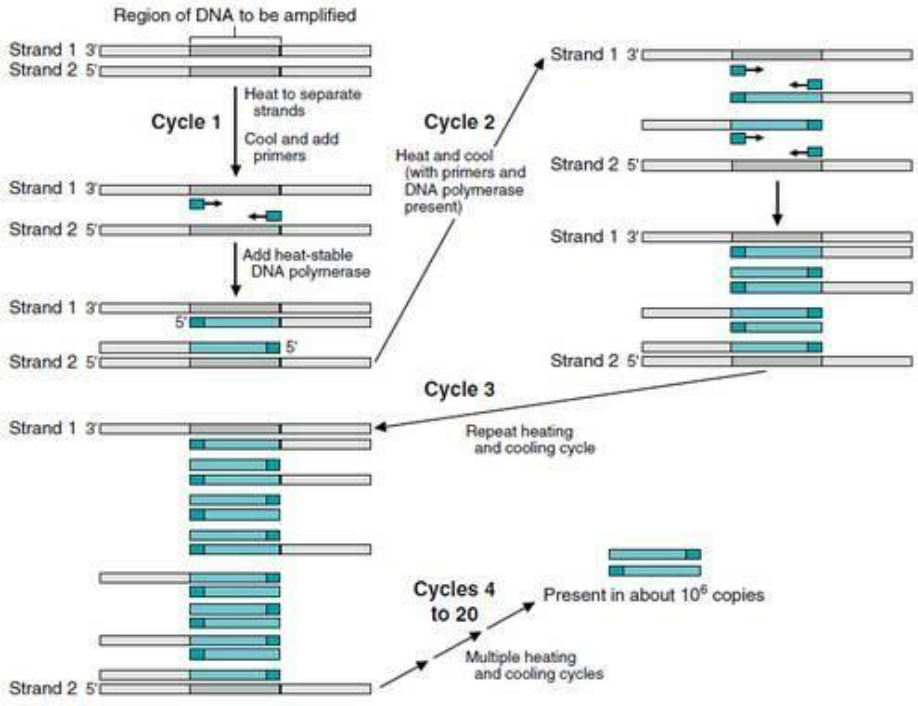

Figure 1.3. Schematic of Polymerase Chain Reaction (RT-qPCR) showing the amplification of target DNA after a number of denaturing and annealing cycles. (Smith et al. 2003).

The RT-qPCR machine itself, or thermocycler, measures DNA or cDNA using fluorometrics. This can be done by either fluorogenic target-specific probes or DNA-binding dye; this thesis uses the latter. A binding-dye binds to double-stranded DNA, or cDNA, which then will emit a fluorescent signal when bound to double-stranded DNA. This dye then binds to both the target cDNA and the background cDNA. The thermal cycler then goes through cycles, doubling the target cDNA until the amount of target cDNA exceeds the amount of background cDNA (see Figure 1.3) (Singh and Pandey 2015).

The cycle number where when the target cDNA exceeds the background cDNA is measured by the RT-qPCR and called the Ct value. A lower Ct value indicates a higher quantity of target

cDNA than a higher Ct value (Livak 2001, Pflaffl 2001). The Ct measurement does not say much by itself, but with the addition of a reference it can relatively quantify the target cDNA (Vandesompele et al. 2002).

Reference Genes

Reference genes are common genes expressed at stable levels throughout intended experimentation. Historically a set of housekeeping genes were frequently used for RT-qPCR including: 18S rRNA, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), elongation factor 1-alpha (EF-1α), polyubiquitin (UBQ), Actin, alpha-tubulin (α-TUB), and beta-tubulin (β-TUB) (Czechowski et al. 2005). Before normalization of data using reference genes, about 90% of papers from 1993 to about 1999 used a single reference gene from this list to perform expression analysis (Suzuki et al. 2000). The standard protocol for using reference genes in RT-qPCR quickly changed after it was realized that using one gene does not give an accurate representation of the expression of a target gene (Czechowski et al. 2005).

Normalization

Normalization of novel candidate reference genes can be established with the use of computer programs to statistically find the most stable genes measured by RT-qPCR. The most common programs currently being used are: Normfinder, BestKeeper, and GeNorm (Kumar et al. 2011). To give an example, the program GeNorm is able rank genes accordingly based on stability of Ct values during experimentation in RT-qPCR. The calculated gene-stability output, or ‘M’ value, essentially describes the variation between two internal reference genes. The ideal internal reference gene variation would be M=0 since stability depends on variation, less stable (or more variation) causes the M value to increase. Assuming all genes vary in expression a certain

amount, to recognize reference genes as stable the M value could not exceed 1.5. Pairwise comparison of each gene across a set of reference genes ultimately determines each reference gene's stability in the array of experimental conditions tested (Vandesompele et al. 2002).

A study by Czechowski and colleagues in 2005 used 25 expressed genes (20 novel and 5 traditional reference genes) to assess the current paradigm on reference gene standards. The study showed that during different biological circumstances *Arabidopsis thaliana* expresses housekeeping genes differently and those traditional reference genes may not accurately show relative quantification of target genes. Thus, different reference genes may be needed depending on the experimental conditions including: the plant species being studied (Vega-Bartol et al. 2013), fruit development, tissue type being studied, abiotic and biotic stresses being introduced (Czechowski 2005).

The paradigm shift in the standards of reference genes used in RT-qPCR caused a large number of subsequent studies to determine optimal reference genes in different experimental conditions. There is now an abundance of information regarding how particular reference genes are optimal for specific experimental conditions. This abundance in information may be able to be used to determine candidate reference genes more systematically, rather than guessing what reference genes might be needed for de novo expression analysis using RT-qPCR.

Study Objective

The goal of this thesis is to identify multiple reference genes for RT-qPCR using existing literature and to experimentally test them in the laboratory. This thesis will require a literature review of primers being used in plant studies that may be viable in Dwarf Mistletoe studies. It will also involve analysis of genetic data and testing outcomes of primer development programs

and theories. After accumulating data, testing of primers will be done using Dwarf Mistletoe RNA, to determine the viability of these methods in the laboratory.

CHAPTER 2. METHODS

Literature Review and Selection of Primers

The current and most useful way of choosing reference genes for RT-qPCR is using recent literature. A literature search is most useful for studies done with common plants such as model organisms (i.e., *Arabidopsis thaliana*) or agricultural species (i.e., *Solanum lycopersicum*) (Kumar et al. 2011), however for *A. americanum* it is perhaps more useful to systematically determine what reference genes could be used in RT-qPCR since there is no optimization yet done in this family of plants.

Literature Used

A total of 57 relevant research articles were used to determine reference genes in *A. americanum*. The criteria for article selection were as follows: 1) articles must have been researching normalization of reference genes for expression analysis in plant(s) using RT-qPCR only; and 2) articles that studied fruit development or an array of different types of studies that were aggregated into one study.

Initially, the total amount of tested reference genes were averaged among all literature articles. Literature that described the most stable genes were tallied based on the most stable value given by the article. Each research article ranked reference genes based on an aggregated M value of all experimental conditions in the study, from lowest M value to highest M value. The most stable reference gene was ranked 1 and had the lowest M value associated with it. Each

candidate reference gene was then tallied for the number of times it was declared as a rank. For example, elongation factor 1 alpha (EF1-α) was tallied in 10 articles as rank 1, 6 articles as rank 2, and so on.

The most common stable genes from the literature were ranked according to a highest weight score. The weight score for each rank was dependent on the total amount of candidate reference genes. Therefore, the reference genes that were most stable had a higher weight score. The total amount of articles listed at a particular rank was multiplied by the total number of candidate reference genes. For each subsequent rank (rank 2, 3, 4 etc.) the total number of candidate reference gene variables was cumulatively lessened by 1, to give subsequent ranks less weight. The total weight score of the gene was the summation of all rank calculations. This system can be represented by the following equation where the WS is the total weight score of the reference gene, N is the number of articles found at a given rank (R) for that reference gene, and G is the average number of reference genes used in all literature.

$$WS = \sum \quad N_1 x \left(G + 1 - (R_1)\right) + N_2 x \left(G + 1 - (R_2)\right) + \cdots N_G \, x \, (G + 1 - (R_G))$$

Figure 2.1. Proposed equation for systematic ranking of reference genes for RT-qPCR given stability based on current literature.

Determination of Candidate Reference Genes

The proposed candidate reference genes to be used for optimization of RT-qPCR in *A. americanum* using the previous methodology can be found in Table 2.2. Each research article varied in number of reference genes tested for normalization averaging to G=12.22 (+/- 5.96553 SD) analyzed reference genes out of a total of 25 possible candidate reference genes. Using systematic determination of candidate reference genes, the top three predicted for future studies

by weight score are: EF1-α, Actin, and Ubiquitin (see Table 2.2). As an example, calculation of weight score sees EF1-α (Figure 2.2).

$$WS_{EF1-\alpha} = 10\,x\,\big((12+1)-1\big) + 6\,x\,(13-2) + 4\,x\,(13-3)\ldots + 1\,x\,(13-12)$$

$$WS_{EF1-\alpha} = 120 + 66 + 40\ldots + 1$$

$$WS_{EF1-\alpha} = 428$$

Figure 2.2. Sample calculation of weight score of Elongation factor 1 alpha (EF1-α) for systematic determination of candidate reference genes to be used in RT-qPCR expression analysis of *A. americanum*).

Primer Development for Quantitative PCR

Aquaporin proteins share common motifs located in the N-terminal cytosolic portion of the protein as well as the six transmembrane helices (Geer et al. 2002). These motif regions are generally more conserved DNA sequences, although they are still subject to single polymorphisms due to translation redundancies. The likelihood that the PIP;2 genetic sequences are similar within the genus of *Arceuthobium* is quite high due to similar putative function in the plant and close evolutionary relationship. Conserved sequences can be useful for developing primers for quantitative polymerase chain reaction (RT-qPCR) within the genus of *Arceuthobium*.

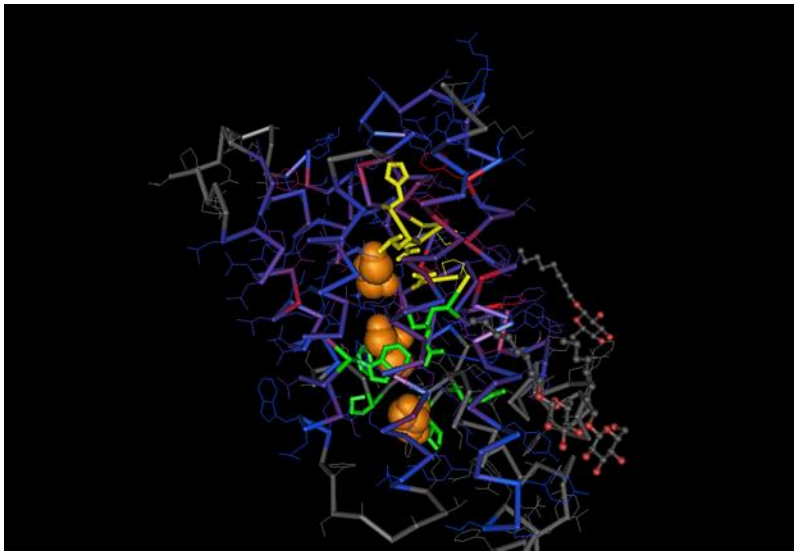

Figure 2.3. Aquaporin protein visualized using Cn3D software. Yellow highlights highly conserved regions of the PIP2 gene, the amphipathic channel, and Asn-Pro-Ala motifs. (Wang et al. 2000).

Primer-designing tools are widely used computerized algorithms to develop primers for genetic analysis, including RT-qPCR. These algorithms consider multiple variables that must be calculated to fit a specific guideline to optimize effectiveness, and ultimately save the time of the researcher. Czechowski and his collegues released a letter indicating most important variables to keep in mind when developing primers (Udvardi et al. 2008). These primers are ideally specific to the gene in question and also follow a set-design criteria: annealing temperature (Tm) = 60 ± 1°C, guanine-cytosine content = 40-60%, amplicon product = 60-150 base pairs, and a primer length of 18-25 base pairs (Udvardi et al. 2008). Primer3 allows detailed parameter conditioning *in silico* (Rozen and Skaletsky 2000). Through time and experimentation, significant improvements have been made as well as conjunct algorithms modified by unassociated organizations such as Primer-BLAST (Koressor and Remm 2007, Untegrasser et al. 2012, Ye et al. 2012).

Primer3 version 4.0.0 and Primer-BLAST designing tools were used to develop RT-qPCR primer pairs for studying PIP;2 expression in *Arceuthobium americanum*. Application of these primers are for practical use in downstream *in vitro* experimentation using SYBR green quantitative assay in the S1000™ thermocycler (Bio-Rad® Laboratories Inc.). Additional steps were included to maximize the efficiency of primer selection, including oligo-analyzing freeware: Beacon Designer (Primer Biosoft International) and to detect possible dimerization of primers.

Identification of Conserved Regions in PIP;2.1 Gene

Aquaporin PIP;2 gene from *Arceuthobium oxycedri* located in the NCBI database (accession number: JN857944.1) was translated into 305 amino acids (see Figure 2.4) and entered into BLASTp algorithm with a maximum target of 100 sequences. The "non-redundant" database was selected for BLASTp, otherwise default settings were used as follows: scoring matrix = BLOSUM62, Gap costs= Existence: 11 Extended:1, and conditional compositional score matrix adjustment was used (Altschul et al. 2005).

A)

ATGGCCAAGGACGTTGAGGCTGGTGCTACCGGCGTCACTGTGAACACCGTCAAGGACTACGAAGACCCTCCGCCGGCTCCTCTGTTCGAGCCACATGAGCTTTT
GAGTTGGTCTTTTTACAGAGCTGCCATCGCGGAGTTCATAGCCACGCTTCTGTTCCTCTACATCACTGTACTCACCGTCATCGGCTACAAAAGCCAGACCGACTCT
AAACTCACGCCTTCCGACCCGTGCGGTGGCGTGGGAGTCCTCGGCATCGCCTGGTCCTTCGGCGGCATGATTTTCGTCCTCGTCTACTGCACCGCCGGCATCTCC
GGGGGGCATATTAACCCGGCGGTGACGTTCGGATTGTTTCTAGCTCGGAAGCTGACGCTGGTGCGAGCCGTGGTGTATATATCGGCTCAGTGCCTGGGAGCGA
TCTGTGGAGTGGGTCTTGTTCGTGCTTTCCAATCTTCCCATTACTTTCGTTACGGCGGCGGAGCCAACATGCTCATGGACGGCTACAGTACCGGCACGGGGCTCG
CCGCCGAGATAATTGGAACTTTCGTGCTAGTCTACACCGTCTTCTCCGCTACCGATCCCAAACGCAGCGCCAGGGATTCCCATGTTCCCGTACTTGCACCGCTTCC
TATTGGGTTCGCAGTGTTCGTGGTTCATCTTGCCACGATCCCCATCACCGGCACGAGCATCAACCCTGCTCGGAGTCTGGGAGCCGCTGTCATATATAATAGAGA
CCAAGAATGGGATGACCAGTGGATTTTCTGGGTAGGGCCATTGTTGGGAGCGGCCATTGCAGCAATATATCATCAATTCATTCTGAGAGCAGGGGCTGTGAAG
GCACTGGGGTCATTCATGGAGTCACCACCCACATCTGATTTTCGGTTCAGTAATTTGATTTACTCANTTATGGAT**TAA**

Translated to

Met A K D V E A G A T G V T V N T V K D Y E D P P P A P L F E P H E L L S W S F Y R A A I A E F I A T L L F L Y I T V L T V I G Y K S Q T D S K L T P S D P
C G G V G V L G I A W S F G G M I F V L V Y C T A G I S G G H I N P A V T F G L F L A R K L T L V R A V V Y I S A Q C L G A I C G V G L V R A F Q S S H Y
F R Y G G G A N M L M D G Y S T G T G L A A E I I G T F V L V Y T V F S A T D P K R S A R D S H V P V L A P L P I G F A V F V V H L A T I P I T G T S I N P
A R S L G A A V I Y N R D Q E W D D Q W I F W V G P L L G A A I A A I Y H Q F I L R A G A V K A L G S F M E S P P T S D F R F S N L I Y S X M D **Stop**

B)

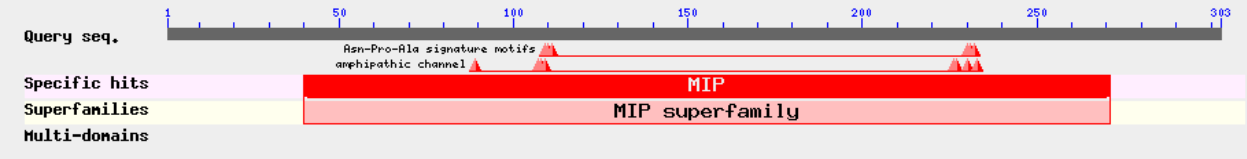

Figure 2.4. A) Aquaporin PIP;2 gene translated DNA to amino acid including conserved sites underlined and highlighted red. Met = Start codon, Stop = TAA (stop codon). B) Conserved regions of PIP;2 gene JN857944.1 shown by BLASTp algorithm conserved regions alignment function.

Primer Development using Primer3 and Primer-BLAST

Primer3 web interface software 4.0.0 was used to develop primers using JN857944.1 gene in FASTA format. The sequence was unaltered from the FASTA file located in the NCBI database and entered into Primer3. General conditions for primers were as follows: Primer size = 18-23 bp (opt 20bp), Primer Tm = 57-62 ± 5 °C (opt 59 °C), GC%=30-70%(opt 50%), and 3' complementarity= <4. Thermodynamic parameters were set to specification indicated by Breslauer et al. (1986). Concentration of monovalent cations were kept as default 50 mM with salt correction standard (Schildkraft and Lifson 1965). All weights were unaltered from default for primer pairing. Primer-BLAST software was set in a similar fashion to that of primer3,

however the thermodynamic parameters and salt correction formulas were standard (SantaLucia 1998). Additional specifications included primers developed specifically in the 150-450 bp,300-600 bp,450-650 bp, and 600-915 bp ranges, as well as primers with forward or reverse primers located in conserved regions.

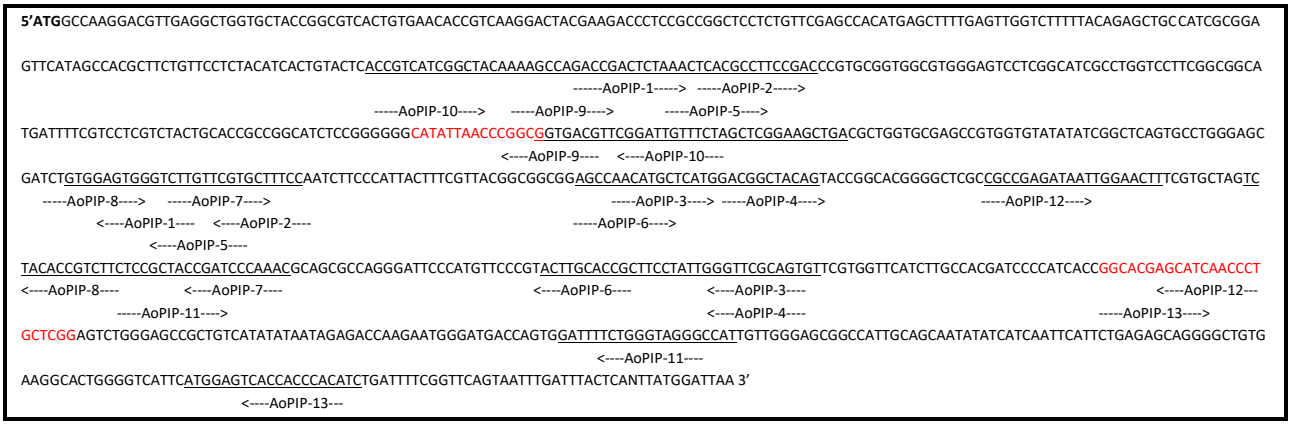

Figure 2.5. Candidate primer pair positions in *Arceuthobium oxycedri* PIP;2 gene. Conserved regions are indicated by boxes and highlighted red. Regions where primers will bind are indicated with an underline.

The Gibbs free energy (ΔG) of primer dimerization was defaulted a limit of -4.00 in cases of cross dimerization, self-dimerization, and 3' end hairpin loop. Amplicon length was set to a range of at least 70 bp to a range of 250 bp long, the maximum of length suggested for the S1000™ thermal cycler manual.

For each program, up to 100 primer pairs were considered given the conditions of each run. However, the top 5 of each run, as suggested by the program, were the only primer pairs considered for this study

Oligonucleotide Analysis of Primers using Beacon Designer 8

Each primer was loaded into Beacon Designer version 8 to determine secondary structures and dimerization probability beyond that calculated in each primer design program. Beacon Designer 8 graphically shows predicted secondary (unwanted) structures and Gibbs free energy (ΔG) for each structure. For each primer pair there was a weight given by adding each structure's ΔG and ranked accordingly, i.e., primer pairs with the lowest ΔG and least number of secondary structures were more favoured.

Isolation and Analysis of RNA

Sample collection

Samples were collected from a population of *Arceuthobium americanum* near Stake Lake, British Columbia (Coordinates via Google map is 50.506141, -120.486170) at 1400-m elevation. A highly infected stand of Lodgepole Pine hosted plenty of samples for summertime collection from 2012 to 2014. On an average basis of twice per week, a branch infected by a female Dwarf Mistletoe was selected through haphazard sampling, clipped at noon, placed in water, and immediately brought to the laboratory for processing. A marker was placed on the tree so that each sample was collected from a separate tree.

Extraction of RNA

To increase the RNA yield for extraction from plant cells, and rupture cells to release mRNA, plant samples were ground into a powder using a mortar and pestle. Equipment used for grinding, as well as forceps and scoopulas, were autoclaved with an exposure time of 15 minutes at 121°C and cleaned prior with ethanol. After equipment was autoclaved, it was thoroughly rinsed with 0.1% diethylpyrocarbonate DEPC-treated water to remove any RNAses still present on the equipment. After equipment was thoroughly cleaned, it was chilled with liquid nitrogen prior to grinding the samples. Whole *Arceuthobium* samples (1.3 grams) were constantly kept in liquid nitrogen after removing them from a -80°C freezer to prevent any thawing. In the mortar, fruits were easily separated from the stems due to the brittle nature of the plant in liquid nitrogen; stems were placed in a separate mortar for grinding. After all fruits and stems were separated, samples were thoroughly ground into a fine dust. Periodically, additional liquid nitrogen was added to replace any evaporated liquid nitrogen in the motor. The pulverized samples were then

placed in a frozen Eppendorf tube with approximately 25-40 powdered fruits or stems (approximately 70 mg ± 36.5mg) in each tube. However, before placing samples back in the -80°C freezer, the open tubes were placed briefly, for approximately 20 seconds, on ice to allow liquid nitrogen to evaporate from the tube, and also for the air inside the tube to warm slightly; this prevented rapid expansion of the air inside the tube, which would cause the tube to explode and destroy the sample. Any frozen samples were kept in a New Brunswick Scientific Ultra Low Temperature Freezer at -80°C equipped with vacuum control.

Fluorometry to determine quantity of RNA isolation was done using Qubit® 2.0 Fluorometer by Invitrogen™ following the manual's protocol. To assess the quality of RNA, it was visualized using agarose gel electrophoresis UV light. Extracted RNA was loaded with at least 1x loading buffer (Life Technologies) into 2.5% agarose gels set in TRIS-Boric acid EDTA (TBE) Bio-Rad (Bio-Rad Laboratories Inc.) nucleic acid electrophoresis boxes. The sample was then running for 30 minutes at 90 volts in a 100mL or 50mL agarose gel, with 10µL or 5µL of GelRed™ (Biotium Inc.), respectively. The chemical GelRed™ is used to bind to the DNA for visualization under UV light.

Gel Electrophoresis of RNA

Gel electrophoresis using agarose gels were used to visualize DNA under UV light. Using Bio-Rad (Bio-Rad Laboratories Inc.) nucleic acid electrophoresis boxes, 1% agarose gels set in TBE buffers were used to run PCR and RT-qPCR products. In a 100mL or 50mL agarose gel, 10µL or 5µL of GelRed™ (Biotium Inc.), respectively, was used to bind to the DNA for visualization under UV light. An electrophoresis power box was set to 90 volts for about 60 minutes for size

resolution of amplicon bands. Each sample was loaded into the gel with at at least 1x Loading Buffer, provided by Life Technologies.

cDNA synthesis

An RNA-DNA hybrid was constructed following the manual's protocol from the Ready-To-Go You-Prime First-Strand Kit provided by GE Healthcare. The second strand synthesis was done with 1 μg of isolated RNA in autoclaved 0.1% distilled DEPC water. RNA was heated to 65°C for 10 minutes, then chilled for 2 minutes on ice and transferred to reaction beads provided by the kit. To prime the polyadenylated tails of the RNA, 0.5 ug of Oligo(dt) primers were added and the solution gently mixed by pipetting up and down. The solution was then incubated for 60 minutes at 37°C and stored at -20°C. Applied Biosystems® reverse transcription kit was also used to synthesize the first strand of mRNA. The protocol as suggested by the operation manual was carried out with 1 μg of each RNA sample. Reverse transcription was performed in BioRad MyCycler™ thermal cycler with a single cycle of the following steps: 25°C for 10 minutes, 37°C for 120 minutes, and 85°C for 5 minutes. Reverse transcriptase was primed with random primers that had the suggested concentrations from the manual, provided from Applied Biosystems®.

Quantitative Polymerase Chain Reaction

Quantitative polymerase chain reactions were done to measure Ct values of reference genes after confirmation of the product using conventional PCR, and subsequent sequencing and alignment. Initial precautions were made to prevent contamination of reagents. Lab space used for PCR, RT-qPCR setup, RT-qPCR amplification, and analysis of PCR products using gel electrophoresis were kept separate from each other while all were maintained and cleaned with 10% bleach before and after use. Dedicated and clean lab coats were used during RT-qPCR experiment

setup, and nitrile powderless gloves were used, and often replaced, to prevent contamination. To prevent contamination of SYBR green Master Mix, it was aliquoted into separate Eppendorf tubes of 270 µL under a biosafety hood, and using sterilized and DNA-free Eppondorf tubes, filter tips, and pipette tips, treated with UV light for 30 minutes before using.

Quality Analysis

Relationships between the quality and the Ct values was measured with correlating Ct and Ratio values. The data was analyzed and Pearson correlation calculations were determined with the R statistics program.

CHAPTER 3. RESULTS

In Vitro Identified Primers

Literature Reviewed Primers

Each research article varied in number of reference genes tested for normalization averaging to G=12.22 (+/- 5.96553 SD) analyzed reference genes out of a total of 25 possible candidate reference genes. Using this systematic determination of candidate reference genes, the top three candidate reference genes predicted for future studies by weight score are: EF1-α, Actin, and Ubiquitin (see Table 3.1). As an example, calculation of weight score sees EF1-α (Eq. 2). The weight scores of potential reference genes of actin and ubiquitin were calculated to be 287 and 249, respectively. The reference gene ranked 12 in this calculation, Adenine phosphoribosyltransferase (APRT), obtained a weight score of 45, whereas the reference gene ranked 13 in the calculation, histone 3, obtained a weight score of 41.

Table 3.1 Proposed candidate reference genes for highest success in optimization of reference genes in *Arceuthobium americanum*.

Protein	**Rank**	**Description** [1,2]
Elongation Factor alpha	1	Transcription elongation [1]
Actin	2	Cytoskeleton component [2]
Ubiquitin	3	Protein-binding and modification [2]
Glyceraldehyde-3-phosphate dehydrogenase	4	Glycolysis-Gluconeogenesis [2]
Alpha Tubulin	5	Structural constituent of cytoskeleton [2]
SAND family protein	6	Membrane protein vesicle traffic and late endocytosis [1]
Beta Tubulin	7	Structural constituent of cytoskeleton [1]
Tonoplastic Iintrinsic protein 41-like family protein	8	Target of rapamycin signaling element [2]
Cyclophilin	9	Protein folding [2]
FBOX family protein	10	Putative role in protein degradation [2]
Protein phosphatase 2A subunit 3	11	Catalytic subunit Protein phosphatase 2A [2]
Adenine phosphoribosyltransferase	12	Purine metabolism [1]
Histone 3	13	DNA packaging[3]
Clathrin Adaptor Proteins	14	Vesicle carrier proteins[3]
Skip 16	15	F-Box family, ubiquitin ligase complex[3]
60s subunit	16	Large ribosomal subunit used in eukaryotic translation[3]
Cell division cycle related protein	17	Kinase related in regulation of cell division[3]
18s subunit	18	Small ribosomal subunit used in eukaryotic translation[3]
YLS8	19	mRNA splicing factor[3]
Zinc Finger Protein	20	Nucleases[3]
Glycine-rich RNA binding protein	21	tRNA production[3]
S-adenosylmethionine decarboxylase	22	Mechanism in cell division[3]
Tryptophan synthase Beta	23	Protein involved in tryptophan metabolism[3]
Sucrose non-fermenting-1 protein kinase	24	Carbon metabolism regulation[3]

DIM1	25	Essential protein for G2/M transition in cell cycle[3]

1- Exposito-Rodriguez et al. 2008.
2- Lilly et al. 2011.
3- Chandna et al. 2012.

Identified Primers from Bioinformatic Methods

Conserved regions of *A. oxycedri* PIP;2 gene were identified using BLASTp by translating JN857944.1 as the amphipathic channel position 324: CATATTAACCCGGCG 3' and position 681: 5'GGCACGAGCATCAACCCTGCTCGG 3' Asn-Pro-Ala motifs which are located between M3 and M7 helices. The translated PIP;2 gene was identified as the Major Intrinsic Protein (MIP) superfamily. Most specific amino acid hits were identified as other Plasma membrane Intrinsic Protein (PIP) proteins in the BLASTp algorithm. No additional significant conserved regions were detected using this method.

Many primers were considered using both Primer3 and Primer-BLAST web-based programs, each of which resolved about 20 viable, different, primer pairs with each program using various thermodynamic and salt correction calculations as discussed in the methods section. Table 3.2 represents of the top primer pairs as suggested by design program. Although all primer pairs that were developed are different, some primer pairs shared very similar sequences varying only by a base pair or two. Primer pairs 42-44 were the result of manual entering of either forward or reverse primer pair of conserved regions into Primer3 conditions as stated in methods section. Interestingly, Primer3 reported that the forward primer 42 3' end had high complementarity with itself of 5 base pairs whereas Beacon Designer projected no self-dimerization or hairpin loop products.

Ranking of candidate primer pairs of the best 10 designed primers of Primer3 and Primer-BLAST resulted in regions of the PIP;2 gene having lower ΔG for candidate primer pairs of either sense or anti-sense orientation (see Figure 2.5). In the region 5'-GTGGAGTGGGTCTTGTTCGTGCTTTCC-3' both forward and reverse primers had relatively low ΔG for secondary structures. Most primer pairs were significantly different save for primer pairs AoPIP3 and AoPIP4 which only varied by a single base pair in the 3' end of the forward primer. Both AoPIP3 and AoPIP4 had the same ranking due to similar ΔG for secondary structures, although they differed in annealing temperature by about 1°C in the forward primer: 61.23°C and 60.24°C for AoPIP3 and AoPIP4, respectively.

Table 3.2 Candidate primer pairs for *in vitro* experimentation of *Arceuthobium americanum* with given ranks. Higher rank indicates a lower ΔG for secondary structures (i.e., primer dimers).

Name	Rank	Primer Pair	Forward Primer	Reverse Primer	Program
AoPIP1	41	34	CCGACTCTAAACTCACGCCTT	AGCACGAACAAGACCCACTC	Primer-BLAST
AoPIP2	40	33	ACTCTAAACTCACGCCTTCCG	AAGCACGAACAAGACCCACT	Primer-BLAST
AoPIP3	38	19	ATGCTCATGGACGGCTACA	ACACTGCGAACCCAATAGGA	Primer3
AoPIP4	38	18	ATGCTCATGGACGGCTACAG	ACACTGCGAACCCAATAGGA	Primer3
AoPIP5	37	36	TCTAAACTCACGCCTTCCGAC	GGAAAGCACGAACAAGACCC	Primer-BLAST
AoPIP6	36	17	AGCCAACATGCTCATGGAC	CAATAGGAAGCGGTGCAAGT	Primer3
AoPIP7	35	29	GGGTCTTGTTCGTGCTTTCC	GTTTGGGATCGGTAGCGGAG	Primer-BLAST
AoPIP8	34	1	GTGGAGTGGGTCTTGTTCGT	AGCGGAGAAGACGGTGTAGA	Primer3
AoPIP9	32	14	ACAAAAGCCAGACCGACTCT	GAAACAATCCGAACGTCACC	Primer3
AoPIP10	32	13	ACCGTCATCGGCTACAAAAG	TCAGCTTCCGAGCTAGAAACA	Primer3

AoPIP11	Manual	42	GTCTACACCGTCTTCTCCGCT	ATGGCCCTACCCAGAAAATC	Primer3
AoPIP12	Manual	43	CGCCGAGATAATTGGAACTT	GAGCAGGGTTGATGCTCGT	Primer3
AoPIP13	Manual	44	GGCACGAGCATCAACCCT	GATGTGGGTGGTGACTCCAT	Primer3

Experimentation *In Vitro*

Testing all primers in Tables 2.2 and 2.3 *in vitro* during PCR and RT-qPCR showed only the aquaporin gene primers having any readings at a Ct value. All other primers that were developed using these methods were only found to produce either primer dimers, or Ct values of very high levels (i.e., 35 or higher).

Analysis of RNA Extraction

The average yield of RNA from Dwarf Mistletoe samples was 100 ng/µL ± 42 ng/µL of purified RNA. There was no significant correlation between weight of sample and extraction yield shown, see Figure 3.1 (r=-0.0579, p=0.6866, n=52).

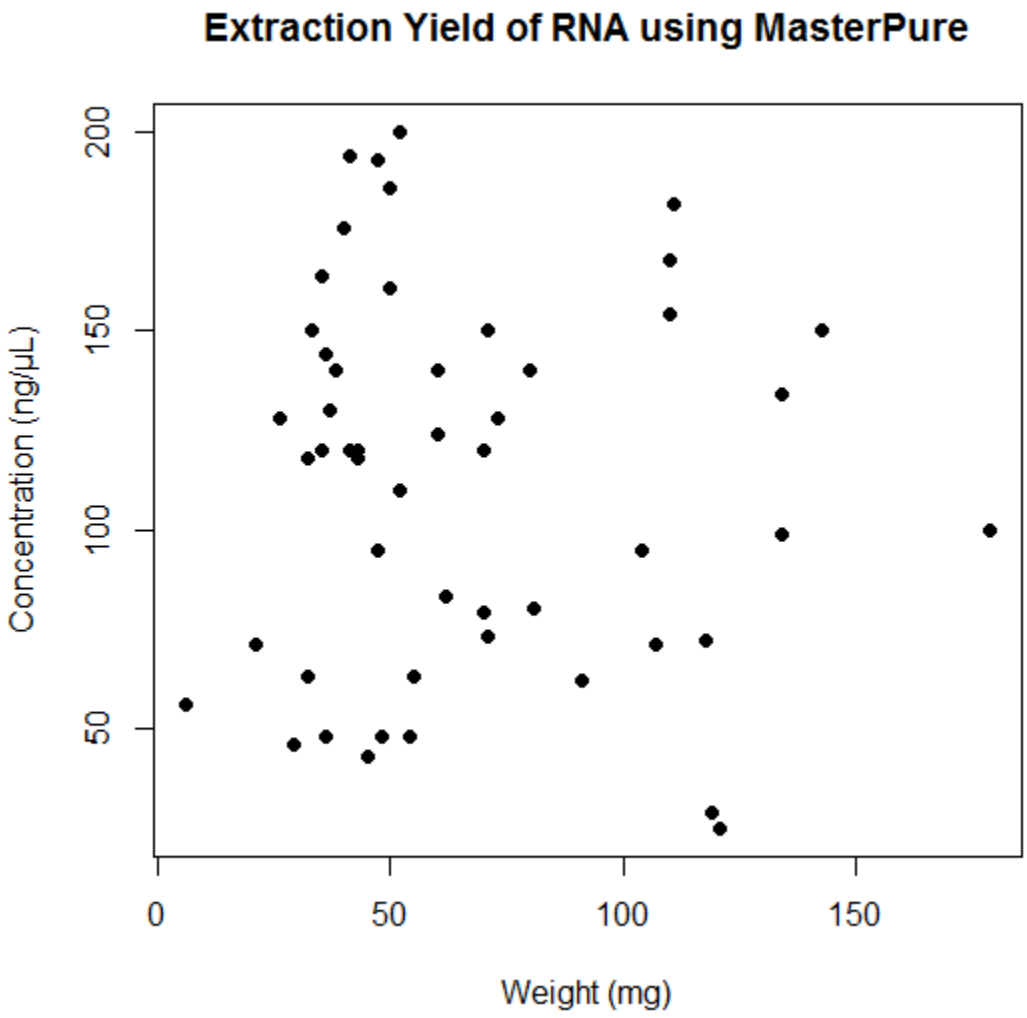

Figure 3.1 Scatter plot of RNA extractions of Dwarf Mistletoe fruit tissue samples using MasturePure™ purification kit.

Bleach-agarose gel electrophoresis shows distinctive bands between 28S and 18S ribosomal RNA bands in intact RNA samples under a UV light. Approximately 10% of the RNA samples showed degradation of ribosomal bands, see Figure 3.2. Samples which were determined to be degraded were not used in cDNA synthesis.

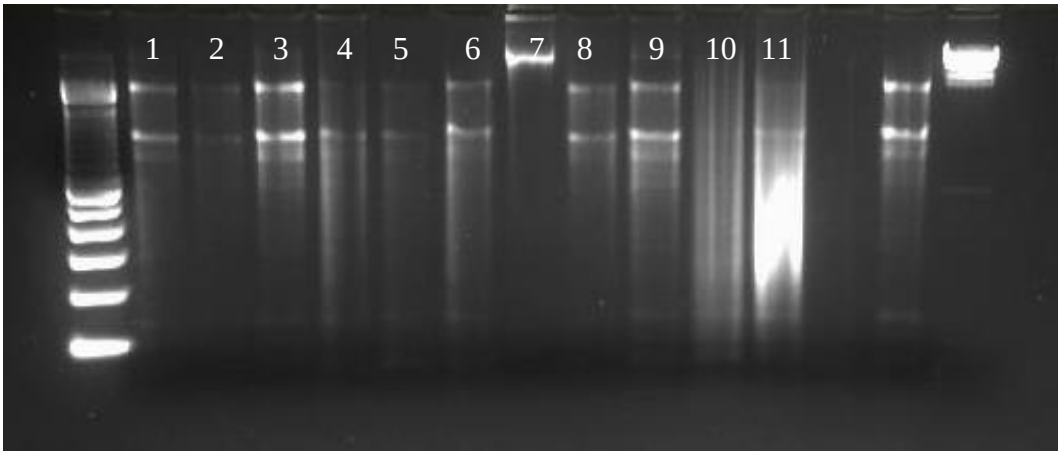

Figure 3.2 Bleach gel electrophoresis of RNA samples. Lanes 10 and 11 show clear signs of degradation.

Qualification of RNA was also done by measuring the intensity of the ribosomal RNA bands with samples showing non-degraded RNA and measuring the absorbance ratios at 260 nm and 280 nm. Intensity ratios between 18S averaged at 1.38 (SD = 0.42, min = 0.6, max = 2.1) when run for 30 minutes on a gel; 28S ribosomal band underwent decay as time progressed so all gel electrophoresis experiments were done with consistent conditions. On average, absorbance ratios measured at 260 nm and 280 nm calculated 1.96 (SD=0.17, min = 1.6, max = 2.2).

Average intensity ratios have changed over time. As shown in Figure 3.3, samples from earlier years tend to have a lower intensity ratio than samples from more recent years. Furthermore, intensity ratios were taken at 30 and 45 mins gel runtime to show any changes in intensities. On average, intensity ratios measured 1.09 (SD = 0.36) at 30 mins and 0.78 (SD = 0.17) at 45 mins.

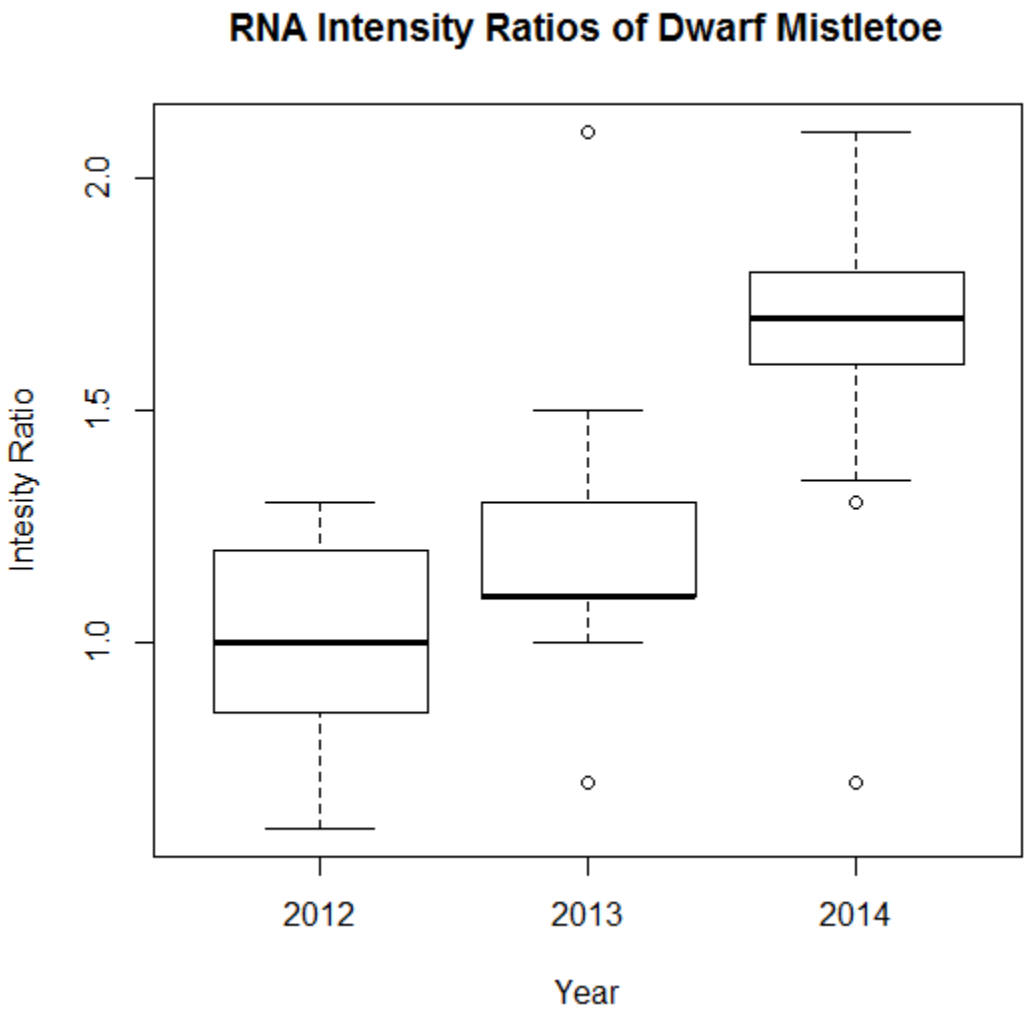

Figure 3.3 Box plot of 18S and 28S ribosomal RNA intensity ratios when viewed under UV light. Samples are grouped into the year they were sampled and extracted. Intensity ratios were taken in July 2015 to show changes, if any, in storage.

An Analysis of Variance (ANOVA) was performed on the threshold values in RT-qPCR to determine if there was a statistical difference between values. The hypothesis in this case was that each Ct value was statistically similar to each other. In the analysis, our alpha value becomes 0.05 and the *F* critical value is 2.65. According to the analysis, we reject the null hypothesis, and

accept that there is some statistical difference between the points within the Ct values of the Aquaporin gene.

Bestkeeper Analysis

Bestkeeper is an Excel macro that analyzes Ct input directly from RT-qPCR and compares the standard deviation between the points and uses a regression analysis, see Table 3.4 (Pfaffl et. al. 2004).

Table 3.4 Bestkeeper Analysis of Aquaporin Ct values showing the coefficient of correction, slope, *P* value, and power to determine stabilization in RT-qPCR.

Coeff of corr	0.43
slope	1.12
P value	0.147
Power	2.18

CHAPTER 4. DISCUSSION

Literature Review

The stability of reference genes is unlikely to be exactly predicted for RT-qPCR. This is especially true when considering the study done by Czechowski et al. (2005) with *A. thaliana*. In that study, EF1-α was shown to be one of the least stable candidate reference genes, exceeding the appropriate stability threshold of M=1.5; however, my paper here has predicted that EF1-α would be the most stable reference gene in *A. americanum* expression analysis – simply due to EF1-α being used more often. Systematic selection of reference genes does not guarantee that most stable reference genes can be predicted, because organic systems are immensely complex;

however, it provides a starting point for any normalization experiment for expression analysis, de novo.

Furthermore, since parasitic flowering plants like *A. americanum* are still not well known genetically, they may have housekeeping genes that are expressed differently from other plants or model organisms. Parasitic flowering plants such as *A. americanum* supplement some nutrition from their host, so their housekeeping genes might be differently expressed during developmental stages. Additionally, novel proteins that are not yet identified may be more stable than the genes presented here, as shown in Czechowski et al. (2005). Inclusion of novel proteins would require additional genetic work to identify in RT-qPCR reactions.

It is also important to note the high standard deviation of ~6 to the 12 average reference genes used in the 57 research articles. More literature could be looked at to increase the number of studies that used reference genes for expression analysis, including reference genes used in microarrays (Lilly et al. 2011). The ranges of the number of reference genes used in literature studies were as low as 7 to as high as 50 in each study.

Non-model Organisms

Dwarf Mistletoe is a non-model organism, meaning that it is not often used in experimentation to answer scientific question. Therefore, there is limited genomic data published. Extra steps must then be taken into consideration when coordinating projects. For certain organisms, the best strategy may be total sequencing and annotation of genes to then develop primers, based on the goal of the experimentation, and if reference genes are required.

Certain organisms, such as viruses, have small genomes which are very easy to sequence cheaply and quickly. Plants have much larger genomes, for example *Arabidopsis thaliana* has approximately 135 Mbp according to The Arabidopsis Information Resource, where the virus COVID-19 has 30kb (Li et al. 2020). The genome, number of genes, and base pairs in Dwarf Mistletoe is not currently published. Therefore, with careful planning, primer development in complicated genomes, like those in plants, can be cost-effective.

The gene expression studies of non-model organisms, like Dwarf Mistletoe, are useful when attempting to better understand other plants with similar qualities uncommon in the plant world, such as hemi-parasitism. Therefore, the efforts needed to develop a gene expression analysis are still justified when considering the knowledge that could be gained from such a study.

Relationship to *Arceuthobium*

It remains unclear, how distantly related plants must be to develop primers. However, *Arceuthobium oxecedri* seems to be at sufficient similarity to produce viable primers to *A. americanum*. If genetic sequences exist for other plants in the same genus, it is more likely to have successful primers developed from that sequence than others (Radomski et al. 2010).

Computational Analysis

The primer pairs AoPIP11, AoPIP12 and AoPIP13, take advantage of the conserved regions of the aquaporin gene (see Figure 2.4). The level of conservation of the sequence between *A. americanum* and *A. oxycedri* in this particular transcript is unknown; therefore, picking a highly conserved region such as this may pose issues with complementarity between other similarly-structured MIP genes. In a sensitive quantitative assay such as RT-qPCR, this may give false

quantitation of PIP;2 by binding to other MIP class genes. Preliminary experiments, which are the norm for RT- RT-qPCR assays, to test primers should show multiple bands, if unspecific binding is occurring during conventional PCR and subsequent gel electrophoresis. Furthermore, if a single amplicon is produced, sequencing can be done to ensure the correct product is achieved with the primer pair. Alternatively, if the PIP;2 transcripts of *A. americanum* and *A. oxycedri* are quite different, then these conserved sequences will yield favourable locations for primer binding, given there is specificity to similar sequences elsewhere in the transcriptome.

The presence of highly similar primer pairs, such as AoPIP3 and AoPIP4, are likely due to limited primer pairs fitting the set conditions in a run. Since each run was done in regions of the gene (i.e., positions: 150-450, 450-750, etc.), it leaves little room for a vast array of 'the best' primer pairs of the gene. Alternatively, the primer-designing algorithm may not 'skip over' outputting similar sequences of primer pairs, even if they only vary by one or two base pairs. A simple way to account for such similar sequences is to increase the primer pair output maximum to a larger number, above the expected number of primer pairs for a given region, and selectively discard extreme similarities.

Primer Development Methodology

Methodology has changed little for primer development. *In silico* development remains an important first step in quantitative analysis using RT-qPCR. Screening of primers can increase efficiency of the wet lab results, which can then optimize primers with additional experimentation, as done in this thesis. The optimization process refers to optimizing the conditions in which the RT-qPCR machine is used (i.e., temperature, concentration, and primer combination) (Bustin and Hugget 2017).

Developed primers are published to a database. To keep consistency and the spirit of open information in the scientific community, primers are added to a larger database so that they can be searched and further studied, if needed. These databases include PrimerBank, qPrimerDB, and PCRdrive, among others (Lu et al. 2018, Burger et al. 2018).

Experimentation

Laboratory Limitations

The inability to grow Dwarf Mistletoe in the lab limits the control over a number of variables during experimentation. In addition, it is highly impractical to collect and extract samples in a similar fashion to other current model organisms. Since this experiment used a method of haphazard collection techniques, it can be used as an indicator for some viability in this method of collection.

Molecular inhibitors are ever present in our world, and even when the utmost care is taken to reduce the presence of these objects. Molecular inhibitors include RNAses and other proteins that may contaminate isolated RNA samples. The inhibitors prevent any one of the multiple steps of polymerase chain reaction, affecting the results of RT-qPCR. It is worth mentioning that there is a possibility that these inhibitors affected RT-qPCR based experimentation. However, quality procedures in this study follow industry qPCR standards set by Bustin et al. 2009.

Specialized sequencing and identification of genes can be done; however, it would be more useful to do these on a metagenomic basis. If done at both levels, it would give a more complete view of the gene mechanism to compare with others and, the greatest confidence that the code

being quantified is the target. At the time of this thesis release, sequencing data can be affordable and available to most laboratories.

In Vitro and *In Vivo*

Maximizing *in vitro* experimentation is important with organisms like Dwarf Mistletoe as it is currently not possible to grow in the lab. This means that samples are limited as they must be collected in the field. As a result, only microliters of samples may be available for experimentation, which might result in insufficient quantities of material for research. Careful planning, in addition to optimizing techniques could reduce this risk.

REFERENCES

Aranda R, Dineen S, Craig R, Guerrieri R, and Robertson J. 2009 Comparison and Evaluation of RNA Quantification Methods Using Viral, Prokaryotic, and Eukaryotic RNA Over a 104 Concentration Range. Analytical Biochemistry. 387: 122-127.

Becker C, Hammerle-Fickinger A, Riedmaier I, and Pfaffl M. 2010 mRNA and microRNA Quality Control for RT-qPCR analysis. Methods. 50(4): 237-43.

deBruyn RAJ, Paetkau M, Ross KA, Godfrey DV, and Ross Friedman C. 2015 Thermogenesis-triggered seed dispersal in Dwarf Mistletoe, Arceuthobium americanum. Nature Communications [Internet]. [cited 2016 Jan 20]; 6, 6262. Available from: https://www.nature.com/articles/ncomms7262 doi:10.1038/ncomms7262.

Broshot L, Larsen L, Tinnin R. 1986. Effects of Arceuthobium americanum on twig growth of Pinus contorta. U.S. Dept. of Agriculture, Forest Service, Pacific Northwest Research Station. Research Note: PNW-RN-453.

Burger F, Angioni M, Russo G, Schad M, Kallarackal J. 2018. PCRdrive: the largest qPCR assay archive to date and endless potential for lab workflow revitalization. BMC Bioinformatics [Internet]. [cited 2020 Nov 11]; 19(447). Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2452-3 doi: https://doi.org/10.1186/s12859-018-2452-3.

Bustin SA, Benes V, Garson JA, Hellemans J, Hugget J, Kubista M, Mueller R, Nolan T, Pfaffl M, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQUE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical Chemistry [Internet]. [cited 2013 Sep 19]; 55(4):611-622 Available from:

https://academic.oup.com/clinchem/article/55/4/611/5631762 doi: 10.1373/clinchem.2008.112797.

Bustin SA, Benes V, Nolan T, and Pfaffl MW. 2005. Quantitative Real-time RT-PCR – a Perspective. J. Mol. Endocrinol. 34(3): 597-601.

Bustin SA, Huggett J. 2017. qPCR primer design revisited. Biomol Detect. Quantif. 14:19-28.

Caixeiro NJ, Lai K, Lee CS. 2016. Quality Assessment and Preservation of RNA from Biobank Tissue Specimens: A Systematic Review. J Clin Pathol [Internet]. [cited 2020 Sep 13]; 69(3): 260-5. Available from: https://jcp.bmj.com/content/69/3/260.long doi: 10.1136/jclinpath-2015-203384.

Chandna R, Augustine R, Bisht NC. 2012. Evaluation of Candidate Reference Genes for Gene Expression Normalization in *Brassica juncea* Using Real Time Quantitative RT-PCR. PLOS ONE [Internet]. [cited 2013 Sep 13]; 7(5):e36918 Available from: https://doi.org/10.1371/journal.pone.0036918 doi: 10.1371/journal.pone.0036918.

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. 2005. Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. Plant Physiology [Internet]. [cited 2013 Sep 13]; 139: 5-17 Available from: http://www.plantphysiol.org/content/139/1/5.full doi: http://dx.doi.org/10.1104/pp.105.063743.

Exposito-Rodriguez M, Borges AA, Borges-Perez A, Perez JA. 2008. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biology [Internet]. [cited 2013 Dec 19]; 8:131. Available from: https://bmcplantbiol.biomedcentral.com/articles/10.1186/1471-2229-8-131 doi: 10.1186/1471-2229-8-131.

Fleige S, Pflaffl MW. 2006. RNA Integrity and the Effect on the Real-time qRT-PCR Performance. Molecular Aspects of Medicine. 27: 126–139.

Geer L, Domrachev M, Lipman D, Bryant SH. 2002. CDART: protein homology by domain architecture. Genome Res [Internet]. [cited 2013 Dec 11]; 12(10): 1619-1623. Available from: https://genome.cshlp.org/content/12/10/1619.long doi: 10.1101/gr.278202.

Hawksworth FG, Hinds TE, Johnson DW, and Landis TD. 1977. Silvicultural control of Dwarf Mistletoe in young Lodgepole Pine stands. For. Ins. and Dis. Manage. Tech. Rep. R2-10. Colorado (CO): U.S. Department of Agriculture, Forest Service, Rocky Mountain Region. 12 p.

Hawksworth FG, Wiens D. 1996. Dwarf Mistletoes: Biology, Pathology, and Systematics: US Department of Agriculture Forest Service Handbook No 709. Washington (WA): US Government Printing Office.

Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, Eveno E, Mueller O, Schroeder A, Auffray C. 2005. Toward Standardization of RNA quality assessment using user-dependant classifiers of microcapillary electrophoresis traces. Nucleic Acids Research [Internet]. [cited 2013 Sep 13] 33(6):e56. Available from: https://academic.oup.com/nar/article/33/6/e56/2401377 doi: 10.1093/nar/gni054.

Ip DW. 1992. Dwarf Mistletoe. Canadian Forest Publications. Forestry Canada, Northwest Region, Northern Forestry Centre, Edmonton, Alberta. Forestry Leaflet 18(11807).

Ishikawa H. 1977. Evolution of Ribosomal RNA. Comp. Biochem. Physiol. B58: 1-7.

Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23(10): 1289-91.

Kumar V, Sharma R, Trivedi PC, Vyas G, Khandewal V. 2011. Traditional and novel references towards systematic normalization of qRT-PCR data in plants. AJCS. 5(11):1455-1468.

Li D, Zhang J, Li J. 2020. Primer design for quantitative real-time PCR for the emerging Coronavirus SARS-CoV-2. Theranostics [Internet]. [cited 2020 Nov 11]; 10(16):7150-7162. Available from: https://www.thno.org/v10p7150.htm doi:10.7150/thno.47649.

Lilly S, Drummond R, Pearson M, and MacDairmid R. 2011. Identification and Validation of Reference Genes for Normalization of Transcripts from Virus-Infected Arabidopsis thaliana. Mol Plant Microbe Interact [Internet]. [cited 2013 Sep 13]; 24(3):294-304. Available from: https://apsjournals.apsnet.org/doi/10.1094/MPMI-10-10-0236?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed doi:10.1094/MPMI -10-10-0236.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25(4):402-8.

Lu K, Li T, He J, Chang W, Zhang R, Liu M, Yu M, Fan Y, Ma J, Sun W, et al. 2018. qPrimerDB: a thermodynamics-based gene-specific qPCR primer database for 147 organisms. Nucleic Acids Research [Internet]. [cited 2020 Nov 11]; 46(D1): D1229–D1236. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5753361/ doi https://doi.org/10.1093/nar/gkx725.

Maurel C, Javot H, Lauvergeat V, Gerbeau P, Taournaire C, Santoni V, and Heyes J. 2002. Molecular physiology of aquaporins in plants. Int Rev Cytol. 215: 105–148.

Nickrent DL, Miguel GA, Martin P, and Mathiasen R. 2004. A Phylogeny of all Species of Arceuthobium (Viscaceae) Using Nuclear and Chloroplast DNA Sequences. American Journal of Botany. 91(1): 125-138.

Pflaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res [Internet]. [cited 2013 Oct 10]; 29(9):e45. Available from: https://academic.oup.com/nar/article/29/9/e45/2384081 doi: 10.1093/nar/29.9.e45.

Pfaffl, MW, Tichopad A, Prgomet C, Neuvians TP. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. Biotechnology Letters [Internet]. [cited 2013 Oct 10]; 26: 509–515. Available from: https://link.springer.com/article/10.1023/B:BILE.0000019559.84305.47 doi: 10.1023/B:BILE.0000019559.84305.47.

Radomski N, Lucas F, Moilleron R, Cambau E, Haenn S, Moulin L. 2010. Development of a Real-Time qPCR Method for Detection and Enumeration of Mycobacterium spp. in Surface Water. Applied and Environmental Microbiology [Internet]. [cited 2014 May 3]; 76(21): 7348-7351. Available from: https://aem.asm.org/content/76/21/7348 doi: 10.1128/AEM.00942-10.

Ross CM. 2006. Viscin cells in the Dwarf Mistletoe Arceuthobium americanum – "green springs" with potential roles in explosive seed discharge and seed adhesion. Davidsonia 17(3): 75-86.

Ross Friedman CM, Ross BN, Martens GD. 2009. Antibodies raised against tobacco aquaporins of the PIP;2 class label the viscin tissue of the explosive Dwarf Mistletoe fruit. Plant Biology 12: 229-233.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for General Users and for Biologist Programmers. Methods in Molecular Biology [Internet]. [cited 2013 Sep 13]; 132: 365-86. Available from: https://link.springer.com/protocol/10.1385%2F1-59259-192-2%3A365 doi: 10.1385/1-59259-192-2:365.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. 1988. Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. Science [Internet]. [cited 2013 Sept 10]; 239(4839):487-9. Available from: https://science.sciencemag.org/content/239/4839/487.long doi: 10.1126/science.2448875.

Salzman RA, Fujita T, Zhu-Salzman K, Hasegawa PM, Bressan RA. 1999. An Improved RNA Isolation Method for Plant Tissues Containing High Levels of Phenolic Compounds or Carbohydrates. Plant. Mol. Biol. Reptr. 17: 11-17.

Shaw DC. Watson DM, Mathiasen RL. 2004. Comparison of Dwarf Mistletoe (Arceuthobium. spp., Viscaceae) in the Western United States with Mistletoes (Amyema spp., Loranthaceae) in Australia – Ecological Analogs and Reciprocal Models for System Management. Australian Journal of Botany. 2004(52): 481-498.

Singh A, Pandey GK. 2015. Primer Design Using Primer Express® for SYBR Green-based Quantitative PCR. Methods Mol Bio [Internet]. [cited 2013 Sep 10]; 1275:153-64. Available from: https://link.springer.com/protocol/10.1007%2F978-1-4939-2365-6_11 doi: 10.1007/978-1-4939-2365-6_11.

Smith C, Lieberman M, Marks A, 2003. Mark's Basic Medical Biology A Clinical Approach, 2ND EDITION. Wiliams & Wilkins. 306 p.

Stellwagen N. 2009. Electrophoresis of DNA in Agarose Gels, Polyacrylamide Gels, and in Free Solution. Electrophoresis. 1: S188-S195.

Suzuki T, Higgins P, Crawford D. 2000. Control Selection for RNA Quantification. Biotechniques. 29:332-337.

Tichopad A, Didier A, and Pfaffl MW. 2004. Inhibition of Real-time RT-PCR Quantification Due to Tissue-specific Contaminants. Molecular Cellular Probes [Internet]. [cited 2013 Sep 20]; 18(1): 45-50 Available from:

https://www.sciencedirect.com/science/article/abs/pii/S0890850803000823?via%3Dihub doi: 10.1016/j.mcp.2003.09.001.

Udvardi MK, Czechowski T, Scheible WR. 2008. Eleven Golden Rules of Quantitative RT-PCR. Plant Cell letter [Internet]. [cited 2013 Oct 10] Available from: http://www.plantcell.org/content/20/7/1736 doi: 10.1105/tpc. 108.061143.

Unger LS. 1992. Dwarf Mistletoes. Forestry Canada, Forest Insect and Disease Survey Leaflet No. 33. 7p.

Untergrasser A, Cutcutache I, Kaµressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Research [Internet]. [cited 2014 May 10]; 40(15):e115. Available from: https://academic.oup.com/nar/article/40/15/e115/1223759 doi: 10.1093/nar/gks596.

USDA, NRCS [Internet]. c2021. Greensboro, NC: National Plant Data Team; [cited 2021 Mar 24]. Available from: http://plants.usda.gov/.

Vandesompele J, dePreter K, Pattyn F, Poppe B, Van Roy N, dePaepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol [Internet]. [cited 2013 Sep 10]; 37: RESEARCH0034. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/gb-2002-3-7-research0034 doi: https://doi.org/10.1186/gb-2002-3-7-research0034.

deVega-Bartol JJ, Santos RR, Simoes M, Miguel C. 2013. Normalizing gene expression by quantitative PCR during somatic embryogenesis in two representative conifer species: Pinus pinaster and Picea abies. Plant Cell Rep [Internet]. [cited 2013 Oct 10]; 32:715–729. Available from: https://link.springer.com/article/10.1007/s00299-013-1407-4 doi: https://doi.org/10.1007/s00299-013-1407-4.

Wang Y, Geer LY, Chappey C, Kans JA, Bryant SH. 2000.Cn3D: sequence and structure views for Entrez. Trends Biochem Sci [Internet]. [cited 2014 Mar 20]; 25(6): 300-2. Available from: https://www.cell.com/trends/biochemical-sciences/fulltext/S0968-0004(00)01561-9?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0968000400015619%3Fshowall%3Dtrue doi: 10.1016/s0968-0004(00)01561-9.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 13:134.