An Intrageneric and Intraspecific Study of Morphological and Genetic Variation in the Neotropical *Compsoneura* and *Virola* (Myristicaceae)

by

Royce Allan David Steeves

A Thesis

Presented to

The University of Guelph

In partial fulfillment of requirements

for the degree of

Doctor of Philosophy

in

Botany

Guelph, Ontario, Canada

ABSTRACT

AN INTRAGENERIC AND INTRASPECIFIC STUDY OF MORPHOLOGICAL AND GENETIC VARIATION IN THE NEOTROPICAL *COMPSONEURA* AND *VIROLA* (MYRISTICACEAE)

Royce Allan David Steeves

University of Guelph, 2011

Advisor:

Dr. Steven G. Newmaster

The Myristicaceae, or nutmeg family, consists of 21 genera and about 500 species of dioecious canopy to sub canopy trees that are distributed worldwide in tropical rainforests. The Myristicaceae are of considerable ecological and ethnobotanical significance as they are important food for many animals and are harvested by humans for timber, spices, dart/arrow poison, medicine, and a hallucinogenic snuff employed in medico-religious ceremonies.

Despite the importance of the Myristicaceae throughout the wet tropics, our taxonomic knowledge of these trees is primarily based on the last revision of the five neotropical genera completed in 1937. The objective of this thesis was to perform a molecular and morphological study of the neotropical genera *Compsoneura* and *Virola*. To this end, I generated phylogenetic hypotheses, surveyed morphological and genetic

diversity of focal species, and tested the ability of DNA barcodes to distinguish species of wild nutmegs.

Morphological and molecular analyses of *Compsoneura.* indicate a deep divergence between two monophyletic clades corresponding to informal sections Hadrocarpa and Compsoneura. Although 23 loci were tested for DNA variability, only the trnH-psbA intergenic spacer contained enough variation to delimit 11 of 13 species sequenced. A morphological and molecular investigation of *Compsoneura capitellata* showed little discrete morphological variation among populations but significant genetic structure among populations.

Phylogenetic analysis of *Virola* also revealed a deep molecular divergence between two clades having numerous contrasting morphologies. In contrast to *Compsoneura*, the trnH-psbA intergenic spacer failed to differentiate the majority of *Virola* species tested. An infraspecific morphological and molecular study of *V. sebifera* and *V. loretensis* showed that each of these species contains morphologically and ecologically discrete sympatric morphotypes that likely represent new species.

In total, this investigation found 5 provisional new species from fewer than 600 collections at biological stations in Ecuador and Peru where these new species were among the most abundant trees in the forest. This suggests that much diversity likely remains to be described in the Myristicaceae and other tropical plant families.

ACKNOWLEDGEMENTS

This thesis wouldn't have been possible without the help of many people. Firstly, I'd like to thank my supervisor, Steve Newmaster, for taking me on 5 years ago as a new masters student, pardon the pun. Your patience and infectious botanical curiosity has been an inspiration to me. Aron Fazekas, thanks for your unrelenting patience with me in the lab and for always having time for those "quick" questions that turn into a lenghty discussion. As for jungle-man John Janovec, thanks for all the great times in the field, your help in getting permits, access to BRIT's plant and human resources and our late-night *Compsoneura* campfire discussions.

I'm also very grateful to all the members of my committee Dr's Rob Hanner, Usher Posluszny, and David Galbraith for your valuable input and guidance throughout these years.

Carol-Anne Lacroix and Ragupathy Sabrumanyam; thanks for sharing your taxonomic and ethnobotanical expertise over the years. I'd also like to thank the following labmates for their input, thoughtfull discussions, help in procrastinating, and most importantly their friendship: Heather Cole, Jillian Bainard, John Gerrath, Troy Mcmullin, Kevin Berg, Kevin Burgess, Jose Maloles, Evan Lusty and Brian Lacey as well as all the other folks that have worked in the herbarium over the years. An extra thanks to Jose and Brian for their help with measuring specimens and preparing figures.

This work couldn't have been completed without all the dedicated field botanists that I've had the pleasure of working with. Derek Piper and Andrea Reid: thanks for your patience with me in the tropics. I know I'm not the easiest person to get along with in the tropical heat while climbing a tree and trying to get the pruning poles together. I also got

to work with some great parataxonomists in my 4 months in the field: Danilo Simba Larco (Ecaudor), Wilson Quizhpe (Ecuador), Angel Balarezo (Peru-BRIT), Pedro Centeno (Peru-BRIT), Benjamin Chambi (Peru-BRIT). Thank you all for the many laughs in the field, your unspeakable tree-climbing/pruning abilities and all your patience with a neotropcial neophytologist like myself. Funding for my research and travel were provided in part byby NSERC, Amazon Conservation Association, and the Latournelle travel grant. David Neil, Renan Valega, Keri McNew, and Tiana Franklin made this research possible by coordinating collection and export permits.

I am deeply indebted to innumerable people that have helped me troubleshoot DNA extraction, PCR, sequencing and other molecular methods over the past few years; namely I'd like to thank the Husband and Crease labs for sharing equipment and Jeff and Angela at the Genomics Facility for their expertise. I'd also like to take the opportunity to my fellow graduate students for their companionship and numerous beer-fueled discussions at the grad lounge. To my girlfriend friend Sarah Baldwin: thank for your love and patience through this whole process, thanks for our long-winded plant genetic discussions and thanks for being you.

Last but not least I'd like to thank my parents, Bob and Diane, for their unrelenting love and support throughout my schooling. Our innumerable forays into the woods and streams of New Brunswick to go hunting, fishing, trapping, berry and fiddlehead picking have spawned a curiosity in nature that I hope will only grow stronger. This thesis is dedicated to you.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

Taxonomy

Ethnotaxonomy

It is likely that the first profession was that of the shaman; the healer, guide to the realm of spirits and maker of medicines. The need for this profession was presumably born of the inherent difficulty and serious nature of plant identification. The shaman may have varying roles in numerous cultures but in most embodiments they are the authority on the use and implementation of the medicinal and religious application of plants. In many cultures these tasks are regarded with such great importance that individuals within traditional cultures may be relieved of the tasks of everyday life from a young age to devote their whole lives to the practice of identifying and implementing plants in medico-religious applications. The act and practice of naming and differentiating species, variants, or cultivars of plants often holds great cultural importance to traditional peoples likely because it has developed from the need to identify the nutritious and intoxicating plants available to them.

Humans have an inherent ability to recognize and classify the diversity of life around them (Newmaster et al. 2006, Diamond and Bishop 1999). Although many species of animals distinguish edible and inedible species from one another based on senses and/or innate aversions to particular foodstuffs, language and tradition have afforded humans the ability to pass on this knowledge to subsequent generations. A unifying trait among human cultures is the development of systems to name and describe the biodiversity in their realm. Although vernacular names may be different, it is interesting that a population of individuals with similar phenotypes is often recognized as

a ‘species’ both in aboriginal cultures and the scientific, or Linnaean, system and aboriginal classifications sometimes recognize more ‘species’ than current scientific nomenclature (Diamond and Bishop 1999).

Taxonomy and Species

Taxonomy is the scientific practice of classifying, identifying, and describing organisms. Although taxonomists are concerned with all organismal levels of classification, but of particular importance is the determination, discovery and description of species. Traditionally, plant taxonomists have used a wide range of characters to help them classify earth’s botanical diversity. Until relatively recently, botanists were largely restricted to studies of gross morphology, anatomy and inter-fertility to classify diversity. Additional characters became available with the use of secondary metabolite analysis and with the advent of chromosome staining botanists have been able to determine and consider chromosome number as a taxonomic character . More recently flow cytometry, allozyme analysis, and DNA-DNA hybridization gave taxonomists more sensitive means of delimiting taxa. However, it was the advent of the polymerase chain reaction (PCR) by Kary Mullis (Mullis and Faloona 1987) that enabled the determination of DNA sequences in a quick, accurate and inexpensive manner, that has revolutionized the field of systematics as we know it today. The additional information afforded by DNA sequence analysis has not only given taxonomists a more characters to help classify taxa, but has allowed systematists to estimate their evolutionary histories with data that is independent of morphology.

There exists well over 20 species concepts in use today however, many biologists will agree that species are lineages of separately evolving populations (de Quieroz 2005).

Species concepts with respect to plants have troubled botanists as phenotypic plasticity, polyploidy, hybridization, and differing methods of reproduction and genetic inheritance makes the standardization of criteria for species recognition difficult and have led some to suggest that plant species do not exist as real entities like animal species (Levin 1979, Bachmann 1998). In spite of these assertions, a recent analysis of crossing data in more than 400 plant and animal genera found that 70% of plant species correspond to reproductively isolated lineages, compared to only 39% of animal species tested (Rieseberg et al. 2006). Although hybridization is known to occur frequently in plants (Golden and Bain 2000, Palme et al. 2003), and is expected to be an important in the generation of new species (Whitney et al 2010), there is emerging evidence that reproductive isolation, genetic and morphological integrity can be maintained in the face of hybridization (Palma-Silva et al. 2011). In a recent study of four morphologically distinct sympatric neotropical Bromeliads Palma-Silva and colleagues (2011) found that 7 haplotypes were shared among species, however, nuclear microsatellite markers showed significant differentiation between species. This differential permeability of chloroplast and nuclear genomes between freely hybridizing species is of prime interest to plant taxonomists as it demonstrates that any given species can largely maintain its nuclear genetic integrity, and therefore morphological integrity despite the capture, or introgression, or another species' chloroplast haplotype. Plant systematists must consider these data as it may result in strong chloroplast versus nuclear gene tree incongruence (Rieseberg and Soltis 1991), and there exists the strong possibility that different loci of the nuclear genome (such as neutral versus non-neutral loci) may exhibit differential permeability between species (i.e. asymmetrical introgression).

Species are the “measuring-stick” with which biologists measure biodiversity. Virtually all major fields of biology rely upon accurate species identification in order to test scientific hypotheses. Therefore, the power of biological inference relies largely upon the accuracy of the measuring stick used, as well as the skill with which it is applied. However, many taxa lack a modern systematic revision, many areas of the world are uncollected, many investigators lack the taxonomic expertise necessary to identify species of interest and most of the world biodiversity has yet to be described (Mora et al 2010). Additionally, taxonomic experts, if there are any for a particular group, are sometimes too occupied with their own studies to perform routine identifications for others. Also, barriers such as language, funding constraints, and manuscript availability may further hinder access to taxonomic information. The specimens themselves may be difficult to identify due to a lack of distinguishing features (such as fertile structures), specimen age, quality and quantity of material. This combination of factors often leads to numerous errors in the course of many investigations where the taxonomy of a group of species is not well studied. Despite its fundamental importance to biological sciences, the field of taxonomy has been in decline in recent years, with a reduction in both the number of investigators as well as financial support (Godfray 2002).

DNA Barcoding

A potential improvement for some of these problems is the implementation of DNA-based identification systems such as DNA barcoding. DNA barcoding is the use of one or several standardized genomic regions as the basis for identification of a species (Hebert et al. 2003a,b). The mitochondrial coding gene cytochrome oxidase I has proven useful in numerous animal groups and algae (Hebert et al. 2004, Hebert et al. 2006,

Saunders 2005) and various barcoding campaigns are quickly building sequence libraries for the worlds flora and fauna (iBOL 2010). Once DNA region(s) suitable for delimitation of species has been found, agreed upon, and a sequence library has been compiled, DNA-based identification systems can be potentially revolutionary for species identification since assignments can be performed on specimens which may be degraded, lacking critical morphological characters, or from small or otherwise unidentifiable pieces of tissue. The ability to identify species from material lacking morphologically differentiating characters represents a drastic transformation in the types of ecological questions that can be addressed. Additional benefits are that it is rapid, largely automatable, accessible to anyone with access to a thermocycler and sequencing equipment (although this could potentially be miniaturized for public use). This automation may ultimately relieve taxonomic experts of the task of routine identifications, giving them more time to pursue questions of more theoretical nature.

Progress towards a universal plant DNA barcode has been largely impeded by the comparatively slow and heterogeneous rate of plastid nucleotide mutation in plants, the desire for universal primers, poor sequence quality in the most discriminating plastid loci (Chase et al. 2007, Fazekas et al. 2009), and the low rates of species discrimination of a single locus (Hollingsworth et al. 2009). Recently, plant taxonomists have proposed the multi-locus combination of plastid coding regions *mat*K *and rbc*L as the proposed loci for a global campaign to barcode land plants (Hollingsworth et al. 2009). These more conserved loci could serve as core regions to identify samples to higher ranks (family and genus) while more variable markers (plastid spacers, nrDNA, and low copy nuclear loci)

will be needed for the identification of some taxa. DNA bardcoding will undoubtedly be more difficult in plants as they are prone to interspecific hybridization, allopolyploidy, introgression, and many species are of relatively recent origin. The asymmetrical introgression of one species' chloroplast genome into another species, also known as chloroplast capture, is well documented by phylogenetic, phylogeographic, and population genetic literature (Rieseberg and Soltis 1991, Golden and Bain 2000, Palme et al. 2003, Palma-Silva et al 2011) and it is expected that this phenomena will confound cpDNA-based species delimitation (such as plant DNA barcoding) even where species limits are rather clear (Hollingsworth et al. 2011.

Phenetic versus Cladistic Analyses in Systematics

Phenetics is method of grouping taxa based upon the amount of similarity or difference, regardless of phylogeny. Phenetic approaches are generally based on molecular or morphological data and seek to group taxa according to overall similarity. When nucleotide sequence data is employed in phenetic approaches the nucleotide differences among two sequences are converted to a numerical value. Then a matrix of these pair-wise values is generated from the nucleotide sequence alignment and a phenetic tree is constructed to group sequences based on similarity.

Cladistic methodologies seek to construct relationships of taxa based on shared derived characters, or synapomorphies. Cladistic methods have largely replaced phenetic means of estimating ancestor-descendant relationships as phenetics is based on summary-statistics (overall similarity), which is widely regarded as inferior to character based cladistic methods of phylogeny estimation (Steel et al. 1988, Farris et al. 1996, Murphy

and Doyle 1998). Cladistic methods on the other hand generally require much more time (hours or days) and often recovers numerous equally probable trees.

This thesis will investigate two main questions with respect to *Compsoneura* and *Virola*: 1) What are the sister-species relationships within these genera and do these clades agree with past sub-generic taxonomic groupings? and 2) Do cryptic species exist in these genera and can DNA loci be used to discriminate species of these trees that are exceedingly difficult to identify? Separate molecular analyses will be used to help answer each of these questions. Bayesian and parsimony based cladistic analyses will be used to estimate evolutionary relationships among taxa (*Compsoneura*=Chapter 1, *Virola*=Chapter 3). Although the use of phenetic analyses with DNA barcoding data has been criticized by some taxonomists (Will and Rubinoff 2004), phenetic methods will be used to assess the ability of highly variable cpDNA loci to discriminate species of these two genera as I am concerned primarily with taxon distinctiveness rather than the reconstruction of sister group relationships in such analyses.

A benefit of analyzing large volumes of sequence data with phenetic methods is that many samples (hundreds if not thousands) can be included from within populations as well as across a species' range in order to survey for undetected speciation. Such dense taxon sampling patterns have revealed potential cryptic speciation events in butterflies, bats, fish, flies, algae, wasps, and amphipods that had gone unnoticed by traditional taxonomic characters (Hebert et al. 2004, Saunders 2005, Witt et al. 2006, Clare et al. 2007, Smith et al. 2007, Hubert et al. 2008, Smith et al. 2008). It is important to note, however, that many of these investigations complemented their molecular characters with additional natural history data such as plant or insect-host-specificity, and geography

(Hebert et al. 2004, Smith et al. 2004, Saunders 2005, Smith et al. 2008). These DNA barcoding studies, as well as population genetic, phylogeography and systemic studies are discovering units that may represent undescribed species faster than these units can be further tested and described. Padial and colleagues (2010) have proposed multiple protocols for recognizing and testing candidate species with integrative data from multiple character sets (molecular, morphology, geography, ecology, etc.) so that these putative taxa can be recognized in some form until they can be fully described. It appears that the sequence data generated by DNA barcoding efforts will only serve to aid systematist in identifying molecular operational taxonomic units (MOTU: as discussed by Blaxter et al. 2005) that can then be treated as candidate species hypotheses to be tested with integrative approaches.

Throughout their evolution as hunter-gatherers, humans have sustained themselves by being able to recognize the biotic elements that can nourish and heal us. Agriculture, a buttress and hallmark of large sedentary societies, has selected choice genetic variants as the basis of improvement for domesticated plants and animals. Without genetic variation such selection is impossible. Therefore, genetic diversity is critical to the welfare of human societies. In our continued search for biological treasures to improve the human condition we will need increasingly sophisticated means to survey the diversity surrounding us. After about 200 years of Linnaean taxonomic pursuits there exists about 1.2 million described species and an estimated 7 million more yet to be discovered, described, and classified (Mora et al. 2011), it can only be hoped that genetic systems of identification will expedite the taxonomic process so that we can more intelligently manage our organismal resources.

Human and geologic history of Central and South America

The neotropical region (the landmass situated between the tropics of Cancer and Capricorn in Central and South America) is home to an estimated 90,000 plant species (Raven 1976, Prance 1977, Thomas 1999), which represents a significant proportion of the world's estimated 235,000-420,000 seed plant species (Govaerts 2003, Scotland and Wortley 2003). Permanent plots and transects conducted in the Amazon basin have routinely found 200 to 300 species per hectare, making the Amazon the world's most diverse forest in terms of tree species richness (Gentry 1988, Pitman et al. 2002).

There are a number of noteworthy geological events that have had a significant effect on the evolution and composition of the neotropical flora. Approximately 130 million years before present (mybp) the African and South American continents began to separate, with complete separation occurring about 80-90 mybp (Raven and Axelrod 1974). The continents of Africa and South America continued to separate during the late Cretaceous (99-65 mybp) while volcanic islands formed between North and South America (Raven and Axelrod 1974). The uplift of the Central American landmass continued into the Miocene (23-5 mybp) culminating in the physical connection of the North and South America an estimated 3 mybp (Keigwin 1978, Marshall 1982). The Andean orogeny began its uplift around 23 mybp but the most intense activity occurred from the middle Miocene (12 mybp) to the early Pliocene (Hoorn et al. 2010). This uplift created the worlds longest current tropical mountain range as well as a large network of lakes and swamps that existed from about 25-10 mybp when the East-West drainage of

the continent was impeded by the forming mountains (Pons and Franceschi 2007, Hoorn et al 2010).

The uplift of the Andean orogeny has likely been a key factor in the generation of neotropical floristic diversity. Gentry (1982, 1988) hypothesized that the relatively recent uplift of the Andean cordillera caused explosive speciation in herbaceous plants and is responsible for the exceedingly high diversity of plant species in the neotropics. Rapid speciation has also been found in tree genera as a molecular analysis of the genus *Inga* found very low level of nucleotide substitutions among species despite morphological differentiation, suggesting recent and rapid speciation in the genus (Richardson et al. 2001). Contemporary analyses that have integrated molecular phylogenetic and geologic data appears to indicate the uplift of the Andes was crucial to the diversification of the neotropical flora primarily through the generation of topographic features, fertile and heterogeneous edaphic conditions, as well as varying precipitation regimes (Hoorn et al. 2010).

The diverse biota of South America has also been impacted in past and present times by its human inhabitants. The first humans are believed to have arrived in South America 12,000-15,000 years ago and they are thought to have contributed to the extinction of numerous large bodied vertebrates through hunting (Janzen 1983). The extinction of these large herbivores likely led to the extinction, or rarefaction, of many species of plants dependent on these megafauna for dispersal and/or germination of their seeds (Janzen and Martin 1981).

Shortly after they arrived in the Neotropics, humans made use of the rich flora around them by domesticating numerous species, making South and Central America an

independent origin of agriculture (Olsen and Schaal 1999, Smith 2006). When Gaspar de Carvajal sailed down the Amazon from 1541-1542 he wrote of large settlements in the region of modern day Manaus, Brazil where some communities stretched uninterrupted for 25 km. However, after studying cultures living at the periphery of the Northwest Amazon basin, anthropologists largely concluded that all pre-Colombian societies were small hunter-gatherer groups clinging to existence and sparsely scattered across the Amazon (Meggers 1971, Hames and Vickers 1982). This was largely based on the notion that great societies, such as those described by Carvajal, required large-scale agriculture that was regarded as impossible on the poor soils of the region. Relatively recently however, scientists came to realize that rich black soils, called terra preta (or black earth), scattered across the Amazon were of human origin (Smith 1980, Glaser et al. 2001), These soils appear to have been created by the generation and application of charcoal to highly oxidized, and therefore poor, soils. It appears that this soil amelioration technique was employed across the Amazon watershed as deposits of terra preta soils have been found from the rivers source to the sea and up many of its tributaries to the North and South (Glaser 2007). The application of charcoal to these soils appears to change the soil microbe community (Grossman et al. 2010) and increase the cation exchange capacity, and therefore nitrogen retention of the soils, leading to improved crop yields (Lehmann et al. 2005, Chan et al. 2007). Increased yields from these soils might have enabled the persistence of large societies (perhaps millions of people) such as those documented by Carvajal (Hackenberger 1999). Assessments of species richness done by Junquiera et al. (2010) found that domesticated and semi-domesticated woody plant diversity is higher on terra preta soils compared to adjacent non-anthropogenic soils. This could mean that the

very composition of the Amazon forest was changed by these societies and that the Amazon rainforest is not the wild "forest primeval" envisioned by many people.

Myristicaceae

The nutmeg family, or Myristicaceae, is composed of sub-canopy to canopy evergreen trees found in virtually any tropical rainforest of the world. The family is currently represented taxonomically by 21 genera and about 500 species (Janovec 2000). The Myristicaceae are ecologically important as species in this family are abundant in tropical rainforests and the family as a whole is often among the top 10 most abundant tree families in these forests (Gentry 1988, Pitman et al. 2001, Pitman et al. 2002).

General Morphology and Cytology

Members of the Myristicaceae have numerous morphological traits that make them easy to identify to the family level in the field. Most species have brown-red bark that is smooth and flakes off in thin slivers or large plates. Perhaps the most striking feature of the family is their characteristic branching pattern whereby the mostly horizontal branches radiate in whorls from the trunk. All leaves of the Myristicaceae are simple, entire, 2 ranked and alternately arranged on branches. Leaf veins are pinnate and free to anastomosing or brochidodromous. Leaf blades may be densely pubescent or totally glabrous. Members of the Myristicaceae are dioecious with the exception of a few monoecious members of the neotropical genus *Iryanthera*. Flowers are unisexual, inconspicuous, small in size (< 4 mm), composed of three sepals (rarely 4) that are fused to varying degrees, and exude a strong fragrance (Smith 1937, Armstrong 1997, Janovec

2000). Female flowers are uni-carpellate and uni-ovulate and they show very little morphological differentiation within the family. Male flowers are composed of 2-60 anthers which are fused to various degrees to a central column. This fused androecium is often referred to as a synandrium and has been of key interest to systematic studies of the family owing to the largely amorphous nature of the leaves and gynoecium (Smith 1937, Warburg 1987, Janovec 2000, Sauquet et al. 2003b). Flowers are borne on racemose or thrysoidal paniculate inflorescences, which themselves originate within leaf axils (Wilde 1991). Once fertilized, the unicarpellate ovary develops into an elliptical fruit called a follicle. The fruit is comprised of a single seed, typically with a ruminate endosperm, covered to various degrees by a fatty white-reddish aril. The arillate seed is encapsulated by a pericarp which dehisces along a single suture in most species at maturity to present the seed to animal dispersers.

Cytological investigations in the family Myristicaceae have largely been restricted to members of *Myristica*. The amount of DNA in an unreplicated haploid nucleus of plants ranges approximately 1000-fold, from 0.1 pg to approximately 125 pg (Bennett et al. 2000). This haploid nuclear complement of DNA, also known as the C-Value, has only been determined for *Myristica dactyloides* and *Myristica fragrans*, which were estimated to contain 1.4 pg and 1.2 pg (or 1,372 and 1176 Mbp) respectively in a haploid complement of their 24 chromosomes (Bennet et al. 2000). Chromosome counts in other taxa have revealed haploid chromosome numbers of 19, 21, 25 and 26 (Mehra and Bawa 1969,Raven 1975). The chromosomes of Myristicaceae are holocentric (Flach 1966), meaning that they do not possess a localized centromere, and consequently kinetochore activity is distributed over the length of the chromosome. Holocentric chromosomes are

rather unusual among angiosperms and have been reported in few families including Cyperaceae, Juncaceae, Melanthiaceae, Cuscutaceae, (Pazy and Plitman 1994, Luceno et al. 1998, Nagaki et al 2005, Hipp et al. 2009). A high degree of variation of in chromosome arrangement in species of *Carex*, which is afforded by holocentric chromosomes, has led to speculation that chromosome rearrangements may have played a significant role in the generation of the extraordinary number of species in Carex (Hipp et al. 2009). An additional chromosomal curiosity is that Myristica fragrans is one of relatively few angiosperms that have been found to have hetergametic sex chromosomes (Ganeshaiah et al. 2000).

Ecology

Wild nutmeg trees are entirely dependent upon animals for both the pollination of their flowers and dispersal of their seeds. The flowers emit a strong sweet fragrance and individual flowers are open for only a few days (Armstrong and Irvine 1989a). Members of the nutmeg family are believed to be pollinated by small generalist pollinators such as beetles (Coleoptera), thrips (Thysanoptera), and flies (Diptera) that consume pollen (Armstrong and Drummond 1986, Armstrong and Irvine 1989b, Armstrong 1997). However, there has yet to be a comprehensive study of pollination agents of the family. Since the female flowers offer no apparent reward for pollinators, it is believed they attract pollinators by mimicking the fragrance of male flowers (Armstrong 1997). The seeds of Myristicaceae are highly prized by birds and mammals for their aril, or seed covering, which is particularly rich in energy-rich fats (Howe 1981). Seed-set in some species may be low; a study by Bullock (1982) found that most female trees of *Compsoneura sprucei* produce less than 10 fruits per fruiting cycle. The nutritious aril

and diurnal opening of the fruit is believed to encourage consumption of seeds by avian and mammalian dispersal vectors. Numerous species of birds such as trogons (*Trogon massena*), motmots (*Baryphthengus martii)*, toucans (*Ramphastos sulfuratus*, *R. swainsonii*), and guans (*Penelope purpurascens*) consume the fruit but typically regurgitate or expel the seeds only minutes after consuming the nutritious aril (Howe and Vande Kerckhove 1981, Howe *et al.* 1985). However, spider monkeys (*Ateles fusciceps*) may act as one of the most important long distance dispersers of nutmeg seeds as they consume the entire seed-aril unit and the seed is excreted intact many hours later (Russo 2003). Additional studies of seed dispersal in *Virola flexuosa* have indicated that human hunting can have an impact on seed removal as nearly 90% of seeds were found to be removed from parent trees in unhunted plots versus only 67% removal in hunted areas (Holbrook and Voiselle 2009). Seed dispersal might be particularly important as Howe et al.(1985) found that 99.2% of *Virola surinamensis* seeds that fell within 45 m of the parent tree succumbed to predation but those which were dispersed to greater distances experienced a 44-fold survival advantage. Additional ecological studies are needed to ascertain whether bats or other large terrestrial mammals are significant dispersers of various species of nutmeg seeds. It is still unknown whether nutmeg trees are diphasic, exactly which species pollinate their flowers, and what effects pollinators and dispersers have on the genetic structure of populations.

Ethnobotany

In addition to their ecological importance, the Myristicaceae are also significant ethnobotanically. The hollowed branches of the neotropical nutmeg genus *Otoba* are used

to construct blowpipes used in hunting (Schultes and Raffauf 1990). The blood red latex of *Virola spp.* is used as part of the mixture for curare; an arrow and dart poison for immobilizing animals (Macrae and Towers 1984). *Virola* latex is also the primary source of tryptamine alkaloids included in yakee or ebena snuff which is employed widely across the Amazon by natives to produce powerful visions for shamanic, religious, and sometimes recreational purposes (Macrae and Towers 1984a and b, Macrae and Towers 1985, Schultes and Raffauf 1990, McKenna et al. 1984). Various preparations of the sap, leaves and bark are used to treat malaria, diarrhea, fungal skin infections, arthritis, and hemorrhoids (Schultes and Raffauf 1990, Lopes et al. 1999). Members of the genus *Virola* are harvested extensively in many South American countries as a source of wood for veneer and timber. In some neotropical countries, exports of *Virola spp.* are rivaled in economic importance only by big-leaf mahogany [*Swietena macrophylla,* Meliaceae] (Macedo and Anderson 1993).

The best known member of the Myristicaceae is the spice bearing species *Myristica fragrans,* the common nutmeg tree. Although believed to have been domesticated only in the last 500 years, *M. fragrans*, has been employed ethnobotanically wherever it has been spread by humans in the world. *Myristica fragrans* is native to a small area known as the Banda Islands of the Maluku Province of modern day Indonesia. The nutmeg tree has been dispersed far beyond its endemic range in the Banda Islands to a pantropical range by people since the late 18th century (Hanna 1978). The nutmeg tree bears a fruit that dehisces when ripe to reveal a single seed covered by a bright red aril, which are known commercially as nutmeg and mace respectively. The primary use of *M. fragrans* is as a flavoring in cooking but many cultures around the world have found it

useful for a myriad of purposes. In its native range the pericarps of nutmeg are made into a sweet snack called ‘pala manis’ or ‘pala gulu’ by repeated soaking in a sugar solution (Gils and Cox 1994). Indigenous Malukans also topically apply nutmeg oil, which gives a warm strengthening feeling, to relieve headaches, stomach aches, diarrhea and flu symptoms (Gils and Cox 1994). Both nutmeg and mace have been used as an aphrodisiac, anti-rheumatoid, anti-malarial, stimulant, and post childbirth tonic (Burkill 1935).

Myristicaceae Systematics

Nutmeg trees made their first appearance in the scientific literature in 1742 when Linnaeus first described the genus *Myristica*. Myristicaceae was established in 1810 by Brown, but the first thorough monographic treatment of the Myristicaceae was not completed until 1856 by De Candolle (cited from Janovec 2000). De Candolle recognized about 90 species worldwide and organized them into 13 sections of the single genus *Myristica*. Bentham and Hooker (1880) retained the monogeneric concept for Myristicaceae and condensed the 13 sections of De Candolle into 7 sections. The most comprehensive work of the pantropical Myristicaeae was done by Warburg (1897, cited from Janovec 2000). Warburg (1897) recognized 240 species organized in 15 genera, the majority of which are still recognized. The last comprehensive treatment of the neotropical Myristicaceae was done by Smith (1937) who recognized the genera *Virola*, *Compsoneura* (split into sections *Coniostele* and *Compsoneura* based on degrees of anther fusion), *Osteophloeum*, *Iryanthera*, and *Dialyanthera* (currently synonymous with *Otoba*), all of which are endemic to South and Central America. Since Smith's (1937)

work, further collections of the Myristicaceae have led to descriptions of both new species and genera (Rodrigues 1989a and b, Janovec and Harrison 2002, Janovec 2002, Janovec and Neill 2002, Sauquet 2003, Sauquet 2004, de Wilde 1991). Taxonomic works on the neotropical Myristicaceae since 1937 have been largely limited to descriptions of new species. The most recent taxonomic work concerning neotropical nutmegs, a monograph of *Compsoneura*, by Janovec (2000) nearly doubled the number of species in the genus and divided the genus into informal sections *Compsoneura* and *Hadrocarpa*, the latter of which was further subdivided into informal sub-groups Atopa and Capitellata.

Myristicaceae are members of the eumagnoliid clade which are among the most primitive of extant angiosperms (Soltis et al. 2000, APG III). Although initial molecular analyses placed Myristicaceae at the base of order Magnoliales making it sister to all remaining families of the order (Soltis et al. 2000, Sauquet et al. 2003), the most recent phylogenetic analysis to date found Magnoliaceae to be sister to all other families including Myristicaceae (Soltis et al. 2011). Molecular phylogenetic analysis completed on the family by Sauquet et al. (2003) and Doyle et al. (2004) found insufficient molecular divergence within the Myristicaceae to make strong phylogenetic inferences of generic relationships. Despite this very weak consensus of intergeneric relationships within the Myristicaceae, there was some support among numerous loci for the placement of the neotropical genus *Compsoneura* as the sister to the rest of the family (Doyle et al. 2004, Sauquet et al. 2003).

Biogeography of the Myristicaceae

Very little is known about the phylogeographic and biogeographic history of the Myristicaceae despite being an ancient group of trees with a pantropical distribution. This paucity of information is largely due to the fact that there is very poor representation of the family in the fossil record compared to other angiosperm families (Doyle et al. 2004). The earliest fossil evidence of the Myristicaceae is from a fossilized wood sample of *Myristicoxylon princeps* from the Paleocene of the African Sahara region. It is unclear, however, as to whether this specimen belongs to a stem or crown group of the family (Boureau 1950). Doyle et al. (2008) document the identification of a Myristicaceous seed from England's London Clay which dates to the Early Eocene (~56-34mya) which is markedly older than the only seed found previously from Miocene (23-5mya) deposits of Germany (Gregor 1977). Additional fossils of seeds (Berry 1929), leaves (Wolfe 1977), flowers (Poinar and Poinar 1999), and pollen (Frederikson 1973, Jan du Chene et al. 1978) of likely Myristicaceae origin have been found but they cannot be reliably distinguished from morphologically similar families (Doyle et al. 2004). The rarity of nutmeg fossils prior to the Paleocene corroborates the molecular clock estimate of Doyle et al. (2004) predicting that the Myristicaceae crown group diversified in the last 15-20 million years. Such a recent diversification is rather surprising as the same markers estimated that the closely related Annonaceae radiated approximately 100-120mya, which correlated well with their rich fossil record (Doyle et al. 2004). However, this relatively young estimation of Myristicaceae radiation could be an artifact of poor fossil representation of the family and/or a marked decrease in the rate of molecular evolution compared to closely related families, leading to a mis-calibration of the molecular clock.

Therefore it is unclear as to whether this low level of molecular divergence is truly the result of a relatively recent diversification in the family or a marked slowdown in molecular evolution resulting in the conservation of DNA sequences. Given the Myristicaceae's pantropical distribution, basal position on the flowering plant phylogeny, and poor long distance dispersal ability, it is unlikely that they radiated in the last 20 million years when the continents were almost in the same positions as today. This leaves us with little knowledge of the evolutionary and phylogeographic patterns of one of the most abundant and diverse plant families of tropical rain forests.

Challenges to Molecular work with the Myristicaceae

A diverse assembly of metabolites produced in plant cells often co-purify during DNA isolation procedures leading to the degradation of genomic DNA and/or the inhibition of PCR. These inhibitors often include but are not restricted to phenols, tannins, polysaccharides, proteins, alkaloids, lignans, flavonoids, and RNA. The Myristicaceae, like many other plants, contain copious amounts of secondary metabolites that are troublesome for molecular investigations. The most notorious are perhaps flavans called myristinins that are found in numerous species of nutmegs and are the most potent inhibitors known of β DNA polymerases and have the added action of autolytic cleavage of dsDNA (Sawadjoon et al. 2002, Deng et al. 2005, Maloney et al. 2005). Additionally, numerous other flavonoids have been isolated and characterized from various genera of the Myristicaceae (Braz Filho et al. 1973, Gottlieb et al. 1973, Gottlieb et al. 1976) which include phenols and tannins, which are known inhibitors of PCR (Wilson 1999, Krause et al. 2001). The inhibitory effects of these contaminating chemicals may be minimized by

the inclusion of chemicals such as Polyvinylpyrillidone, phenacylthiazolium bromide, and sodium sulfite in lysis buffers during DNA extraction (Ivanova et al. 2008, Poinar et al. 1998, Whitlock et al. 2008). If inhibitory chemicals remain in purified DNA extracts then the inclusion of PCR facilitating chemicals such as trehalose, PVP, tween 20, PEG, and BSA may be necessary to achieve amplification of DNA sequences (Pierpont 1969, Sasaki et al. 1998, Horne et al. 2004, Speiss et al. 2004, Barabwal et al. 2003,). Thus, overcoming the challenges associated with nucleic acid manipulation in the Myristicaceae may contribute to enhancing molecular protocols for DNA extraction, amplification, and sequencing vital to studying many species of plants.

Why work with the Myristicaceae?

Biologists depend largely on morphological taxonomic descriptions to identify the species encountered in their studies. Examinations of the composition and diversity of tropical rainforests have been the focus of many contemporary tropical plant ecologists, however, they are faced with a flora that is exceedingly diverse and difficult to identify. Previous investigations of the Myristicaceae have revealed that the mis-identification rate of Myristicaceae species typically exceeds 50% in herbarium samples and 25% in ecological plots (J. Janovec personal communication, Steeves personal observations). A DNA based identification system would aid tropical plant identification, however there have been few tests of DNA barcoding regions in a diverse group with many pairs of morphologically similar and recently evolved sister species (Gao et al. 2009, Gonzales et al. 2009, Spooner et al. 2009). The Myristicaceae, and *Compsoneura* and *Virola* in particular, provide a rigorous test of DNA-based identification for plants as they are

diverse, morphologically difficult to identify, contain many sister species, and the family is known to have exceedingly slow rates of molecular evolution (Sauquet et al. 2003).

The primary goal of this thesis was to explore morphological and molecular diversity among and within species of *Virola* and *Compsoneura*; two genera of the neotropical Myristicaceae endemic to Central and South America. This was accomplished by surveying DNA sequence variation of many species within each of these genera and by selecting species complexes from each genus to provide an assessment of genetic variation within these focal species. The objective of the generic level investigations was to generate phylogenetic estimates of relationships of the respective species in hopes of elucidating general trends of morphological evolution, biogeography, and illuminating cases of high genetic variability that may indicate cryptic and/or incipient speciation events. The main objective of surveying DNA sequence variability within the selected species was to see if the high morphological variability exhibited by these taxa was the result of phenotypic plasticity, cryptic/incipient speciation, or infraspecific genetic variation. This thesis reveals that there are many species yet to be described in one of the most common tree families of the neotropics and that the taxonomic understanding of the Amazon and other tropical forests is arguably in its infancy. If we are to attempt to understand this botanically diverse region from an ecological, evolutionary, biogeographical, or any other biological standpoint we must first finely dissect its most fundamental units: the species of which it is composed.

Literature cited

Armstrong, J.E., & Drummond, B.A. 1986. Floral biology of *Myristica fragrans* Houtt. (Myristicaceae), the nutmeg of commerce. *Biotropica* **18**: 32-38.

Armstrong, J. E., & Irvine, A. K. 1989a. Floral biology of *Myristica insipida* (Myristicaceae), a distinctive beetle pollination syndrome. *American Journal of Botany* **76**: 86–94.

Armstrong, J.E. & Irvine, A.K. 1989b. Flowering, sex ratios, pollen-ovule ratios, fruit set, and reproductive effort of a dioecious tree, *Myristica insipida* (Myristicaceae), in two different rainforest communities. American Journal of Botany **76**: 74-85.

Armstrong, J. E. 1997. Pollination by deceit in nutmeg (*Myristica insipida*, Myristicaceae): floral displays and beetle activity at male and female trees. *American journal of botany* **84**: 1266-1274.

Bachmann, K. 1998. Species as units of diversity: an outdated concept. Theory *Bioscience* **117**: 213-230.

Bennett, M.D., Bhandol, P., & Leitch, I. 2000. Nuclear DNA amounts in Angiosperms and their modern uses-807 new estimates. *Annals of Botany* **86**: 859-909.

Bentham, G., & Hooker, J.D. 1880. *Genera plantarum,* London: Reeve and Co.

Berry, E. 1929. Early Tertiary fruits and seeds from Belen, Peru. *Johns Hopkins University Studies in Geology* **10**: 137-180.

Blaxter, M. L. 2004. The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**: 669.

Boureau, E. 1950. Étude paléoxylologique du Sahara (IX). Sur un Myristicoxylon princeps n. gen., n. sp., du Danien d'Asselar (Sahara soudanais). *Bulletin du Muséum National d'Histoire Naturelle* **2**: 523-528.

Braz Filho, R., Frota Leite, M., & Gottlieb, O. 1973. Constitutions of diarylproppanoids from *Virola multinervia*. *Phytochemistry* **12**: 417-419.

Bullock, S. H. 1982. Population structure and reproduction in the neotropical dioecious tree *Compsoneura sprucei*. *Oecologia* **55**: 238–242.

Chan, K. Y., Van Zwieten, L., Meszaros, I., Downie, A., & Joseph, S. 2008. Using poultry litter biochars as soil amendments. *Australian Journal of Soil Research* **46**: 437–444.

Clare, E. L., Lim, B. K., Engstrom, M. D., Eger, J. L., & Hebert, P. D. 2007. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. *Molecular Ecology Notes* **7**: 184–190.

Cole, R. J. 2009. Postdispersal seed fate of tropical montane trees in an agricultural landscape,
Southern Costa Rica. *Biotropica* **41**: 319–327.

Deng, J., Starck, S. R., Li, S., & Hecht, S. M. 2005. (+)-Myristinins A and D from *Knema elegans*, which Inhibit DNA Polymerase β and Cleave DNA. *Journal of Natural Products* **68**: 1625-1628.

de Queiroz, K. 2005. A unified concept of species and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences* **56**: 196.

Diamond, J., & Bishop, K. D. 1999. Ethno-ornithology of the Ketengban people, Indonesian New Guinea. In *Folkbiology*. MIT Press (pp. 17-45).

Doyle, J. A., Manchester, S. R., & Sauquet, H. 2008. A seed related to Myristicaceae in the early Eocene of Southern England. *Systematic Botany* **33**: 636–646.

Doyle, J. A., Sauquet, H., Scharaschkin, T., & Le Thomas, A. 2004. Phylogeny, molecular and fossil dating, and biogeographic history of Annonaceae and Myristicaceae (Magnoliales). *International Journal of Plant Science* **165**: S55–S67.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D., & Kluge, A. G. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12**: 99–124.

Fazekas, A. J., Kesanakurti, P. R., Burgess, K. S., Percy, D. M., Graham, S. W., Barrett, S. C., Newmaster, S. G., Hajibabael, M., & Husband, B. C. 2009. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Molecular Ecology Resources* **9**: 130–139.

Felsenstein, J. 1984. Distance methods for inferring phylogenies: A justification. *Evolution* **38**: 16-24.
Felsenstein, J. 1987. Estimation of hominoid phylogeny from a DNA hybridization data set. *Journal of molecular evolution* **26**: 123–131.

Flach, M. 1966. Diffuse centromeres in a dicotyledenous plant. *Nature* **5030**:1369-1370.

Frederiksen, N. 1973. New mid-Tertiary spores and pollen grains from Mississippi and Alabama. *Tulane Studies in Geology and Paleontology* **10**: 65–86.

Ganeshaiah, K.N., Ravishankar, K.V., Anand, L., Shibu, M.P., Shaanker, U. 2000. Identification of sex-specific DNA markers in the dioecious tree, nutmeg (*Myristica fragrans* Houtt.). *Plant Genetic Resources Newsletter* **121**:59–61

Gao, T., Sun, Z., Yao, H., Song, J., Zhu, Y., Ma, X., & Chen, S. 2009. Identification of Fabaceae plants using the DNA barcode matK. *Planta Medica* **4**: 1-3.

Gentry, A. H. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* **69**: 557–593.

Gentry, A. H. 1988. Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 156-159.

Glaser, B. 2007. Prehistorically modified soils of central Amazonia: a model for sustainable agriculture in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**: 187.

Glaser, B., Haumaier, L., Guggenberger, G., & Zech, W. 2001. The 'Terra Preta' phenomenon: a model for sustainable agriculture in the humid tropics. *Naturwissenschaften* **88**: 37-41.

Godfray, H. C. 2002. Challenges for taxonomy. *Nature* **417**: 17–19.

Golden, J. L., & Bain, J. F. 2000. Phylogeographic patterns and high levels of chloroplast DNA diversity in four *Packera* (Asteraceae) species in southwestern Alberta. *Evolution* **54**: 1566–1579.

Gonzalez, M. A., Baraloto, C., Engel, J., Mori, S. A., Pétronelli, P., Riéra, B., Roger, A., Thébaud, C., & Chave, J. 2009. Identification of Amazonian trees with DNA barcodes. *PLoS One* **4**: e7483.

Gottlieb, O. 1976. Neolignans from *Virola carinata*. *Phytochemistry* **15**: 773-774.

Gottlieb, O., Loureiro, A., Dos Santo Carneiro, M., & Da Rocha, A. 1973. Distribution of diarylpropanoids in Amazonian *Virola* species. *Phytochemistry* **12**: 1830.

Govaerts, R. 2003. How many species of seed plants are there?-a response. *Taxon* **52**: 583-584.

Gregor, H. 1977. Subtropische Elemente im europäischen Tertiär II (Fruktifikationen). *Paläontologische Zeitschrift* **51**: 199–226.

Grossman, J. M., O'Neill, B. E., Tsai, S. M., Liang, B., Neves, E., Lehmann, J., & Thies, J. E. 2010. Amazonian anthrosols support similar microbial communities that differ distinctly from those extant in adjacent, unmodified soils of the same mineralogy. *Microbial Ecology* **60**: 192-205.

Hackenberger, M. J., Petersen, J. B., & Neves, E. G. 1999. Village size and permanence in Amazonia: two archaeological examples from Brazil. *Latin American Antiquity* **10**: 353-376.

Hames, R. B., & Vickers, W. T. 1982. *Adaptive responses of native americans*. New York: Academic press.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* **270**: 313-321.

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences* **270**: S96-S99.

Hebert, P. D. N., Stoeckle, M. Y., Zemlack, T. S., & Francis, C. M. 2004. Identification of birds through DNA barcodes. *PLOS Biology* **2**: 1657-1663.

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

Hipp, A.L., Rothrock, P.E., & Roalson, E.H. 2009. The evolution of Chromosome arrangements in *Carex* (Cyperaceae). *Botanical Review* **75**: 96-109.

Holbrook, K. M., & Loiselle, B. A. 2009. Dispersal in a Neotropical tree, *Virola flexuosa* (Myristicaceae): Does hunting of large vertebrates limit seed removal? *Ecology* **90**: 1449–1455.

Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., Van Der Bank, M., Chase, M. W., Cowan, R. S., Erickson, D. L., Fazekas, A. J., & others. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences* **106**: 12794-12797.

Hollingsworth, P. M., Graham, S. W., & Little, D. P. 2011. Choosing and using a plant DNA barcode. *PloS one* **6**: e19254.

Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., Sanmartin, I., Sanchez-Meseguer, A., Anderson, C. L., Figueiredo, J. P., Jaramillo, C., Riff, D., Negri, F. R., Hooghiemstra, H., Lundberg, J., et al. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330**: 927-931.

Howe, H. F. 1981. Dispersal of a neotropical nutmeg (*Virola sebifera*) by birds. *The Auk* **98**: 88–98.

Howe, H. F., & Vande Kerckhove, G. A. 1981. Removal of wild nutmeg (*Virola surinamensis*) crops by birds. *Ecology* **62**: 1093–1106.

Howe, H. F., Schupp, E. W., & Westley, L. C. 1985. Early consequences of seed dispersal for a neotropical tree (*Virola surinamensis*). *Ecology* **66**: 781–791.

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., Watkinson, D., Dumont, P., Curry, A., Bentzen, P., & others. 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One* **3**: e2490.

Ivanova, N. V., Fazekas, A. J., & Hebert, P. D. N. 2008. Semi-automated, membrane-based protocol for DNA isolation from plants. *Plant Molecular Biology Reporter* **26**: 186-198.

Jan du Chene, R., Onyike, M., & Sowunmi, M. 1978. Some new Eocene pollen of the Ogwashi-Asabe formation, South-Eastern Nigeria. *Rev Esp Micropaleontol* **10**: 285–322.

Janovec, J. P. 2002. *Compsoneura diazii* (Myristicaceae), a new species from the Rio Cenepa area of Northwestern Peru. *Novon* **12**: 366–368.

Janovec, J. P., & Harrison, J. S. 2002. A morphological analysis of the *Compsoneura sprucei* complex (Myristicaceae), with a new combination for the Central American species *Compsoneura mexicana*. *Systematic Botany* **27**: 662–673.

Janovec, J. P., & Neill, A. K. 2002. Studies of the Myristicaceae: an overview of the *Compsoneura atopa* complex, with descriptions of new species from Colombia. *Brittonia* **54**: 251–261.

Janovec, J. P. 2000. A systematic study of Compsoneura (A. DC.) Warb., A Neotropical member of the nutmeg family. Texas A&M University Dissertation: 1-359.

Janzen, D. H. 1983. The Pleistocene hunters had help. *The American naturalist* **121**: 598–599.

Junqueira, A. B., Shepard, G. H., & Clement, C. R. 2010. Secondary forests on anthropogenic soils in Brazilian Amazonia conserve agrobiodiversity. *Biodiversity and Conservation* **19**: 1933-1961.

Keigwin, L. D. 1978. Pliocene closing of the Isthmus of Panama, based on biostratigraphic evidence from nearby Pacific Ocean and Caribbean Sea cores. *Geology* **6**: 630.

Krause, D., Smith, W., & McSweeney, C. 2001. Extraction of microbial DNA from rumen contents containing plant tannins. *Biotechniques* **31**: 294-298.

Lehmann, J., Pereira da Silva, J., Steiner, C., Nehls, T., Zech, W., & Glaser, B. 2003. Nutrient availability and leaching in an archaeological Anthrosol and a Ferralsol of the Central Amazon basin: fertilizer, manure and charcoal amendments. *Plant and Soil* **249**: 343–357.

Levin, D. A. 1979. The nature of plant species. *Science* **204**: 381-384.

Lopes, N. P., & Kato, M. J. 1999. Antimalarial use of volatile oil from leaves of *Virola surinamensis* (Rol.) Warb. by Waiãpi Amazon Indians. *Journal of ethnopharmacology* **67**: 313–319.

Luceno, M., Vanzela, A.L.L., & Guerra, M. 1998. Cytotaxonomic studies in Brazilian *Rhynchospora* (Cyperaceae), a genus exhibiting holocentric chromosomes. *Canadian Journal of Botany* **76**:440-449.

Macedo, D. S., & Anderson, A. B. 1993. Early ecological changes associated with logging in an Amazon floodplain. *Biotropica* **25**: 151-163.

MacRae, W. D., & Towers, G. H. 1985. Non-alkaloidal constituents of *Virola elongata* bark. *Phytochemistry* **24**: 561–566.

Macrae, W. D., & Towers, G. H. 1984a. *Justicia pectoralis*: a study of the basis for its use as a hallucinogenic snuff ingredient. *Journal of ethnopharmacology* **12**: 93–111.

Macrae, W., & Towers, G. 1984b. An ethnopharmacological examination of *Virola elongata* bark: A South American arrow poison. *Journal of ethnopharmacology* **12**: 75–92.

Maloney, D. J., Deng, J., Starck, S. R., Gao, Z., & Hecht, S. M. 2005. (+)-Myristinin A, a naturally occurring DNA polymerase β inhibitor and potent DNA-damaging agent. *Journal of the American Chemical Society* **127**: 4140-4141.

Marshall, L. G., Webb, S. D., Sepkoski, J. J., & Raup, D. M. 1982. Mammalian evolution and the great American interchange. *Science* **215**: 1351-1357.

McKenna, D. J., Towers, G. H. N., & Abbott, F. S. 1984. Monoamine oxidase inhibitors in South American hallucinogenic plants part 2: Constituents of orally-active Myristicaceous hallucinogens. *Journal of Ethnopharmacology* **12**: 179–211.

Meggers, B. J. 1971. *Amazonia: Man and nature in a counterfeit paradise*. Chicago: Aldine.

Mehra, P.H., & Bawa, K.S. 1969. Chromosomal evolution in tropical hardwoods. *Evolution* **23**: 466-481.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., & Worm, B. 2011. How many species are there on the earth and in the oceans. *Public Library of Science-Biology* **9**: e1001127.

Mullis, K. B., & Faloona, F. A. 1987. Specific synthesis of DNA in Vitro via a Polymerase-Catalyzed chain reaction. *Methods in Enzymology* **155**: 335-350.

Nagaki, K., Kasjihara, K., & Murata, M. 2005. Visualization of diffuse centromeres with centromere-specific histone H3 in the holocentric plant *Luzula nivea*. *The plant cell* **17**:1886-1893.

Olsen, K. M., & Schaal, B. A. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 5586.

Padial, J.M., Miralles, A., De la Riva, I., and Vences, M. 2010. The integrative future of taxonomy. *Frontiers in Zoology* **7**:16.

Palma-Silva, C., Wendt, T., Pinheiro, F., Barbará, T., FAY, M. F., Cozzolino, S., & Lexer, C. Sympatric bromeliad species (*Pitcairnia spp.)* facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* 20: 3185–3201

Palmé, A. E., Semerikov, V., & Lascoux, M. 2003. Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity* **91**: 465–474.

Pazy, B., & Plitmann, U. 1994. Holocentric chromosome behaviour in *Cuscuta* (Cuscutaceae). *Plant Systematics and Evolution* **191**:105-109.

Pierpoint, W. S. 1969. o-Quinones formed in plant extracts. Their reaction with bovine serum albumin. *Biochemical Journal* **112**: 619.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* **82**: 2101–2117.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2002. A comparison of tree species diversity in two upper Amazonian forests. *Ecology* **83**: 3210–3224.

Poinar, G. J., & Poinar, R. 1999. *The amber forest: a reconstruction of a vanished world*. Princeton, N.J.: Princeton University Press.

Poinar, H. N. 1998. Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281**: 402-406.

Pons, D., & Franceschi, D. D. 2007. Neogene woods from western Peruvian Amazon and palaeoenvironmental interpretation. *Bulletin of Geosciences* **82**: 343-354.

Prance, G. T. 1977. Floristic Inventory of the Tropics: Where Do We Stand? *Annals of the Missouri Botanical Garden* **64**: 659–684.

Raven, P.H. 1975. The bases of Angiosperm phylogeny: Cytology. *Annals of the Missouri Botanical Garden* **62**: 724-764.

Raven, P. H. 1976. Systematics and plant population biology. *Systematic Botany* **1**: 284–316.

Raven, P. H., & Axelrod, D. I. 1974. Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* **61**: 539-673.

Richardson, J. E. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* **293**: 2242-2245.

Rieseberg, L.H., & Soltis, D.E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**: 65-84.

Rieseberg, L.H., Wood, T.E., & Baack, E.J. 2006. The nature of plant species. Nature 440: 524-527.

Roch, S. 2010. Toward extracting all phylogenetic information from matrices of evolutionary distances. *Science* **327**: 1376-1379.

Rodrigues, W. A. 1989a. A new Venezuelan *Virola* (Myristicaceae). *Annals of the Missouri Botanical Garden* **76**: 1163-1164.

Rodrigues, W. A. 1989b. Two new neotropical species of *Compsoneura* (Myristicaceae). *Brittionia* **4**: 160-163.

Russo, S. E. 2003. Responses of dispersal agents to tree and fruit traits in *Virola calophylla* (Myristicaceae): implications for selection. *Oecologia* **136**: 80-87.

Saunders, G. W. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**: 1879.

Sauquet, H. 2003. Androecium diversity and evolution in Myristicaceae (Magnoliales), with a description of a new Malagasy genus, *Doyleanthus* gen. nov. *American Journal of Botany* **90**: 1293-1305.

Sauquet, H. 2004. Systematic revision of Myristicaceae (Magnoliales) in Madagascar, with four new species of Mauloutchia. *Botanical Journal of the Linnean Society* **146**: 351–368.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W., & Le Thomas, A. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* **142**: 125–186.

Sawadjoon, S., Kittakoop, P., Kirtikara, K., Vichai, V., Tanticharoen, M., & Thebtaranonth, Y. 2002. Atropisomeric myristinins: Selective COX-2 inhibitors and antifungal agents from *Myristica cinnamomea*. *The Journal of Organic Chemistry* **67**: 5470-5475.

Scotland, R. W., & Wortley, A. H. 2003. How many species of seed plants are there? *Taxon* **52**: 101–104.

Smith, A. C. 1937. The American species of Myristicaceae. *Brittonia* **2**: 393-510.

Smith, B. D. 2006. Eastern North America as an independent center of plant domestication. *Proceedings of the National Academy of Sciences* **103**: 12223.

Smith, M. A., Rodriguez, J. J., Whitfield, J. B., Deans, A. R., Janzen, D. H., Hallwachs, W., & Hebert, P. D. 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proceedings of the National Academy of Sciences* **105**: 12359.

Smith, M. A., Wood, D. M., Janzen, D. H., Hallwachs, W., & Hebert, P. D. N. 2007. DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera,Tachinidae) are not all generalists. *Proceedings of the National Academy of Sciences* **104**: 4967-4972.

Smith, N. J. H. 1980. Anthrosols and human carrying capacity in Amazonia. *Annals of the Association of American Geographers* **70**: 553–566.

Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savolainen, V., Hahn, W. H., Hoop, S. B., & Fay, M. F. 2000. Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. *Botanical Journal of the Linnean Society* **133**: 381–461.

Soltis, D. E., Smith, S. A., Cellinese, N., Wurdack, K. J., Tank, D. C., Brockington, S. F., Refulio-Rodriguez, N. F., Walker, J. B., Moore, M. J., Carlsward, B. S., & others. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* **98**: 704.

Spooner, D. M. 2009. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany* **96**: 1177-1189.

Steel, M. A., Hendy, M. D., & Penny, D. 1988. Loss of information in genetic distances. *Nature* **336**: 118.

Thomas, W. W. 1999. Conservation and monographic research on the flora of Tropical America. *Biodiversity and Conservation* **8**: 1007–1015.

Trueman, J. W. 1998. Reverse successive weighting. *Systematic Biology* **47**: 733.

Van Gils, C., & Cox, P. A. 1994. Ethnobotany of nutmeg in the Spice Islands. *Journal of Ethnopharmacology* **42**: 117–124.

Warburg, O. 1897. Monographie der Myristicaceen. *Nova Acta Acad. Caes. Leop.-Carol* **68**: 1-680.

Whitlock, R., Hipperson, H., Mannarelli, M., & Burke, T. 2008. A high-throughput protocol for extracting high-purity genomic DNA from plants and animals. *Molecular Ecology Resources* **8**: 736–741.

Wilde, W. J. J. O. D. 1991. The genera of the Myristicaceae as distinguished by their inflorescences and the description of a new genus: *Bicuiba*. *Beitr. Biol. Pflanzen* **66**: 95-125.

Will, K. W., & Rubinoff, D. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* **20**: 47–55.

Will, K., Mishler, B., & Wheeler, Q. 2005. The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* **54**: 844-851.

Wilson, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology* **63**: 3741-3751.

Witt, J. D., Threloff, D. L., & Hebert, P. D. 2006. DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. *Molecular Ecology* **15**: 3073–3082.

Wolfe, J. 1977. Paleogene floras from the Gulf of Alaska region. *U.S. Geological Survey Professional Paper* **997**: 1–108.

Chapter 1

A PHYLOGENETIC AND MORPHOLIGICAL ANALYSIS OF *COMPSONEURA*

Abstract

Compsoneura currently comprises 21 described species of canopy to sub-canopy Myristicaceae trees native to tropical rainforests below 1500 m elevation in South and Central America. *Compsoneura* species are difficult to identify due to their similar looking leaves and small (1-4mm), largely amorphous, flowers. In this chapter infrageneric relationships are estimated using molecular phylogenetic analyses employing one chloroplast (trnH-psbA) and two nuclear markers (*AT103* and *AGT1*) and both maximum parsimony and Bayesian inference. Phylogenetic analyses revealed a deep evolutionary rift within *Compsoneura* corresponding to informal sections Compsoneura and Hadrocarpa proposed by previous morphological and biogeographical studies. Phylogenetic analyses also revealed that mountain building has likely been an important generator of speciation in the taxa studied. A distance based molecular analysis was also conducted with increased taxon sampling to test DNA-based identification using the highly variable locus trnH-psbA. Distance analyses revealed that all but two of 13 *Compsoneura* species included can be identified by one or more nucleotide polymorphisms. The molecularly indistinguishable species do not share a common range, making molecular identification of all tested species possible with a combination of sequence and location data. Herein I present molecular and morphological evidence supporting the sub-generic classification of Janovec (2000) and discuss the results with respect to sub-generic classification, morphology, and biogeography.

Introduction

The Myristicaceae, or nutmeg family, consists of 21 genera and more than 500 species of dioecious, and rarely monoecious, canopy to sub canopy trees distributed worldwide in tropical rainforest environments (Smith 1937, Wilde 1991, Janovec 2000). The Myristicaceae diverged early in the radiation of the angiosperms and previous molecular studies has placed them as the sister to all members of the order Magnoliales (APG III 2003, Qiu et al. 2006). However, the most recent and comprehensive molecular phylogenetic analysis employing 17 genes places Magnoliaceae as sister to all other members of the Magnoliales, including Myristicaceae (Soltis et al 2011). Nutmeg trees are readily recognized in the field by their combination of some or all of the following diagnostic features: (1) Myristicaceous branching (branches originating in a whorled fashion in a single plane and at relatively regular intervals on the trunk); (2) frequent presence of pink to blood-red sap with a bitter, astringent taste; (3) tiny urceolate to cup-shaped flowers with three to four tepals; (4) reduced to elongated filament column terminated with free to fused anthers dehiscing to the outside; (5) a one-seeded fruit that typically dehisces longitudinally into two valves; and (6) the presence of an aril that is usually red, or white. Numerous studies have reported the significant ecological importance of the Myristicaceae in wet, lowland tropical forests of Central and South America, Africa, Madagascar, India, and Asia (Gentry 1982, Pascal and Pelissier 1996, Poulsen et al. 1996, Spichiger et al. 1996). The Myristicaceae also hold considerable economic and cultural importance as the source of the commercial nutmeg spice (*Myristica fragrans*) and DMT-containing hallucinogens derived from *Virola spp*. in the

Amazon basin (Chagnon 1971, Mckenna et al. 1984, Schultes 1984, Van Gils and Cox 1994). Many members of the family are also valued for their wood, which is used in rural house construction as well as fine carpentry that reaches an international market (Macedo and Anderson 1993). Additionally, plant extracts of members of the nutmeg family have been investigated for their pharmacological potential, including, but not limited to, aphrodisiac and anti-cancer effects (Sawadjoon et al. 2002, Maloney et al. 2005, Tajuddin et al. 2005, Nguyen et al. 2010, Patro et al. 2010, Pusztai et al. 2010).

Compsoneura is one of six neotropical genera of Myristicaceae, which comprises 21 described dioecious species that inhabit rainforest environments below 1500m elevation in South and Central America (Smith 1937, Wilde 1991, Janovec 2000). *Compsoneura* is easily differentiated from other neotropical members of the family by virtual absence of pubescence on the leaf lamina and subparallel tertiary nerves (some taxa of Annonaceae and Icacinaceae [*Discophora spp.*] also have similarly arranged nerves and often confused with *Compsoneura spp.*). Since species of *Compsoneura* share similar leaf morphologies, identification relies heavily upon characteristics of the small flowers (1–4 mm), particularly characteristics of the androecium (Warburg 1897, Smith 1937, Armstrong and Tucker 1986, Janovec 2000, Janovec and Harrison 2002). Flowers are only present on adult trees for a few months every year (Bullock 1982, Armstrong 1997), which can make identification of female and vegetative specimens of these dioecious trees exceedingly difficult. Herbarium studies of the genus *Compsoneura* (Janovec 2000) and studies of the neotropical Myristicaceae in ecological plots by Steeves (2008, personal observations) encountered mis-identification rates of approximately 50% and 25% respectively.

Despite the cultural, ecological, and economic importance of the Myristicaceae throughout the wet tropics of the world, our systematic knowledge of the Myristicaceae is mostly based on the last, grand synthesis of the family by Warburg (1897). In Monographie der Myristicaceen (1897), which was rooted in prior studies by De Candolle (1856), Warburg described 240 species and 15 genera, including four of the six neotropical genera recognized today. Among the genera elevated by Warburg to the generic rank was *Compsoneura*, which was based on de Candolle's (1856) *Myristica* section *Compsoneura*. Both agreed that *Compsoneura* as a group of similar species was characterized by tertiary leaf venation parallel to the secondary veins, an orange fruit pericarp dehiscing longitudinally, an entire red aril, and mottled seeds. With the eight specimens he had available at the time, Warburg recognized five species of *Compsoneura* and divided them into two sections: (1) section *Eucompsoneura* (Warb.) with free anthers and (2) section *Coniostele* (Warb.) with fused anthers. Based on a study of 70 specimens available at the time from Central and South America, Smith (1937) described five additional species of *Compsoneura* and synonymized one (*C. capitellata* and *C. tessmannii* as the former), but adopted Warburg's sections *Eucompsoneura* and *Coniostele*. Since that time, the neotropical Myristicaceae has been the focus of only a few studies focused on new species and regional floristics. In 1956 Smith described an intriguing new species named *Virola atopa* from the Pacific coastal region of Colombia but later placed this species in *Compsoneura* once staminate specimens were made available by the explorations of Richard Evans Schultes (Smith 1956). Rodrigues (1989) described two new species of *Compsoneura*, bringing the total to 11 species. These preceding studies of the 19th and 20th centuries have been plagued by the availability of

few samples from across the geographical range of *Compsoneura* as well as a paucity of differentiating morphological characters.

In recent years, increased collecting efforts and the advent of molecular analyses has made more specimens and characters available for taxonomic investigation. Contemporary studies by Janovec and co-workers over the past 10 years have examined over 3000 specimens of the genus *Compsoneura* for studies concerning population genetics, morphological evolution and taxonomy (Janovec 2000, Janovec and Harrison 2002, Janovec and Neill 2002, Newmaster et al. 2008). Janovec and Neill (2002) described two new species in the *Compsoneura atopa* complex and differentiated this species group from other *Compsoneura* based on the following combination of characters: (1) Secondary venation conspicuously brochidodromous (the secondary veins strongly anastomosing near the laminar margins); (2) Trichomes stellate to dendritic, often forming a ferrugineous tomentum on the young stem, leaf node, leaf lamina, petiole, external perianth surface, ovary, and pericarp surface; (3) Anthers free, their apices slightly incurved; (4) Oil cells abundant in the perianth and androecium, visible under low magnification and conspicuous when viewed in longitudinal or cross section; (5) Pericarp ligneous, strongly rostrate, and carinate with longitudinal ridges running from base to near apex; (6) Aril deeply lacinate or reduced, white to tan when fresh, tan to light brown when dry; (7) Testa brown, muricate, often with reticulate, vein-like impressions.

In the most recent systematic treatment of *Compsoneura*, Janovec (2000) opted to refer to informal sections Compsoneura and Hadrocarpa, which were based on multiple vegetative and reproductive characters, as opposed to Warburg's (1897) sectional

divisions based solely on free or fused anthers (hereafter non-italicized “Hadrocarpa” and “Compsoneura” will refer to these informal sections sensu Janovec 2000). Although morphological evidence indicated that Hadrocarpa and Compsoneura were distinct evolutionary lineages, Janovec (2000) and Janovec and Neill (2002) lacked molecular support. Such molecular evidence has been long in coming due to the difficulties of extracting, amplifying, and sequencing members of the Myristicaceae that posses potent DNA polymerase inhibitors (Sawadjoon et al. 2002, Maloney et al. 2005). These problems have been particularly difficult to overcome in *Compsoneura s.l.* as the vast majority of specimens are from old collections, which were not made with molecular investigations in mind. Additionally, molecular work of the family has been plagued by low rates of molecular evolution in loci traditionally used for phylogenetics (Sauquet et al. 2003).

The genus *Compsoneura* provides fertile ground for testing taxonomic hypotheses and DNA based identification methodologies as they are ecologically and ethnobotanically important, likely contains sister species of relatively recent origin (Janovec & Harrison 2002), and also has additional new cryptic species (Janovec et al. in preparation). The objectives of this research were (1) to test the sub-generic classifications of Warburg (*Eucompsoneura* and *Coniostele*) and Janovec (Compsoneura and Hadrocarpa) with molecular phylogenetic analyses and morphological comparisons and (2) to assess the ability of DNA sequences to differentiate species of *Compsoneura sensu lato*. Molecular phylogenetic analyses were performed using trnH-psbA, the most rapidly evolving cpDNA loci surveyed in a DNA barcoding study of *Compsoneura* (Newmaster et al. 2008), as well as two low copy nuclear loci (*AT103* and *AGT1*).

Materials and Methods

Taxon sampling and outgroup selection

Many species of neotropical *Compsoneura* are represented by one or very few herbarium specimens, most of which were collected over 20 years ago. It is well known that DNA extraction can be particularly difficult from such herbarium material (Savolainen et al. 1995, Lister et al. 2008, Telle and Thines 2008). This problem is exacerbated in members of the nutmeg family as they contain high levels of potent DNA polymerase inhibiting secondary metabolites and DNA degrading chemicals (Sawajadoon et al. 2002, Maloney et al. 2005). Due to the rare nature and poor condition of many collections, it was not possible to include all species of *Compsoneura* in molecular analyses. Forty-six collections representing 13 species of *Compsoneura* were used in the molecular analyses. Some sequences used in this analysis were retrieved from Genbank from previous studies of *Compsoneura* (Newmaster et al. 2008). Since DNA sequences were not retrieved from all collections nor all loci, the number of collections and species used in the phylogenetic and distance analyses differ. All collections used in molecular analyses were identified according to the most current taxonomic treatment (Janovec 2000) and all taxa have accessions archived at one or more of the following herbaria: Botanical Research Institute of Texas (BRIT), Ontario Agricultural College (OAC), New York Botanical Gardens (NYBG), Missouri Botanical Gardens (MO), and the National Herbarium of the Ecuadorian Museum of Natural Sciences (QCNE)[Table 1.1].

Since phylogenetic relationships remain unresolved in the Myristicaceae (Sauquet et al. 2003), members of *Iryanthera*, *Otoba* and *Virola* were selected as outgroups for phylogenetic and distance-based analyses.

DNA extraction, Amplification and Sequencing

Total genomic DNA was extracted from leaf tissue of silica-dried or herbarium specimens using the Macherey-Nagel Nucleospin II plant Kit. Lysis buffer 1 was used according to the manufactures' instructions with the exception of an increase of the post homogenization incubation period to 1hr (from 10 min) and the addition of 20 mM N-Phenacylthiazolium Bromide, which has been found to improve amplification of recalcitrant samples (Poinar et al. 1998, Asif and Cannon 2005).

Initially a suite of chloroplast (accd, atpF-H, matK, psbK-I, rbcL, rpoB, rpoC1, trnH-psbA, trnL-F, trnS-G, UPA, and ycf5) and nuclear (AGT1, APG1, AT103, EIF3E, GI, GS, ITS, IGS, PHYA, PHYC, and sqd1) loci were amplified to investigate their utility for systemic investigations of *Compsoneura*. However, only a small subset (trnH-psbA, AT103, and AGT1) consistently produced a single banded PCR product with high quality sequence traces and were sufficiently variable for systematic investigations.

Species not represented by trnH-psbA accessions in GenBank from Newmaster et al. (2008) were PCR amplified and sequenced using the primers trnH2 (5'-CGCATGGTGGATTCACAATCC-3'; Tate and Simpson 2003) and psbAF (5'GTTATGCATGAACGTAATGCTC-3'; Sang et al. 1997). PCR was performed in a 20 µl volume using 0.4 µl of Phire II polymerase (Finnzymes) with 1X Phire II reaction buffer (with 1.5mM MgCl), 0.2 mM of each DNTP, 0.2 µM of each primer and 2.0 µg of BSA (Kreader 1996). Cycling conditions entailed an initial denaturation step of 1 min at 98°; 35 cycles of 98° for 5 s, 64° for 5 s, 72° for 10 s: and a final elongation step of 72°

for 1 min followed by a 4° hold. Phire II was used to amplify trnH-psbA as it is robust to the inhibitors contained in nutmeg extracts and as it is a fusion-based polymerase which has been found to reduce stuttering in sequences containing homopolymer regions such as the trnH-psbA intergenic spacer (Fazekas et al. 2010).

Nuclear loci were amplified using the primers AT103F (5'-CTTCAAGCCMAAGTTCATCTTCTA-3'; Li et al. 2008), AT103R (5'-TTGGCAATCATTGAGGTACATNGTMACATA-3';), AGT1-MYR-F (5'-GGGCATTGACGTAGCTTTGACAGG-3'; this thesis), and AGT1-MYR-R (5'-GTGCAGTTCTTCAAGCCCCAAGC-3'; this thesis). Nuclear loci were amplified with 0.5U of AmpliTaq Gold (Applied biosystems) DNA polymerase in a 20 µl reaction containing 1X reaction buffer, 2.5 mM MgCl, 8% W/V Polyethylene glycol (Zimmerman and Harrison 1987, Sasaki et al. 2006) , 0.2 M trehalose (Speiss et al. 2004), 2 µg BSA, 0.2 mM each DNTP, and 0.2 µM of each primer.

Amplification products were sequenced directly using the same primers employed in PCR. Cycle sequencing reactions were performed in a 10.5 µL reaction volume containing 0.5 µL of BigDye terminator mix v3.1, 1.88 µL of 5x sequencing buffer (Applied Biosystems), 1.0 µM of primer and 0.5 µL of PCR product. Thermal cycling parameters were 96° for 2 min; 30 cycles of 96° for 30s, 56-60° (primer dependant) for 15s, and 72° for 4 min; and a 4° hold. Cycle sequencing reactions were cleaned using sephadex columns (Cat.no. S5897; Sigma-Aldrich, St. Louis, MO, USA) and the samples were run on an ABI 3730 sequencer (Applied Biosystems).

Alignment, phylogenetic and distance analysis

Sequence contigs were assembled and edited using Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA). DNA sequences were aligned using the default settings of the ClustalW algorithm (Thompson et al. 1994) in Bioedit (Hall 1999) and adjusted manually. Alignments can be found in Appendix 1 at the end of this thesis. A 4bp inversion believed to belong to the loop structure of the 3'UTR (untranscribed region) of the psbA gene (Storchova and Olsen 2007) was omitted from data matrices due to the high likelihood of homoplasious inversions. Gaps in the alignments were coded using Simple Indel Coding (SIC) (Simmons and Ochoterena 2000) for Bayesian analyses and Modified Complex Indel Coding for distance-based and parsimony analyses using Indelcoder (Muller 2006).

A previous study of *Compsoneura* by Newmaster et al. (2008) found the trnH-psbA intergenic spacer to be the most useful of numerous proposed chloroplast barcode loci. I amplified and sequenced 5 additional species of rarely collected *Compsoneura* from herbarium material and generated a Neighbour-Joining (NJ) uncorrected p-distance tree using PAUP 4.0 (Swafford 2002) to further explore the ability of this locus to discriminate species of *Compsoneura s.l.* with increased taxon sampling. Additionally, intra and interspecific distances were calculated using pair-wise uncorrected p-distance in Mega 3.1 (Kumar et al. 2008) for species with multiple sequences per taxon to investigate whether there exists a barcoding gap (Hebert et al. 2004a, Hebert et al. 2004b) in these taxa. Uncorrected p-distances were calculated as there existed a very low level of nucleotide substitution among taxa and therefore it was not deemed necessary to emloy a substitution model in calculating distance estimates.

An incongruence length difference test (ILD: Farris et al. 1995) was executed in PAUP (Swafford 2002) to determine whether the three loci employed in this study contained significant incongruence. The ILD was performed using 100 replicates of a heuristic search strategy, MAXTREES set at 100, 10 random addition sequence replicates holding 5 trees at each step, MULTREES option in effect, and tree bisection-reconnection (TBR) branch swapping. No significant incongruence was detected (p-value=0.17) and individual gene trees generated by Bayesian inference showed no significant topological discrepancies among the major clades, therefore all three loci were concatenated in a total evidence approach for Bayesian inference and parsimony analyses.

Maximum parsimony analyses were performed in PAUP 4.0b10 (Swafford 2002) and each nucleotide position was treated as an unordered character and all positions were equally weighted. The single heuristic search was performed with 1000 replicates of random addition sequence, holding 10 trees at each step, TBR branch swapping with the MULTREES option not in effect and MAXTREES set to 5000. Branch support was estimated using a heuristic search, 5000 bootstrap (BS) replicates, 10 random addition sequence replicates holding 1 tree at each step, TBR branch swapping, and a maximum number of trees set at 50,000 with MULTREES not in effect.

Phylogenetic trees were also generated using Bayesian inference with the program MrBayes (Ronquist and Huelsenbeck 2003). Nucleotide substitution models were selected using Mrmodeltest (Nylander 2004) using the Akaike Information Criterion. Mrmodeltest selected the following substitution models for the three loci: trnH-psbA= HKY+G, *AT103*=GTR+G, *AGT1*=GTR+G. Ten million generations were performed

using 4 chains and 2 runs with trees being sampled every 100 generations. Log-likelihood values stabilized after 2.5 million generations; therefore the consensus tree and posterior probabilities (PP) were estimated using a burn-in of 25,000 trees. Since posterior probabilities are largely considered to be overestimations of support for clades (Douady et al. 2003, Alfaro and Holder 2006, Yang and Rannala 2010), especially when employing concatenated data sets (Suzuki et al. 2002), any nodes with less than 0.75 posterior probability were collapsed on the phylogram using treegraph2 (Stover and Muller 2010).

Additionally, gross morphological differences and similarities between the informal sections Hadrocarpa and Compsoneura were observed in the field, herbaria and literature (Janovec 2000, Janovec and Harrison 2002). Differentiating traits of the leaves, flowers, inflorescence, fruit and indument of all these parts were synthesized from the most recent taxonomic treatment (Janovec 2000) and subsequent species descriptions (Janovec and Harrison 2002, Janovec and Neil 2002).

Results

DNA sequencing and alignment

Table 1.1 presents specimens included in various genetic analyses that were successfully amplified for one or more loci along with their Barcode of Life Database (BOLD) process identification numbers that can be used to access sequence and specimen info online at www.barcodinglife.org. Genbank accession numbers are included in Table 1.1 for specimens whose trnH-psbA sequences were retrieved from Newmaster et al. (2008). The trnH-psbA spacer alignment for phylogenetic analyses consisted of 342

nucleotides, 39 variable characters and 18 parsimony informative characters. The AGT1 gene alignment consisted of 710 nucleotides, 59 variable and 34 parsimony informative sites while the AT103 gene alignment consisted of 412 nucleotides, 31 variable characters and 22 parsimony informative sites. The concatenated alignment contained 1464 characters of which 129 were variable and 74 parsimony informative nucleotide positions. The outgroups contributed 89 variable and 29 parsimony informative characters. The MCIC algorithm employed in indelcoder coded 14 separate indel events.

Phylogenetic Analyses

A total of 14 collections representing 8 species of *Compsoneura* were included in the phylogenetic analyses as high-quality DNA was not extracted from all samples. Trees constructed from individual loci varied slightly in their overall topology, mostly due to the markers varying information content, but in all cases the monophyly of informal sections Hadrocarpa and Compsoneura was recovered and supported by high (1.0) posterior probabilities (data not shown). Additionally, no significant incongruence was detected with an ILD test, therefore the loci were combined in an all evidence approach. The 50% majority-rule Bayesian consensus tree with posterior probabilities is shown in Figure 1.1. The most parsimonious phylogram estimated by PAUP recovered a nearly identical topology (Figure 1.2). A single most parsimonious tree was found with a length of 187, a consistency index (CI) of 0.85, a homoplasy index (HI) of 0.15, and a retention index (RI) of 0.88. Both Bayesian and parsimony analyses of the three regions revealed strong support of the monophyly of *Compsoneura* s.l. as measured by posterior probabilities (1.0) and bootstrap support (99). There was also strong support (PP=1.0,

BS=99) for the nodes separating informal sections Hadrocarpa and Compsoneura. However, it appears that fused anthers have evolved multiple times in *Compsoneura* (Figures 1.1 and 1.2), making Warburg's sections *Eucompsoneura* and *Coniostele* paraphyletic in nature. Additionally, there was a considerable branch length separating Hadrocarpa and Compsoneura compared to the relatively short branches within each group, indicating a significant amount of evolutionary divergence between these two informal sections.

Morphology

Species of the two sections share a few morphological similarities such as tertiary veins growing more or less perpendicular to the midvein, an entire endosperm, and generally glabrous leaves. However, thirteen contrasting morphological characters were found to differentiate Hadrocarpa from Compsoneura (Table 1.2).

Distance analysis

This study was able to obtain sequences from 5 additional species of *Compsoneura* than were analysed by Newmaster et al. (2008). These 5 species are relatively rare and poorly collected members of section Hadrocarpa (sensu Janovec 2000), which were collected between 1972 and 1996. All species with the exception of *C. cuatrecasasi* and *C. carchifolia* can be identified by one or more single nucleotide polymorphisms (Figure 1.3). It is also interesting to note that *C. capitellata, C. sprucei*, and *C. mexicana* exhibit a high degree of molecular variation relative to other species. Intraspecific distances were found to sometimes be larger than interspecific distances and

therefore there was no barcoding gap found in the taxa tested (Figure 1.3). Additionally, both sections Hadrocarpa and Compsoneura were found to be monophyletic in this analysis and within section Hadrocarpa, species from Western Colombia and Panama formed their own clade due to a 9-bp indel as well as a shared single nucleotide polymorphism.

Discussion

Taxonomy is most biologically relevant when it seeks to classify and identify organisms with respect to their evolutionary history. This study represents the most comprehensive molecular analysis of the genus *Compsoneura* to date. Previous work by Newmaster et al. (2008) assessed levels of molecular variation of different cpDNA markers in the genus on a more restricted suite of taxa but did not conduct a phylogenetic analysis. Janovec (2000) investigated the microsatellite allele variation amongst populations of South and Central American *Compsoneura sprucei* which, together with a morphometric analysis, led to the resurrection of *C. mexicana* (Janovec and Harrison 2002). With the use of markers from the chloroplast and nuclear genomes, my analyses revealed a significant molecular divergence that is mirrored by morphological adaptations of sections Hadrocarpa and Compsoneura (sensu Janovec 2000) that I feel are sufficient to warrant the recognition of these sections as separate respective genera. The following discussion will address this as well as implications of the distance analysis.

Phylogenetic Analyses

Previous studies of *Compsoneura* have highlighted numerous morphological and biogeographical differences between informal sections Hadrocarpa and Compsoneura but did not include molecular analyses (Janovec 2000). In this study I present the first phylogenetic estimation of a genus of the Myristicaceae, thus providing some illumination of the relationships between members of *Compsoneura*. Very high posterior probabilities (1.0) and bootstrap support (99) was found on the nodes that separate sections Compsoneura and Hadrocarpa, indicating they are monophyletic. It is also interesting to note that both phylogenetic analyses indicated that Warburg's sections *Eucompsoneura* (free anthers) and *Coniostele* (fused anthers) are paraphyletic, as fused anthers appear to have evolved multiple times within *Compsoneura*. Additionally, similar branch lengths were recovered separating the two sections of *Compsoneura* as were found between *Otoba-Compsoneura* and *Otoba-Iryanthera*. This suggests that the two sections have been on diverging evolutionary paths for a substantial amount of time. The two sections also have distinct biogeographic patterns. Members of Compsoneura are widespread in the Amazon basin, the Western coast of the Andes as well as Central America. Members of Hadrocarpa on the other hand primarily reside on the Western side of the Andes in Ecuador, Colombia and Panama with the vast majority of species being endemic to Western Colombia.

It thus appears that the informal sections Hadrocarpa and Compsoneura represent a more natural sub-generic classification of the genus compared to Warburg's (1897) sections *Eucompsoneura* and *Coniostele* when both molecular and morphological data are considered. There exists so many dichotomous morphological and molecular characters among section Hadrocarpa and Compsoneura, that I suggest that the former be

elevated to the generic rank so as to recognize the morphological and molecular distinctiveness of these clades. I believe that this is important as end-users of taxonomy use field characters not only to recognize species, but also to help group collections and as an aid to memory. If recognized formally, I proposed that the genus be named Hadrocarpa (derived from the greek hadros, or thick, and carpa, or seed/fruit) as large fruits with thick ligneous pericarps are characteristic of these trees.

The Myristicaceae remain poorly understood from a phylogeographic perspective due to a paucity of data and low rates of nucleotide substitution that has plagued previous investigators (Sauquet et al. 2003). The only molecular phylogenetic analysis of the family conducted by Sauquet et al. (2003) was unable to recover well supported relationships among genera of the family despite using some of the most variable loci commonly used in low-level angiosperm phylogenetics (trnL-trnF, ndhF, trnK, and matk). I elected to use *Iryanthera* as an outgroup to my analyses as they are the only dioecious (assumed ancestral state) members of the family and *Iryanthera* leaves appear to be the most similar of any Myristicaceae to the sister of the family, *Idiospermum australiense* (Calycanthaceae). Initial efforts were made to use *Idiospermum australiense* as an outgroup, however, difficulty in amplification and alignment of sequences prevented its inclusion. Despite this uncertainty of rooting and therefore character polarity, there remain numerous interesting and well supported relationships that have been revealed with this analysis.

The well supported *C. mutisii-C. mexicana-C. excelsa* clade was recovered in all trees made from the three individual loci (data not shown) and is of particular phylogeographic interest. It appears that the species of this clade arose through a series of

allopatric speciation events initiated by the uprise of the Andean Orogeny. These speciation events would likely have been initiated by the rise of the Andean range resulting in the separation of current *C. mutisii* from its Amazonian progenitor populations (likely *C. ulei* or *C. sprucei*). This event was followed by dispersal of an ancestor of *C. mutisii* to Central America (likely via birds, their known dispersal vector) to a forming Central American archipelago or after the formation of the Panamanian Isthmus (Raven and Axelrod 1974). After this event, the up-rise of a mountain range in present day Costa Rica separated populations on the East and West side of the forming cordillera. This vicariance event likely led to the differentiation of the species we know as *C. excelsa* and *C. mexicana* respectively. It is also interesting to note that a morphological cladistic analysis of *Compsoneura* by Janovec (2000) found that *C. ulei* consistently grouped within the *mutisii-excelsa-mexicana clade*, a group which shares the trait of fused or partially fused anthers on the filament column. The somewhat low support (PP=0.81, BS=51) on the node separating *C. ulei* and *C. debilis* from the other members of the genus may be due to the fact that the most variable locus used in this study, *AGT1*, was not successfully sequenced for *C. ulei*. Recovery of *AGT1* or other variable low copy nuclear markers may provide increased resolution and support of infrageneric relationships within section Compsoneura. Although the position of *C. ulei* is not well supported, it has a fused anther column like *C. mutisii* and *C. excelsa* but the rest of the Compsoneura and Hadrocarpa clades have free anthers. The phylogenetic estimation created here requires multiple conversions from the state of free to fused anthers. Great importance has been placed on the androecium in morphological cladistics since it is one of the few variable characters of *Compsoneura* (Janovec 2000) as well as

other genera (Sauquet et al. 2003). If the phylogenetic estimation presented here is found to be robust with future data, it may mean that fusion of the androecium is potentially homoplasious.

Although there were relatively few species included from section Hadrocarpa, due to the old age of most collections and difficulty in extracting high quality DNA for the amplification of low copy nuclear sequences, there remains a few interesting trends to note. The first is the apparent paraphyletic nature of the cpDNA of *C. capitellata*. Due to difficulties in amplifying herbarium material, the *C. atopa* specimen used in this analysis came from the most recent collection which was found growing in a coffee plantation in the vicinity of Comuna 24 de Mayo, Ecuador, in 2000. This tree also constitutes one of only a few collections of *C. atopa* in the Amazonian basin as it is endemic to Colombia West of the Andes. Despite numerous attempts to locate additional individuals, there have been only two collections made of this species east of the Andes to date. Therefore it is possible that this tree might be one of the last remnants of a refugial Amazonian population of *C. atopa* and may have hybridized to some extent with *C. capitellata*. Additional data from chloroplast and nuclear markers will be required to test this hypothesis as the most variable marker (AGT1) was not recovered for *C. atopa* in this study. Additionally there appears to be a great deal of molecular divergence within the species *C. capitellata* which will be discussed below and in a subsequent chapter.

A robust phylogenetic analysis will require known outgroups, however, this is difficult due to the unresolved phylogeny of the Myristicaceae (Sauquet et al. 2003) and the fact that the family likely diverged from sister taxa about 100 mybp. This large amount of evolutionary time may lead to a great deal of homoplasy and long stem

lengths, potentially leading to long branch attraction in the highly variable markers necessary to resolve intergeneric relationships of the Myristicaceae. Although it cannot be assured that this is a properly rooted phylogeny, the topology recovered here, however, makes intuitive sense based on the apparent pattern of trans-andean speciation of some taxa of *Compsoneura*. Additionally, the 3 markers employed in this study had substantially more variation than others used by previous studies and may be informative for future studies of Myristicaceae phylogeny.

Distance Analysis

The second objective of this study was to perform an analysis of the ability of a chloroplast locus to distinguish members of *Compsoneura*, including rare collections and samples already sequenced by Newmaster et al. (2008). The morphological uniformity, slow rates of molecular evolution, and dearth of molecular markers from previous investigators has made the taxonomy of this group particularly difficult. The inclusion of 5 additional taxa to the 8 taxa studied by Newmaster et al. (2008) resulted in a trnH-psbA NJ tree where all but one species pair (*C. carchifolia* and *C. cuatrecasasi*) are differentiated by unique haplotypes. These two species species are easily distinguished morphologically by dissimilar leaf and flower morphologies and also have disjunct distributions.

Although most species are rather morphologically coherent, there appears to be a relatively large divergence among the sequences of different populations of some species (*C. sprucei*, *C. mutisii*, *C. mexicana*, and *C. capitellata* for example). This pattern may be due to a number of non-mutual exclusive phenomena. The first possibility is that these

divergent populations may represent cryptic species. This is very possible as recent molecular studies of tropical insects (Hebert et al. 2004b, Smith et al. 2007, Smith et al. 2008) have revealed previously unnoticed alpha diversity. It is highly possible that there exists undescribed species of neotropical Myristicaceae considering the cryptic morphological nature of these trees and the wide range of climatological and edaphic regimes present across the Amazon basin. If these populations are not continuous it is also possible that these geographically isolated haplotypes represent restricted gene flow between discontinuos populations. Since the chloroplast genome is inherited from the maternal parent and spread via seed, this pattern could be generated by restricted seed dispersal and/or restricted seed establishment between populations. If these genetic polymorphisms are indeed an indicator of intra-specific haplotype diversity and not cryptic speciation it will prove difficult to unambiguously identify new collections by means of DNA barcodes as the genetic distance within a species may be equal or greater than that between currently recognized species. Despite these problems, trnH-psbA will be a very useful tool to biologists to help identify sterile, herbarium and fragmentary material of *Compsoneura* since morphological means of identification may lead to mis-identification rates of 25-60% (J. Janovec personal communication, R. Steeves personal observations). Additional studies with increased population-level sampling are recommended to sample haplotype diversity across the geographic range of these species; of particular interest will be the study of the more hypervariable taxa such as *C. sprucei*, *C. mexicana*, and *C. capitellata*.

Althougth the primary function of the distance analysis was to assess the ability of the trnH-psbA locus to discriminate species of *Compsoneura*, there are also some

relationships on the tree (Figure 1.3) of biogeographical significance worthy of discussion. Two members of Janovec's (2000) informal subgroup Atopa (*C. rigidifolia* and *C. sp1*) form a clade (Choco-Panama clade Figure 1.3) with the exception of the groups namesake, *C. atopa*. Likewise, members of Janovec's (2000) informal subgroup Capitellata included in this analysis (*C. capitellata*, *C. cuatrecasasii*,*C. diazii, and C. sp2*) are also paraphyletic in this tree. Members of subgroup Atopa are endemic to Western Colombia and Panama and those of subgroup Capitellata are found in both the East and Western slopes of the Andes. Members of subgroup Atopa and Capitellata have a number of vegetatitive and reproductive morphological tratits which distinguish each respective subgroup (Janovec 2000) but it appears that the similarity of their trnH-psbA haplotypes is primarily correlated to geographic region (Amazonian or Choco-Panama) rather than morphological subgroupings (sensu Janovec 2000). This apparent incongruence of morphological affinities of taxa and their haplotypes similarities is possibly due to convergent morphological evolution but is more likely a result of cpDNA paraphyly due to introgression of the maternally inherited cpDNA. Chloroplast haplotype introgression between species, also known as chloroplast sharing or chloroplast capture, is well documented in plants by phylogenetic and phylogeographic studies and appears to be relatively common (Rieseberg and Soltis 1991, Golden and Bain 2000, Palme et al. 2003, Palma-Silva et al. 2011). This phenomena may also account for high levels of intraspecific genetic distance found in plant plastid regions (Hollingsworth et al 2011). In the future it would be desirable to compare nuclear and plastid phylogenies of *Compsoneura* (subgroups Atopa and Capitellata in particular) to establish whether

introgression or convergent morphological evolution are likely the cause of discrepancy between morphological and molecular data.

Tables

Table 1.1 *Compsoneura* and outgroup taxa used in all analyses. Collector(s), Collections number (Coll. #), code of Herbaria storing (Herb.), Country of origin, decimal degrees latitude (Lat.), decimal degrees longitude (Long.), Barcode of Life Database sequence process ID (BOLD #), and genbank accession numbers are listed if applicable.

Species	Collector (s)	Coll. #	Herb.	Country	Lat.	Long.	BOLD #	Genbank #
C. atopa	J.P. Janovec and W. Quizhpe	1374	NY	Ecuador	-0.47	-77.26	RSMYR002-11	EU090622.1
C. capitellata	R.A.D. Steeves	527	OAC	Peru	-13.23	-70.78	RSMYR003-11	
C. capitellata	R.A.D. Steeves	551	OAC	Peru	-13.23	-70.78	RSMYR004-11	
C. capitellata	J.P. Janovec	835		Peru	-6.15	-76.17	RSMYR005-11	EU090623.1
C. capitellata	J.P. Janovec	855		Peru	-3.52	-73.15	RSMYR006-11	EU090624.1
C. capitellata	J.P. Janovec	872		Peru	-3.52	-73.15	RSMYR007-11	EU090625.1
C. capitellata	J.P. Janovec	875		Peru	-3.52	-73.15	RSMYR008-11	EU090626.1
C. capitellata	J.P. Janovec and W. Quizhpe	889		Ecuador	-1.04	-77.37	RSMYR009-11	EU090627.1
*C. sp.*1	C. Aulestia et al.	848	QCNE	Ecuador	0.55	-78.32	RSMYR010-11	
C. cuatrecasasi	A.H. Gentry and A. Juncosa	40682	MO	Colombia	3.15	-77.25	RSMYR011-11	
C. diazii	C. Diaz et al.	7644	MO	Peru	-5.03	-78.22	RSMYR012-11	
C. rigidifolia	G. Mcpherson	10008	Duke	Panama	9.15	-79.30	RSMYR013-11	
*C. sp.*2	A.H. Gentry and E. Renteria	23826	MO	Colombia	5.30	-76.33	RSMYR014-11	
C. debilis	M.L. Kawasaki	190	CAS	Brazil	0.18	-68.40	RSMYR015-11	EU090628.1
C. debilis	P.E. Berry et al.	6172	MO	Venezuela	2.52	-67.18	RSMYR016-11	EU090629.1
C. debilis	P.E. Berry	7209		Venezuela			RSMYR017-11	EU090630.1
C. debilis	R.L. Liesner and G. Carnevali	22972	MO	Venezuela	1.51	-67.03	RSMYR018-11	EU090631.1
C. excelsa	J.P. Janovec and R. Aguilar	636	NY	Costa Rica	8.43	-83.12	RSMYR019-11	EU090632.1
C. excelsa	J.P. Janovec and R. Aguilar	666	NY	Costa Rica	8.44	-83.28	RSMYR020-11	EU090633.1
C. excelsa	J.P. Janovec	668	NY	Costa Rica	8.26	-83.22	RSMYR021-11	EU090634.1
C. excelsa	J.P. Janovec	669	NY	Costa Rica	8.26	-83.22	RSMYR022-11	EU090635.1
C. excelsa	J.P. Janovec	671	NY	Costa Rica	8.26	-83.22	RSMYR023-11	EU090636.1
C. mexicana	V. Tzub	007					RSMYR024-11	EU090637.1
C. mexicana	J.P. Janovec	354	TAMU	Costa Rica	10.25	-83.29	RSMYR025-11	EU090638.1
C. mexicana	J.P. Janovec	362	TAMU	Costa Rica	10.25	-83.29	RSMYR026-11	EU090639.1

Species	Collector (s)	Coll. #	Herb.	Country	Lat.	Long.	BOLD	GB
C. mexicana	J.P. Janovec and A. Neill	696	NY	Belize	16.23	-89.08	RSMYR027-11	EU090640.1
C. mexicana	J.P. Janovec and A. Neill	701	NY	Belize	16.23	-89.08	RSMYR028-11	EU090641.1
C. mexicana	J.P. Janovec and A. Neill	719	NY	Belize	16.12	-89.03	RSMYR029-11	EU090642.1
C. mexicana	J.P. Janovec and A. Neill	720	NY	Belize	16.12	-89.03	RSMYR030-11	EU090643.1
C. mexicana	J.P. Janovec and J.A. Janovec	757	NY	Belize	16.20	-89.10	RSMYR031-11	EU090644.1
C. mexicana	J.P. Janovec	1283					RSMYR032-11	EU090645.1
C. mutisii	J.P. Janovec and Quizhpe	911	NY	Ecuador	1.03	-78.32	RSMYR033-11	EU090646.1
C. mutisii	J.P. Janovec and Quizhpe	913	NY	Ecuador	1.03	-78.32	RSMYR034-11	EU090647.1
C. mutisii	J.P. Janovec and Quizhpe	914	NY	Ecuador	1.03	-78.32	RSMYR035-11	EU090648.1
C. mutisii	J.P. Janovec	1290					RSMYR036-11	EU090649.1
C. mutisii	J.P. Janovec	1295					RSMYR037-11	EU090650.1
C. sprucei	J.P. Janovec and A. Pena Cruz	812	NY	Peru	-6.06	-76.11	RSMYR038-11	EU090651.1
C. sprucei	J.P. Janovec and A. Pena Cruz	817	NY	Peru	-6.06	-76.11	RSMYR039-11	EU090652.1
C. sprucei	J.P. Janovec and A. Pena Cruz	821	NY	Peru	-6.06	-76.11	RSMYR040-11	EU090653.1
C. sprucei	J.P. Janovec and W. Quizhpe	884	NY	Ecuador	-1.04	-77.37	RSMYR041-11	EU090654.1
C. sprucei	J.P. Janovec and W. Quizhpe	887	NY	Ecuador	-1.04	-77.37	RSMYR042-11	EU090655.1
C. sprucei	J.P. Janovec and W. Quizhpe	903	NY	Ecuador	-1.04	-77.37	RSMYR043-11	EU090656.1
C. ulei	E. Lima and A. Silva	088	NY	Brazil	-3.50	-49.42	RSMYR044-11	EU090657.1
C. ulei	C.R. Sperling et al.	6192	US	Brazil	-5.49	-50.32	RSMYR045-11	EU090658.1
C. ulei	M. Nee	42644	NY	Brazil	-2.25	-59.54	RSMYR046-11	EU090659.1
O. glycycarpa	R.A.D. Steeves	546	OAC	Peru	-13.23	-70.78	RSMYR047-11	
O. parvifolia	R.A.D. Steeves	598	OAC	Peru	-13.23	-70.78	RSMYR048-11	
V. surinamensis	R.A.D. Steeves	078	OAC	Peru	-13.23	-70.78	RSMYR049-11	
I. juruensis	R.A.D. Steeves	451	OAC	Peru	-13.23	-70.78	RSMYR050-11	
I. laevis	R.A.D. Steeves	460	OAC	Peru	-13.23	-70.78	RSMYR051-11	

Table 1.2 Unambiguous morphological character transformations that differentiate informal sections Compsoneura and Hadrocarpa.

Compsoneura	**Morphological Character**	**Hadrocarpa**
Drying green, sometimes brown	Leaves	Drying brown, rarely olive or brownish-green
Free, arching (venation eucamptodromous)	Secondary leaf veins	Anastomosing marginally, at least in part (venation weakly to strongly brochidodromous)
Free from above the base, semi-connate, or connate	Anthers	Free from the base
Present	Glandular trichomes on inner flower perianth, filaments, and filament columns	Absent or rarely present in one species (*C. diazii*)
Absent	T-shaped or stellate trichomes on vegetative and reproductive organs	Present
Absent, or few and inconspicuous	Floral oil cells	Abundant and conspicuous
Elliptic to ovate-elliptic	Fruit shape	Sub-globose, elliptic-ovate, to strongly rostrate
Orange, thin, fleshy	Fruit pericarp at maturity	Green to brown, thick, fortified, ligneous
Always smooth	Fruit pericarp surface	Sulcate to strongly furrowed, or carinate to strongly ridged
Strongly dehiscent along one longitudinal line	Fruit pericarp dehiscence	Indehiscent to partially or "forced" dehiscent
Glabrous	Fruit pericarp pubescence	Pubescent throughout or at least in part; hairs t-shaped to irregularly stellate
Thick, entire, somewhat fleshy, bright red to scarlet when mature	Aril	Thin, entire to deeply laciniate from base, or rudimentary, reduced, and sometimes seemingly absent; white when mature but rarely seen because the fruit pericarps are largely indehiscent

Smooth, gray, black to purple mottled	Seed testa	Rough to rugose, weakly furrowed to veined, brown, non-mottled

Figures

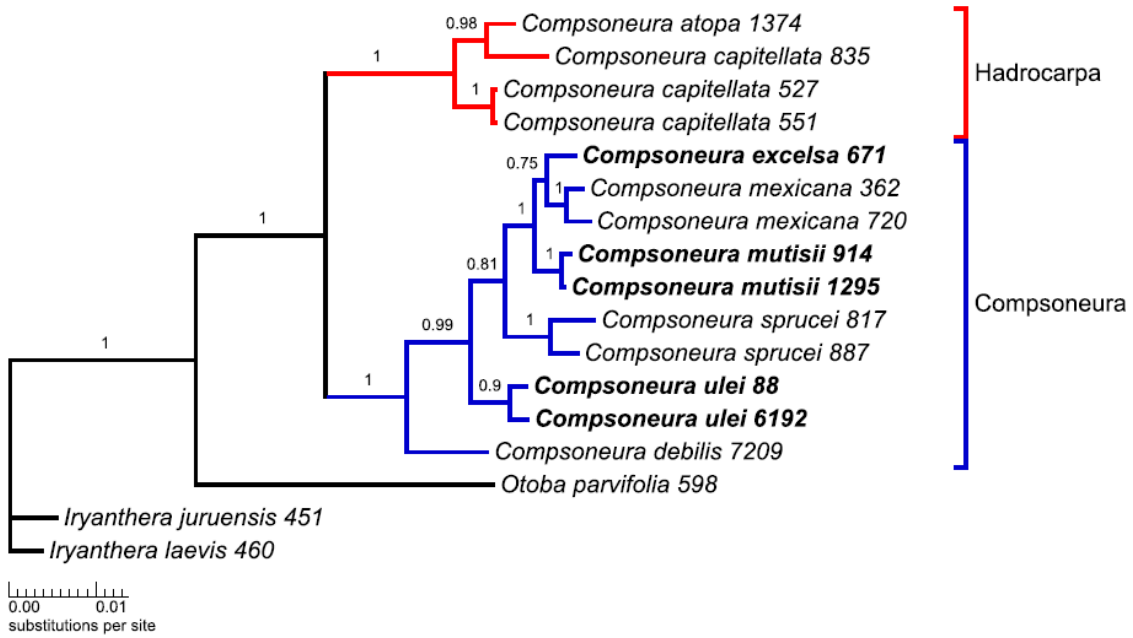

Figure 1.1 Phylogram of consensus tree created with sequences of trnH-psbA, AT103 and AGT1 sequences of 8 species of *Compsoneura*. Numbers on branches indicate Bayesian posterior probabilities of nodes and numbers following taxon names are the collection numbers of herbarium samples from which DNA was extracted (see table 1.1 for collection information). Branch colours indicate sections Hadrocarpa (red) and Compsoneura (blue). Bolded taxa possess fused anthers. Branch lengths are proportional to distance.

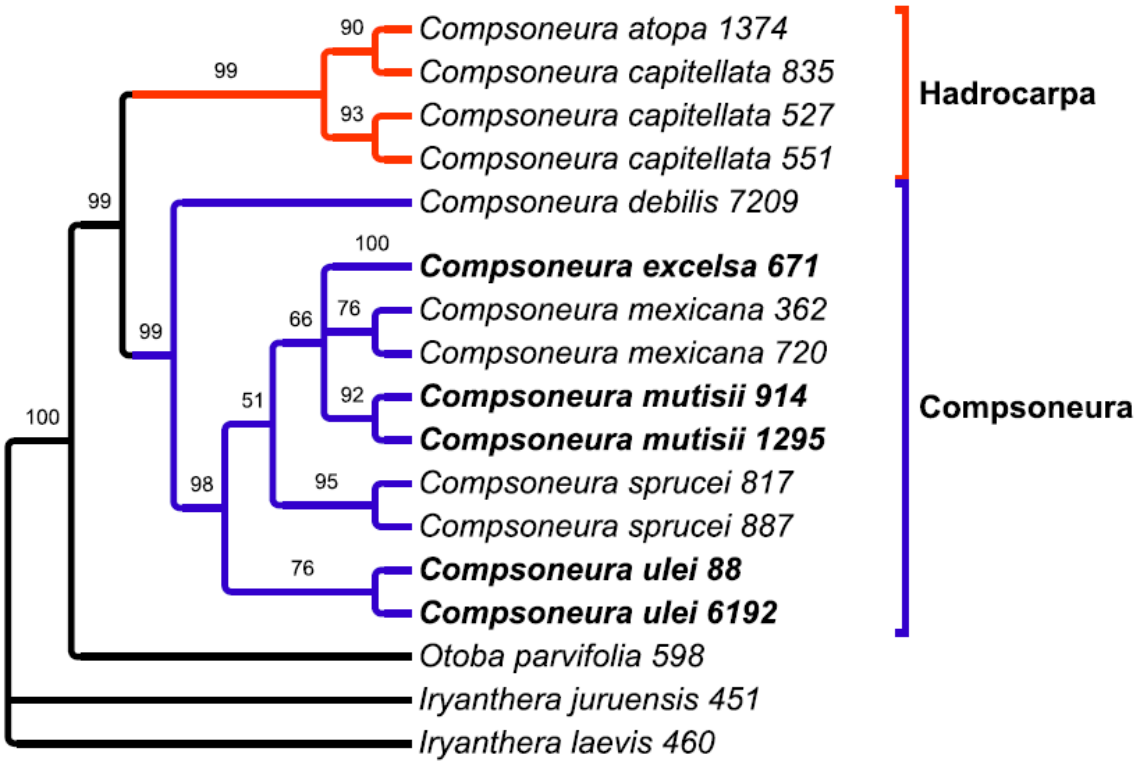

Figure 1.2 Phylogram of the single most parsimonious tree (Length=187) constructed from a concatenated alignment of trnH-psbA, AGT1 and AT103 data sets (CI=0.85, RI=0.88, HI=0.15). Numbers on branches indicate bootstrap values and numbers following taxon names are the collection numbers of herbarium samples from which DNA was extracted (see table 1.1 for collection information). Branch colours indicate sections Hadrocarpa (red) and Compsoneura (blue). *Compsoneura* taxa in bold font possess fused anthers and non-bolded taxa possess free anthers.

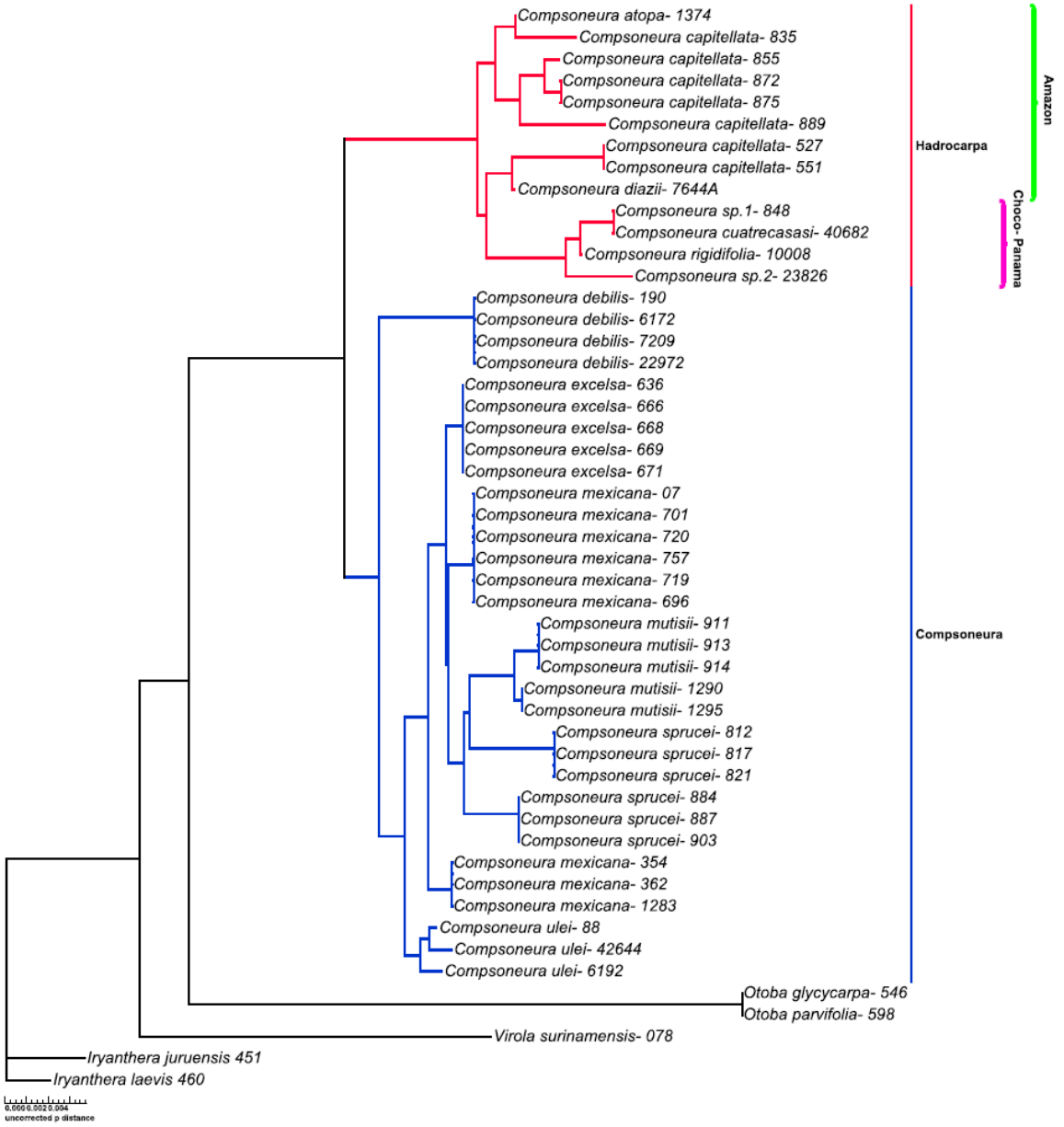

Figure 1.3 Neighbour joining (uncorrected p-distance) tree constructed using trnH-psbA data. Red branches indicate members of section Hadrocarpa and blue branches indicate section Compsoneura. Numbers indicate collection numbers of herbarium samples from which DNA was extracted (also see table 1.1). Geographical zones (Amazonian or Choco-Panama) of various species of Hadrocarpa is indicated by lines on the right. Branch lengths are proportional to distance.

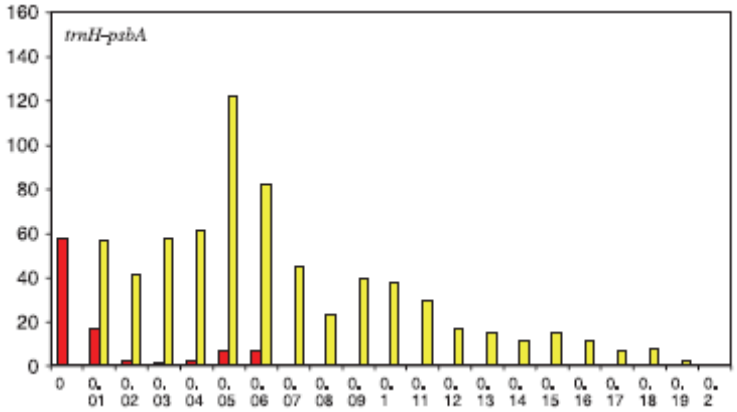

Figure 1.4 Histograms of the number of pairwise intraspecific (red bars) and interspecific divergence distances estimates (x-axes) among all *Compsoneura* samples included in the neighbour joining analysis. The number of pairwise comparisons is on the y-axis and the pairwise p-distance for the trnH-psbA region is on the x-axis.

Literature Cited

Alfaro, M. E., & Holder, M. T. 2006. The posterior and the prior in Bayesian phylogenetics. *Annual review of ecology, evolution, and systematics* **37**: 19-42.

APG III. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II.

Armstrong, J. E., & Tucker, S. C. 1986. Floral development in *Myristica* (Myristicaceae). *American Journal of Botany* **73**: 1131–1143.

Armstrong, J. E. 1997. Pollination by deceit in nutmeg (*Myristica insipida*, Myristicaceae): floral displays and beetle activity at male and female trees. *American journal of botany* **84**: 1266-1274.

Asif, M. J., & Cannon, C. H. 2005. DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*). *Plant Molecular Biology Reporter* **23**: 185–192.

Bullock, S. H. 1982. Population structure and reproduction in the neotropical dioecious tree *Compsoneura sprucei*. *Oecologia* **55**: 238–242.

Candolle, A. D. 1856. Myristicaceae. *Prodromus Systematis Naturali Vegetablis* **14**: 187-208.

Chagnon, N. A., Le Quesne, P., & Cook, J. M. 1971. Yanomamo¨ Hallucinogens: Anthropological, Botanical, and Chemical Findings. *Current Anthropology* **12**: 72–74.

Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F., & Douzery, E. J. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution* **20**: 248-254.

Farris, J. S., Kallersjo, M., Kluge, A. G., & Bult, C. 1995. Testing significance of incongruence. *Cladistics* **10**: 315-319.

Fazekas, A. J., Steeves, R., & Newmaster, S. G. 2010. Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques* **48**: 277–285.

Gentry, A. H. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* **69**: 557–593.

Golden, J. L., & Bain, J. F. 2000. Phylogeographic patterns and high levels of chloroplast DNA diversity in four *Packera* (Asteraceae) species in southwestern Alberta. *Evolution* **54**: 1566–1579.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series*. (pp. 95–98).

Hebert, P. D. N., Stoeckle, M. Y., Zemlack, T. S., & Francis, C. M. 2004. Identification of birds through DNA barcodes. *PLOS Biology* **2**: 1657-1663.

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

Hollingsworth, P. M., Graham, S. W., & Little, D. P. 2011. Choosing and using a plant DNA barcode. *PloS one* **6**: e19254.

Janovec, J. P., & Harrison, J. S. 2002. A morphological analysis of the *Compsoneura sprucei* complex (Myristicaceae), with a new combination for the Central American species *Compsoneura mexicana*. *Systematic Botany* **27**: 662–673.

Janovec, J. P., & Neill, A. K. 2002. Studies of the Myristicaceae: an overview of the *Compsoneura atopa* complex, with descriptions of new species from Colombia. *Brittonia* **54**: 251–261.

Janovec, J. P. 2000. A systematic study of *Compsoneura* (A. DC.) Warb., A Neotropical member of the nutmeg family. Texas A&M University Dissertation: 1-359.

Kreader, C. A. 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology* **62**: 1102-1106.

Kumar, S., Nei, M., Dudley, J., & Tamura, K. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in bioinformatics* **9**: 299-306.

Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H., & Varotto, C. 2008. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* **24**: 727–745.

Lister, D. L., Bower, M. A., Howe, C. J., & Jones, M. K. 2008. Extraction and amplification of nuclear DNA from herbarium specimens of emmer wheat: a

method for assessing DNA preservation by maximum amplicon length recovery. *Taxon* **57**: 254–258.

Macedo, D. S., & Anderson, A. B. 1993. Early ecological changes associated with logging in an Amazon floodplain. *Biotropica* **25**: 151-163.

Maloney, D. J., Deng, J., Starck, S. R., Gao, Z., & Hecht, S. M. 2005. (+)-Myristinin A, a naturally occurring DNA polymerase β inhibitor and potent DNA-damaging agent. *Journal of the American Chemical Society* **127**: 4140-4141.

McKenna, D. J., Towers, G. H. N., & Abbott, F. S. 1984. Monoamine oxidase inhibitors in South American hallucinogenic plants part 2: Constituents of orally-active Myristicaceous hallucinogens. *Journal of Ethnopharmacology* **12**: 179–211.

Müller, K. 2006. Incorporating information from length-mutational events into phylogenetic analysis. *Molecular phylogenetics and evolution* **38**: 667–676.

Newmaster, S. G., Fazekas, A. J., Steeves, R. A. D., & Janovec, J. 2008. Testing candidate plant barcode regions in the Myristicaceae. *Molecular Ecology Resources* **8**: 480–490.

Nguyen, P. H., Le, T. V. T., Kang, H. W., Chae, J., Kim, S. K., Kwon, K., Seo, D. B., Lee, S. J., & Oh, W. K. 2010. AMP-activated protein kinase (AMPK) activators from *Myristica fragrans* (nutmeg) and their anti-obesity effect. *Bioorganic and Medicinal Chemistry Letters* **20**: 4128-4131.

Nylander, J. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Palma-Silva, C., Wendt, T., Pinheiro, F., Barbará, T., FAY, M. F., Cozzolino, S., & Lexer, C. Sympatric bromeliad species (*Pitcairnia spp.)* facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* 20: 3185–3201

Palmé, A. E., Semerikov, V., & Lascoux, M. 2003. Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity* **91**: 465–474.

Pascal, J. P., & Pelissier, R. 1996. Structure and floristic composition of a tropical evergreen forest in south-west India. *Journal of Tropical Ecology* **12**: 191–214.

Patro, S. P., Tyagi, M., Saha, J., & Chattopadhyay, S. 2010. Comparative nuclease and anti-cancer properties of the naturally occurring malabaricones. *Bioorganic and Medicinal Chemistry* **18**: 7043-7051.

Poinar, H. N. 1998. Molecular coproscopy:dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281**: 402-406.

Poulsen, A., Nelson, I., Tan, S., & Balsev, H. 1996. A quantitative inventory of trees in one hectare of mixed dipterocarp forest in Temburong, Bnmei Darussalam. In *Tropical rainforest research: current issues*. London: Kluwer Academic Publishers (pp. 139-150).

Pusztai, R., Abrantes, M., Sherly, J., Duarte, N., Molnar, J., & Ferreira, M. J. 2010. Antitumor-promoting Activity of Lignans: Inhibition of Human Cytomegalovirus IE Gene Expression. *Anticancer research* **30**: 451-454.

Qiu, Y-L., Li.,L., Hendry, T.A., Li, R., Taylor, D.W., Issa, M.J., Ronen, A.J., Vekaria, M.L., & White, A.M. 2006. Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes. *Taxon* **55**:837-856.

Raven, P. H., & Axelrod, D. I. 1974. Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* **61**: 539-673.

Rieseberg, L.H., & Soltis, D.E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**: 65-84.

Rodrigues, W. A. 1989. Two new neotropical species of *Compsoneura* (Myristicaceae). *Brittionia* **4**: 160-163.

Ronquist, F., & Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574.

Sang, T., Crawford, D. J., & Stuessy, T. F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). *American Journal of Botany* **84**: 1120-1136.

Sasaki, Y., Miyoshi, D., & Sugimoto, N. 2006. Effect of molecular crowding on DNA polymerase activity. *Biotechnology Journal* **1**: 440–446.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W., & Le Thomas, A. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* **142**: 125–186.

Savolainen, V., Cuénoud, P., Spichiger, R., Martinez, M. D., Crèvecoeur, M., & Manen, J. F. 1995. The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. *Plant Systematics and Evolution* **197**: 87–98.

Sawadjoon, S., Kittakoop, P., Kirtikara, K., Vichai, V., Tanticharoen, M., & Thebtaranonth, Y. 2002. Atropisomeric myristinins: Selective COX-2 inhibitors and antifungal agents from *Myristica cinnamomea*. *The Journal of Organic Chemistry* **67**: 5470-5475.

Schultes, R. 1981. Iconography of the New World Hallucinogens. *Arnoldia* **41**: 80-125.

Simmons, M., & Ochoterena, H. 2000. Gaps as characters in sequence-based phylogenetic analysis. *Systematic Botany* **49**: 369-381.

Smith, A. C. Studies of South American plants: XV. *American Journal of Botany* **43**: 573-577.

Smith, A. C. 1937. The American species of Myristicaceae. *Brittonia* **2**: 393-510.

Smith, M. A., Rodriguez, J. J., Whitfield, J. B., Deans, A. R., Janzen, D. H., Hallwachs, W., & Hebert, P. D. 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proceedings of the National Academy of Sciences* **105**: 12359-12364.

Smith, M. A., Wood, D. M., Janzen, D. H., Hallwachs, W., & Hebert, P. D. N. 2007. DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera,Tachinidae) are not all generalists. *Proceedings of the National Academy of Sciences* **104**: 4967-4972.

Soltis, D. E., Smith, S. A., Cellinese, N., Wurdack, K. J., Tank, D. C., Brockington, S. F., Refulio-Rodriguez, N. F., Walker, J. B., Moore, M. J., Carlsward, B. S., & others. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* **98**: 704.

Speiss, A., T. Mueller, & Ivell, R. 2004. Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of Taq polymerase by the disaccharide trehalose. *Clinical chemistry* **50**: 1256-1259.

Spichiger, R., Loizeau, P., Latour, C., & Barriera, G. 1996. Tree species richness in a southwestern Amazonian forest. *Candollea* **51**: 559-577.

Štorchová, H., & Olson, M. S. 2007. The architecture of the chloroplast psbA-trnH non-coding region in angiosperms. *Plant Systematics and Evolution* **268**: 235-256.

Stöver, B. C., & Müller, K. F. 2010. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC bioinformatics* **11**: 1-9.

Suzuki, Y., Glazko,G.V., & Nei, M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences* **99**: 16138-16143.

Swafford, D. 1999. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4 . Sunderland : Sinauer Associates.

Tajuddin, S. A., Latif, A., Qasmi, I., & Amin, K. 2005. An experimental study of sexual function improving effect of *Myristica fragrans* Houtt.(nutmeg). *BMC Complementary and Alternative Medicine* **5**: 16-22.

Tate, J. A., & Simpson, B. B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* **28**: 723–737.

Telle, S., & Thines, M. 2008. Amplification of cox2 (~620 bp) from 2 mg of Up to 129 Years Old Herbarium Specimens, Comparing 19 Extraction Methods and 15 Polymerases (R. DeSalle, Ed.). *PLoS ONE* **3**: e3584.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**: 4673-4680.

Van Gils, C., & Cox, P. A. 1994. Ethnobotany of nutmeg in the Spice Islands. *Journal of Ethnopharmacology* **42**: 117–124.

Warburg, O. 1897. Monographie der Myristicaceen. *Nova Acta Acad. Caes. Leop.-Carol* **68**: 1-680.

Wilde, W. J. J. O. D. 1991. The genera of the Myristicaceae as distinguished by their inflorescences and the description of a new genus: *Bicuiba*. *Beitr. Biol. Pflanzen* **66**: 95-125.

Yang, Z., & Rannala, B. 2010. Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences* **107**: 9264-9269.

Zimmerman, S. B., & Harrison, B. 1987. Macromolecular crowding increases binding of DNA polymerase to DNA: an adaptive effect. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 1871-1875.

Chapter 2

A MORPHOLOGICAL AND MOLECULAR INVESTIGATION OF THE *COMPSONEURA CAPITELLATA* (A. DC) Warb. COMPLEX

Abstract

Members of *Compsoneura* are canopy-subcanopy trees native to tropical rainforests of Central and South America. *Compsoneura capitellata* is a common and widespread member of the Myristicaceae family found in both cloud and lowland rainforests of Northwestern Amazonia. The flowers, fruits and leaves of members of the *Compsoneura capitellata* complex exhibit a high degree of morphological variation but the species has not received careful study since 1937 when few specimens were available and even the fruits were unknown. Many specimens of *C. capitellata* have since been collected from the entire species' range. The aim of this study was to investigate patterns of morphological and genetic diversity in the *C. capitellata* complex. Morphological analyses of leaf, perianth and androecial characters revealed little differentiation among populations, however, leaf characters exhibited some differentiation according to habitat (montane versus lowland populations). In contrast, DNA sequencing of 60 *C. capitellata* individuals, from 7 populations, with the trnH-psbA intergenic spacer revealed 9 haplotypes, with each population containing only haplotypes unique to that locale. I speculate that the apparent morphological continuum and contrasting genetic sub-division across the geographic landscape is potentially a result of restricted seed dispersal, historic anthropogenic use, cryptic speciation, or a combination of these phenomena. Herein I

present *C. morona-santiagoensis*, a provisional new species endemic to the sandstone substrates of Cutucu and the Cordillera del Condor of Ecuador and suggest the resurrection of *C. tessmannii* (Markgraf), a large-leaved taxa from the mountainous cloud forests of Peru.

Introduction

The Myristicaceae, or nutmeg family, is a pantropical assemblage of canopy to subcanopy trees comprised of 21 genera and about 500 species. These trees are native to cloud and lowland tropical rainforests where they are often among the most diverse and abundant tree families of those environments (Gentry 1988, Pitman et al. 2001, Pitman et al. 2002). Six genera, *Bicuiba*, *Compsoneura*, *Iryanthera*, *Ostephloem*, *Otoba*, and *Virola* are native to the Central and South America. A recent investigation of *Compsoneura* (chapter 1 this thesis) revealed evidence of a deep evolutionary divergence between informal sections Hadrocarpa and Compsoneura (sensu Janovec 2000) supported by numerous morphological and molecular characters. Members of *Compsoneura* are typified by having mostly glabrous leaves with brochidodromous or eucamptodromous secondary veins, tertiary venation perpendicular to the midvein and are endemic to Central and South America.

Compsoneura capitellata is a common and widespread dioecious tree species that inhabits old growth cloud and lowland rainforests of Northwest Amazonia (Brazil, Colombia, Ecuador, and Peru) from 100 m-1500 m elevation above sea level (asl). *Compsoneura capitellata* was one of the first species of *Compsoneura* described by de Candolle (1856), however it has been studied little since that time. The flowers, fruits and leaves of *Compsoneura capitellata* exhibit a high degree of morphological diversity but

the species has not received careful study since the last monographic treatment by Smith (1937). At that time very few specimens were available for study (n=11) from only a few localities and the only fruit sample available (Ducke 19576, B, K, S, US, Utr) was deemed so atypical (i.e. a thick woody pericarp and large non-arillate seed) that it couldn't belong to Myristicaceae (Smith 1937), although the fruit in question was that of *C. capitellata* (Janovec 2000). Smith had very few fertile specimens to study and elected to ascribe vegetatively diverse collections from the Northwest Amazon as *C. capitellata* and synonymized the previously described *C. tessmannii* (Markgraf), following the monospecific concept of Warburg (1897).

Many collections of *C. capitellata* have been made since the last taxonomic treatment, which has expanded the species' known range and revealed a considerable degree of morphological variation. The chartaceous to thick coriaceous, simple, and alternate leaves of this species are elliptic-ovate in shape, being 2-13 cm wide (at half length) and 4-50 cm long with the smallest leaves generally occurring in the lowlands and the larger-leafed varieties generally occurring at higher elevations. The bark exhibits a wide range of colours and textures; from pale grey to brown to red and peeling off in large brittle plates or long fibrous slivers. The staminate flower is typically composed of 3-4 tepals forming a globose to cupuliform flower 1.4-2.5 mm long and 1.1-1.8 mm wide with an androecium composed of 5-12 anthers which are 0.4-0.7 mm in length. Fruits of this species are typically 2-6 cm long and 2-5 cm wide. The fruit is composed of a greenish, woody, indehiscent or forced-dehiscent pericarp which is 3-7 mm thick, glabrous to ferruginous-tomentellous, green to reddish brown in colour, and smooth, muricate, or linearly to irregularly sulcate in texture. A thin white aril, which is

sometimes greatly reduced, typically covers the large dull brown seed (1.5-4 cm), which is comprised of an entire, white, endosperm. *Compsoneura capitellata* is often confused with *C. sprucei* collections as *C. capitellata* is only reliable differentiated by its thick, woody pericarp, white aril, and sometimes by the thickness of the leaves and more strongly arcuate secondary veins compared to C. sprucei.

While other members of Myristicaceae have been the subject of ecological investigations (Howe and Vande Kerkhove 1981, Bullock 1982, Cole 2009, Holbrook and Loiselle 2009), members of *Compsoneura* have been the subject of few ecological studies (Bullock 1982). *C. capitellata* has yet to be the focus of any detailed ecological studies and thus mechanisms of pollination, vectors of dispersal, and habitat preference are virtually unknown (Janovec personal communication). The large geographic and altitudinal range that this species inhabits translates into a wide range of light regimes, temperatures, precipitation, and substratum encountered by populations. Although the pollinators of *C. capitellata* are not known, the small (1-4mm) size of the yellow flowers and casual observations suggests that small generalist pollinators such as thrips (Thysanoptera) and beetles (Coleoptera) pollinate these dioecious trees (personal observation) as has been found in the nutmeg of commerce, *Myristica fragrans* (Armstrong and Irvine 1989). Informal questioning of local foresters has failed to identify a known disperser of the large, indehiscent, white arillate fruits.

More intensive collection and specimen-based studies of herbarium material are needed to improve our understanding of tropical species of plants (Tobler et al. 2007, Thomas 1999). Examination of field and herbarium specimens by J. Janovec and R. Steeves revealed a relatively high degree of leaf and fruit morphological variation in *C. capitellata* compared to other congeners. Such specimen studies quickly lead to the realization that our

understanding of the *C. capitellata* complex is hampered by lack of both morphological and nucleotide sequence data. In fact, no intensive survey of floral morphological patterns and genetic variation had ever been conducted in this group, or *Compsoneura* in general. Although the tiny flowers of *C. capitellata* have been broadly characterized by a syndrome of free anthers and presence of a filament column, variation in relative size, shape, and positioning of anthers and filament columns was observed in field and herbaria. However, these androecial characteristics have never been studied with quantitative techniques.

Levels of gene flow, inbreeding, and genetic differentiation among tropical plants are of prime interest to tropical ecologists, conservationists and taxonomists. Although the Northwest Amazon, to which *Compsoneura capitellata* is endemic, is considered to be the most botanically diverse region of the world (Prance 1977, Gentry 1982) it is rather poorly understood from a genetic perspective, especially with respect to the Myristicaceae.

Previous molecular investigations of *Compsoneura* have revealed considerable nucleotide sequence variation in *C. capitellata* (Newmaster et al. 2008, Chapter 1 of this thesis). The aim of this study was to examine the morphological and molecular variation of this species in order to investigate whether genetic isolation exists between populations. If genetic diversity is greater among than within populations then it is more likely that reproductive isolation has occurred and morphological diversity may represent cryptic speciation.

Materials and Methods

Morphological data sampling

In order to quantify morphological differences, measures of 10 leaf, 7 androecium and 9 perianth dimensions were made (Figure 2.1) from collections for many locations

across *C. capitellata*'s range. Leaf measurements were made by measuring two leaves of herbarium samples with a ruler or by measuring photographs of pressed samples with a digital ruler in the program TPSdig2 (Rohlf 2006). Likewise, perianth and androecial metrics were measured using the Miniscale (Bioquip Inc., 2000), a miniature 5 mm ruler demarcated in 0.10 mm increments that can be used to gather quantitative data under a microscope, or from high resolution digital images of androecia which were captured using a Hirox microscope and associated software (Hirox-USA, Hackensack, NJ). Minimum and maximum measurements of mature structures on each specimen for the 26 metrics were recorded to document the range of leaf sizes within a population. Since species of *Compsoneura* are dioecious, only a subset of the total number of specimens possessed perianths and/or staminate flowers. Of the 268 specimens available for this study (Table 2.1), only 41 possessed staminate flowers. SEM micrographs and light microscope images were also compiled to demonstrate variation in androecial characters.

Morphological data analysis

A principal component analysis (PCA) was performed on leaf morphometric data using Canoco 4.5 (ter Braak, 1998) to identify the length of the ordination axis and unimodal ordination model was applied (Correspondence Analysis, CA). The relationship between quantitative leaf characters was analysed via nonmetric multi-dimensional scaling (NMS; Kruskal 1964; Primer 2002). In NMS, the Bray-Curtis distance measure was used because of its robustness for both large and small scales on the axes (Minchin 1987). Data were standardized by species maxima and two-dimensional solutions were appropriately chosen based on plotting a measure of fit ('stress') to the number of dimensions. Stress represents distortion in the data and a stress value over 0.2 is high

enough that the results are invalidated (Primer 2002). One thousand iterations were used for each NMS run, using random start coordinates. The first two ordination axes were rotated to enhance interpretability with the different axes. As an independent check, detrended correspondence analysis (DCA; ter Braak 1998) was used to evaluate the NMS classification. A Pearson-Correlation analysis was performed to investigate which characters contributed the most to the differentiation of the samples.

DNA extraction, amplification, sequencing and alignment

Like many members of the Myristicaceae, PCR-amenable genomic DNA of *C. capitellata* is very difficult to obtain from old collections or even fresh herbarium material that hasn't been rapidly dried and maintained in a desiccated state (personal observations). Almost all species of *Compsoneura* dry a dark brown colour, presumably due to the oxidation of secondary metabolites released from vacuoles as the leaves dry. Therefore DNA was sampled mostly from recent collections (< ~10 years old) and/or had been immediately placed in silica gel upon collection in the field. In attempts to capture the inherent morphological and genetic variability encountered in this species, seven populations were sampled from a wide area of the species' range in Ecuador (n=4) and Peru (n=3) spanning 1550 km and 1200 m of elevation. One specimen each of *C. atopa* and *C. diazii* were included in analyses as a previous genetic study (Chapter 1 of this thesis) has found a close relationship with *C. capitellata*. Figure 2.2 indicates collection locations and elevations of sampled populations.

Total genomic DNA was extracted from leaf tissue of silica-dried or herbarium specimens using the Macherey-Nagel Nucleospin II plant Kit. Lysis buffer 1 was used

according to the manufactures' instructions with the exception of the elongation of the post homogenization incubation period to 1hr (from 10 minutes) and the addition of 20mM N-Phenacylthiazolium Bromide which has been found to improve amplification of recalcitrant samples (Poinar et al. 1998, Asif and Cannon 2005).

Initially a suite of chloroplast (matK, psbK-I, rbcL, rpoB, rpoC1, trnH-psbA, UPA, and ycf5) and nuclear (AGT1, AT103, ITS, IGS, PHYA, PHYC, and sqd1) loci were amplified to investigate their utility for infra-specific studies of *C. capitellata*. However, only trnH-psbA was both consistently amplified and contained numerous polymorphisms. Five *C. capitellata* trnH-psbA accessions were retrieved via GenBank from Newmaster et al. (2008). An additional 55 samples were PCR amplified and sequenced using the primers trnH2 (5'-CGCATGGTGGATTCACAATCC-3'; Tate and Simpson 2003) and psbAF (5'GTTATGCATGAACGTAATGCTC-3'; Sang et al. 1997) or the custom designed primer pair trnH-Myrist (5'-TTGATCCACTTGGCTACATCC-3'; this thesis) and psbA-Myrist (5'- GACCTAGCTGCTGTTGAAGC-3':this thesis). PCR was performed in a 20 µl volume using 0.4 µl of Phire II polymerase (Finnzymes) with 1X Phire II reaction buffer (with 1.5mM MgCl), 0.2 mM of each DNTP, 0.2 µM of each primer and 2.0 µg of BSA (Kreader 1996). Cycling conditions entailed an initial denaturation step of 1 min at 98°; 35 cycles of 98° for 5 s, 64° for 5 s, 72° for 10 s: and a final elongation step of 72° for 1 min followed by a 4° hold. Phire II was used to amplify trnH-psbA as it has been found to be robust to the inhibitors contained in nutmeg extracts and as it is a fusion-based polymerase which has been found to reduce stuttering in regions containing homopolymer sequences such as the trnH-psbA intergenic spacer (Fazekas et al. 2010).

Amplification products were sequenced directly using the primers trnH-Myrist and psbA-Myrist. Cycle sequencing reactions were performed in a 10.5 μL reaction volume containing 0.5 μL of BigDye terminator mix v3.1, 1.88 μL of 5x sequencing buffer (Applied Biosystems), 1.0 μM of primer and 0.5 μL of PCR product. Thermal cycling parameters were 96° for 2min; 30 cycles of 96° for 30s, 56° for 15s, and 72° for 4min; and a 4° hold. Cycle sequencing reactions were cleaned using sephadex columns (Cat.no. S5897; Sigma-Aldrich, St. Louis, MO, USA) and the samples were run on an ABI 3730 sequencer (Applied Biosystems).

Sequence contigs were assembled and edited using Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA). DNA sequences were aligned using the default settings of the ClustalW algorithm (Thompson et al. 1994) in Bioedit (Hall 1999) and adjusted manually. Three of four nucleotides of a loop structure believed to be the 3' untranscribed region (UTR) of the psbA gene (Storchova and Olsen 2007) was omitted from the data matrix. Stem-loop secondary structures of DNA have a high likelihood of homoplasious inversions, even at low taxonomic ranks (Golenberg et al. 1993, Graham et al. 2000, Shaw et al. 2005). DNA sequence alignments can be found in appendix 2. Gaps in the alignments were coded using Modified Complex Indel Coding for distance analyses using Indelcoder (Muller 2006). All sequences generated for the molecular analyses have been deposited in the Barcode of Life Database (BOLD) [www.barcodinglife.org] and their BOLD process identification numbers (which can be used online to access sequence data, images, as well as additional specimen information) are found in Table 2.3.

Statistical Parsimony analyses

In attempts to estimate the phylogeographic history of *C. capitellata*,, haplotype networks were constructed by statistical parsimony (Templeton et al. 1992, Templeton 1998) using the program TCS (Clement et al. 2000), which calculates the number of differences between haplotypes that are due to a sequence of single mutations at each site. Networks were calculated with TCS as implemented by ANeCA (Panchal 2007) by treating gaps as a 5th character state and with a 95% confidence interval. Haplotype networks may be preferable to traditional phylogenetic methods at the intraspecific level because population level data often violate many of the assumptions made by traditional tree-building methods, such as Maximum Parsimony, Maximum Likelihood and Bayesian Analyses (Posada and Crandall 2001). In comparison, networks are able to utilize haplotypic data that contains low levels of genetic divergence, ancestral haplotypes, multiple mutational variants from a given haplotype within the population, and reticulation that may be caused by recombination or hybridization (Templeton et al. 1992; Posada and Crandall 2001). Since loops are not theoretically possible in haplotype networks of non-recombining regions such as plastid DNA, the connection deemed the most unlikely connecting haplotypes of a loop was removed.

Results

Morphology

Leaf length in *C. capitellata* varied between 3.6 and 51.7 cm long and the width at half length ranged from 1.7-13.5 cm. The ordination analyses of quantitative leaf metrics did not reveal that any *C. capitellata* regional populations were distinct when all regions are considered. However, there are many specimens from Amazonas, and Cusco, Peru

that were found to cluster together on the NMS ordination (Figure 2.3). Pearson correlation analyses revealed that the X-axis is strongly correlated with 8 measures of leaf length and width (Figure 2.3; Table 2.2) and the y-axis was mostly correlated to leaf vein number and space (Figure 2.3; Table 2.2). In order to investigate whether populations isolated within Andean valleys and at high elevations exhibit morphological affinities of their leaves, collections were coded as either Montane or Lowland and plotted using the NMS ordination values (Figure 2.4). It was evident from this plot that there is a general trend for montane populations to have larger leaves compared to lowland populations (Figure 2.5).

Anther number ranged from 4-9 per androecium and the individual mature anthers were 0.4-1.6 mm long. Much of this variation in androecial characters is demonstrated in the SEM and light microscope images of Figure 2.6. Ordination analyses revealed three distinct clusters that were due to variation in minimum and maximum anther number, which was variable within almost all regions sampled. Pearson correlation analyses showed that the x-axis was strongly correlated to max/min anther number, length and cluster width while the y-axis was mostly correlated to those same measures as well as anther sac length (Figure 2.7; Table 2.2).

Perianths of *C. capitellata* were 0.4-2.0 mm at their widest, 1.1-3.2 mm long and were on petioles 0.5-3.0 mm long and 0.1-1.0 mm wide. No particular groupings were evident from the ordination, but two peripheral points (from Peru-Huanuco and Ecuador-Napo populations) belonged to the only female flowers inlcuded the analysis (Figure 2.8). Pistillate perianths also had the thickest pedicels at 1.0 mm, presumably owing to the mechanical stresses placed upon the pedicel during fruit development.

The fruits of the *C. capitellata* complex were found to be relatively consistent in size but highly variable in pericarp texture. Pericarps ranged from green to ferruginous, glabrous to tomentellous, and smooth to sulcate or warty in appearance (Figure 2.9). Although few fruits were available for study, pericarp morphology appeared to be consistent within populations.

Molecular analyses

Bi-directional trnH-psbA intergenic spacer sequences were obtained for 59 *Compsoneura capitellata* individuals representing 7 populations and single specimens of *C. atopa* and *C. diazii*. Aligned sequences were 302 bp in length and contained 13 variable sites including one indel originating from an A/T homopolymer run 9-12 bp in length. Table 2.3 documents the species determinations, collection numbers, locations and Barcode of Life Database process ID numbers of collections used in the molecular analyses.

The haplotype network estimated by TCS at a 95% confidence level revealed 10 haplotypes (Figure 2.10). Each haplotype was restricted to a particular sampling location and no evidence was found of haplotypes being shared among two populations. The only locality that showed intra-population variability, Loreto, had three haplotypes that varied by only by either a single nucleotide polymorphism or a single base insertion in the longest homopolymer tract. A branch connecting the Zamora-Chinchipe and Morona-Santiago populations was removed as the sole difference between these three populations is a GAAG, GAAA, or CTTC (5'-3') sequence of the apical loop of the 3' UTR (untranslated region) and the most probable mutational pattern for this region is for GAAG↔CTTC through an inversion and GAAG↔GAAA through a single transition.

Likewise a branch connecting the Morona-Santiago population to an unsampled/ancestral haplotype was removed as it was more likely to have originated from an extant and nearby population than an unsampled haplotype.

Discussion

This study represents the first intraspecific morphological and molecular investigation of the genus *Compsoneura*, representing 266 and 60 specimens examined for each respective analysis. Although the findings of this investigation raise many questions about the taxonomic status of the *C. capitellata* complex, I will hypothesize causes of the patterns of morphological and molecular variation observed across the species' range and provide an account of a provisional new and resurrected species from the *Compsoneura capitellata* complex.

Morphology

Since leaves are always present on male, female and juvenile trees they should be very important taxonomically, however, they are notoriously similar in morphology among species and even genera of Myristicaceae. Despite great differences in leaf size among some populations of *C. capitellata*, the ordination analyses indicate that leaf metrics alone cannot unambiguously differentiate populations when specimens from the entire range of the species are included. However, there was a general trend observed where montane populations showed larger leaves and fewer veins compared to their lowland conspecifics. Many of these montane specimens are from the Peruvian Andean provinces of Cusco, Amazonas and Huanuco and Junin. These populations are not only at higher altitudes (500-1000 m asl), but are typically separated from the lowland Amazon

basin by small mountain ranges. Even relatively small mountain ranges and associated changes in altitude are considered to be important generators of speciation in the tropics (Janzen 1967) so the contrasting sizes of leaves in montane and lowland may be of taxonomic significance. The leaves of trees from Morona-Santiago, Ecuador were not distinctive in size from other populations of *C. capitellata* but were distinct in the fact that their laminas were domed between the secondary veins rather than plane as in all other specimens examined.

Characteristics of the androecium have traditionally been regarded as the most important taxonomic characters of *Compsoneura* (Smith 1937, Janovec 2000). Although the ordination of androecial characters has three distinct groups, these groupings are caused mainly by anther number differences, with each cluster containing most populations. It is possible that these clusters are a result of infraspecfic variation in androecial characters rather than incipient speciation but anther number is rarely characteristic of *Compsoneura* species (Janovec 2000).

Characteristics of the perianth have been given little taxonomic attention with respect to Myristicaceae species but multivariate analyses have never been performed to verify their taxonomic utility in the family. Although this analysis was primarily infraspecific, it appears to confirm that quantitative perianth characters show little variation and form a continuum and are unlikely to be of great taxonomic utility.

Even though many additional specimens have been collected since Smith's (1937) monograph of *Compsoneura*, few populations of *C. capitellata* have mature staminate/pistillate flowers as well as fruits available for analysis. The relatively few fruits available for study were wildly divergent in pericarp morphologies. Fruit specimens

from the San Martin population were smooth, green and glabrescent, those of Loreto were Ferruginous-tomerntellous and longitudinally sulcate, while a single fruit found in the Pastaza population was ferruginous-tomentellous and reticulated-furrowed with a warted appearance. Although it appears that fruit morphology is conserved within populations, so few fruits are available for study that it is not evident whether this holds true for all populations. Future collections of *C. capitellata* should make deliberate efforts to collect and document multiple fruits within populations

Genetic analyses

Compsoneura atopa* and *Compsoneura diazii

C. atopa and *C. diazii* were included in this investigation since previous phylogenetic investigations of *Compsoneura* had found these taxa to be similar to *C. capitellata* with respect to their trnH-psbA haplotypes. Perhaps one of the most surprising revelations of the molecular sequence data was that this *C. atopa* specimen differed by only a single nucleotide polymorphism (SNP) when compared to three populations of Ecuadorian *C. capitellata* and by two SNP's from a Peruvian *C. capitellata* population. *C. atopa* is a large tree (35 m or more tall) that has been collected extensively on the Western slopes of the Andes in Colombia and Ecuador but is known from only two collections east of the Ecuadorian Andes. *C. atopa* differs from *C. capitellata* in a number of morphological traits including the possession of brochidodromous venation, persistent abaxial leaf pubescence, secondary vein number (12-27 versus 4-12), a dense indument of the perianth, apical anther spurs, and the presence of bracteoles on inflorescences. The differences between these two species are so numerous and pronounced that Janovec (2000) subdivided the informal section Hadrocarpa into

subgroups Atopa and Capitellata, which were named after these species. Therefore, it was truly surprising to find that *C. atopa* differed by only 1 or 2 SNP's from 4 different populations of *C. capitellata*, especially considering that these SNP's are primarily due to hairpin-loop inversions or homopolymer run indels. It is also interesting that this specimen's haplotype is most similar to *C. capitellata* from southeastern Ecuador (120-608 km distant) rather than the nearby *C. capitellata* population (only 38 km distant). A previous investigation (Chapter 1 of this thesis) that employed the trnH-psbA spacer for some members of subgroup Atopa (*C. atopa*, *C. rigidifolia,* and *C. sp1*) found that *C. atopa* (J. J. Janovec 1374), *C. cuatrecasasi* (AHG 40682), and *C. sp* 2 (AHG 23826) are paraphyletic with respect to other members of their subgroups. The two *C. atopa* individuals that have been found east of the Andean Cordillera may either represent a population that predates the uplift of the Andes or a more contemporary migration event. The seed source for such a dispersal event could have come from the western slopes of Colombia or the 20 km wide, 1874 m asl, Las Cruces mountain pass which connects the Choco region of Colombia with the lowland Amazon via the Magdalena valley. Although other members of subgroup Atopa have been collected in the Magdalena Valley, this latter suggestion seems unlikely as *C. atopa* has yet to be collected there. It is also interesting to note that the seeds of *C. atopa* are boiled and roasted by indigenous peoples of the Baudo region of Colombia (La Rotta 1985) so it is conceivable that humans may have played a role in the dispersal of this species across the Andes. Plastid haplotype sharing, or chloroplast capture (Hollingsworth et al 2011), is well document in many plant taxa (Rieseberg and Soltis 1991, Golden and Bain 2000, Palma et al 2003) even in the presence strong morphological dissimilarity among species (Palma-Silva et al. 2011).

It is therefore possible that haplotype similarity among the morphologically dissimilar *C. capitellata*, *C. atopa,* and *C. diazii* observed in this study is due to relatively recent hybridization between these species resulting in plastid introgression. If the low collection rates of *C. atopa* east of the Andes are an indicator of its population density then hybridization with the more numerous *C. capitellata* trees may have been more likely. In light of these findings, I plotted the collection locations of *C. atopa* on a map of precipitation regimes in South America (Figure 2.11) to see if there was any obvious trend. It appears that *C. atopa* is restricted to areas of high precipitation (>4000mm/year). Therefore future collections in unexplored regions of Amazonian Ecuador and Colombia may find additional populations of this species. Future investigations should seek additional collections of *C. atopa* from throughout its western and eastern Andean range as well as more variable chloroplast and multiple bi-parentally inherited markers in order to investigate the origin of Amazonian *C. atopa*.

Compsoneura diazii was separated from the *C. capitellata* complex as a new species by Janovec (2002) so it is of little surprise that *C. diazii* is nested well within the haplotype network. It is somewhat surprising, however, that the *C. diazii* haplotype was most similar to the Pastaza, Ecuador population of *C. capitellata* rather than the less geographically distant populations of *C. capitellata* from San Martin, Peru, and South Western Ecuador (Zamora-Chinchipe and Morona-Santiago).

Compsoneura capitellata

All sampled populations of *C. capitellata* appear to be fixed for one or very few haplotypes. Only one sampled population, Loreto, exhibited SNP's within its population but each of these other haplotypes consisted of a single SNP relative to the most common

haplotype. This is somewhat surprising given that relatively high levels of within population diversity has been found in previous studies of tropical tree cpDNA using restriction polymorphisms (Caron et al. 2000, Latouche-Halle et al. 2003). However, the low variability found in this study was not unexpected given that Sauquet et al. (2003) found very low levels of nucleotide substitution in cpDNA coding and non-coding regions amongst genera of Myristicaceae. Once the loops in the haplotype network were removed it is evident that the most northerly populations (Napo, Pastaza, Morona-Santiago, Zamora-Chinchipe and San Martin) have the fewest nucleotide differences between them as compared to the more isolated populations (at least with respect to sampling patterns) of Loreto and Cusco. Given the apparent fixation of haplotypes for each population it is difficult to tell if these populations represent a large panmictic assemblage of populations of a single morphologically and molecularly diverse species or numerous isolated populations and/or distinct undescribed species. These questions are particularly difficult given the low levels of molecular variation available and missing morphological characters for many populations, and few distinguishing characters when present.

Of central importance to understanding and interpreting patterns of molecular and morphological variation is the natural history of the species of study. Such high levels of population differentiation can be easily generated in maternally inherited DNA, such as plastid DNA, if propagule dispersal is restricted. The dispersal agents of these large-seeded trees remains unknown to date but it is possible that these trees were dispersed by now extinct large mammals such as the giant ground sloths or gomphotheres (Janzen and Martin 1982). The fruits of *C. capitellata* are covered by a thick, woody, and indehiscent

to forced-dehiscent pericarp that protects a single large seed (2-4 cm in diameter) covered by a thin, white and fatty aril (Janovec 2000). These fruits fit many of the criteria of the Pleistocene megafaunal dispersal syndrome outlined by Janzen and Martin (1982). The largest mammals currently found in the range of *C. capitellata* that could act as potential dispersal agents are white-tailed and Brocket's deer (*Mazama spp.)*, tapir (*Tapirus spp*.), as well as spider (*Ateles spp*.), howler (*Alouatta spp*.) and wooly monkeys (*Lagothrix spp*.) but it seems highly unlikely that any of these species would act as effective long distance dispersers as it would necessitate them swallowing the 2-4 cm seeds, which is likely too large for them to consume intact. This hypothesis could be tested with greenhouse studies to see whether scarification and the application of stomach-like acids have a positive effect on the germination potential of *C. capitellata* as well as field studies to investigate the rate of seed removal, and predation versus dispersal.

There remains one other extant indigenous mammal capable of dispersing the seeds of these trees: *Homo sapiens*. Humans have also likely had a great impact on the Amazonian flora and fauna for the last 10,000 years (Janzen 1983, Roosevelt et al. 1996, Paz-Riviera and Putz 2009). It has been found relatively recently that a charcoal enriched soil called *Terra Pret*a, which is found in large areas (perhaps 20% of the Brazilian Amazon) of Brazil adjacent Amazon river, was created by humans between 600-8700 years BP (Smith 1980, Liang et al. 2008). *Terra Preta* vastly changes the soil microbiota (Kim et al. 2007, Grossman et al. 2010) and increases a soils cation-exchange capacity which helps to trap nitrogen within the soil and release it and other nutrients to plants (Liang et al. 2006, Chan et al. 2007). It is believed that the advent of *Terra Preta* may have enabled the development of large human populations on the relatively poor oxisols

of the Amazon via a pre-Colombian agricultural revolution (Smith 1980, Hackenberger et al. 1999) but these societies may have collapsed due to introduced diseases. The rich endosperm of *Compsoneura mexicana* seeds is commonly eaten by children in Central America (Janovec 2000). It is therefore plausible that such ancient societies may have actively cultivated and spread *C. capitellata* seeds for their fatty arils and rich endosperm which is reported to taste similar to Brazil nuts (*Bertholettia excelsa*). If this occurred it is possible that the cultivation of these trees has impacted its distribution and genetic structure.

A large degree of population subdivision may be expected in a non-recombining genome of a shade tolerant, old growth forest species such as *C. capitellata*. Previous studies of tropical (Hamilton 1999) and temperate broadleaved trees (Demesure et al. 1996, Petit et al. 2002) have reported a large degree of fixation of cpDNA haplotypes within populations even at small spatial scales. Demesure et al.(1996) found a large degree of population differentiation in European Beech (*Fagus sylvatica*) which was attributed to the last glaciation but there was also a widespread haplotype that connected all populations. Most analyses of cpDNA structure in the Neotropics have focused on investigating the effects of forest fragmentation on population genetic structure and have therefore been concerned with relatively small spatial scales. Additionally, population subdivision may be easily achieved with cpDNA as its effective population is a quarter that of nuclear markers and selection may be especially strong in non-recombining genomes dense in mRNA and protein-coding genes.

It is also possible that these populations are exchanging genes via pollen but this would not be evident using maternally inherited markers. Although possible, I think this

is highly unlikely as the flowers are likely visited only by small thrips and beetles which are unlikely to travel great distances. More field work is needed to ascertain whether *C. capitellata* exists as discreet or continuous breeding populations across their range. Additional and more variable molecular markers will also be needed in the future to help elucidate whether seed dispersal, pollen flow, or other factors are restricting gene flow between these populations. It would also be of great interest to investigate whether similar genetic patterns exists for other taxa as this may indicate that such patterns are the result of ecological speciation, human cultivation and/or phylogeographic processes.

Compsoneura capitellata inhabits the largest geographical and altitudinal range of any species in the genus, spanning 1700 km of latitude, 1200 km longitude, and 1200 m elevation. These trees can be found growing in white sands (San Martin), sandstone ridges (Morona-Santiago), and red oxisols (much of its range). These studies have revealed a large degree of variation in leaf, flower, and fruit morphologies but most of these values form a continuum with respect to geographical populations. The highest degree of morphological and genetic diversity was found in the topographically and edaphically diverse Southwestern region of Ecuador.

Conclusion

With these data it appears that the the *C. capitellata* complex exists as a morphologically diverse group with pronounced genetic discontinuities between populations. Perhaps more variable markers and increased sampling of the Andes to the Amazon region will help answer whether these populations are actively exchanging genes via seed and/or pollen dispersal. It would also be of considerable interest to determine whether the larger leaved montane populations are more genetically similar to

local lowland populations or more distant populations of similar morphology and altitude (i.e. have montane varieties arisen multiple times in the Andes). Given the large number of nucleotide differences of *C. capitellata* in the mountains of Cusco, Peru and this population's atypical large leaf morphologies and affinities to Markgraf's *C. tessmannii*, I propose the provisional resurrection of *C. tessamannii*. Further genetic evidence would be desirable to investigate whether these two large-leaved trees are indeed genetically related. Additionally, collections from the sandstone ridges of the Cordillera del Condor and Cutucu formations of Ecuador appear to represent a new taxon that is provisional described below as *Compsoneura morona-santiagoensis*. *C. morona-santiagoensis* is unique within the *C. capitellata* complex in having domed leaf lamina in between the secondary veins but in other characters of the leaves, flowers and fruits it is similar to other members of the *C. capitellata* complex. It is possible that some of the other taxa with morphological oddities such as the Pastaza, Ecuador population that possesses a warted pericarp and distinct haplotype represents new species, however the description of these taxa as new species requires additional morphological and molecular data in order to establish the degree to which these populations are related.

Provisional Taxonomic descriptions:

The following information summarizes a provisional combination (*C. morona-santiago*) and provisional resurrection (*C. tessmanii*) of taxa from the *C. capitellata* complex.

Compsoneura morona-santiago J. Janovec & R. Steeves, sp. nov. -TYPE: ECUADOR. Morona-Santiago: Taisha, 2°225' S, 77°31 W, 500 m, 31 Jan 1962, *P.C. D. Cazalet and T. D. Pennington 7602* (holotype: NY; isotype: US).

Tree to 15 m. **Bark** unknown. **Sap** red from inner bark. **Branches** teret to subterete, 0.2-0.6 cm diam., rugose to longitudinally striate, brown, densely to sparsely ferrugineous-tomentellous throughout when young, glabrescent to minutely tomentellous in leaf axils and throughout with age, the hairs short-stalked to sessile, 2-branched, the upper branch with a concave surface. **Leaves** simple, alternate, spirally arranged on upper trunk, distichous on branches, **Petioles** stout, subterete, slightly canaliculate, 0.9-1.5 x 0.1-0.2 cm, brown to nearly black, glabrous to glabrescent , the hairs short stalked to sessile, 2-branched; **Blades** elliptic to slightly obovate-elliptic, 7.5-21.1 cm long, 2.2-6.5 cm wide at ½ length, 2.5-7.6 cm wide at ½ the length, 2.3-6.8 cm at ¾ length, thick chartaceous to think coriaceous, the adaxial surface drying olive-brown to dark brown, dull to glossy, glabrescent to sparsely ferrugineous-tomentellous when young, especially at the base, glabrous to glabrescent throughout with age, the abaxial surface drying light brown to brown, dull, glabrescent to sparsely ferrugineous-tomentellous when young, specially at base, glabrous to glabrescent throughout with age, the hairs minute, sessile, short-stalked, 2-branched, the base acute to broadly to weakly acute, the margins conspicuously revolute, the apex weakly to strongly caudate, the acumen 0.7-1.1 cm long; **Venation** with costa raised and glabrous to glabrescent at base adaxially, prominently keeled and glabrescent or sparsely tomentellous abaxially, especially toward base, the secondary nerves acute-ascending, eucamptodromous, distinct near margins, 7-12 per side, spaced 0.6-3.5cm, slightly impressed, same color as lamina glabrous adaxially, prominent, same color as lamina, glabrous to glabescent abaxially, the tertiary veins conspicuous, semi-parallel, and semi-perpendicular to secondary veins. **Staminate**

inflorescence axillary, paired, paniculate, 5.5-8.0 x 0.8-1.2 cm, sparsely surface drying olive-brown to dark brown, dull to glossy, glabrescent to sparsely ferrugineous-tomentellous to glabrescent with age, the hairs 2-branced; **Rachis** densely alternate clusters of about 6-9 per branchlet, arising slightly unilaterally from branchlet pedicels, up to 0.15-0.28 x 0.02 x 0.03 cm; **Pistillate inflorescence** unknown. **Staminate flower** buds long ovate to long elliptic-ovate. **Staminate buds** unknown. **Staminate perianth** elongate vasiform, 0.23-0.32 x 0.15-0.2 cm, coriaceous to thick-coriaceous, yellow to orange when fresh, drying brown, glabrous adaxially, densely to minutely ferrugineous-tomentellus abaxially, the hairs short-stalked or sessile, 2-breanched, tube 0.6-0.10 cm long, lobes 3-4, lanceolate to ovate-lanceolate to narrowly deltoid, 0.16-0.24 x 0.13- 0.18 cm, erect from base to apex, the apices acute. **Androecium** 0.07-0.15 x 0.12 cm long, the filament column 0.02-0.04 cm long, 0.01-0.05 cm at base. 0.01-0.04 cm at apex, dark, solid, oily, glabrous, the anthers 6-7, distinct, erect from base, 0.05-0.07cm long. The bases rounded, oil cells conspicuous in the anther connective, the apices strongly acute. **Pistillate flowers**: known only from persistent perianth subtending fruit, thin coriaceous, drying brown, glabrous adaxially, densely to thinly ferrugineous-tomentellus abaxially, the hairs 2-branched. **Mature Fruits** 1-5 per infructescence, sub-globose to widely transversely elliptic, 2.9-4.5 x 3.0-5.0 cm, minutely to densely, muricate, slightly rugose and green when fresh, drying brown, glabrous to glabrescent; **Pericarp** minutely to densely muricate, glabrescent to sparsely ferrugineous-tomentellous, strongly ligneous, 0.2-0.3 cm thick, appearing indehiscent or partially deshiscent from base; **Seed** with an entire, thin aril, white when fresh, yellowish when dry; **Testa** dull, dark brown, somewhat warty to slightly muricate; **Endosperm** white, entire.

Distribution. This species is a member of the Capitellata subgroup (sensu Janovec 2000), with brown drying leaves, weakly anastomosing secondary veins, and sub-globose, ligneous fruit with a softly sulcate or furrowed pericarp. Elongate, narrow, densely flowered paniculate inflorescences characterize this species. Based on anthers which are strongly erect-ascending from base to apex, this species is distinct from *C. capitellata* and may be related to *C. diazii*. In leaf morphology it is similar to *C. capitellata*, with brown lamina and weakly anastomosing veins. *Compsoneura diazii* is easily differentiated on the basis of strongly anastomosing veins. This species is endemic to southeastern Ecuador where it has been collected on numerous occasions from moist premontane to cloud forest exclusively on soils of sandstone origin. In Morona-Santiago it was collected from around the municipality of Taisha and the in the Cordilllera de Cutucu on the road between Mendez and Morona. In Zamora-Chinchipe it was collected along the road between Nudo de Sabanillo and Vallodolid as well as near the town of Quimi. Sandstone substrate is prevalent in this region of southeastern Ecuador and, with future field work, may prove to be the limiting biogeographic factor for this species.
Etymology: This species name is derived from the province of Morona-Santiago, Ecuador, where this species was first collected.

Compsoneura tessmannii **(**Markgraf) , Notizbl. Bot. Gart. Berlin 9: 964. 1926. TYPE: PERU. Loreto: basin of Rio Marañon from Iquitos upstream on the way to the mouth of Rio Santiago at Pongo de Manseriche, ca. 3^{o}50' S, 77^{o}40' W, 23 Sep 1924, *Tessmann 4108* (holotype: B; fragment of holotype: F; photos of holotype: F, H, MO, NY; isotype fragment: F, NY).

Tree to 18 m tall and 30 cm dbh. **Bark** reddish in colour, furrowed and peels off in long thin fibrous slivers. **Sap** profuse and clear-red. **Branchlets** terete to subterete, 0.3-0.8 cm wide, rugose to longitudinally striate when dried, brown-green, lenticellate, glabrescent to minutely tomentellous in leaf axils and glabrous throughout with age, the hairs short-stalked to sessile, 2-branched. **Leaves** simple, alternate, thick-coriaceous , distichous on branches, leaf buds ferruginous but soon glabrous upon leaf expansion; **petioles** stout, subterete, canaliculate, 0.8-2.6 cm long and 0.3-0.6 cm wide, brown to nearly black when dry, glabrous to glabrescent , the hairs short stalked to sessile, 2-branched; **blades** elliptic-oblong, lamina 14.8-30.0 cm long, 4.5-11.5 cm wide at ¼ length, 3.5-13.6 cm wide at ½ width, 4.9-12.0 cm wide at ¾ length, base broadly cuneate to acute and mucronate to cuspidate at apex, conspicuously revolute at margins, drying light to dark brown or infrequently greenish brown, glabrescent when young but soon glabrous, adaxial surface deep green and glossy, abaxial surface lighter green with a dull shine; **venation** costa glabrous and raised above, prominently keeled and glabrous-glabrescent below; secondary nerves 6-11 per side, spaced by 1.1-2.5 cm, arcuate ascending, eucamptodromous, partially anastomosing near margin. **Staminate inflorescences** axillary, paniculate 3-12 cm long, 2-4 cm wide, rachis 4-25 alternate branched, 7-15 cm long, vestiture of staminate inflorescences densely ferruginous-tomentellous, hairs 2-branched. **Staminate flower buds** green and globose. **Staminate flowers** subglobose to globose, arranged in clusters of 4-12 per inflorescence branch, pedicels slender, 0.1-0.13 cm wide and 0.29-0.42 cm long; **perianth** coriacous, yellow fresh, drying dark brown, 0.56-0.61cm long, 0.44-0.5 cm wide; **tepal lobes** 3 (rarely 2-4), ovate-deltoid, 0.43-0.45

long, 0.32-0.37 wide at base, **perianth tube** 0.14-0.17 cm long; perianth staminate vestiture glabrous within, densely ferruginous-tementellous outside, hairs 2-branched. Androecium 0.2-0.22 cm long, filament column 0.09-0.15 cm long, 0.1-0.11 cm wide at apex, 0.1-0.12 cm wide at base, the anthers 6-10, free and spreading from the base, slightly incurved above, 0.2-0.22 cm long, acute-obtuse at apex, rounded at base**.** **Pistillate inflorescences** unknown. **Fruits** (immature) 1-3 per infructescence, globose, tepals slightly accrescent; **Pericarp** red tomentose to glabrescent with age, sulcate to rugose with a warty appearance, presumably indehiscent or forced dehiscent; **Seed** unknown; **Endosperm** unknown

Distribution: Receently collected from the vicinity of Quincemil, Peru in the mountains of the Madre de Dios watershed. The dried leaf lamina of the Quincemil population has a strong resemblance to type specimens of *C. tessmannii* (Markgraf) which was described in 1926 as a new species of with affinities to capitellata from Pongo de Manseriche, Iquitos, Peru but was synonymized as *C. capitellata* by Janovec (2000). Further morphological and genetic studies are desireable to determine whether *C. tessmannii* populations of Amazonas and Cusco provinces represent disjunct conspecific taxa.

Additional specimens examined: **Peru. CUSCO:** Quispicanchi Province, District of Camanti, Community of Quincemil, -13.23, -70.78, 500-1200 m, 06/26-07/04, 2008, R. Steeves and P. Centeno 527(OAC); R. Steeves and P. Centeno 531 (OAC); R. Steeves and P. Centeno 535 (OAC); R. Steeves and P. Centeno 538 (OAC); R. Steeves and P. Centeno 548 (OAC); R. Steeves and P. Centeno 550 (OAC); R. Steeves and P. Centeno 551 (OAC); R. Steeves and P. Centeno 556 (OAC); R. Steeves and P. Centeno 558

(OAC); R. Steeves and P. Centeno 562 (OAC); R. Steeves and P. Centeno 563 (OAC); R. Steeves and P. Centeno 568 (OAC); R. Steeves and P. Centeno 569 (OAC); R. Steeves and P. Centeno 570 (OAC); R. Steeves and P. Centeno 571 (OAC); R. Steeves and P. Centeno 572 (OAC); R. Steeves and P. Centeno 577 (OAC); R. Steeves and P. Centeno 591 (OAC); R. Steeves and P. Centeno 593 (OAC); and R. Steeves and P. Centeno 600 (OAC).

Tables

Table 2.1 Specimens used in vegetative and floral ordinations. Herbarium code and accession number, collectors, collection number (Coll.#), country, decimal degrees latitude (DD Lat.), decimal degrees longitude (DD Long.), and elevation (Elev.) are listed.

Herbarium	Collector(s)	Coll. #	Country	DD Lat.	DD Long.	Elev.
QCNE-88196	Aulestia, M.	2260	Ecuador	0.65	-76.43	250
	Schultes, R. E., and I. Cabrera	14144	Colombia	0.12	-71.25	250
US 2171356	Schultes, R. E., and I. Cabrera	14144	Colombia	0.12	-71.25	250
QCNE-64227	Palacios, W., et al.	9197	Ecuador	0.08	-76.18	230
US 2171510	Schultes, R. E., and I.	15107	Colombia	0.07	-71.12	300
US 2171510	Schultes, R. E., and I. Cabrera	15107	Colombia	0.07	-71.12	300
US 2220060	Schultes, R. E., and I. Cabrera	15107	Colombia	0.07	-71.12	300
H	Schultes, R. E., and I. Cabrera	15107	Colombia	0.07	-71.12	300
US 2220060	Schultes, R. E., and I. Cabrera	15107	Colombia	0.07	-71.12	300
H	Schultes, R. E., and I. Cabrera	15107	Colombia	0.07	-71.12	300
NY	Schultes, R. E., and I. Cabrera	15107	Colombia	0.07	-71.12	300
QCNE-64522	Palacios, W., et al.	9226	Ecuador	0.00	-76.18	230
QCNE-63507	Palacios, W., et al.	8989	Ecuador	0.00	-76.18	230
QCNE-64198	Palacios, W., et al.	9311	Ecuador	0.00	-76.18	230
NY	Cer¢n, C., and F. Hurtado	4138	Ecuador	-0.10	-76.18	200
QCNE-44552	Cer¢n, C., and J. Ayala	9497	Ecuador	-0.13	-76.37	240
MO-5020318	Gudi, O, E.	118	Ecuador	-0.33	-77.08	250
QCNE-45495	Rubio, D.	280	Ecuador	-0.33	-77.08	250
MO3806254	Espinoza, S.	83	Ecuador	-0.33	-77.08	250
QCNE 39714	Gudi, O, E.	118	Ecuador	-0.33	-77.08	250
NY	Rubio, D.	280	Ecuador	-0.33	-77.00	250
MO 3794978	Rubio, D.	280	Ecuador	-0.33	-77.00	250
MO 05030983	Aulestia, M.	1651	Ecuador	-0.42	-76.58	250
QCNE 84139	Aulestia, M.	1651	Ecuador	-0.42	-76.58	250
MO 4063871	Palacios, W.	2310	Ecuador	-0.43	-77.10	300
NY	Palacios, W.	2310	Ecuador	-0.43	-77.10	300
QCNE 19906	Palacios, W.	2310	Ecuador	-0.43	-77.10	300
NY	Neill, D.	7150	Ecuador	-0.43	-77.02	250
MO 3482889	Neill, D.	7150	Ecuador	-0.43	-77.02	250
QCNE 55032	Palacios, W., et al.	7586	Ecuador	-0.48	-75.53	230
MO 4063877	Cer¢n, C., and W. Palacios	3029	Ecuador	-0.50	-77.02	250

NY	Ollgaard, B., et al.	57050	Ecuador	-0.53	-76.37	300
QCNE 87119	Dik, A., and C. Aulestia	1436	Ecuador	-0.55	-76.52	250
QCNE-75762	Aulestia, M., and G. Grefa	269	Ecuador	-0.55	-76.50	250
QCNE-53357	Korning, J., and K. Thomsen	47562	Ecuador	-0.55	-76.37	370
QCNE-53352	Korning, J., and K. Thomsen	47556	Ecuador	-0.55	-76.37	250
F 2612343	Palacios, W., et al.	7798	Ecuador	-0.55	-75.53	230
F 2162341	Palacios, W., et al.	7812	Ecuador	-0.55	-75.53	230
QCNE 54636	Palacios, W., et al.	7798	Ecuador	-0.55	-75.53	230
QCNE 54624	Palacios, W., et al.	7812	Ecuador	-0.55	-75.53	230
QCNE/MO	Freire, E., P. Cerda	23	Ecuador	-0.63	-77.45	690
MO 05030981	Aulestia, M.	2746	Ecuador	-0.65	-76.43	250
QCNE-88049	Aulestia, M.	2446	Ecuador	-0.65	-76.43	250
MO 05030982	Aulestia, M.	2713	Ecuador	-0.65	-76.43	250
QCNE-88239	Aulestia, M.	2497	Ecuador	-0.65	-76.43	250
QCNE-88319	Aulestia, M.	2497	Ecuador	-0.65	-76.43	250
QCNE-88464	Aulestia, M.	2746	Ecuador	-0.65	-76.43	250
QCNE-88658	Aulestia, M., and T. Ahue	2909	Ecuador	-0.68	-76.45	260
QCNE/MO	Alvarado, A.	390	Ecuador	-0.70	-77.33	810
Mo 05030980	Aulestia, M., and T. Ahue	2909	Ecuador	-0.78	-76.45	260
QCNE-8642	Dik, A., and J. Andi	953	Ecuador	-0.83	-76.35	270
MO 05030962	Dik, A., and J. Andi	953	Ecuador	-0.83	-76.35	270
QCNE-86816	Dik, A., and R. Enomenga	1145	Ecuador	-0.83	-76.30	250
QCNE-86694	Dik, A.	1232	Ecuador	-0.85	-76.25	260
MO 3654088	Cer¢n, C., et al.	4123	Ecuador	-0.92	-76.18	200
MO 3654089	Cer¢n, C., et al.	4138	Ecuador	-0.92	-76.18	200
US 3129061	Cer¢n, C., and F. Hurtado	4140	Ecuador	-0.92	-76.18	200
QCNE-17872	Cer¢n, C., and F. Hurtado	4140	Ecuador	-0.92	-76.18	200
QCNE 28641	Cer¢n, C., and F. Hurtado	4123	Ecuador	-0.92	-76.18	200
MO 3654090	Cer¢n, C., and F. Hurtado	4140	Ecuador	-0.92	-76.18	200
QCNE-28645	Cer¢n, C., and F. Hurtado	4138	Ecuador	-0.92	-76.18	200
MO 4063873	Cer¢n, C., and F. Hurtado	3876	Ecuador	-0.92	-76.18	200
QCNE-84478	Aulestia, M., and O. Gonti	1983	Ecuador	-0.92	-76.15	250
QCNE 84462	Aulestia, M.	1859	Ecuador	-0.92	-76.15	250
QCNE 86988	Dik, A.	1455	Ecuador	-0.93	-76.22	248
QCNE-88822	Pitman, N.	587	Ecuador	-0.93	-76.22	250
QCNE-86996	Dik, A.	1447	Ecuador	-0.93	-76.22	248
NY	Palacios, W.	2409	Ecuador	-0.95	-76.22	230
QCNE-28524	Palacios, W.	2409	Ecuador	-0.95	-76.22	230
MO 3597372	Palacios, W.	2409	Ecuador	-0.95	-76.22	230
QCNE/MO	Aulestia, M., and A. Omehuat	3246	Ecuador	-0.98	-76.20	235
QCNE/MO	Aulestia, M., and B. Quihuinamo	3111	Ecuador	-0.98	-76.20	237
MO	Aulestia, M., and B. Quihuinamo	3207	Ecuador	-0.98	-76.20	235
QCNE 096476	Aulestia, M., and B. Quihuinamo	3111	Ecuador	-0.98	-76.20	237
QCNE 096452	Aulestia, M., and B. Quihuinamo	3207	Ecuador	-0.98	-76.20	235

QCNE 87099	Dik, A., and T. Ahue	1563	Ecuador	-1.00	-76.18	250
Harvard	Schultes, R. E., and F. Lopez	10200	Brazil	-1.00	-69.50	150
QCNE-87276	Dik, A., and T. Ahue	1598	Ecuador	-1.03	-76.18	249
MO 4221156	Cer¢n, C., and C. Iguago	5599	Ecuador	-1.07	-77.60	450
QCNE 46430	Cer¢n, C., et al.	8728	Ecuador	-1.07	-77.60	400
MO 4066077	Cer¢n, C., et al.	8728	Ecuador	-1.07	-77.60	400
QCNE-46737	Cer¢n, C., et al.	8728	Ecuador	-1.07	-77.60	400
NY	Zak, V.	4161	Ecuador	-1.25	-76.92	320
MO 4210196	Zak, V.	4161	Ecuador	-1.25	-76.92	320
QCNE-21871	Zak, V.	4161	Ecuador	-1.25	-76.92	320
Harvard	Garcia-Barriga, H.	14769	Colombia	-1.25	-69.63	240
NY	Tipaz, G., et al.	571	Ecuador	-1.42	-77.33	400
F 2157840	Hurtado, F.	2987	Ecuador	-1.42	-77.33	400
QCNE/MO	Palacios, W.	12164	Ecuador	-1.47	-77.43	500
QCNE-49315	Gudi, O, E.	930	Ecuador	-1.57	-77.42	580
QCNE 80470	Palacios, W.	11379	Ecuador	-1.77	-78.00	900
Mo 2820305	Diaz, C.	1380	Peru	-2.50	-75.75	200
F 668926	Klug, G.	2130	Peru	-2.87	-75.25	180
Harvard	Klug, G.	2130	Peru	-2.87	-75.25	180
Harvard	Klug, G.	2130	Peru	-2.88	-75.25	180
US 1456778	Klug, G.	2130	Peru	-2.88	-75.25	180
NY	Klug, G.	2130	Peru	-2.88	-75.25	180
MO1039604	Klug, G.	2130	Peru	-2.88	-75.25	180
NY	Cid Ferreira, C. A., et al.	8462	Brazil	-2.88	-67.75	180
H	Schultes, R. E., and G. A. Black	8166	Brazil	-3.00	-69.00	100
US 1989313	Schultes, R. E., and G. A. Black	8166	Brazil	-3.00	-69.00	100
MO 4257892	Pipoly, J.	13212	Peru	-3.33	-72.92	400
MO 3826656	Vasquez, R.	11810	Peru	-3.33	-72.92	130
MO 4229666	Vasquez, R., and N. Jaramillo	16072	Peru	-3.33	-72.92	140
F1868820	Rimachi Y., M.	3043	Peru	-3.50	-73.07	100
NY	Rimachi Y., M.	3043	Peru	-3.50	-73.07	100
MO 2825596	Rimachi Y., M.	3043	Peru	-3.50	-73.07	100
MO 3630620	Vasquez, R., et al.	5164	Peru	-3.50	-72.83	106
NY	Vasquez, R., et al.	5164	Peru	-3.50	-72.83	106
F 2034668	Vasquez, R., et al.	5164	Peru	-3.50	-72.83	106
F 1312427	Fr¢es, R. L. de	20820	Brazil	-3.50	-68.95	200
US 2343268	Fr¢es, R. L. de	34864	Brazil	-3.50	-68.95	200
MO 2188259	Ducke, A.	19576	Brazil	-3.50	-68.95	200
NY	Ducke, A.	314	Brazil	-3.50	-68.95	200
NY 3097	Fr¢es, R. L. de	20813	Brazil	-3.50	-68.95	200
MO 2188260	Ducke, A.	23693	Brazil	-3.50	-68.95	200
F 1230229	Ducke, A.	561	Brazil	-3.50	-68.95	200
H	Ducke, A.	1486	Brazil	-3.50	-68.95	100
NY	Ducke, A.	23693	Brazil	-3.50	-68.95	200
MO 1255784	Ducke, A.	561	Brazil	-3.50	-68.95	200
US 1516503	Ducke, A.	23693	Brazil	-3.50	-68.95	200
Harvard	Ducke, A.	143	Brazil	-3.50	-68.95	200
NY-2728	Ducke, A.	1486	Brazil	-3.50	-68.95	100
NY	Ducke, A.	19576	Brazil	-3.50	-68.95	200

F 1486643	Ducke, A.	1486	Brazil	-3.50	-68.95	200
NY	Ducke, A.	561	Brazil	-3.50	-68.95	200
NY	Fr¢es, R. L. de	20820	Brazil	-3.50	-68.95	200
MO 4257890	Pipoly, J., et al.	12632	Peru	-3.58	-73.25	110
NY	Klug, G.	180	Peru	-3.58	-73.25	100
F 6184000	Williams, L.	3786	Peru	-3.58	-73.25	120
NY	Tessman, G.	5101	Peru	-3.58	-73.25	120
F 624285	Klug, G.	180	Peru	-3.58	-73.25	100
MO 3869563	Vasquez, R., and N. Jaramillo	9676	Peru	-3.80	-73.42	122
MO 3042849	Tunqui, S.	189	Peru	-3.83	-77.67	200
MO 2892231	Huashikat, V.	1786	Peru	-3.83	-77.67	200
MO 2813382	Huashikat, V.	1035	Peru	-3.83	-77.67	180
MO 2892223	Huashikat, V.	2012	Peru	-3.83	-77.67	200
MO 3042848	Tunqui, S.	301	Peru	-3.83	-77.67	200
MO 4257893	Vasquez, R., et al.	15809	Peru	-3.87	-73.25	180
MO 4257891	Vasquez, R.	14517	Peru	-3.87	-73.25	200
MO 3630613	Vasquez, R., et al.	5880	Peru	-3.87	-73.25	200
MO 2734836	Rimachi Y., M.	3303	Peru	-3.88	-73.63	170
FLAS 186349	Rimachi Y., M.	4532	Peru	-3.88	-73.63	170
NY	McDaniel, S., and M. Rimachi	21069	Peru	-3.88	-73.63	160
NY	Rimachi Y., M.	3303	Peru	-3.88	-73.63	170
MO 2427881	Rimachi Y., M.	2849	Peru	-3.88	-73.63	160
MO 2734836	Rimachi Y., M.	3303	Peru	-3.88	-73.63	160
NY	Rimachi Y., M.	4532	Peru	-3.88	-73.63	160
NY	Rimachi Y., M.	165	Peru	-3.88	-73.63	150
US 3177354	Rimachi Y., M.	4532	Peru	-3.88	-73.63	160
NY	Vasquez, R., et al.	5354	Peru	-3.92	-73.92	0
MO 3842988	Vasquez, R., et al.	12195	Peru	-3.92	-73.67	150
Tex	Vasquez, R., et al.	5354	Peru	-3.92	-73.58	130
MO 3584403	Vasquez, R., et al.	5354	Peru	-3.92	-73.58	130
MO 2204538	Kayap, R.	662	Peru	-4.00	-78.00	375
NY	McDaniel, S.	29641	Peru	-4.00	-73.25	200
F2027320	Vasquez, R., et al.	7296	Peru	-4.17	-72.00	116
NY	Vasquez, R., et al.	7296	Peru	-4.17	-72.00	116
MO 3628224	Vasquez, R., et al.	7296	Peru	-4.17	-72.00	116
QCNE 54324	Palacios, W., et al.	8654	Ecuador	-4.27	-78.70	930
NY	Prance, G. T., et al.	23868	Brazil	-4.47	-71.68	300
MO 2725124	Prance, G. T., et al.	23868	Brazil	-4.47	-71.68	200
MO 2323593	Ancuash, E.	218	Peru	-4.50	-78.12	333
MO 2205012	Ancuash, E.	142	Peru	-4.50	-78.12	33
NY-00066801	Tessmann, G.	4108	Peru	-4.50	-77.45	500
F-1022586	Tessmann, G.	4108	Peru	-4.50	-77.42	500
MO -4929828	Vasquez, R., et al.	24158	Peru	-4.55	-78.20	600
MO	Vasquez, R.	24158	Peru	-4.55	-78.20	600
F 2157471	Cid Ferreira, C. A., et al.	9949	Brazil	-4.55	-71.67	100
NY	Cid Ferreira, C. A., et al.	9949	Brazil	-4.55	-71.67	100
US 3290254	Cid Ferreira, C. A., et al.	9949	Brazil	-4.55	-71.67	200
MO 2435914	Ancuash, E.	424	Peru	-4.63	-78.13	440
NY	Ancuash, E.	424	Peru	-4.63	-78.13	440
MO 3032654	Ancuash, E.	424	Peru	-4.63	-78.13	440

MO-5096929	Jaramillo, N., and K. Katip	797	Peru	-4.92	-78.32	320
MO-5096932	Vasquez, R., et al.	19932	Peru	-4.92	-78.32	550
MO-5096933	Vasquez, R., et al.	19956	Peru	-4.92	-78.32	550
MO-5096931	Vasquez, R., et al.	20016	Peru	-4.92	-78.32	550
MO-4929838	Vasquez, R., et al.	20363	Peru	-4.92	-78.32	320
MO-5096930	Jaramillo, N.	869	Peru	-4.92	-78.32	320
MO-4929829	Vasquez, R., et al.	23886	Peru	-5.05	-78.33	600
MO-4929833	Vasquez, R., et al.	21416	Peru	-5.05	-78.33	380
MO-429825	Diaz, C., et al.	4172	Peru	-5.05	-78.33	310
MO	Diaz, C.	4128	Peru	-5.05	-78.33	310
MO-4929827	Diaz, C., et al.	4128	Peru	-5.05	-78.33	310
MO	Diaz, C., et al.	4172	Peru	-5.05	-78.33	310
MO-4929832	Vasquez, R., et al.	24895	Peru	-5.05	-78.33	450
	Vasquez, R., et al.	24895	Peru	-5.07	-78.33	450
MO0429834	Vasquez, R., et al.	24643	Peru	-5.25	-78.37	900
MO	Vasquez, R., et al.	24643	Peru	-5.25	-78.37	900
F-1706096	Vigo, J. S.	6439	Peru	-6.03	-75.88	540
MO 5103767	Vigo, J. S.	6519	Peru	-6.03	-75.88	532
US 2703971	Vigo, J. S.	6519	Peru	-6.03	-75.88	532
MO 5103768	Vigo, J. S.	6439	Peru	-6.03	-75.88	540
F 1706102	Vigo, J. S.	6519	Peru	-6.03	-75.88	532
US 3260273	Knapp, S., and J. Mallet	8466	Peru	-6.25	-76.28	200
MO 3632210	Knapp, S., and J. Mallet	8466	Peru	-6.25	-76.28	200
NY	Vigo, J. S.	4596	Peru	-8.00	-76.62	500
H	Vigo, J. S.	4596	Peru	-8.00	-76.62	500
F 1871036	Vigo, J. S.	4596	Peru	-8.00	-76.62	500
Tex	Vigo, J. S.	5566	Peru	-8.12	-76.52	600
F 1874381	Vigo, J. S.	5566	Peru	-8.12	-76.52	600
Duke 311051	Vigo, J. S.	5566	Peru	-8.12	-76.52	600
NY	Vigo, J. S.	5566	Peru	-8.12	-76.52	600
F 1898708	Rodriguez, L. T.	03	Peru	-9.00	-76.00	650
NY	MacBride, J. F.	5065	Peru	-9.55	-75.90	1170
F 536109	MacBride, J. F.	5065	Peru	-9.55	-75.90	1170
NY	Killip, E. P., and A. C. Smith	26053	Peru	-10.87	-73.75	1100
F 616725	Killip, E. P., and A. C. Smith	26053	Peru	-10.87	-73.75	1100
MO-4657673	Betancur, J., et al.	4316	Colombia	1.30	-78.13	1325
MO-2637120	Renteria, E., et al.	28	Colombia	7.38	-73.83	320
MO-2665544	Renteria, E., et al.	14	Colombia	5.58	-73.50	320
H	Soejarto, D.	2782	Colombia	7.50	-74.92	700
H	Soejarto, D.	2782	Colombia	7.50	-74.92	700
H	Soejarto, D.	2782	Colombia	7.50	-74.92	700
MO 2716725	Gentry, A. H., and E. Renteria	23826	Colombia	5.50	-76.55	50
NY	Gentry, A. H., and E. Renteria	23826	Colombia	5.50	-76.55	50
NY	Cazalet, P. C. D., and T. D. Pennington	7602	Ecuador	-2.38	-77.52	500
US 2406117	Cazalet, P. C. D., and T. D. Pennington	7602	Ecuador	-2.38	-77.52	500
NY	van der Werff, H., and W. Palacios	10331	Ecuador	-2.78	-77.83	800
MO 4066528	van der Werff, H., and W. Palacios	10331	Ecuador	-2.78	-77.83	800

NY	van der Werff, H., and W. Palacios	9393	Ecuador	-4.25	-79.20	2000
OAC	Steeves, R. et al	RS 328	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 328	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 332	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 334	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 334	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 336	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 338	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 346	Ecuador	-3.57	-78.45	1200
OAC	Steeves, R. et al	RS 271	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 271	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 274	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 278	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 278	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 280	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 280	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 284	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 284	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 293	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 294	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 295	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 295	Ecuador	-1.04	-77.37	450
OAC	Steeves, R. et al	RS 364	Ecuador	-1.78	-77.83	900
OAC	Steeves, R. et al	RS 364	Ecuador	-1.78	-77.83	900
OAC	Steeves, R. et al	RS 370	Ecuador	-1.78	-77.83	900
OAC	Steeves, R. et al	RS 406	Ecuador	-1.78	-77.83	900
OAC	Steeves, R. et al	RS 418	Ecuador	-1.78	-77.83	900
OAC	Steeves, R. et al	RS 406	Ecuador	-1.78	-77.83	900
OAC	Steeves, R. et al	RS 527	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 531	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 535	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 535	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 538	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 538	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 538	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 538	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 551	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 562	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 566	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 568	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 570	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 571	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 571	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 572	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 572	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 577	Peru	-13.24	-70.78	1000
OAC	Steeves, R. et al	RS 600	Peru	-13.24	-70.78	1000

Table 2.2 CA analysis of 48 quantitative morphological metrics for *C. capitellata* taxa. Bolded Pearson correlation (P Corr.) values indicate the metrics most significant to the ordination (** = p value < 0.01).

Metric	X-axis		Y-axis	
	P Corr.	Sig. (2-tailed)	P Corr.	Sig.(2-tailed)
Stem and leaf				
Stem width-min.	0.261	0.000	0.060	0.329
Stem width-max	0.127	0.038	-0.032	0.607
Petiole length-min	**0.713****	0.000	0.083	0.175
Petiole length-max	0.671	0.000	0.141	0.021
Leaf Petiole width-min	0.585	0.000	-0.080	0.194
Leaf Petiole width-max	0.653	0.000	-0.036	0.558
Lamina Length-min	**0.919****	0.000	-0.136	0.026
Lamina Length-max	**0.857****	0.000	0.081	0.187
Leaf width ¼ length-min	**0.924****	0.000	-0.046	0.454
Leaf width ¼ length-max	**0.918****	0.000	0.126	0.040
Leaf width ½ length-min	**0.941****	0.000	-0.046	0.459
Leaf width ½ length-max	**0.904****	0.000	0.135	0.027
Leaf width ¾ length-min	**0.945****	0.000	-0.030	0.624
Leaf width ¾ length-max	**0.883****	0.000	0.140	0.023
Leaf vein number-min	0.604	0.000	**-0.454****	0.000
Leaf vein number-max	0.544	0.000	**-0.464****	0.000
Leaf vein space-min	0.463	0.000	**0.475****	0.000
Leaf vein space-max	0.325	0.000	**0.487****	0.000
Leaf acumen length-min	0.467	0.000	0.004	0.951
Leaf acumen length-max	0.443	0.000	-0.043	0.487
Androecial				
Anther number min	**-0.790****	0.000	0.502	0.000
Anther number max	**-0.736****	0.000	**0.556****	0.000
Anther length min	**-0.593****	0.000	**-0.684****	0.000
Anther length max	**-0.593****	0.000	**-0.684****	0.000
Anther sac length min	-0.419	0.000	**-0.608****	0.000
Anther sac length max	-0.496	0.000	**-0.517****	0.000
Anther cluster width min	**-0.576****	0.000	**-0.555****	0.000
Anther cluster width max	**-0.590****	0.000	**-0.535****	0.000
Filament column length min	**-0.545****	0.000	**-0.609****	0.000
Filament column length max	**-0.573****	0.000	-0.566	0.000
Filament column apical width min	-0.440	0.000	-0.251	0.030
Filament column apical width max	-0.408	0.000	-0.238	0.040
Filament column basal width min	-0.482	0.000	-0.195	0.093
Filament column basal width max	-0.515	0.000	-0.144	0.217
Perianth				
Flower pedicel length min	0.572	0.000	**0.700****	0.000
Flower pedicel length max	0.659	0.000	**0.638****	0.000
Flower pedicel width min	0.329	0.036	-0.337	0.031
Flower pedicel width max	0.345	0.027	-0.391	0.011
Perianth thickness min	0.174	0.277	-0.374	0.016
Perianth thickness max	0.230	0.148	**-0.405****	0.009
Tepal lobe length min	**0.881****	0.000	-0.235	0.139
Tepal lobe length max	**0.881****	0.000	-0.236	0.138
Perianth tube length min	0.529	0.000	-0.028	0.863
Perianth tube length max	0.737	0.000	-0.014	0.929
Perianth total length min	**0.915****	0.000	-0.211	0.184
Perianth total length max	**0.948****	0.000	-0.183	0.253

Perianth width at widest point min	**0.878****	0.000	0.049	0.760
Perianth width at widest point max	**0.856****	0.000	0.026	0.872

Table 2.3 Collection numbers (Coll. #), Country, Population name, decimal degrees latitude (DD Lat.), decimal degrees longitude (DD long.), elevation above sea level (Elev a.s.l.), and Barcode of Life Database process identification numbers (BOLD. process ID) for specimens used in molecular analyses. Collectors of samples were J. Janovec (JPJ), Camila Diaz (CDiaz), and R. Steeves (RADS). Specimens of *C. atopa* and *C. diazii* are indicated by a "-Ca" and "Cd" proceeding the collection number. All other collections represent *C. capitellata*.

Coll. #	Country	Population name	DD Lat.	DD Long.	Elev. a.s.l.	BOLD process ID
JPJ1374-Ca	Ecuador	C. atopa-Orellana	-0.47	-77.26	400m	RSMYR002-11
RADS332	Ecuador	Morona-Santiagoensis	-3.57	-78.45	1200m	RSMYR053-11
RADS334	Ecuador	Morona-Santiagoensis	-3.57	-78.45	1200m	RSMYR054-11
RADS336	Ecuador	Morona-Santiagoensis	-3.57	-78.45	1200m	RSMYR055-11
RADS338	Ecuador	Morona-Santiagoensis	-3.57	-78.45	1200m	RSMYR056-11
RADS346	Ecuador	Morona-Santiagoensis	-3.57	-78.45	1200m	RSMYR057-11
JPJ893	Ecuador	Napo	-1.04	-77.37	450m	RSMYR058-11
JPJ894	Ecuador	Napo	-1.04	-77.37	450m	RSMYR059-11
JPJ899	Ecuador	Napo	-1.04	-77.37	450m	RSMYR009-11
JPJ900	Ecuador	Napo	-1.04	-77.37	450m	RSMYR060-11
JPJ902	Ecuador	Napo	-1.04	-77.37	450m	RSMYR061-11
RADS271	Ecuador	Napo	-1.04	-77.37	450m	RSMYR062-11
RADS274	Ecuador	Napo	-1.04	-77.37	450m	RSMYR063-11
RADS278	Ecuador	Napo	-1.04	-77.37	450m	RSMYR064-11
RADS280	Ecuador	Napo	-1.04	-77.37	450m	RSMYR065-11
RADS284	Ecuador	Napo	-1.04	-77.37	450m	RSMYR066-11
RADS293	Ecuador	Napo	-1.04	-77.37	450m	RSMYR067-11
RADS294	Ecuador	Napo	-1.04	-77.37	450m	RSMYR068-11
RADS295	Ecuador	Napo	-1.04	-77.37	450m	RSMYR069-11
RADS364	Ecuador	Pastaza	-1.78	-77.83	1000m	RSMYR070-11
RADS370	Ecuador	Pastaza	-1.78	-77.83	1000m	RSMYR071-11
RADS406	Ecuador	Pastaza	-1.78	-77.83	1000m	RSMYR072-11
RADS409	Ecuador	Pastaza	-1.78	-77.83	1000m	RSMYR073-11
RADS418	Ecuador	Pastaza	-1.78	-77.83	1000m	RSMYR074-11
RADS427	Ecuador	Pastaza	-1.78	-77.83	1000m	RSMYR075-11
JPJ1542	Ecuador	Zamora-Chinchipe	-4.29	-78.63	1000m	RSMYR076-11
JPJ1543	Ecuador	Zamora-Chinchipe	-4.29	-78.63	1000m	RSMYR077-11
JPJ1544	Ecuador	Zamora-Chinchipe	-4.29	-78.63	1000m	RSMYR078-11
JPJ1545	Ecuador	Zamora-Chinchipe	-4.29	-78.63	1000m	RSMYR079-11
RADS527	Peru	Cusco	-13.24	-70.78	1000m	RSMYR003-11

RADS531	Peru	Cusco	-13.24	-70.78	1000m	RSMYR080-11
RADS535	Peru	Cusco	-13.24	-70.78	1000m	RSMYR081-11
RADS538	Peru	Cusco	-13.24	-70.78	1000m	RSMYR082-11
RADS550	Peru	Cusco	-13.24	-70.78	1000m	RSMYR083-11
RADS551	Peru	Cusco	-13.24	-70.78	1000m	RSMYR004-11
RADS562	Peru	Cusco	-13.24	-70.78	1000m	RSMYR084-11
RADS566	Peru	Cusco	-13.24	-70.78	1000m	RSMYR085-11
RADS568	Peru	Cusco	-13.24	-70.78	1000m	RSMYR086-11
RADS569	Peru	Cusco	-13.24	-70.78	1000m	RSMYR087-11
RADS570	Peru	Cusco	-13.24	-70.78	1000m	RSMYR088-11
RADS571	Peru	Cusco	-13.24	-70.78	1000m	RSMYR089-11
RADS572	Peru	Cusco	-13.24	-70.78	1000m	RSMYR090-11
JPJ843	Peru	Loreto	-3.48	-74.25	100m	RSMYR093-11
JPJ844	Peru	Loreto	-3.48	-74.25	100m	RSMYR094-11
JPJ855	Peru	Loreto	-3.52	-73.15	100m	RSMYR006-11
JPJ860	Peru	Loreto	-3.52	-73.15	100m	RSMYR095-11
JPJ862	Peru	Loreto	-3.52	-73.15	100m	RSMYR096-11
JPJ863	Peru	Loreto	-3.52	-73.15	100m	RSMYR097-11
JPJ872	Peru	Loreto	-3.52	-73.15	100m	RSMYR007-11
JPJ873	Peru	Loreto	-3.52	-73.15	100m	RSMYR098-11
JPJ874	Peru	Loreto	-3.52	-73.15	100m	RSMYR099-11
JPJ875	Peru	Loreto	-3.52	-73.15	100m	RSMYR008-11
JPJ827	Peru	San Martin	-6.15	-76.17	250m	RSMYR101-11
JPJ829	Peru	San Martin	-6.15	-76.17	250m	RSMYR103-11
JPJ830	Peru	San Martin	-6.15	-76.17	250m	RSMYR104-11
JPJ831	Peru	San Martin	-6.15	-76.17	250m	RSMYR105-11
JPJ832	Peru	San Martin	-6.15	-76.17	250m	RSMYR106-11
JPJ833	Peru	San Martin	-6.15	-76.17	250m	RSMYR107-11
JPJ834	Peru	San Martin	-6.15	-76.17	250m	RSMYR108-11
JPJ835	Peru	San Martin	-6.15	-76.17	250m	RSMYR005-11
CDiaz7644-Cd	Peru	Bagua	-5.03	-78.22	800m	RSMYR012-11

Figures

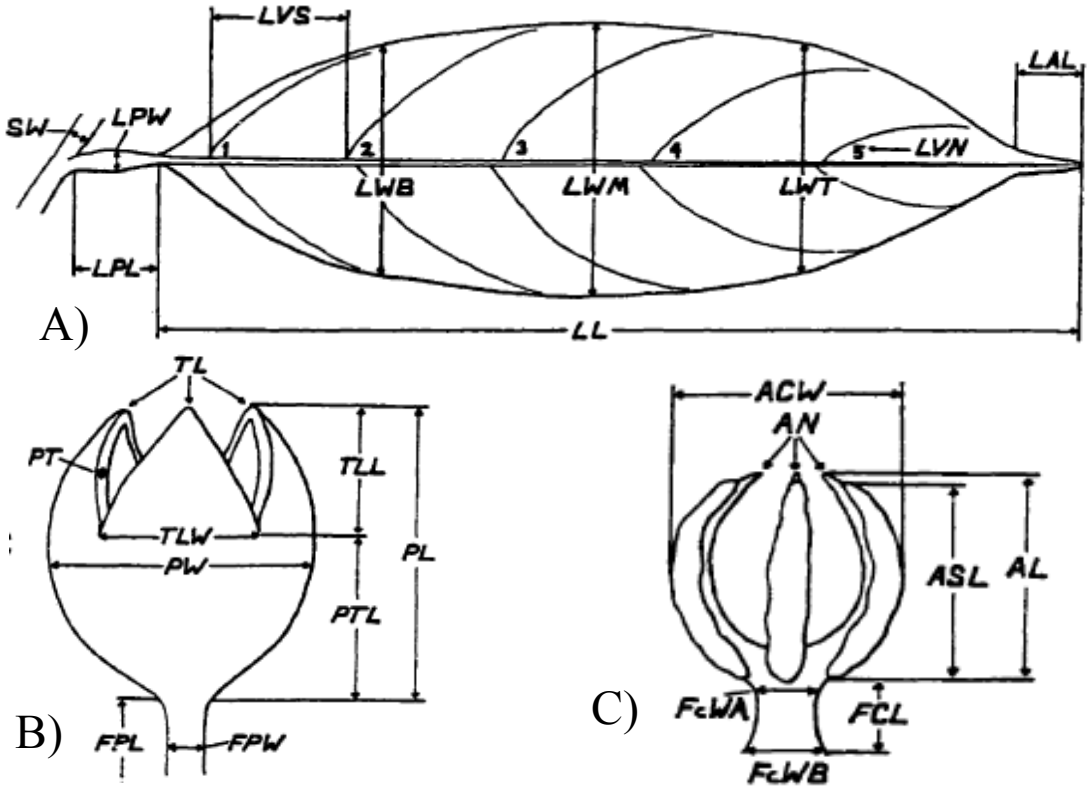

Figure 2.1 Illustration of morphological metrics: A) Vegetative data points: SW=Stem Width, LPL=Petiole Length, LPW=Petiole Width, LVS=space between secondary veins, LVN=Vein Number, LL=Lamina Length, LWB= Leaf Width at ¼ length, LWM=Leaf Width at ½ length, LWT=Leaf Width at ¾ length, LVN=Leaf Vein Number, LAL=Leaf Acumen Length. B) Perianth data points: TL=Tepal Lobe number, PT=Perianth Thickness, TLW=Tepal Lobe Basal Width, PW=Perianth Width at widest point, FPL=Flower Pedicel Length, FPW=Flower Pedicel Width, TLL=Perianth Tube Length, PTL= Perianth tube length, PL=Perianth total Length. C) Androecium data points: AN=Anther Number, AL=Anther Length, ASL=Anther Sac Length, ACW=Anther Cluster Width, FCL=Filament Column Length, FCWA=Filament Column Apical Width, FCWB= Filament Column Basal Width. (modified from Janovec 2000).

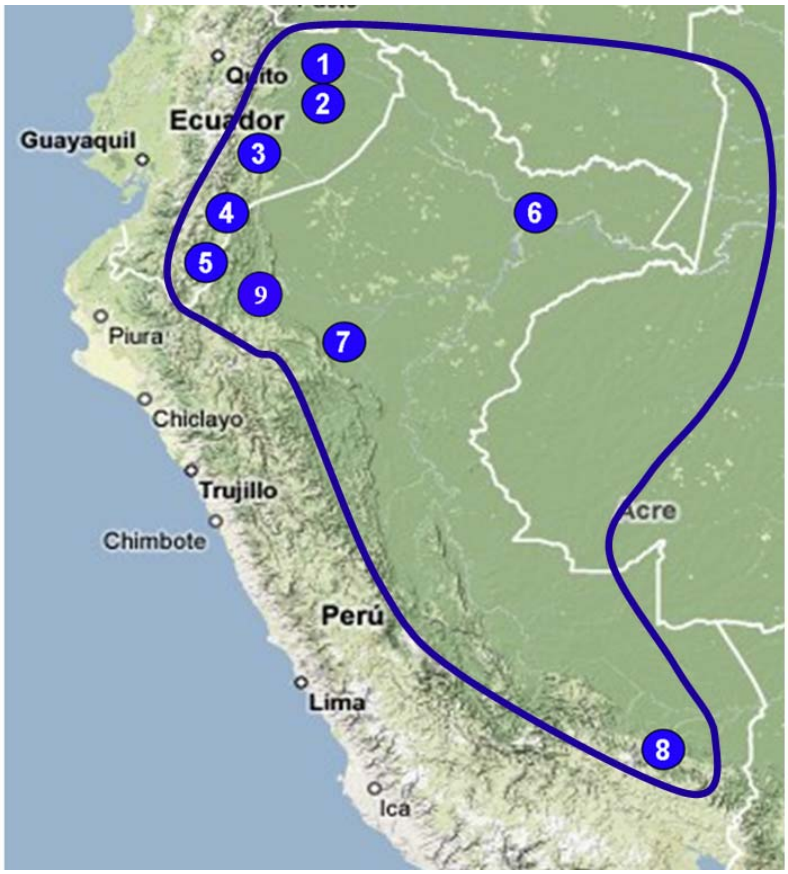

Figure 2.2 Sampling localities of *Compsoneura* taxa for genetic analyses: 1) *C. atopa*, Ecuador, Orellana. 2) *C. capitellata*, Ecuador, Napo. 3) *C. capitellata*, Ecuador, Pastaza, 4) *C. capitellata*, Ecuador, Morona-Santiago. 5) *C. capitellata*, Ecuador, Zamora-Chinchipe. 6) *C. capitellata*, Peru, Loreto. 7) *C. capitellata*, Peru, San Martin. 8) *C. capitellata*, Peru, Cusco. 9) *C. diazii*, Peru, Bagua. The area inside the blue line indicates the known range of *C. capitellata*.

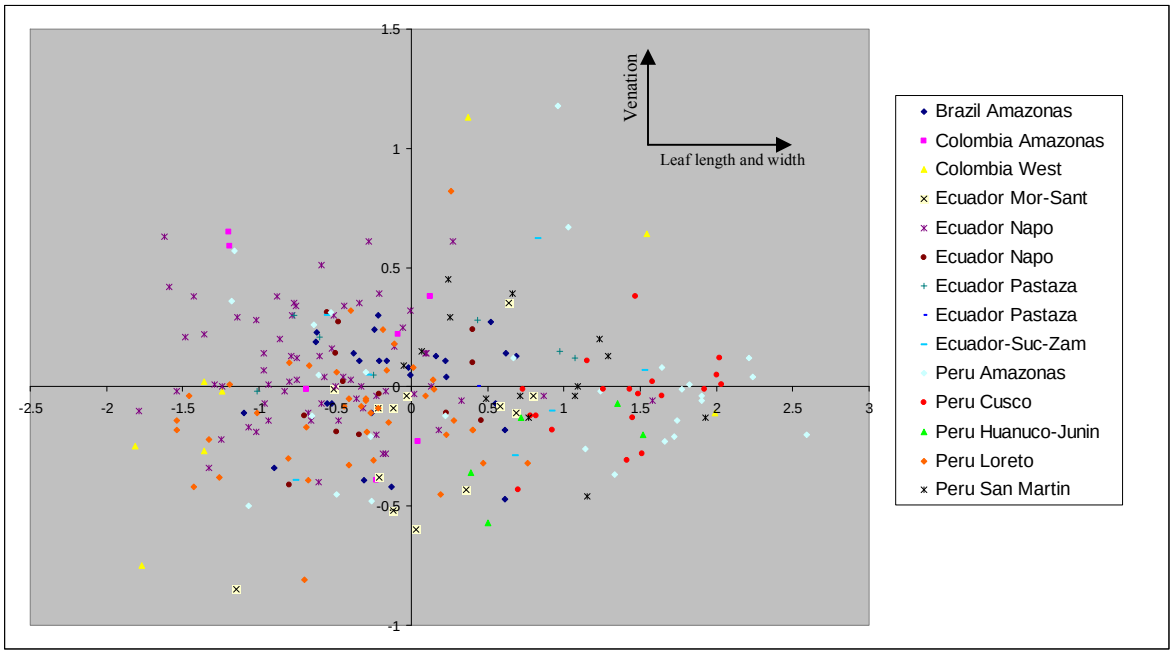

Figure 2.3 NMS ordination of quantitative leaf data for 268 specimens of *C. capitellata*. Stress value 0.09. Legend indicates the country and province or region the specimens originated from. Arrows show the morphological characters most significant to the axes as revealed by the Pearson correlation analysis.

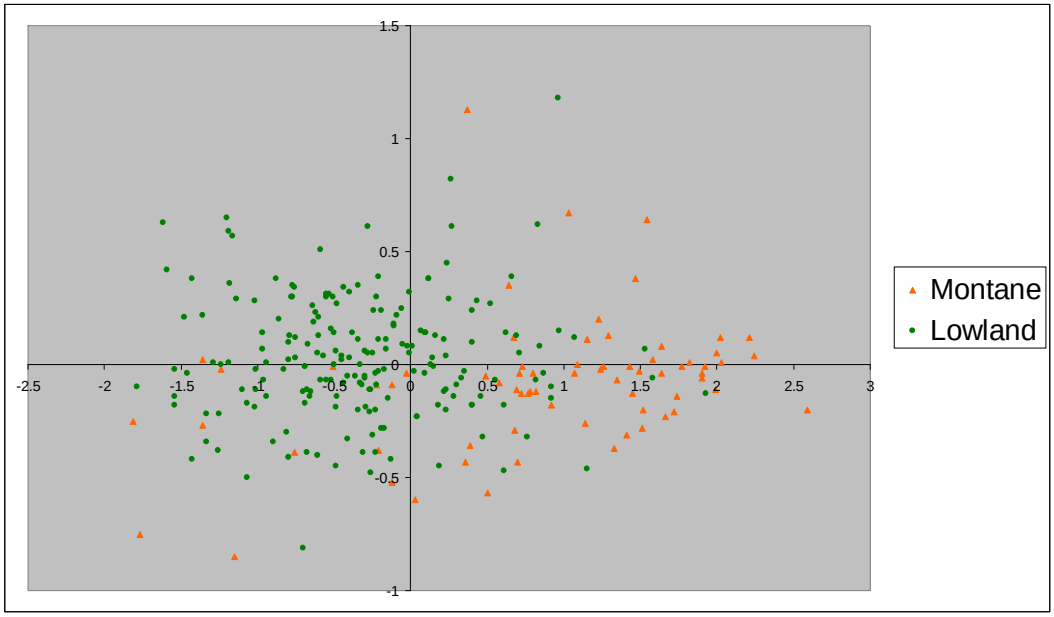

Figure 2.4 NMS ordination of quantitative leaf data for 268 specimens of the *C. capitellata* complex with specimens classified as montane and lowland. Stress value 0.09.

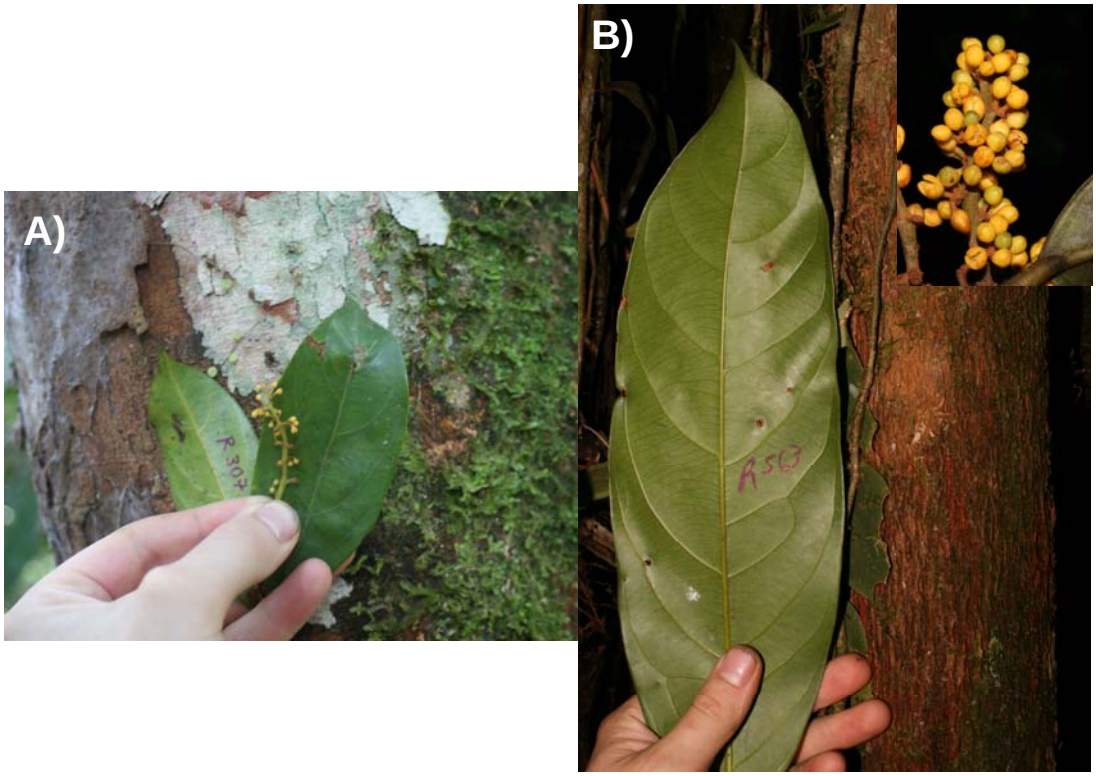

Figure 2.5 An example of bark, leaf and flower morphological trait variation in *C. capitellata* among (A) lowland populations (Ecuador, Napo) and (B) montane populations (Peru, Cusco).

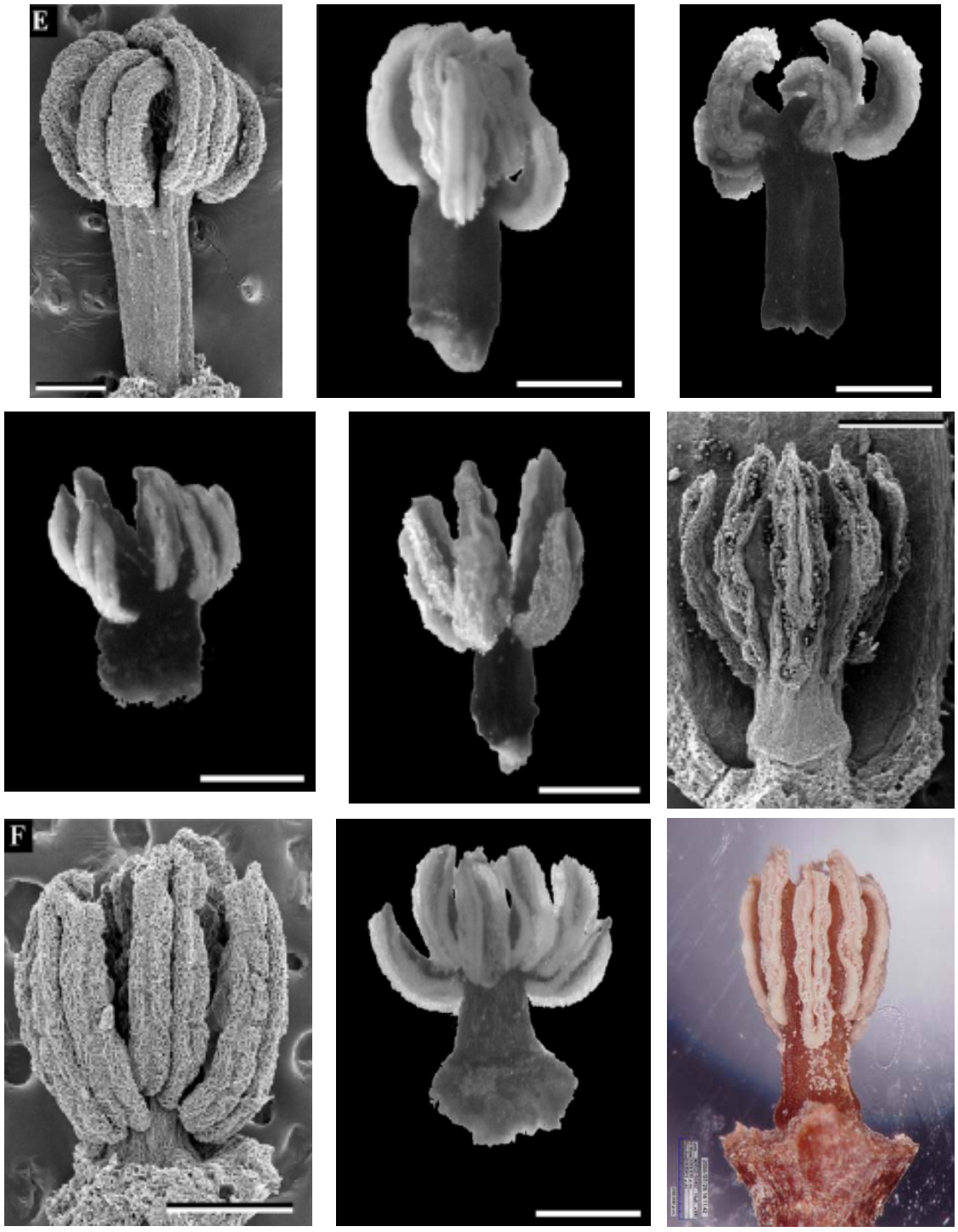

Figure 2.6 S.E.M. and light micrographs of androecial morphology of 9 staminate collections from the *C. capitellata* complex of Brazil, Ecuador and Peru. A) R. L. Froes 34648, Brazil (scale bar= 250 µm) B) A. Ducke, 561, Brazil (scale bar= 500 µm) C) A. Ducke 23693, Brazil (scale bar= 500 µm) D) M. Aulestia 2713, Ecuador (scale bar= 500 µm) E) M. Aulestia 1651, Ecuador (scale bar= 500 µm) F) P. C. D. Cazalet and T. D. Pennington 7602, Ecuador (scale bar= 500 µm) G) E. P. Killip and A. C. Smith 26053,

Peru (scale bar= 500 µm) H) E. Ancuash 424, Peru (scale bar= 500 µm) I) R. S. Steeves 600, Peru (scale bar= 200 µm).

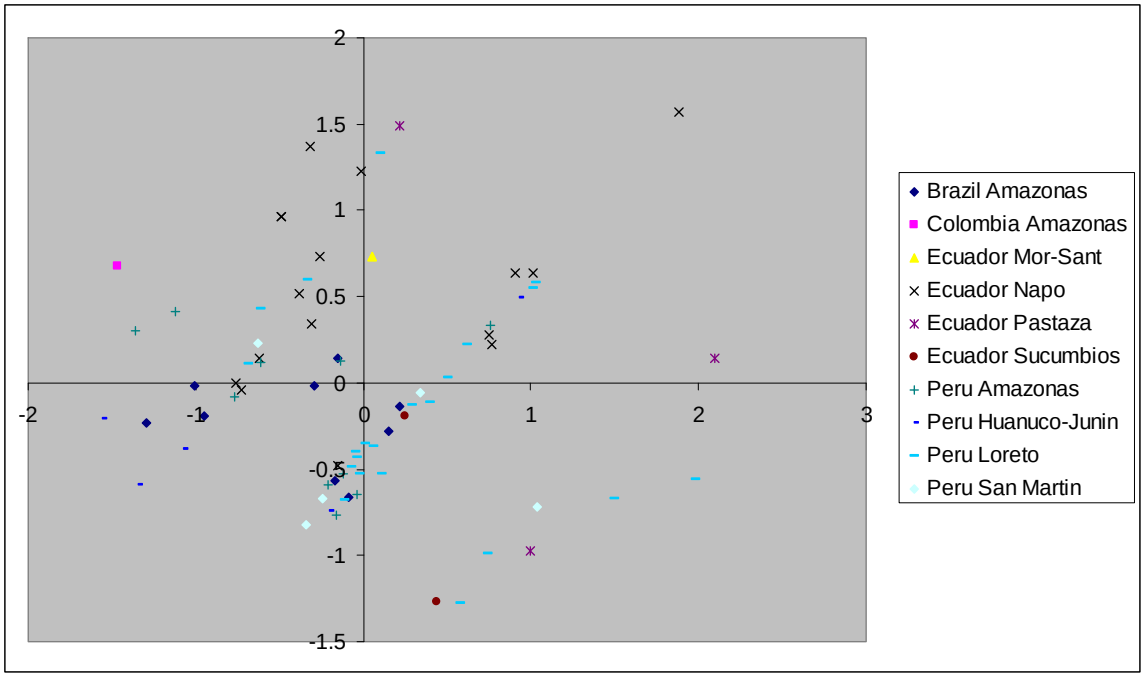

Figure 2.7 NMS ordination of 14 androecial characters for 75 samples. The legend indicates the country and region of origin for the samples. Stress=0.13.

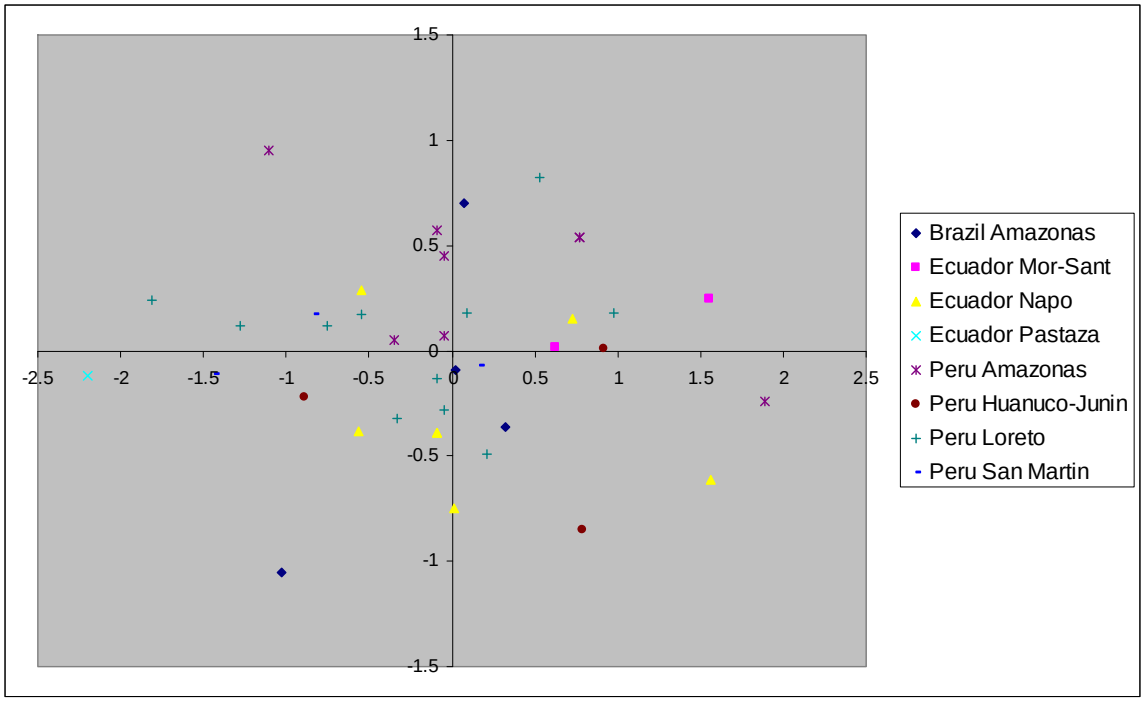

Figure 2.8 NMS ordination of 7 perianth characters for 41 samples. Legend indicates the country and region of origin for the samples used. Stress=0.10.

Figure 2.9 Fruit variation in *C. capitellata*: A) Intact glabrous fruit type found in San Martin, Peru (Top: J. Janovec 836) and intact sulcate fruit of Loreto, Peru (bottom: J. Janovec 843) B) Dissected fruit from Pastaza province, Ecuador (R. Steeves 427); interior of pericarp on left, seed exterior and interior in middle, and warted external surface of pericarp on the right.

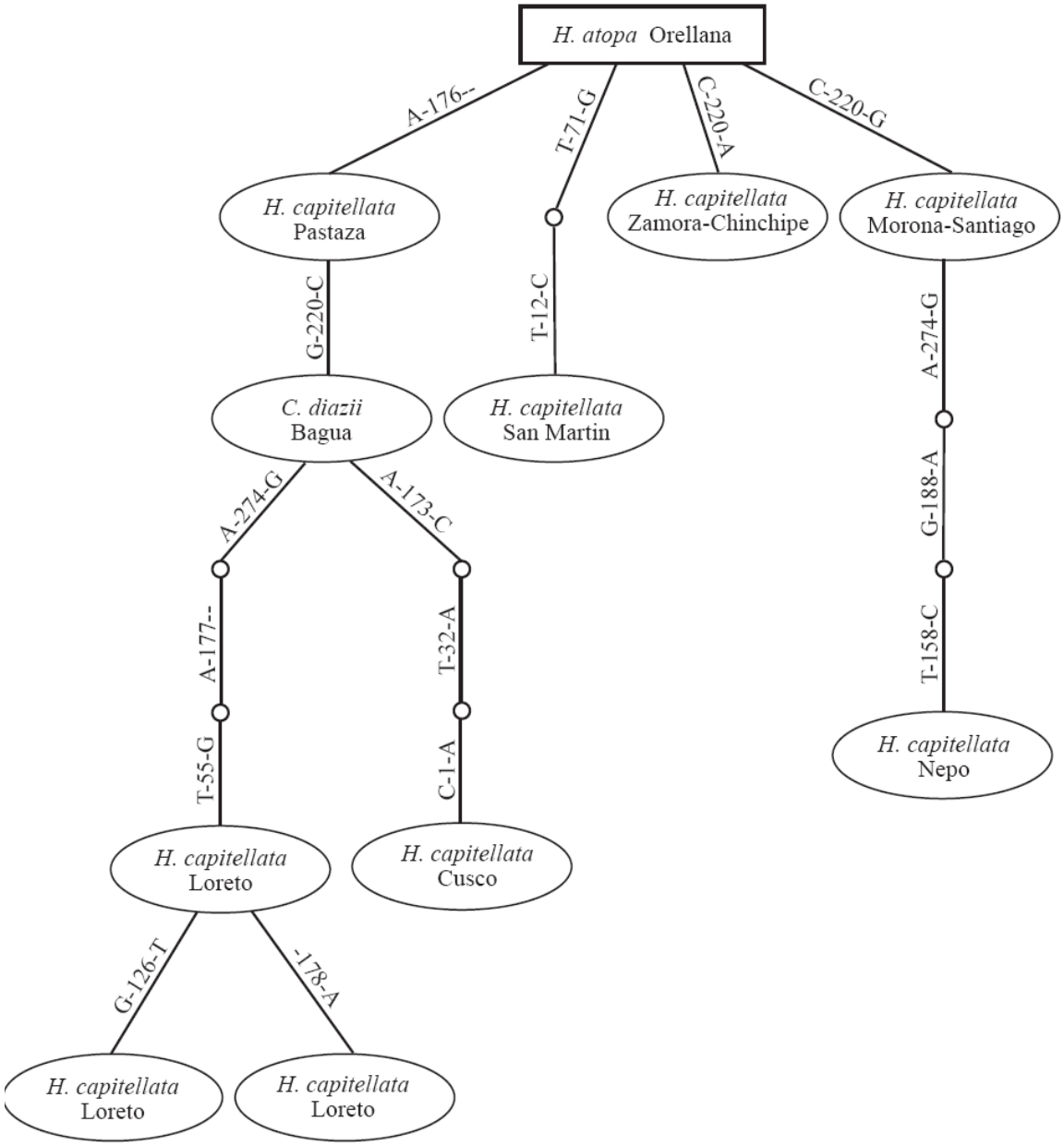

Figure 2.10 trnH-psbA haplotype network for 7 populations of *C. capitellata* and 1 sample each of *C. atopa* and *C. diazii*. Nucleotide changes, and their positions are shown on branches between nodes.

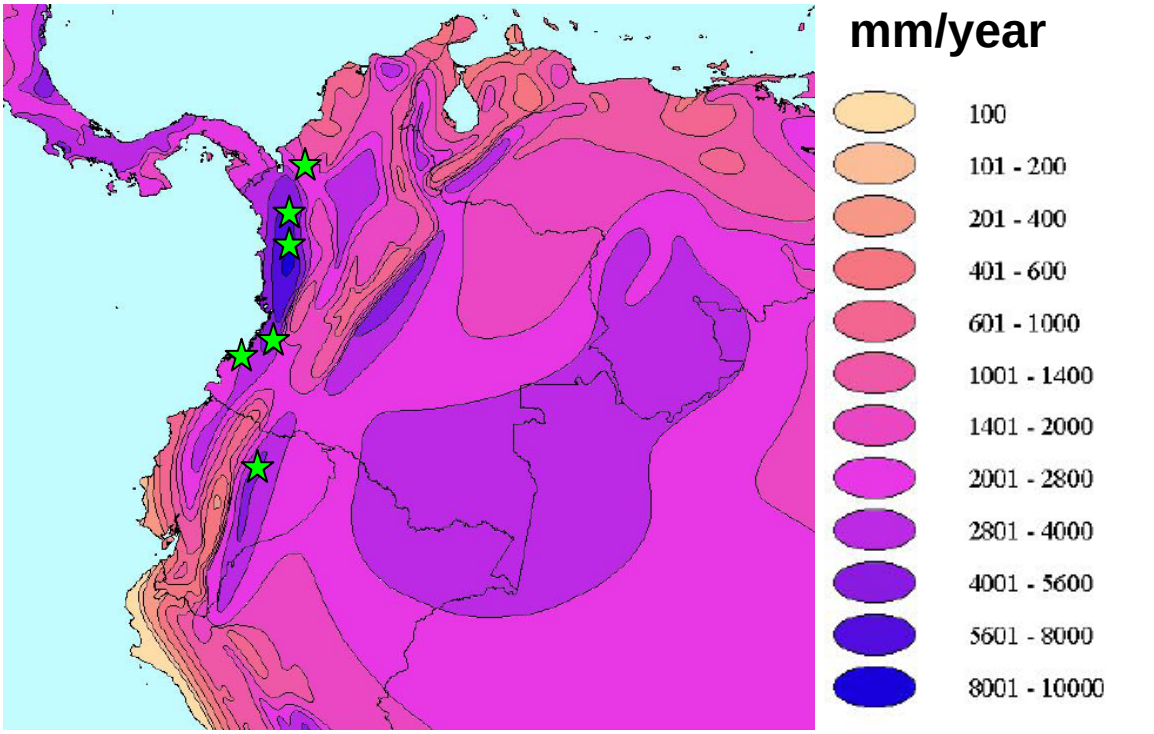

Figure 2.11 Precipitation regimes in South America and collection localities of *C. atopa*.

Literature Cited

Armstrong, J. E., & Irvine, A. K. 1989. Floral biology of *Myristica insipida* (Myristicaceae), a distinctive beetle pollination syndrome. *American Journal of Botany* **76**: 86–94.

Asif, M. J., & Cannon, C. H. 2005. DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*). *Plant Molecular Biology Reporter* **23**: 185–192.

Bullock, S. H. 1982. Population structure and reproduction in the neotropical dioecious tree *Compsoneura sprucei*. *Oecologia* **55**: 238–242.

Candolle, A. D. 1856. Myristicaceae. *Prodromus Systematis Naturali Vegetablis* **14**: 187-208.

Caron, H., Dumas, S., Marque, G., Messier, C., Bandou, E., Petit, R. J., & Kremer, A. 2000. Spatial and temporal distribution of chloroplast DNA polymorphism in a tropical tree species. *Molecular Ecology* **9**: 1089–1098.

Chan, K. Y., Zwieten, L., Meszaros, I., Downie, A., & Joseph, S. 2007. Agronomic values of greenwaste biochar as a soil amendment. *Australian Journal of Soil Research* **45**: 629-634.

Clement, M., Posada, D., & Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1659.

Cole, R. J. 2009. Postdispersal seed fate of tropical montane trees in an agricultural landscape, Southern Costa Rica. *Biotropica* **41**: 319–327.

Demesure, B., Comps, B., & Petit, R. J. 1996. Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* **50**: 2515-2520.

Fazekas, A. J., Steeves, R., & Newmaster, S. G. 2010. Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques* **48**: 277–285.

Gentry, A. H. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* **69**: 557–593.

Gentry, A. H. 1988. Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 156-159.

Golden, J. L., & Bain, J. F. 2000. Phylogeographic patterns and high levels of chloroplast DNA diversity in four *Packera* (Asteraceae) species in southwestern Alberta. *Evolution* **54**: 1566–1579.

Golenberg, E. M., Clegg, M. T., Durbin, M. L., Doebley, J., & Pow Ma, D. 1993. Evolution of a noncoding region of the chloroplast genome. *Molecular phylogenetics and evolution* **2**: 52-64.

Graham, S. W., Reeves, P. A., Burns, A. C., & Olmstead, R. G. 2000. Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. *International Journal of Plant Science* **161**: S83–S96.

Grossman, J. M., O'Neill, B. E., Tsai, S. M., Liang, B., Neves, E., Lehmann, J., & Thies, J. E. 2010. Amazonian anthrosols support similar microbial communities that differ distinctly from those extant in adjacent, unmodified soils of the same mineralogy. *Microbial Ecology* **60**: 192-205.

Hackenberger, M. J., Petersen, J. B., & Neves, E. G. 1999. Village size and permanence in Amazonia: two archaeological examples from Brazil. *Latin American Antiquity* **10**: 353-376.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series*. (pp. 95–98).

Hamilton, M. B. 1999. Tropical tree gene flow and seed dispersal-deforestation affects the genetic structure of the surviving forest fragments. *Nature* **401**: 129-130.

Holbrook, K. M., & Loiselle, B. A. 2009. Dispersal in a Neotropical tree, *Virola flexuosa* (Myristicaceae): Does hunting of large vertebrates limit seed removal? *Ecology* **90**: 1449–1455.

Hollingsworth, P. M., Graham, S. W., & Little, D. P. 2011. Choosing and using a plant DNA barcode. *PloS one* **6**: e19254.

Howe, H. F., & Vande Kerckhove, G. A. 1981. Removal of wild nutmeg (*Virola surinamensis*) crops by birds. *Ecology* **62**: 1093–1106.

Janovec, J. P. 2000. A systematic study of *Compsoneura* (A. DC.) Warb., A Neotropical member of the nutmeg family. Texas A&M University Dissertation: 1-359.

Janovec, J. P. 2002. Compsoneura diazii (Myristicaceae), a new species from the Rio Cenepa area of Northwestern Peru. *Novon* **12**: 366–368.

Janzen, D. H. 1983. The Pleistocene hunters had help. *The American naturalist* **121**: 598–599.

Janzen, D. H. 1967. Why mountain passes are higher in the tropics. *The American Naturalist* **101**: 233-249.

Janzen, D. H., & Martin, P. S. 1982. Neotropical anachronisms: the fruits the gomphotheres ate. *Science* **215**: 19-27.

Kim, J. S., Sparovek, G., Longo, R. M., De Melo, W. J., & Crowley, D. 2007. Bacterial diversity of terra preta and pristine forest soil from the Western Amazon. *Soil Biology and Biochemistry* **39**: 684–690.

Kreader, C. A. 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology* **62**: 1102-1106.

Kruskal, J. B. 1964. Nonmetric multidimensional scaling: a numerical method. *Psychometrika* **29**: 115–129.

La Rotta, C. 1985. *Estudio etnobotanico de las especies utilizadas por la comunidad indigena Emera del alto rio Baudo. Mecanografido. Fundacion Segunda Expedicion Botanica. Bogota, Colombia.* Bogota Colombia: Mecanografido. Fundacion Segunda Expedicion Botanica.

Latouche-Hallé, C., Ramboer, A., Bandou, E., Caron, H., & Kremer, A. 2003. Nuclear and chloroplast genetic structure indicate fine-scale spatial dynamics in a neotropical tree population. *Heredity* **91**: 181–190.

Liang, B., Lehmann, J., Solomon, D., Sohi, S., Thies, J. E., Skjemstad, J. O., Luizao, F. J., Engelhard, M. H., Neves, E. G., & Wirick, S. 2008. Stability of biomass-derived black carbon in soils. *Geochimica et Cosmochimica Acta* **72**: 6069–6078.

Liang, B., Solomon, D., Grossman, J. M., O'Neill, B., Skjemstad, J. O., Thies, J., Luizao, F. J., J. Petersen, & Neves, E. G. 2006. Black carbon increases cation exchange capacity in soils. *Soil Science Society of America Journal* **70**: 1719-1730.

Minchin, P. R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. *Plant Ecology* **69**: 89–107.

Müller, K. 2006. Incorporating information from length-mutational events into phylogenetic analysis. *Molecular phylogenetics and evolution* **38**: 667–676.

Newmaster, S. G., Fazekas, A. J., Steeves, R. A. D., & Janovec, J. 2008. Testing candidate plant barcode regions in the Myristicaceae. *Molecular Ecology Resources* **8**: 480–490.

Palma-Silva, C., Wendt, T., Pinheiro, F., Barbará, T., FAY, M. F., Cozzolino, S., & Lexer, C. Sympatric bromeliad species (*Pitcairnia spp.)* facilitate tests of

mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* 20: 3185–3201

Palmé, A. E., Semerikov, V., & Lascoux, M. 2003. Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity* **91**: 465–474.

Panchal, M. 2007. The automation of nested clade phylogeographic analysis. *Bioinformatics* **23**: 509-510.

Paz-Rivera, C., & Putz, F. E. 2009. Anthropogenic soils and tree distributions in a lowland forest in Bolivia. *Biotropica* **41**: 665–675.

Petit R.J., Csaikl, U.M., Bordacs, S., Burg, K., Coart, E., et al. 2002. Chloroplast DNA variation in European white oaks - Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management* **156**: 5–26.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2002. A comparison of tree species diversity in two upper Amazonian forests. *Ecology* **83**: 3210–3224.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* **82**: 2101–2117.

Poinar, H. N. 1998. Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281**: 402-406.

Posada, D., & Crandall, K. A. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**: 37–45.

Prance, G. T. 1977. Floristic inventory of the tropics: Where do we stand? *Annals of the Missouri Botanical Garden* **64**: 659–684.

Rieseberg, L.H., & Soltis, D.E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**: 65-84.

Rohlf, F. J. 2006. *tpsDig version 2.10, Department of Ecology and Evoluation*. State University of New York at Stony Brook, New York.

Roosevelt, A. C., Lima da Costa, M., Lopes Machado, C., Michab, M., Mercier, N., Valladas, H., Feathers, J., Barnett, W., Imazio da Silveira, M., Henderson, A., Silva, J., Chernoff, B., Reese, D. S., Holman, J. A., Toth, N., et al. 1996. Paleoindian cave dwellers in the Amazon: The peopling of the Americas. *Science* **272**: 373-384.

Sang, T., Crawford, D. J., & Stuessy, T. F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). *American Journal of Botany* **84**: 1120-1136.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W., & Le Thomas, A. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* **142**: 125–186.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E., & Small, R. L. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**: 142-166.

Smith, A. C. 1956. Studies of South American plants: XV. *American Journal of Botany* **43**: 573-577.

Smith, A. C. 1937. The American species of Myristicaceae. *Brittonia* **2**: 393-510.

Smith, N. J. H. 1980. Anthrosols and human carrying capacity in Amazonia. *Annals of the Association of American Geographers* **70**: 553–566.

Štorchová, H., & Olson, M. S. 2007. The architecture of the chloroplast psbA-trnH non-coding region in angiosperms. *Plant Systematics and Evolution* **268**: 235-256.

Tate, J. A., & Simpson, B. B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* **28**: 723–737.

Templeton, A. R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* **7**: 381–397.

Templeton, A. R., Crandall, K. A., & Sing, C. F. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619-633.

ter Braak, C. 1998. *Canoco 4 Centre for Biometry.* Wageningenm, The Netherlands.

Thomas, W. W. 1999. Conservation and monographic research on the flora of Tropical America. *Biodiversity and Conservation* **8**: 1007–1015.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**: 4673-4680.

Tobler, M., Honorio, E., Janovec, J., & Reynel, C. 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: an example of two neotropical plant families (Moraceae and Myristicaceae) in Peru. *Biodiversity and Conservation* **16**: 659-677.

Warburg, O. 1897. Monographie der Myristicaceen. *Nova Acta Acad. Caes. Leop.-Carol* **68**: 1-680.

Chapter 3

A MOLECULAR PHYLOGENETIC INVESTIGATION OF VIROLA Aublet

Abstract

The Myristicaceae are a widespread and specios family of tropical trees with great ethnobotanical and ecological importance, yet they remain to be understood from a phylogenetic perspective. *Virola* is a genus comprised of about 60 species of canopy and sub-canopy trees endemic to the neotropics whose fruits are important food for many vertebrates and sap is revered by numerous native tribes as a source of medicine and hallucinogenic snuff. The objective of this study was to estimate a phylogenetic tree for to illuminate infrageneric relationships among *Virola* taxa from Northwestern South America. An additional objective of this study was to test whether the trnH-psbA spacer could differentiate species of *Virola* and whether DNA sequence data indicates undescribed species. Although levels of DNA sequence divergence were low, phylogenetic hypotheses generated using both Bayesian and Parsimony methodologies supported similar topologies and showed a great deal of divergence amongst two groups in the genus (Multinervae and Sebiferae). Phylogenetic analyses also supported the recognition of 3 new provisional species. The most variable plant DNA barcode known for the Myristicaceae, the trnH-psbA intergenic spacer, failed to differentiate the majority of species included in this study.

Introduction

The Myristicaceae are a family of flowering plants consisting of 21 genera and about 500 species found worldwide in tropical rainforests. The family is best known as the source of the spices nutmeg and mace, which are produced from the fruit of the Asian species *Myristica fragrans*. Myristicaceae species are small to large trees (5-55 m tall) that hold significant ecological importance throughout most of the lowland wet tropical forests of the world (Gentry 1982, Pascal and Pelissier 1996, Pitman et al. 2002, Pitman et al. 2008). The family has been represented in molecular systematics studies (Chase et al. 1993, Qiu et al. 1993, Sauquet et al. 2003, Soltis et al. 2011), which have confirmed traditional hypotheses of classification of the Myristicaceae as belonging to order Magnoliales of subclass Magnoliidae. Unfortunately, few studies have addressed species diversity and distribution of the Myristicaceae within genera (Sauquet et al. 2003, Sauquet 2004), at the level of species (Janovec and Harrison 2002), populations (Degen et al. 2004), and individual specimens. Because of their abundance and diversity in tropical moist forests of the world, the Myristicaceae are an ideal family for studying speciation, evolution, and biogeography of lowland tropical forests but such investigations are difficult without well resolved phylogenies.

Virola is a genus of approximately 60 described species that are endemic to the lowland and cloud rainforests of Central and South America. Species of *Virola* are vegetatively characterized by having profuse red latex (rarely green-yellow), simple, entire, alternate, and regularly spaced linear-ovate to obovate leaves that differ from other neotropical genera by the presence of dendritic or stalked-sessile stellate hairs on their

abaxial surface. Secondary veins are prominent, pinnate, and arcuate-ascending but sometimes diverge towards the apex. Like all other Myristicaceae their panicle-raceme-like inflorescences are born in leaf axils. The flowers of these dioecious trees are very small (1-4mm in diameter) and are composed of 3-4 tepals which are often covered in a dendritic or stellate pubescence. The fruits of *Virola* are composed of a dehiscent pubescent-glabrous pericarp which opens at maturity to reveal a single globose-elliptic seed which is covered in a bright orange-red fat-rich aril that is attractive to numerous large vertebrate species.

Ecology

Due to their ubiquitous nature and position as one of the 5-10 most abundant fruit-producing tree genera of South American rainforests (Gentry 1988, Pitman et al. 2001, Pitman et al. 2002, Pitman 2008), *Virola spp.* have been extensively studied from an ecological perspective. For example, investigation of sex ratios in *Virola* trees have revealed a male bias within smaller age classes (< 30 cm DBH) of *Virola* trees (Ackerly et al. 1990, Queenborough et al. 2003), which is typical of dioecious tropical trees (Armstrong and Irvine 1989, Opler and Bawa 1978). The arils of *Virola spp.* are energy-dense and highly nutritious. Analyses of *Virola surinamensis* fruit arils have found a 63% lipid content, 3% protein and 9% carbohydrate (Howe and Vande Kerkhove 1981) while a similar study of *V. sebifera* found comparable percentages of lipid (53.7%), protein (7.1%) and carbohydrates (8.4%) (Howe 1981). Numerous studies have found that birds are by far the most speciose and abundant agents of seed dispersal for this genus (*V. sebifera*; Howe 1981, *V. surinamensis*; Howe and Vande Kerkhove 1981, *V. surinamensis*; Howe et al. 1985, *V. flexuosa*; Holbrook and Loiselle 2009). Although

birds are undeniably important dispersers of *Virola* seeds, they often regurgitate seeds directly under the parent tree after stripping the nutritious aril from the seeds with their gizzards (Howe 1981, Howe et al. 1985). It is therefore notable that two studies by Russo (1992, 2003) found that even though *V. calophylla* fruits were consumed by 17 species of birds, 82% (1992) and 92% (2003) of all seeds dispersed by vertebrates were removed by spider monkeys (*Ateles paniscus*). These primates typically consume the seed and defecate it many hours later, thus acting to more efficiently disperse the seeds away from the parent tree. This may be especially important as it has been found that 99.96% of *V. surinamensis* and 30-35% of *V. nobilis* seeds that fell within 45 m of a parent tree were consumed by invertebrate pre-dispersal seed predators (Howe et al. 1985, Howe 1993) and studies of *V. sebifera* found low to high rates (5-96%) of seed predation in Panama (Howe 1981). These studies seem to support the conformity of *Virola* species to the Janzen-Connell model (Janzen 1970, Connell 1971). However, investigations of *V. michellii*, *V. bicuhyba* and *V. kwatae* have found high seed survival with little to no seed predation by invertebrates or no decrease in seed predation with increasing distance from the mother tree (Forget et al. 2000, Zipparo and Morellato 2005, Forget and Cuijpers 2008). Additionally, it has been found that fruit removal in *V. flexuosa* by dispersal agents is more efficient in non-hunted versus hunted forests (89.4% versus 66.8%) in the forests of Amazonian Ecuador (Holbrook and Loiselle 2009). Despite these extensive investigations of seed dispersal/predation in various species of *Virola*, little is known of of their pollination biology, what may act as herbivores/parasites, or whether these ecologically important trees are truly dioecious or diphasic in nature (alternating male and female depending on life stage).

Ethnobotany

Virola spp. also hold considerable ethnobotanical importance to both industrial and traditional societies of Central and South America. *Virola surinamensis*, known as ucuuba in the Amazonian basin, grows in the lowland floodplains of the Amazon basin from the Foothills of the Andes to the Amazonian delta in near monospecific stands and are valued for their lumber, which is in such abundant supply that it is rivaled in economic importance only by big-leafed mahogany [*Swietenia macrophylla*] (Macedo and Anderson 1993). The fatty seeds of *V. surinamensis* and *V. sebifera* are also harvested for their abundant and aromatic waxes which are used in the production of candle waxes (Williams 1960). Members of the genus are also widely employed by natives across their range for a multitude of uses. Perhaps one of the most widespread applications of *Virola* is as a hallucinogenic snuff. Traditional cultures of the Northwest Amazon purify the red latex derived from the bark of various *Virola* species to make a powdered snuff rich in various tryptamine alkaloids (primarily N,N-Dimethyltryptamine and 5-Methoxydimethyltryptamine), which purportedly produces powerful visions to aid in the teaching of medicine, divination of evil spirits, communication with other shamen, and healing of the sick (Macrae and Towers 1984a and b, Chagnon 1971, McKenna et al. 1984, Bennett and Alarcon 1994, Schultes and Holmstedt 1968, Schultes 1981, Schultes and Raffauf 1990, Prance 1972). Interestingly, similar preparations of the sap can be used as arrow/dart poisons to immobilize animals (Mckenna et al. 1984, MaCrae and Towers 1984b). Although these tryptamines are rendered inert by ingestion via the enzyme monoamine oxidase (MAO), some tribes have combined potent MAO inhibitors with the resin to make it orally active as an entheogen (Mckenna et al. 1984). *Virola* sap is also

used to decrease pain and inflammation, as an anti-fungal skin treatment, to prevent and treat dental caries, bladder/stomach ailments, thrush, marlaria, and is given as a snuff to hunting dogs to increase their ability to smell (Davis and Yost 1983, Schultes and Raffauf 1991, Beloz 1992, Bennett and Alarcon 1994, Roumy et al. 1997). The limited pharmacological investigations performed on *Virola* species have found them efficacious as oral hallucinogens as well as anti-malarial and anti-microbial agents (McKenna et al. 1984, Roumy et al. 1997, Lopes et al. 1999).

Taxonomy

Despite being the most species-rich genus of neotropical Myristicaceae, *Virola* has received little taxonomic attention. *Virola* was first described by Aublet (1775) as a genus endemic to Central and South America. The first comprehensive treatment of the Myristicaceae was performed by De Candolle (1856) who organized 90 species into 13 sections of the single genus *Myristica*. De Candolle grouped taxa we today recognize as *Virola* into sections *Virola* and *Sychoneura* based upon length of the filament column, anther apiculation, and leaf venation. Warburg (1897) relegitimated the generic concept of *Virola* in his thorough study of the Myristicaceae where he proposed its subdivision into sections *Oxyanthera* and *Amblyanthera* on the basis of the ratio of filament column to anther ratios. The genus was further divided into 6 unranked species groups (*Mollissimae*, *Sebiferae*, *Calophyllae*, *Rugulosae*, *Surinamenses*, and *Subsessilis*) in Smith's (1937) monograph of the neotropical Myristicaceae. Although these sections were arguably more natural and utilitarian owing to the inclusion of vegetative traits as well as characters of the pistillate inflorescences, some species belonged to multiple groups. The last comprehensive treatment of *Virola* was performed by Rodrigues (1980)

but is somewhat incomplete due to its restriction to Brazilian members of the genus. Walker and Walker (1979) examined pollen morphology in 30 species of *Virola* which were found to be highly similar in terms of size and shape but divided them into tree groups based upon differences in exine sculpturing patterns. Additional palynological comparisons among genera of Myristicaceae have shown that *Virola* along with the other American, Asian and African genera possess monosulcate boat-shaped pollen with a reticulate-rugulate tectum and a columnate infratectum (Sauquet and le Thomas 2003). Since 1980, taxonomic investigations of *Virola* have been primarily alpha-taxonomic in nature and about 60 species are currently recognized in the genus (Rodrigues 1989, Sabatier 1997, Jaramillo et al. 2000, Rodrigues 2002). Paleobotanical studies of *Virola* have been difficult as the genus is represented in the fossil record by a single cluster of flowers from the Dominican amber (Poinar and Poinar 1999).

In spite of its abundance and diversity in the Neotropics, *Virola* has been underrepresented in modern molecular studies of the Myristicaceae. *Virola* was represented by three species in the only molecular systematic study of Myristicaceae to date, which used chloroplast coding and non-coding intergenic spacers to estimate the evolutionary relationships among the 21 genera. This investigation failed to find well supported relationships among genera, largely due to low levels of nucleotide substitution among loci sampled (Sauquet et al. 2003). This phylogenetic uncertainty has hindered further studies of character evolution and biogeography of this pantropical family. Using RAPD's, Degen et al.(2001) studied *V. michellii*'s spatial genetic structure in comparison to 7 other tropical tree species and found significant genetic structure at small spatial scales (50-300m), which they hypothesized was due to aggregated seed dispersal from

monkeys excreting ingested seeds. Additionally, microsatellite markers have also been developed for *V. flexuosa* and *V. surinamensis* (Holbrook et al. 2006, Draheim et al. 2009) but their application in wide-ranging studies have yet to be published.

The objectives of this investigation were to 1) reconstruct infrageneric evolutionary relationships in order to test previous sub-generic classifications and 2) to test the ability of chloroplast DNA to discriminate species of *Virola*. Although poor in terms of resolution in some clades, a phylogenetic tree was generated for the genus and relationships are discussed with respect to ecology, morphology, and biogeography. Additionally, three new provisional species of *Virola* from the neotropics are inferred from molecular phylogenetic analysis and morphological distinctiveness.

Materials and Methods

Taxon Sampling

Virola is well represented in herbarium collections owing to their abundance in neotropical rainforests and intensive ecological investigations. These samples are often soaked in alcohol upon collection or dried immediately in the field. Most *Virola spp.* dry dark brown upon collection and DNA from samples more than a year old are generally not amenable to DNA extraction and amplification (personal observation), likely due to abundant secondary metabolites that have the ability to degrade DNA and inhibit PCR through direct enzymatic inhibition and/or the generation of non-bypassable crosslinking via maillard reactions (Poinar 1998, Sawadjoon 2002). Initial efforts to extract whole genomic DNA from *Virola* herbarium specimens were largely ineffectual. Therefore, new herbarium and DNA collections were made in Peru and Ecuador so that DNA of

sufficient quality and quantity could be extracted from specimens. Fresh DNA collections were made by John Janovec or the author by putting young and clean (i.e. visibly free of epiphytes) leaf tissue immediately on silica gel upon specimen collection in order to minimize DNA degradation. Conventional herbarium vouchers were also made of all collections. In all, 45 new collections representing 14 species of *Virola* were included in phylogenetic analyses. Greater taxon sampling was not possible due to the poor condition of DNA retrieved from herbarium specimens. Since this is the first phylogenetic investigation of the genus, it is unclear as to whether this particular suite of species represents the phylogenetic breadth of the genus. Additional sequences and species were included in distance analyses from tropical DNA barcoding campaigns (Gonzalez et al. 2009, Kress et al. 2009). All collections used in molecular analyses were identified according to the most current taxonomic treatment (Smith 1937) and all specimens employed have accessions archived at one or more of the following herbaria: Botanical Research Institute of Texas (BRIT), Ontario Agricultural College (OAC), and the National Herbarium of the Ecuadorian Museum of Natural Sciences (QCNE). Figures 3.1-3.14 show vegetative and reproductive structures of the 14 taxa collected. It was observed in the field that there were big-leafed and small-leafed morphotypes of *V. sebifera* and *V. loretensis*. Consequently, collections have been annotated with a BL (Big-Leaf) or SL (Small-Leaf) to indicate to which of the morphotypes they belong. A more thorough analysis and discussion of the taxonomic implications of these morphotypes is provided in chapter 4 of this thesis. Collection information of the specimens used in this study, including Barcode Of Life Database (BOLD) process identification numbers, can be found in Table 3.1. Since phylogenetic relationships

remain largely unresolved in the Myristicaceae (Sauquet et al. 2003), 2 members of *Iryanthera*, and *Compsoneura* were selected as outgroups for phylogenetic analysis as they are hypothesized to have close relationships to *Virola* (Sauquet et al. 2003).

Morphology

In addition to sterile material (that was photographed, pressed, and dried) fruits and flowers were dried and/or preserved in FAA (Formalin-Acetic Acid-Alcohol) for morphological investigations when they were available. A data matrix of quantitative vegetative and reproductive morphological metrics was assembled for 8 species available from Smith (1937) and other species were not included due to a lack of data on leaf, flower and fruit metics. Measurements of 10 quantitative vegetative metrics (Figure 3.15) were made from digital photographs using tpsDig 2.0 (Rohlf 2006) for use in species descriptions when provisional new species were collected.

DNA extraction, Amplification, Sequencing and Alignment

Total genomic DNA was extracted from leaf tissue of silica-dried or herbarium specimens using the Macherey-Nagel Nucleospin II Plant Kit (Macherey-Nagel, Duren Germany). Lysis buffer 1 was used according to the manufactures' instructions with the exception of an increase of the post homogenization incubation period from 10 minutes to 1hr and the addition of 20 mM N-Phenacylthiazolium Bromide, which has been found to result in improved PCR amplification of recalcitrant samples (Poinar et al. 1998, Asif and Cannon 2005).

Due to the low levels of nucleotide substitution in the Myristicaceae (Sauquet et al. 2003), the most variable *Virola* loci available were targeted by screening available chloroplast sequences from genbank (matK, psbK-I, rbcL, rpoB, rpoC1, trnH-psbA, trnL-F) and from a preliminary screening of the most variable nuclear markers from studies of *Compsoneura* (Chapter 2 of this thesis)[AGT1 and AT103]. Only trnH-psbA, AT103, and AGT1 consistently produced a single PCR product that yielded high quality sequence traces and were sufficiently variable for systematic investigations.

Taxa not represented by trnH-psbA accessions in GenBank from Newmaster et al. (2008) were PCR amplified and sequenced using the primers trnH2 (5'-CGCATGGTGGATTCACAATCC-3'; Tate and Simpson 2003) and psbAF (5'GTTATGCATGAACGTAATGCTC-3'; Sang et al. 1997). PCR was performed in a 20 µl volume using 0.4 µl of Phire II polymerase (Finnzymes) with 1X Phire II reaction buffer (with 1.5mM MgCl), 0.2 mM of each DNTP, 0.2 µM of each primer and 2.0 µg of BSA (Kreader 1996). Cycling conditions entailed an initial denaturation step of 1 min at 98°C; 35 cycles of 98°C for 5 s, 64°C for 5 s, 72°C for 10 s: and a final elongation step of 72°C for 1 min followed by a 4° hold. Phire II was used to amplify trnH-psbA as it is robust to the inhibitors contained in nutmeg extracts (personal observations) and as it is a fusion-based polymerase which has been found to reduce stuttering in sequences containing homopolymer regions such as the trnH-psbA intergenic spacer (Fazekas et al. 2010).

Nuclear loci were amplified using the primers AT103F (5'-CTTCAAGCCMAAGTTCATCTTCTA-3'; Li et al. 2008), AT103R (5'-TTGGCAATCATTGAGGTACATNGTMACATA-3'; this thesis), AGT1-MYR-F (5'-

GGGCATTGACGTAGCTTTGACAGG-3'; this thesis), and AGT1-MYR-R (5'-GTGCAGTTCTTCAAGCCCCAAGC-3'; this thesis). Nuclear loci were amplified with 0.5U of Ampli*Taq* Gold (Applied biosystems) DNA polymerase in a 20 µl reaction containing 1X reaction buffer, 2.5 mM MgCl, 8% W/V Polyethylene glycol (Zimmerman and Harrison 1987, Sasaki et al. 2006) , 0.2 M trehalose (Speiss et al. 2004), 2 µg BSA, 0.2 mM each DNTP, and 0.2 µM of each primer.

Amplification products were sequenced directly using the same primers employed in PCR. Cycle sequencing reactions were performed in a 10.5 µL reaction volume containing 0.5 µL of BigDye terminator mix v3.1, 1.88 µL of 5x sequencing buffer (Applied Biosystems), 1.0 µM of primer and 0.5 µL of PCR product. Thermal cycling parameters were 96° for 2 min; 30 cycles of 96° for 30s, 56-60° (primer dependant) for 15s, and 72° for 4 min; and a 4° hold. Cycle sequencing reactions were cleaned using sephadex columns (Cat.no. S5897; Sigma-Aldrich, St. Louis, MO, USA) and the samples were run on an ABI 3730 sequencer (Applied Biosystems).

Sequence contigs were assembled and edited using Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA). DNA sequences were aligned using the default settings of the ClustalW algorithm (Thompson et al. 1994) in Bioedit (Hall 1999) and adjusted manually. Alignments can be found in Appendix 3. Gaps in the alignments were coded using Simple Indel Coding (SIC) (Simmons and Ochoterena 2000) for Bayesian analyses and Modified Complex Indel Coding for distance analyses using Indelcoder (Muller 2006).

Cladistic analyses

An incongruence length difference test (ILD: Farris et al. 1994) was executed in PAUP (Swafford 2002) to determine whether the three loci employed in this study contained significant incongruence. The ILD was performed using 100 replicates of a heuristic search strategy, MAXTREES set at 100, 10 random addition sequence replicates holding 5 trees at each step, MULTREES in effect, and tree bisection-reconnection (TBR) branch swapping. No significant incongruence was detected (p-value=0.98) therefore all three loci were concatenated in a total evidence approach in Bayesian and Parsimony analyses.

Phylogenetic trees were generated using Bayesian inference with the program MrBayes (Ronquist and Huelsenbeck 2003). Nucleotide substitution models were first selected using Mrmodeltest (Nylander 2004) which uses the Akaike Information Criterion to assess the significance of adding parameters to the different models of evolution employed by MrBayes 3.1 (Ronquist and Huelsenbeck 2003). Mrmodeltest selected the following substitution models for the three loci: trnH-psbA= F81, AT103=GTR, and AGT1=HKY+I. The three genetic loci were assigned to separate partitions with their corresponding models of nucleotide substitution and the indel data was given a fourth partition and assigned the model "restriction data" as recommended by Ronquist and Huelsenbeck (2003). Ten million generations were performed using 4 chains and 2 runs with trees being sampled every 100 generations. Log-likelihood values stabilized after 2.5 million generations; therefore the consensus tree and posterior probabilities were estimated using a burn-in of 25,000 trees. Since posterior probabilities are largely considered to be overestimations of clade support (Douady et al. 2003, Alfaro and Holder 2006, Yang and Rannala 2010), especially when employing concatenated

data sets (Suzuki et al. 2002), all nodes with less than 0.75 posterior probability were collapsed on the cladogram using treegraph2 (Stover and Muller 2010).

Maximum parsimony analyses were performed in PAUP 4.0b10 (Swafford 2002) and each nucleotide position was treated as an unordered character and all positions were equally weighted. A "two step" analysis was performed as it may permit the investigation of more optimal tree topologies compared to a "one step" analysis (Davis et al. 2004). The first heuristic search was performed with 1000 replicates of random addition sequence holding 10 trees at each step, TBR branch swapping with the MULTREES option not in effect with MAXTREES set to 5000. Using these trees in memory from the initial analysis, a second heuristic search was performed using the same parameters with a MAXTREES of 50 000 and the MULTREES option in effect. Branch support was estimated using a heuristic search, 5000 bootstrap replicates, 10 random addition sequence replicates holding 1 tree at each step, TBR branch swapping, and a maximum number of trees set at 50,000 with MULTREES not in effect.

Distance analyses

Full length trnH-psbA sequences of 13 species of *Virola* were amplified from whole-genomic DNA extracts and 5 species' sequences were retrieved from Genbank accessions (*Virola kwatae*: FJ039018.1;*Virola michelli*: GQ428669.1, FJ039021.1, FJ039020.1, FJ039019.1;*Virola multicostata*: GQ428671.1, GQ428670.1; *Virola multiflora*: GQ982401.1;*Virola nobilis*: GQ982402.1). These sequences were then used to generate a Neighbour-Joining uncorrected p-distance tree using PAUP 4.0 (Swafford 2002) to explore the ability of this locus to discriminate species of *Virola*. Uncorrected p-

distance was used as there existed very low levels of nucleotide substitution among taxa and it was therefore not thought to be beneficial to include a model of molecular substitution patterns, which is employed by other distance calculation methods. An 11 bp tract of homopolymer A/T repeat was omitted from distance analyses due to the high likelihood of homoplasious indels.

Results

Morphology

Morphological trait data was retrieved from Smith (1937) for a total of 8 of the 14 species included in the phylogenetic analyses. These data are presented in Table 3.2 and the species are broken up into informal groups Sebiferae and Multinervae (see "Cladistic Analysis" below) in order to show contrasting morphological traits of two clades from phylogenetic analyses. Many values were very comparable amongst the two clades, however, leaf vein number was greater in members of Multinervae (20-60 veins) than Sebiferae (9-28), Multinervae also had thicker pericarps (1-7 mm) than did members of Sebiferae (0.3-1.3 mm) and the arils of Multinervae taxa are generally lacinate to nearly the base while those of Sebiferae taxa are only lacinate for half the length of the seed.

DNA sequencing and Alignment

Bi-directional sequence pherograms were obtained for 41 *Virola* samples and 4 outgroup members for the trnH-psbA intergenic spacer. The trnH-psbA alignment for the distance analysis included 9 additional sequences retrieved from Genbank and was 363 base pairs (bp) long with 49 variable sites, 23 of which were provided from outgroup

taxa. These sequences varied from 162-339 bp in length. The trnH-psbA alignment for phylogenetic analysis was 357 bp in length and contained 42 variable and 30 parsimony informative sites, with outgroups responsible for 22 and 11 of these sites respectively.

AGT1 bi-directional sequences were recovered from 42 *Virola* and all 4 outgroup taxa. The aligned AGT1 data matrix was 728 bp in length with 122 variable sites of which 89 were parsimony informative. Outgroups provided 38 variable and 21 parsimony informative sites. The short coding section (264 bp) of AGT1 contained no stop codons.

AT103 sequences were recovered from all 49 samples (45 Virola taxa and 4 outgroup members). The aligned AT103 data matrix was 438 bp in length with 42 variable sites of which 27 were parsimony informative. Outgroup taxa were responsible for 16 variable and 8 parsimony informative sites. AT103 Sequences varied from 428-438 bp in length.

The concatenation and alignment of all three loci for phylogenetic analyses was 1517 bp in length and contained 206 (13.6%) variable sites, of which 146 (9.6%) were parsimony informative. The 4 outgroup taxa contributed 76 variable and 40 parsimony informative characters.

Cladistic Analysis

The first heuristic search of the concatenated data retrieved 2856 equally parsimonious trees (Length=245 steps). The second heuristic search yielded 50,000 equally parsimonious trees (Length=245 steps, consistency index [CI]=0.910, retention index [RI]=0.979, and rescaled retention index [RC]=0.891). The semi-strict consensus tree of the 50,000 most parsimonious trees was almost identical in topology to that of the

Bayesian analysis. Therefore, parametric bootstrap support measures were plotted on the Bayesian cladogram.

The Bayesian consensus tree and posterior probabilities of the two MCMC runs estimated in MrBayes with proportional branch lengths is presented in Figure 3.16 along with bootstrap support values. Since the Bayesian inference and parsimony topologies are virtually the same, the results of the two will be discussed together. Since two large clades were evident on the resulting tree, the clades were informally named Multinervae and Sebiferae for the sake of discussion.

Distance analyses

Four outgroup and 41 full length *Virola* trnh-psbA sequences (the same sequences as used in phylogenetic analyses) were recovered from fresh, silica gel dried, tissue and an additional 9 sequences were retrieved from Genbank. An uncorrected p-distance phylogram is presented in Figure 3.17. Of the 18 aligned species of *Virola*, only 6 (*V. sp. RADS1*, *V. caducifolia*, *V. multicostata*, *V. michellii*, *V. pavonis*, and *V. kwatae*) form distinct monophyletic clades on the phylogram. All but one sample of *V. multinervia* formed a clade. All 8 species of the Sebiferae clade possessed a single haplotype.

Discussion

Cladistic analysis and Morphology

This study is the first molecular systematic investigation of *Virola* and the first systematic study of this species-rich and widely distributed genus since Rodrigues'

investigation of its Brazilian members (1980). Otto Warburg (1897) presented the first cladistic hypothesis of the Myristicaceae based on morphological characters but even contemporary attempts to hypothesize relationships within the family using both morphological and molecular data have failed to find well supported relationships, mainly due to the conservative nature of morphological and molecular characters and/or recent diversification of the family (Sauquet et al 2003). Smith (1937) noted that previous attempts by De Candolle, Warburg and Ducke to form groups within *Virola* have relied mainly on characteristics of the androecium that were, at times, continuous. This heavy reliance on characters of the male flower also made classification of female or otherwise sterile specimens difficult. Therefore, Smith (1937) sought groupings supported by vegetative as well as male and female reproductive characters so that specimens without flowers and fruits could be more easily keyed to groups with major morphological affinities. Since Smith's primary objective was identification and not estimating ancestor-descendant relationships, some species were placed in multiple groups.

Although there are few species that form well supported monophyletic clades, this investigation recovered numerous clades above the species-level within *Virola* supported by high bootstrap and posterior probability values. Since node support inferred from Bayesian posterior probabilities are generally viewed as overestimations (Suzuki et al. 2002, Douady et al. 2003), a conservative approach was taken in this analysis by collapsing nodes with relatively low support (<0.75); the resulting tree is nearly identical to relationships inferred from the more conservative approach of parametric bootstrapping.

Both tree building approaches revealed cladistic support for Smith's (1937) sub-generic groups that were included in this analysis. The first is Smith's Rugulosae, containing *V. multinervia* and *V. flexuosa*, which was recovered in my phylogenetic analysis, Smith also included *V. loretensis* which is not placed in this clade on the phylogeny so it appears that this clade is paraphyletic with respect to molecular data. Surinameses, *sensu* Smith (1937), encompasses *V. surinamensis* and *V. pavonis*, also formed a monophyletic clade. Although there is relatively limited taxon sampling in this analysis, it appears that there is considerable degree of support for the groupings of Smith (1937).

The genus *Virola* is divided into two well supported clades that I have elected to informally name "Multinervae" and "Sebiferae". These two clades have numerous contrasting morphological traits as well as some ecological and ethnobotanical distinctions. Multinervae is composed of tall emergent canopy trees that typically have more numerous secondary leaf veins compared to the sub-canopy to canopy species of Sebiferae (15-60 versus 9-28 respectively). The pericarps of Multinervae (1-7 mm) are also typically thicker than those of Sebiferae (0.3-1.3 mm). Furthermore, taxa of Multinervae are have an aril that is lacinate nearly to its base while those of Sebiferae are lacinate for half the length of the seed. Although incomplete quantitative data exists, it has also been observed in the field that members of Multinervae generally have larger, globose seeds while those of Sebiferae are generally smaller and more ellipsoid in shape. Additionally, 4 species (*V. flexuosa*, *V. multinervia*, *V. pavonis*, and *V. surinamensis*) of Multinervae included in this analysis were studied by Walker and Walker (1979) and classified as having type II pollen while 5 species (*V. calophylla*, *V. elongata*, *V.*

loretensis, *V. peruviana*, and *V. sebifera*) of the Sebiferae clade were classified as having type I pollen.

Although some species in this analysis formed well supported clades, there was little to no phylogenetic signal in members of the Sebiferae group, although a previous investigation of the neotropical genus *Compsoneura* (Chapter 1 of this thesis) found relatively well resolved relationships among many species with the same three loci. It is unclear as to whether the differing levels of phylogenetic signal between these informal subgroups of *Virola* are due to incomplete taxon sampling, nucleotide substitution rate heterogeneity, incomplete lineage sorting, a recent origin of Sebiferae species, or some combination of these. Since the sub-generic classification of *Virola* has been rather unstable due to a low degree of morphological trait variation, it is hoped that increased sampling of taxa, loci, and morphological characters in the future will enable a more stable classification rooted in the evolutionary histories of these species.

The Multinervae and Sebiferae clades exhibit numerous contrasting traits that likely have ecological and evolutionary origins. The Multinervae clade, as circumscribed here, is composed of tall canopy trees with globose seeds and relatively thick pericarps. Although speculative, it is possible that these adaptations are a result of co-evolution with primate and large bird species that act as dispersal agents as numerous studies have found members of this clade to be primarily dispersed by toucans and large primates (Julliot 1996, Juliotte and Sabatier 1993, Holbrook and Loiselle 2009). Their thick, astringent, pericarps may be adapted to prevent pre-dispersal seed consumption by these larger dispersers and/or discourage ovipositioning of invertebrate seed predators that have been found to infect large percentages of seeds of some species if this clade (Howe et al. 1985,

Howe 1993). The leaves of these canopy trees generally have longer and thinner leaves with more secondary veins/unit of length than members of Sebiferae (see Table 3.2). These characters may possibly be an adaptation to deal with the increased stresses of wind and rain experienced in the canopy. Another dichotomy previously mentioned is the type II pollen of Multinervae (sensu Walker and Walker 1979) opposed to the type I pollen of Sebiferae (Walker and Walker 1979), however it is unclear as to what implications, if any, this difference in pollen morphology has to pollination ecology and evolution of *Virola*. One more anecdotal difference between these two clades is that members of Sebiferae are often employed as hallucinogenic snuffs by natives while members of Multinervae are rarely used (Schultes and Raffauf 1990), possibly due to differences in the quality and/or quantity of their alkaloidal constituents or perhaps the increased ease with which the bark is removed from the smaller sub-canopy trees of Sebiferae.

After careful investigation of *Virola spp. in situ* in Ecuador and Peru as well as extensive herbarium investigations (BRIT, KEW, QCNE, LOJA, QAP, and USM), it was suspected that the taxa included in this study contained a minimum of three new species. This phylogenetic analysis gives further support for the distinction of these species but it is difficult to justify formal description with such incomplete taxon sampling. Additionally, reproductive material is lacking from these 3 putative new species, therefore their morphological characters are incomplete and their descriptions are provisional. These putative new species will be referred to in the following discussion according to their taxon names currently registered on the Neotropical Myristicaceae

project of the Barcode of Life Database (www.barcodinglife.org). The "RADS" portion of the species name refers to the initials of the author.

Working outwards from the base of the phylogeny, the first novel species is *V. sp.RADS4*; a species found west of the Andes near Camarones, Ecuador on the coastal plain in a seasonally dry forest. *Virola sp.RADS4* is a tall emergent tree (50 m, 30 cm dbh) that usually possesses stilt roots when mature and was found growing near streams in an otherwise dry forest that contained cacti, an unusual habitat for *Virola*. It appears to have affinities to *V. duckei* but differs in having fewer secondary veins (10-17 versus 25-37) and less pubescence (essentially glabrous versus ferruginous-tomentose). This species also differs from *V. aequatorialis* (Jaramillo et al. 2000), another species of the coastal plains of Ecuador, in not having a winged petiole and having less than 20 veins per leaf. The red sap of this species was also self-administered for a recalcitrant ringworm (*Tinea sp.*) infection of the skin and found to be highly efficacious.

Virola sp.RADS3 is also proposed as a new species with affinities to *V. calophylloidea*. *Virola sp.RADS3* has only been found in the Cordillera del Condor of Ecuador growing at an elevation of 1200 m asl. Due to its deep reddish sessile pubescence, *V. sp.RADS3* was initially identified in a permanent ecological plot as *Otoba parvifolia*. It appears that this taxon does not represent a range extension of *V. calophylloidea* of the Rio Negro of Brazil, owing to its rich reddish pubescence and the fact that it was growing on a rare sandstone substrate home to many highly endemic species (Rogers 2002, Ulloa and Neil 2006, Clarke et al. 2010, Janovec et al. in prep.).

The third novel species discovered in the course of this analysis is *V. sp.RADS1*. This species has been collected solely in the Madre de Dios watershed of Peru in terra

firme forests from the foothills of the Andes (750 m asl) and the lowland Amazon (200 m asl). This species had been identified in ecological plots as *V. pavonis* due to its whitish abaxial lamina colour and superficially similar leaf venation. However, these molecular analyses clearly show that it forms a well supported monophyletic clade and has no close genetic affinity to *V. pavonis*. *Virola sp.RADS1* can be easily differentiated vegetatively from *V. pavonis* in having fewer leaf veins (11-15 versus 15-20) and conspicuously undulating edges of the leaf lamina.

Distance analysis

Of the 18 species of *Virola* for which trnH-psbA sequences were retrieved, 6 (*V. sp.RADS4*, *V. caducifolia*, *V. multicostata*, *V. michelii*, *V. pavonis*, and *V. kwatae*) had haplotypes distinct from other species (i.e. had diagnostic characters). All members of the Sebiferae clade had a large deletion in the most variable region of the trnH-psbA spacer and exhibited identical haplotypes with their homopolymer runs omitted, making species delimitation impossible within this clade with this loci. However, only 4 species pairs in the Multinervae clade shared haplotypes. One specimen of *V. multinervia* had an identical haplotype to *V. flexuosa; V. multiflora* and *V. nobilis* also shared haplotypes. Of the remaining 7 species, *V. surinamensis* had two paraphyletic haplotypes. This apparent paraphyly of conspecific sequences may be due to undetected speciation or population divergence as the two *V. surinamensis* sequence clusters on the distance tree represent populations in both Peru and Ecuador. However, chloroplast haplotype introgression, or chloroplast capture, is expected to occur frequently in plant species (Rieseberg and Soltis 1991, Hollingsworth et al 2011) and it is important to keep this in mind before delimiting species based solely on haplotype data. Since many species were represented by fewer

than 4 specimens from restricted geographical areas, it remains unclear whether chloroplast haplotype introgression is commonplace in this genus but given the emerging evidence of widespread plastid introgression in other plant families (Golden and Bain 2000, Palme et al. 2003, Palma-Silva et al. 2011) it is likely that plastid-based DNA identification will only be confounded by increased taxon sampling of *Virola*. Despite this possibility, the trnH-psbA spacer may represent an alternative means of identifying collections (particularly juveniles or difficult material such as pollen or roots) of the Multinervae clade as they are often mis-identified in ecological plots and there are usually few species at any one locale. Since numerous other proposed DNA barcoding loci have exhibited little variation in initial trials and introgression and paraphyly are likely to plague species discrimination, efforts should be focused on finding variable low-copy nuclear loci that can more effectively discriminate species of these ecologically dominant trees.

Descriptions of provisional species

Virola sp.RADS1 R. Steeves sp. nov.-TYPE: Peru. Madre de dios: Los Amigos Biological Research Station -12.57 DD Lat, -70.10 DD Long, 250 m, 17 Jun 2008, R. Steeves and A. Balarezo 500 (holotype: OAC)

Tree to 20 m tall and 40 cm dbh. **Bark** grey in colour, tough and ridged on young trees but growing smoother with age. **Sap** profuse and clear-red. **Branchlets** terete to subterete, 1.5-4 mm wide, bark smooth, green and may be covered in an evanescent

pubescence, the hairs short-stalked to sessile, stellate. **Leaves** simple, alternate, thin-chartaceus, leaf buds ferruginous but soon glabrous upon leaf expansion; **petioles** stout, subterete, canaliculate, 0.5-1.2 cm long and 1-2.5 mm wide, brown when dry, glabrous to glabrescent, the hairs short stalked to sessile, stellate; **blades** elliptic-oblong, lamina 12.7-17.0 cm long, 3.25-4.8 cm wide at ¼ length, 3.3-5.5 cm wide at ½ width, 1.9-4.1 cm wide at ¾ length, base attenuate to rounded and acuminate at apex, drying dark brown, glabrescent, adaxial surface whitish with a waxy appearance, abaxial surface dark green with a glossy appearance; **venation** costa glabrous and raised above, prominently keeled and glabrescent below; secondary nerves 11-16 per side, spaced by 7-9 mm, arcuate ascending, eucamptodromous, partially anastomosing near margin. **Staminate inflorescences** unknown **Pistillate inflorescences** unknown. **Fruits** (nearly mature) 1-3 per infructescence, globose, pedicel 0.9-1.0 cm long and 2.5-3 mm wide; **Pericarp** green and glabrescent with age, smooth and ferruginous-tomentose, 1.2-1.5 cm wide and 1.5-2 cm long and 0.8-1.2 mm thick, presumably dehiscent when mature; **Seed** 1-1.2cm wide, 1.3-1.5cm long; **Testa** striated by aril; **Aril** lacinate to base; **Endosperm** unknown.

Distribution and notes: Collected only from the Madre de Dios watershed in Peru from elevations of 150-750 m. *V. sp.RADS1*'s leaf colour, indument and venation has many affinities to *V. pavonis* and *V. surinamensis* but differs in having an acuminate apex, fewer than 16 secondary veins and a comparatively thinner pericarp. When formalized this species will be named after the forester and conservationist Aldo Leopold whose early 20th century essay "The River of the Mother of God" about the Madre de Dios River has inspired me and many others to explore wild places.

Additional specimens examined: **Peru. CUSCO:** Quispicanchi Province, District of Camanti, Community of Quincemil, -13.23, -70.78, 500-1200 m, 06/26-07/04, 2008, R. Steeves, P. Centeno, D. White, and K. Ward 561 (OAC); **Madre de Dios:** Manu province, Los amigos biological station, Confluence of Los Amigos and Madre de Dios rivers, -12.57, -70.1, 200-250 m, 06/7-17/2008, R. Steeves and A. Balarezo 431 (OAC), R. Steeves and A. Balarezo 432 (OAC), R Steeves and A. Balarezo 447 (OAC), R. Steeves and A. Balarezo 450 (OAC), R. Steeves and A. Balarezo 470 (OAC), R. Steeves and A. Balarezo 500 (OAC).

Virola sp.RADS4 R. Steeves-TYPE: Ecuador. Manabi: Diez de Augusto, Camarones, -0.083 DD Lat, -80.16 DD Long, 80 m, 28 April 2008, R. Steeves, A. Reid and D. Simba 205 (holotype: OAC; QCNE)

Tree to 50 m tall and 30 cm dbh. **Bark** grey in colour, tough and ridged. **Sap** profuse and clear-red. **Branchlets** terete to subterete, 0.27-0.42 cm wide, bark smooth, green and may be covered in an evanescent pubescence, the hairs short-stalked, stellate. **Leaves** simple, alternate, thin-chartaceus, leaf buds ferruginous but soon glabrescent upon leaf expansion; **petioles** stout, subterete, canaliculate, 0.48-1.22 cm long and 0.26-0.44 cm wide, brown when dry, glabrous to glabrescent, the hairs short stalked, stellate; **blades** elliptic-oblong, lamina 11.5-19.7 cm long, 3.0-7.0 cm wide at ¼ length, 3.4-7.6 cm wide at ½ width, 3.2-6.9 cm wide at ¾ length, base attenuate to rounded and acuminate at apex, drying dark brown, glabrescent, adaxial surface dull green when fresh, abaxial surface dark green with a glossy appearance; **venation** costa glabrous and raised above, prominently keeled and glabrescent below; secondary nerves 11-17 per side, spaced by

0.35-0.7 cm, arcuate ascending, eucamptodromous, partially anastomosing near margin. **Staminate inflorescences** unknown **Pistillate inflorescences** unknown. **Fruits** (immature) 1-3 per infructescence, elliptic, pedicel 0.9-1.0 cm long and 02.5-3 mm wide; **Pericarp** green and glabrescent with age, smooth and ferruginous-tomentose, 1.2-1.5 cm wide and 1.5-2 cm long and 0.8-1.2 mm thick, presumably dehiscent when mature; **Seed** 1-1.2cm wide, 1.3-1.5cm long; **Testa** striated by aril; **Aril** lacinate to base; **Endosperm** unknown.

Distribution and notes: Collected only from dry coastal forests 20 km south of Pedernales, Ecuador at the Lalo Loor research station near Camarones where it was one of the most common canopy tree species. This species has morphological affinities to *V. duckei* from which it differs in having only 10-17 secondary nerves compared to the 25-37 typical of *V. duckei* (Smith 1937).

Additional specimens examined: **Ecuador. Manabi:** Diez de Augusto, -0.083 DD Lat, -80.16 DD Long, 80 m, 04/26-29/2008, R. Steeves, A. Reid and D. Simba 205 (holotype: OAC! ; QCNE!), 2008, R. Steeves, A. Reid and D. Larco 188 (OAC), R. Steeves, A. Reid and D. Larco 189 (OAC), R. Steeves, A. Reid and D. Larco 190 (OAC), R. Steeves, A. Reid and D. Larco 191 (OAC), R. Steeves, A. Reid and D. Larco 192 (OAC), R. Steeves, A. Reid and D. Larco 193 (OAC), R. Steeves, A. Reid and D. Larco 194 (OAC), R. Steeves, A. Reid and D. Larco 195(OAC), R. Steeves, A. Reid and D. Larco 196 (OAC), R. Steeves, A. Reid and D. Larco 197 (OAC), R. Steeves, A. Reid and D. Larco 198 (OAC), R. Steeves, A. Reid and D. Larco 199 (OAC), R. Steeves, A. Reid and D. Larco 200 (OAC), R. Steeves, A. Reid and D. Larco 201 (OAC), R. Steeves, A.

Reid and D. Larco 202 (OAC), R. Steeves, A. Reid and D. Larco 203 (OAC), R. Steeves, A. Reid and D. Larco 204 (OAC), R. Steeves, A. Reid and D. Larco 206 (OAC), R. Steeves, A. Reid and D. Larco 207 (OAC), R. Steeves, A. Reid and D. Larco 208 (OAC), R. Steeves, A. Reid and D. Larco 209 (OAC), R. Steeves, A. Reid and D. Larco 210 (OAC), R. Steeves, A. Reid and D. Larco 211 (OAC), R. Steeves, A. Reid and D. Larco 212 (OAC), R. Steeves, A. Reid and D. Larco 213 (OAC), R. Steeves, A. Reid and D. Larco 214 (OAC), R. Steeves, A. Reid and D. Larco 215 (OAC).

Virola sp.RADS3 R. Steeves-TYPE: Ecuador. Zamora-Chinchipe, Quimi, -3.59 DD Lat, -78.50 DD Long, 1200 m, 17 May 2008, R. Steeves, W. Quizhpe and D. Simba 335 (holotype: OAC; QCNE)

Tree to 30 m tall and 25 cm dbh. **Bark** grey in colour, tough and ridged-smooth. **Sap** profuse and clear-red. **Branchlets** terete to subterete, 4-7 mm wide, bark smooth, grey and may be covered in an evanescent pubescence, the hairs short-stalked sessile, stellate. **Leaves** simple, alternate, thin-chartaceus, leaf buds ferruginous, persistenly ferruginous-tomentellous; **petioles** stout, subterete, canaliculate, 1.27-1.74 cm long and 3-5 mm wide, brown when dry, glabrescent, the hairs sessile, stellate; **blades** lanceolate-oblong, lamina 12.3-19.2 cm long, 4.7-5.9 cm wide at ¼ length, 5.3-6.6 cm wide at ½ width, 4.4-5.56 cm wide at ¾ length, base rounded and acuminate at apex, drying dark brown, glabrescent, adaxial surface reddish-white when fresh, abaxial surface dark green with a glossy appearance when fresh, dark brown when dry; **venation** costa glabrous and raised above, prominently keeled and glabrescent below; secondary nerves 9-12per side, spaced by 6.4-

7.8 mm, arcuate ascending, eucamptodromous, partially anastomosing near margin. **Staminate inflorescences** unknown **Pistillate inflorescences** unknown. **Fruits** unknown.

Distribution and notes: Collected only from the slopes of the Cordillera del Condor near Quimi, Ecuador on the site of a proposed Canadian gold mine. Two trees were found growing at 1200 m elevation in a permanent plot situated on soil of sandstone origin. This species was originally identified as *Otoba parvifolia* in a permanent plot owing to its faint secondary veins and reddish adaxial leaf colour. This species has morphological affinities to *V. calophylloidea* but differs in its reddish pubescence, distribution, and its unique habitat where many highly endemic species are found. When formalized, this species will be named after the native Shuar people that inhabit the area where these trees are found.

Additional specimens examined: **Ecuador. Zamora-Chinchipe:** -3.59 DD Lat, -78.50 DD Long, 1200 m, 17 May 2008, R. Steeves, W. Quizhpe and D. Simba 339 (OAC).

Tables

Table 3.1 Species[SL and BL indicate small leaf and big leaf morphotypes of *V. loretensis* and *V. Sebifera*], collectors, collection numbers (Coll. #), Barcode of Life Database process ID number (BOLD ID), and decimal degree co-ordinates (latitude= DD Lat., longitude= DD long.) of taxa collected for this study.

Species	Collector(s)	Coll. #	BOLD ID	DD Lat.	DD Long.
C. capitellata	Steeves, R. et al	RS 551	RSMYR004-11	-13.24	-70.78
C. debilis	Berry, P	PB 7209	RSMYR017-11		
I. juruensis	Steeves, R. et al	RS 451	RSMYR050-11	-12.57	-70.10
I. laevis	Steeves, R. et al	RS 460	RSMYR051-11	-12.57	-70.10
V. caducifolia	Janovec, J.	JJ 847	RSMYR109-11		
V. calophylla	Steeves, R. et al	RS 430	RSMYR110-11	-12.57	-70.10
V. calophylla	Steeves, R. et al	RS 454	RSMYR111-11	-12.57	-70.10
V. calophylla	Steeves, R. et al	RS 481	RSMYR112-11	-12.57	-70.10
V. calopylla	Steeves, R. et al	RS 511	RSMYR113-11	-12.61	-69.20
V. dixonii	Steeves, R. et al	RS 225	RSMYR114-11	-1.04	-77.37
V. elongata	Steeves, R. et al	RS 502	RSMYR115-11	-12.61	-69.20
V. elongata	Steeves, R. et al	RS 437	RSMYR116-11	-12.57	-70.10
V. elongata	Steeves, R. et al	RS 494	RSMYR117-11	-12.57	-70.10
V. flexuosa	Steeves, R. et al	RS 522	RSMYR118-11	-12.61	-69.20
V. flexuosa	Steeves, R. et al	RS 595	RSMYR119-11	-13.24	-70.78
V. flexuosa	Steeves, R. et al	RS 442	RSMYR120-11	-12.57	-70.10
V. sp.RADS1	Steeves, R. et al	RS 510	RSMYR121-11	-12.61	-69.20
V. sp.RADS1	Steeves, R. et al	RS 561	RSMYR122-11	-13.24	-70.78
V. sp.RADS1	Steeves, R. et al	RS 432	RSMYR123-11	-12.57	-70.10
V. sp.RADS1	Steeves, R. et al	RS 500	RSMYR124-11	-12.57	-70.10
V. loretensis	Steeves, R. et al	RS 516	RSMYR125-11	-12.61	-69.20
V. loretensis	Steeves, R. et al	RS 526	RSMYR126-11	-12.61	-69.20
V. loretensis	Steeves, R. et al	RS 480	RSMYR127-11	-12.57	-70.10
V. multinervia	Steeves, R. et al	RS 107	RSMYR128-11	-12.57	-70.10
V. multinervia	Steeves, R. et al	RS 108	RSMYR129-11	-12.57	-70.10
V. multinervia	Steeves, R. et al	RS 543	RSMYR130-11	-13.24	-70.78
V. multinervia	Steeves, R. et al	RS 549	RSMYR131-11	-13.24	-70.78
V. multinervia	Steeves, R. et al	RS 429	RSMYR132-11	-12.57	-70.10
V. multinervia	Steeves, R. et al	RS 466	RSMYR133-11	-12.57	-70.10
V. multinervia	Steeves, R. et al	RS 350	RSMYR134-11	-3.57	-78.45
V. pavonis	Steeves, R. et al	RS 216	RSMYR135-11	-1.04	-77.37
V. peruviana	Janovec, J.	JJ 772	RSMYR136-11		
V. sebifera-BL	Steeves, R. et al	RS 584	RSMYR137-11	-13.24	-70.78
V. sebifera-SL	Steeves, R. et al	RS 507	RSMYR138-11	-12.61	-69.20
V. sebifera-SL	Steeves, R. et al	RS 534	RSMYR139-11	-13.24	-70.78
V. sebifera-SL	Steeves, R. et al	RS 552	RSMYR140-11	-13.24	-70.78
V. loretensis-SL	Steeves, R. et al	RS 483	RSMYR141-11	-12.57	-70.10

V. sp.RADS3	Steeves, R. et al	RS 335	RSMYR142-11	-3.57	-78.45
V. sp.RADS3	Steeves, R. et al	RS 339	RSMYR143-11	-3.57	-78.45
V. sp.RADS4	Steeves, R. et al	RS 213	RSMYR144-11	-0.08	-80.17
V. sp.RADS4	Steeves, R. et al	RS 214	RSMYR145-11	-0.08	-80.17
V. surinamensis	Steeves, R. et al	RS 501	RSMYR146-11	-12.61	-69.20
V. surinamensis	Steeves, R. et al	RS 428	RSMYR147-11	-12.57	-70.10
V. surinamensis	Steeves, R. et al	RS 489	RSMYR148-11	-12.57	-70.10
V. surinamensis	Steeves, R. et al	RS 82	RSMYR149-11	-12.57	-70.10
V. surinamensis	Steeves, R. et al	RS 83	RSMYR150-11	-12.57	-70.10
V. surinamensis	Steeves, R. et al	RS 84	RSMYR151-11	-12.57	-70.10
V. surinamensis	Steeves, R. et al	RS 248	RSMYR152-11	-1.04	-77.37
V. surinamensis	Steeves, R. et al	RS 324	RSMYR153-11	-3.57	-78.45

Table 3.2 Morphological trait values of 8 species of *Virola* taken from species descriptions of Smith (1937). Minimum and maximum trait values are given as ranges. NDA indicates that there was no data available.

	Sebiferae				**Multinervae**			
	V. elongata	*V. sebifera*	*V. calophylla*	*V. loretensis*	*V. multinervia*	*V. flexuosa*	*V. surinamensis*	*V. pavonis*
Height (m)	25	40	4 to 10	4 to 10	30	30	25	23
Leaf petiole length (mm)	5-16	10-25	7-20	3-12	4-15	2-7	2-9	4-13
Leaf lamina length (cm)	12-32	15-47	20-55	15-35	25-45	5-11	10-22	8-21
Leaf width at half length (cm)	4-11	6-15	7-24	4-10	8-16	1.7-4	2-5	2-6.5
Leaf vein number	9-20	10-28	12-27	15-26	40-60	40-50	16-30	15-20
Inflorescence length (cm)	4-18	8-23	12-30	to 25	15-20	4-9	7-17	7-15
Flowers per branch (#)	2-8	3-10	4-10	5-10	20-50	10-15	5-20	3-8
Anther length (mm)	0.6-1.6	0.9-2	1-1.7	1-1.4	to 0.9	0.8-0.9	1.3-1.9	1.1-1.5
Anther number	3(4-6)	3(4-5)	3	3	3	3	3	3
Fruit length (mm)	11-16	10-19	NDA	13-22	20-30	NDA	13-21	25-50
Fruit width (mm)	8-12	7-14	NDA	13-22	15-25	NDA	11-18	15-23
Pericarp thickness (mm)	0.3-1.3	0.5-1	NDA	<.5mm	1.5-4 mm	NDA	1 to 2	2 to 7
Aril	1/2	1/2	1/2	1/2	lacinate to	lacinate to	lacinate to	lacinate

	lacinate	lacinate	lacinate	lacinate	base	base	base	to base

Figures

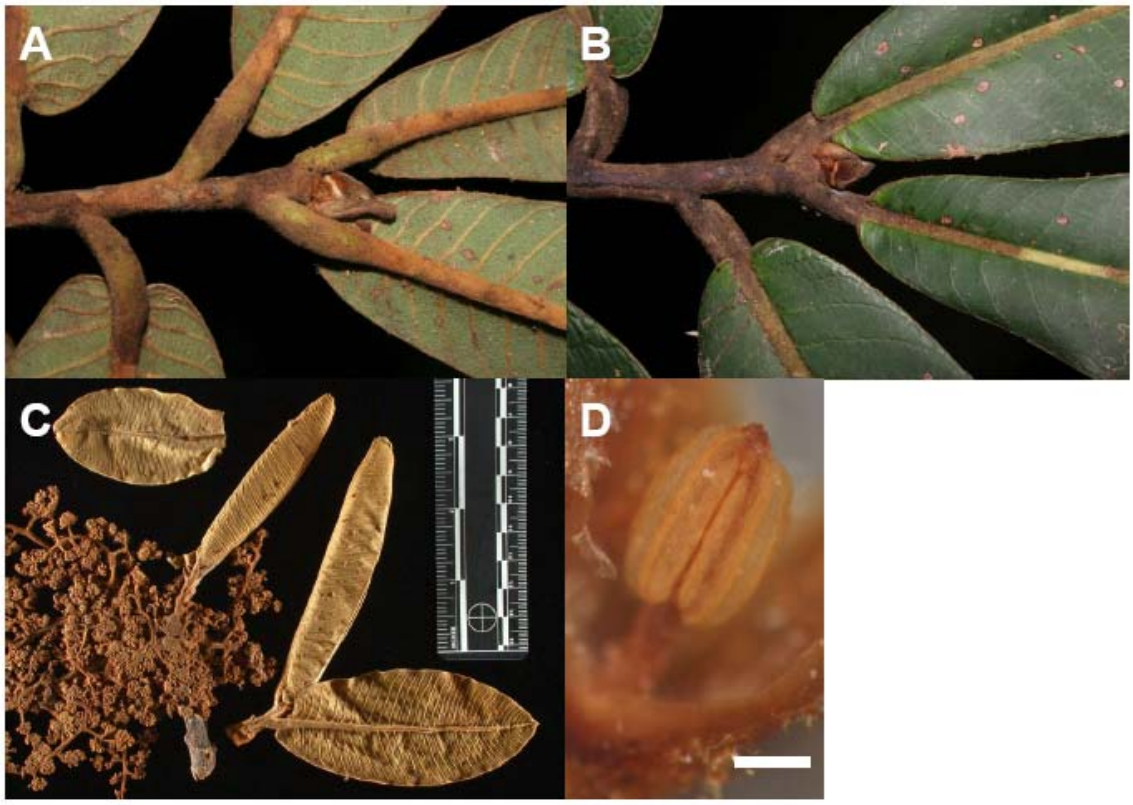

Figure 3.1 *Virola flexuosa*: A) Abaxial leaf surfaces. B) Adaxial leaf surfaces. C) Herbarium specimen with inflorescence. D) Staminate flower showing androecium (scale bar= 0.5 mm).

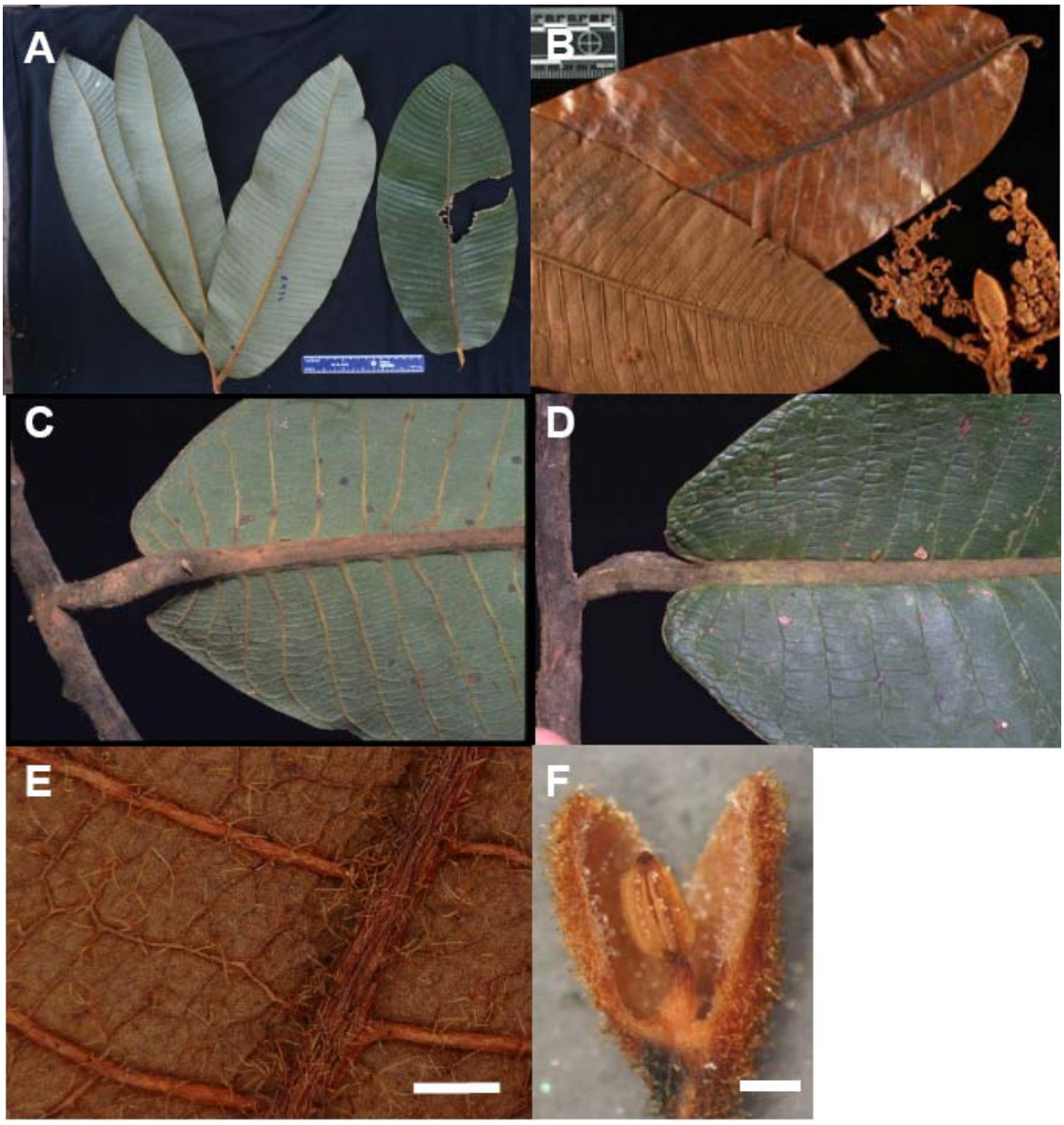

Figure 3.2 *Virola multinervia*: A) Abaxial and adaxial leaf surfaces from juvenile tree (15 cm ruler). B) Herbarium specimen with immature inflorescence. C) Close-up of adaxial surface of mature tree. D) Close-up of adaxial surface of mature tree. E) Abaxial leaf pubescence (scale bar= 1 mm). F) Flower with one tepal dissected to show androecium (scale bar= 0.5 mm).

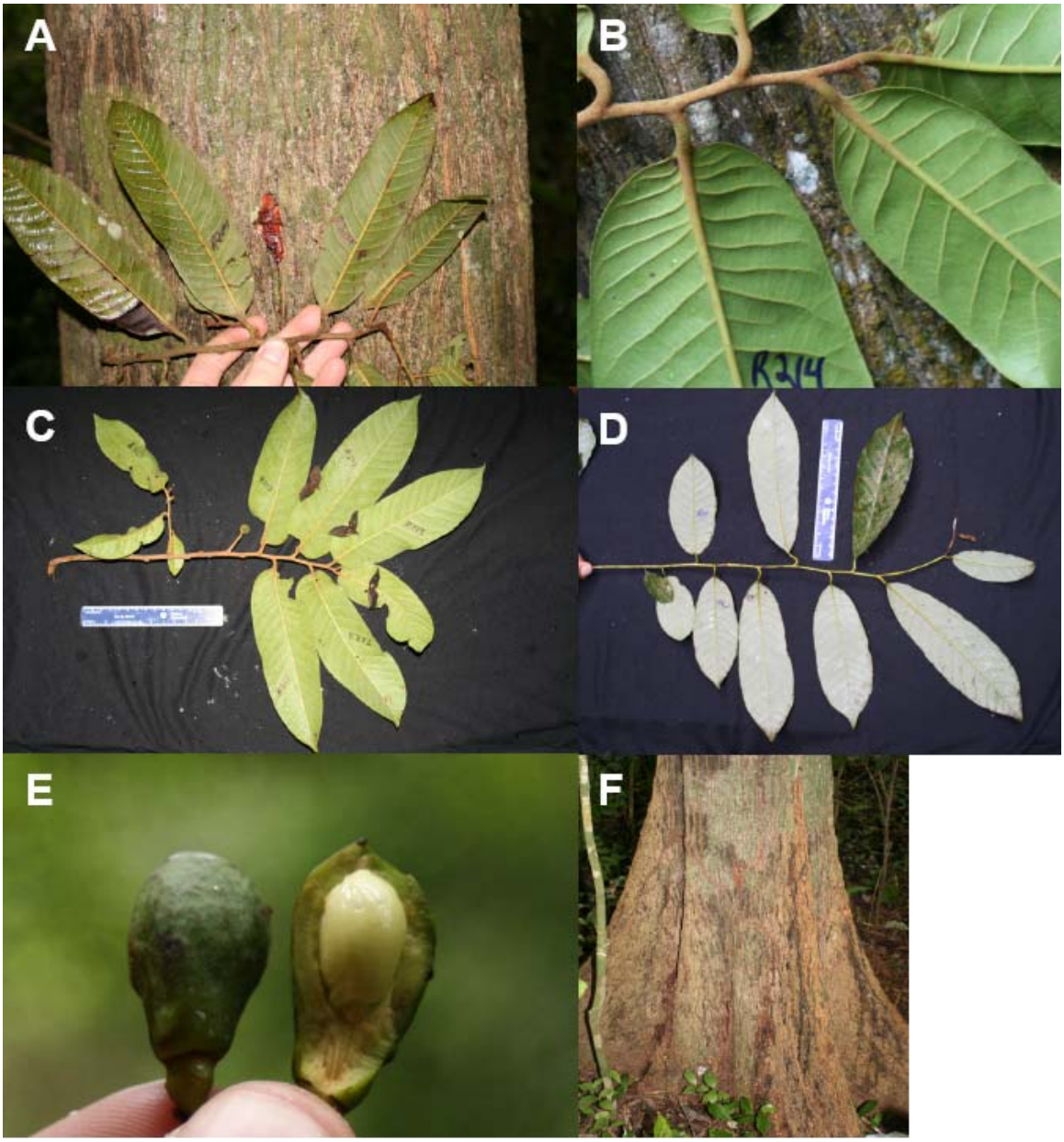

Figure 3.3 *Virola sp.RADS4* A) Bark, leaves and red sap. B) Abaxial leaf surface showing glabrescent surface and secondary veins. C) Branchlet of mature tree (15 cm ruler). D) Branchlet of immature tree (15 cm ruler). E) Nearly mature fruit dissected to show outer glabrescent pericarp and nearly entire aril. F) Buttressed base of mature tree.

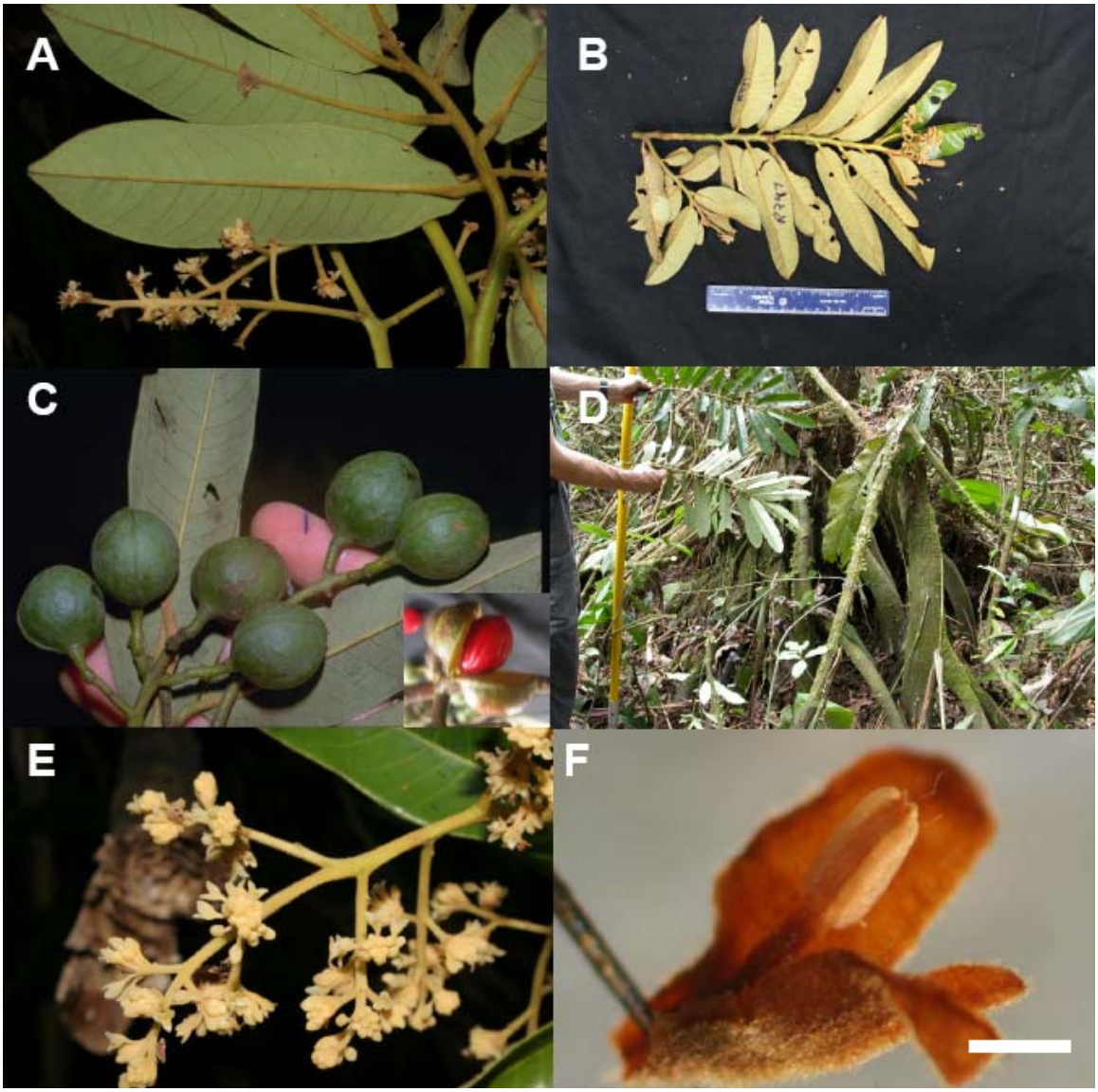

Figure 3.4 *Virola surinamensis*: A) Branch with inflorescence and abaxial leaf surface. B) Branch of mature tree with inflorescences. C) Undehisced, nearly mature fruits and dehisced fruit showing red aril (inset). D) Branches and stilt roots typical of this species . E) Close-up of inflorescence. F) Staminate flower with petal dissected to show androecium (scale bar= 0.5 mm).

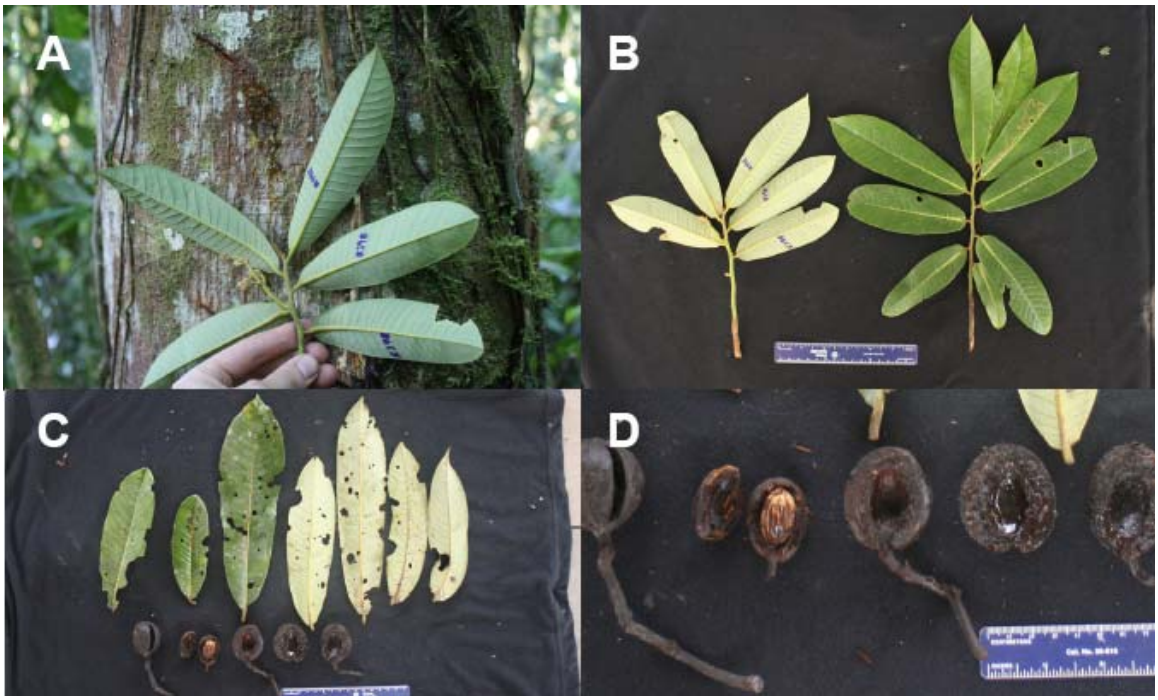

Figure 3.5 *Virola pavonis*: A) Bark, leaves, inflorescence, and greenish latex. B) Specimens showing adaxial and abaxial leaf surfaces (15 cm ruler). C) Leaves and pericarps (15 cm ruler). D) Close-up of pericarps and seed testa.

Figure 3.6 *Virola sp.RADS3* specimen showing reddish abaxial pubescence and glabrous glossy green adaxial surface (15 cm ruler).

Figure 3.7 *Virola calophylla*: A) Leaves with inflorescence. B) Bark and leaves of mature tree. C) Nearly mature florescence. D) Close-up of inflorescence with ants. E) Abaxial leaf pubescence (scale bar= 2 mm). F) Indehiscent and dehiscent mature fruit with lacinate aril.

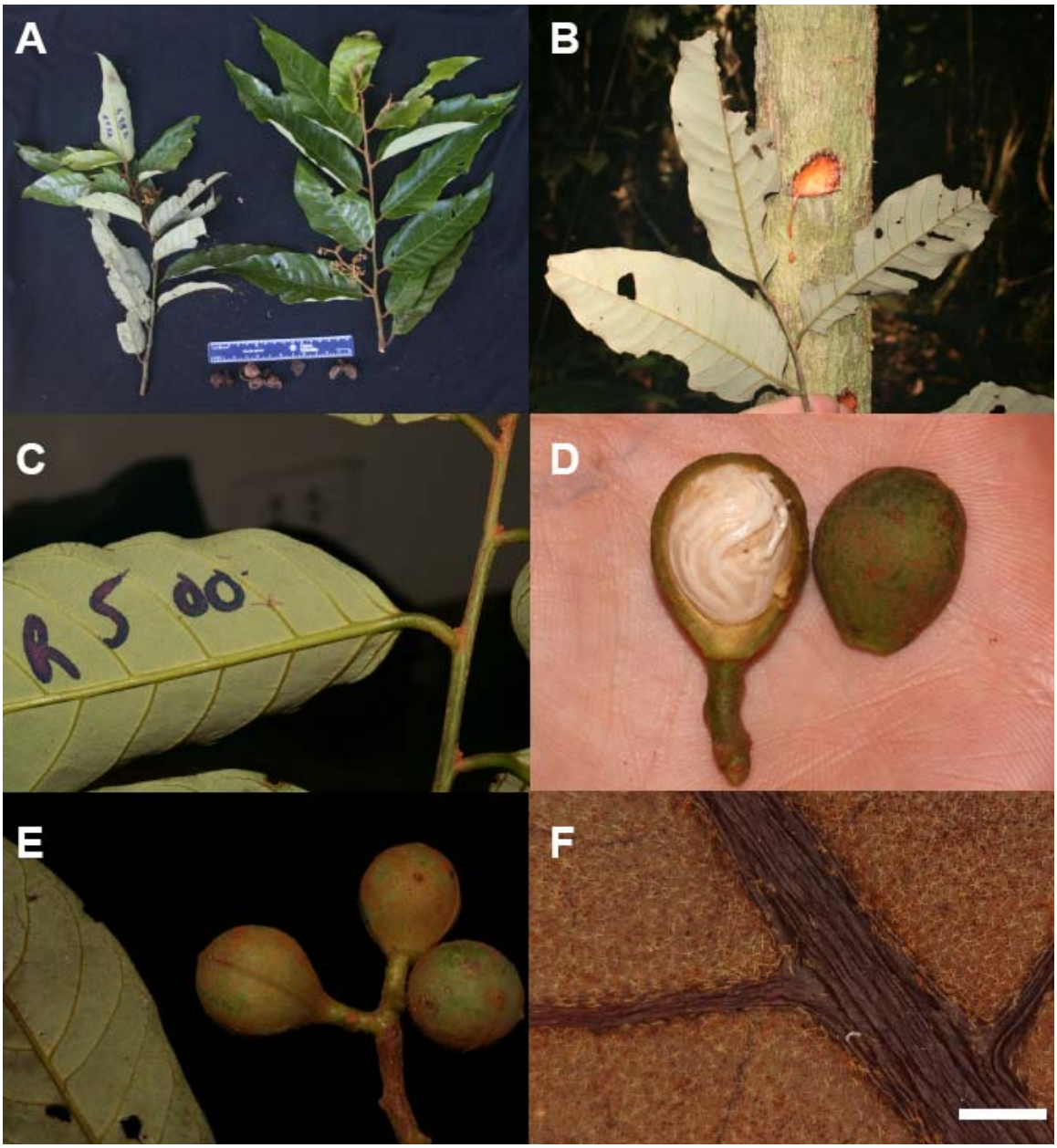

Figure 3.8 *Virola sp.RADS1*: A) Leaves with inflorescence and pericarps (15 cm ruler). B) Bark showing latex and leaves. C) Close-up of leaf and branch showing undulating leaf margins. D) Dissected fruit showing deeply lacinate aril and red-pubescent pericarp. E) Leaf and nearly mature un-dehisced fruits. F) Adaxial leaf pubescence, primary and secondary veins (Scale bar= 1 mm).

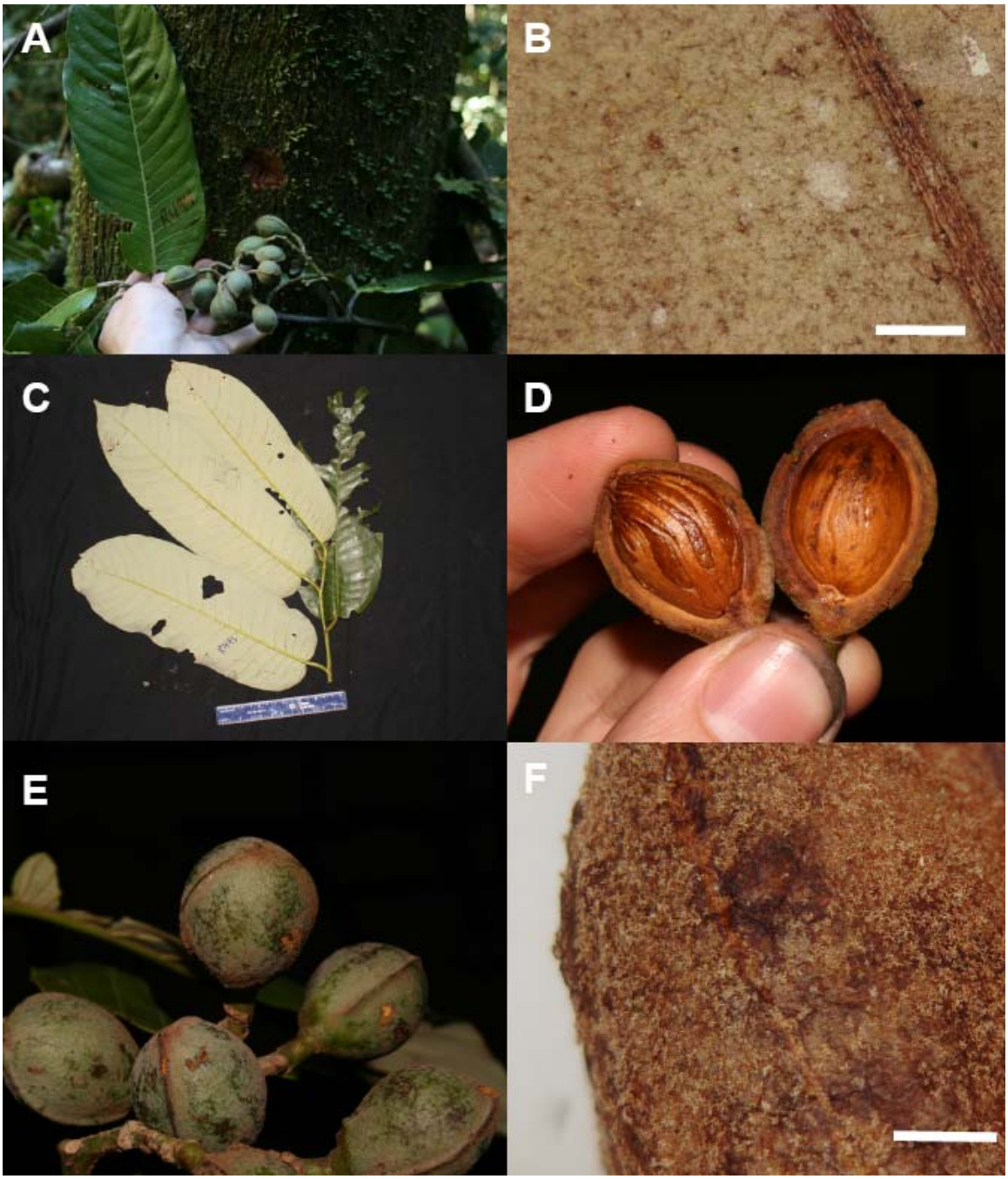

Figure 3.9 *Virola elongata*: A) Leaves, infructescence and bark. B) Adaxial leaf pubescence and secondary vein (scalebar= 1 mm) C) Abaxial and abaxial leaf surfaces (15 cm ruler). D) Dissected mature fruit with deeply lacinated aril (semi-dried). E) Undehisced nearly mature fruits. F) Close-up of fruit pubescence (scale bar= 1 mm).

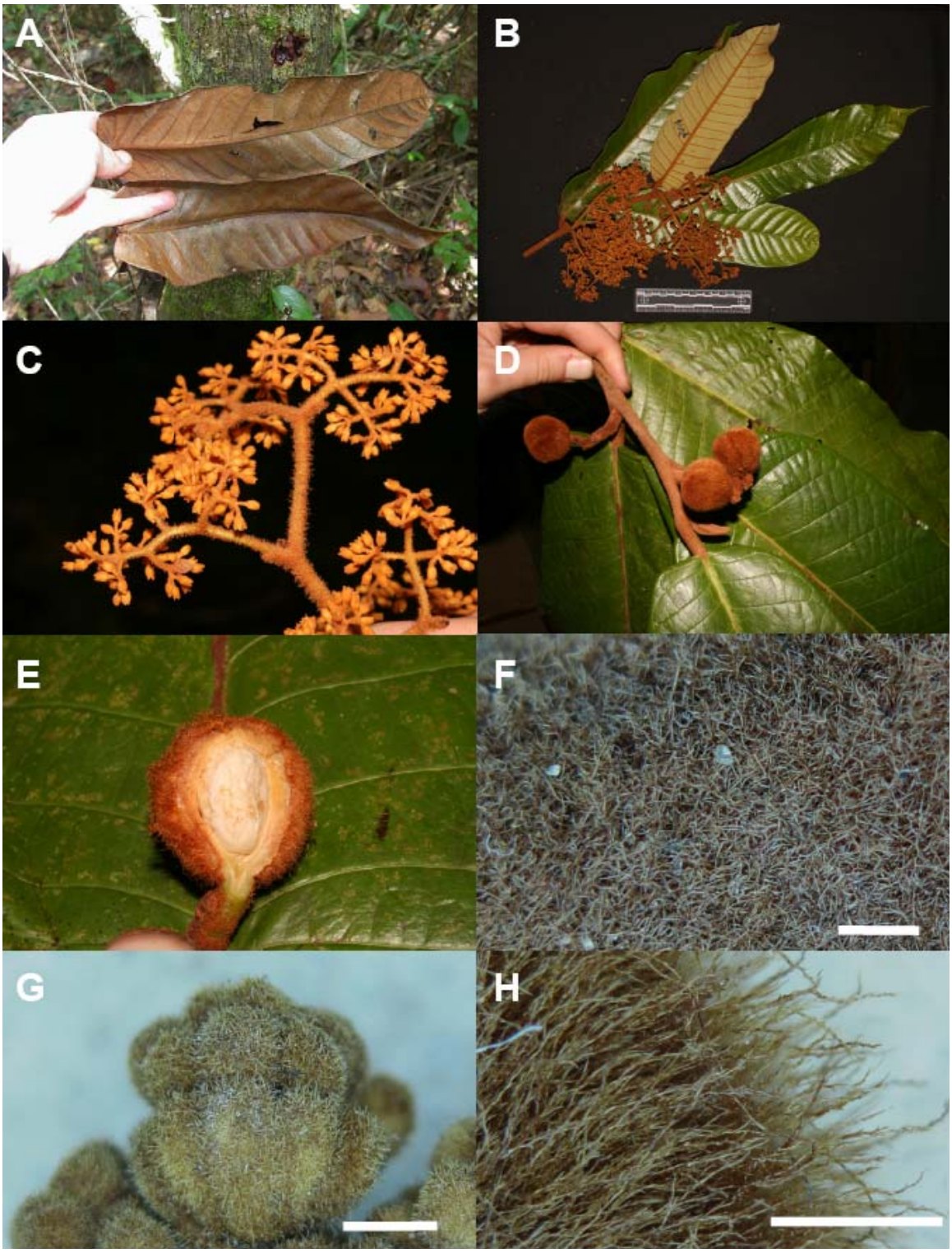

Figure 3.10 *Virola loretensis-BL* (Big Leaf morphotype): A) Bark, deep red latex and dried (dropped) leaves. B) Abaxial, adaxial leaf surfaces of sample with inflorescences (15 cm ruler). C) Close-up of inflorescence. D) Branch with mature fruits. E) Dissected

immature fruit showing dense pubescence. F) Adaxial leaf pubescence. G) Developing flower buds (scale bar=1 mm). H) Close-up of fruit indument (scale bar= 1 mm).

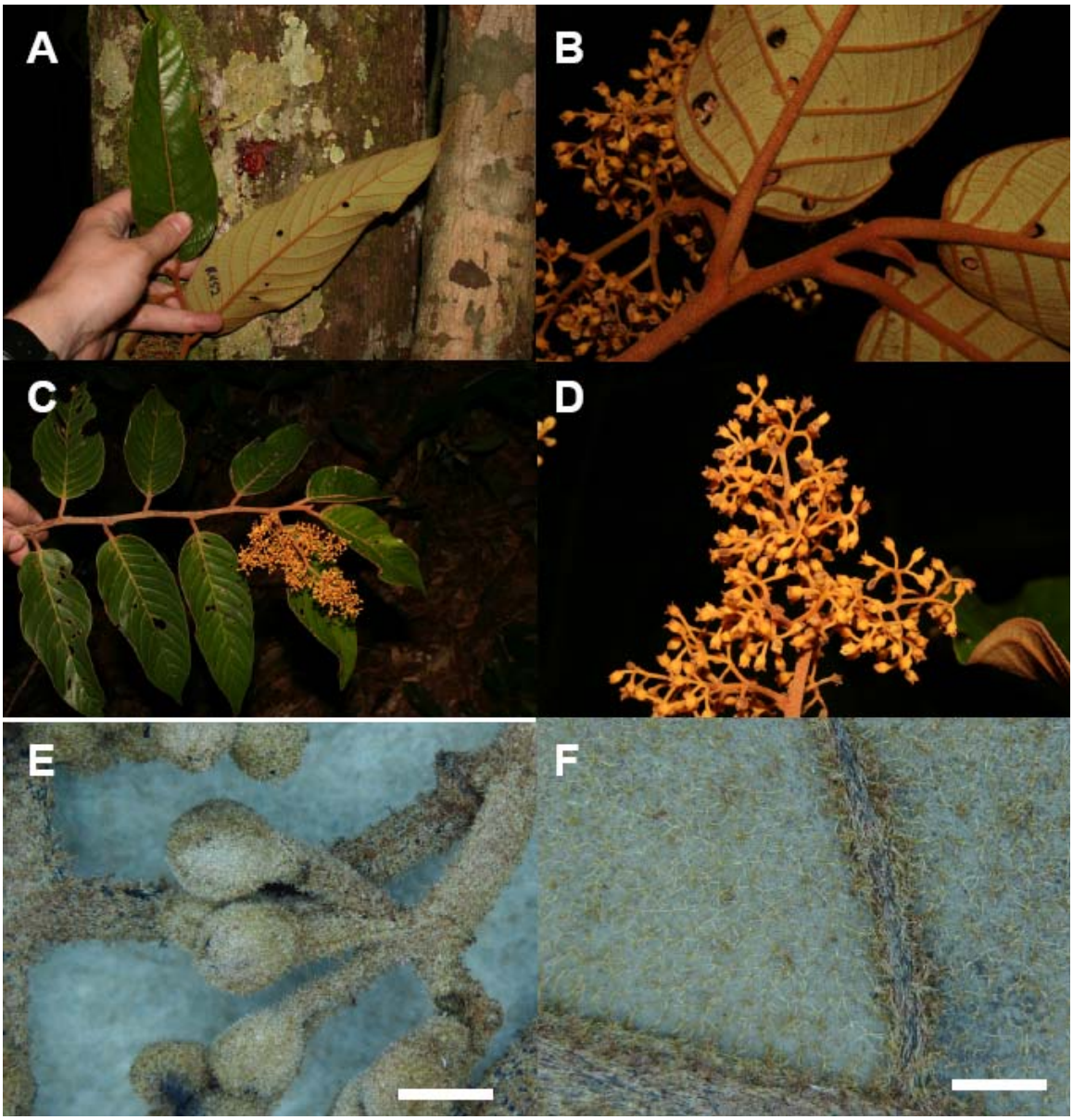

Figure 3.11 *Virola loretensis-SL* (Small Leaf morphotype): A) Bark, leaves and red latex. B) Adaxial leaf surfaces with inflorescence. C) Branch with inflorescence. D) Close-up of mature inflorescence. E) Flower pubescence (scale bar= 1 mm). F) Adaxial leaf pubescence (scale bar= 1 mm).

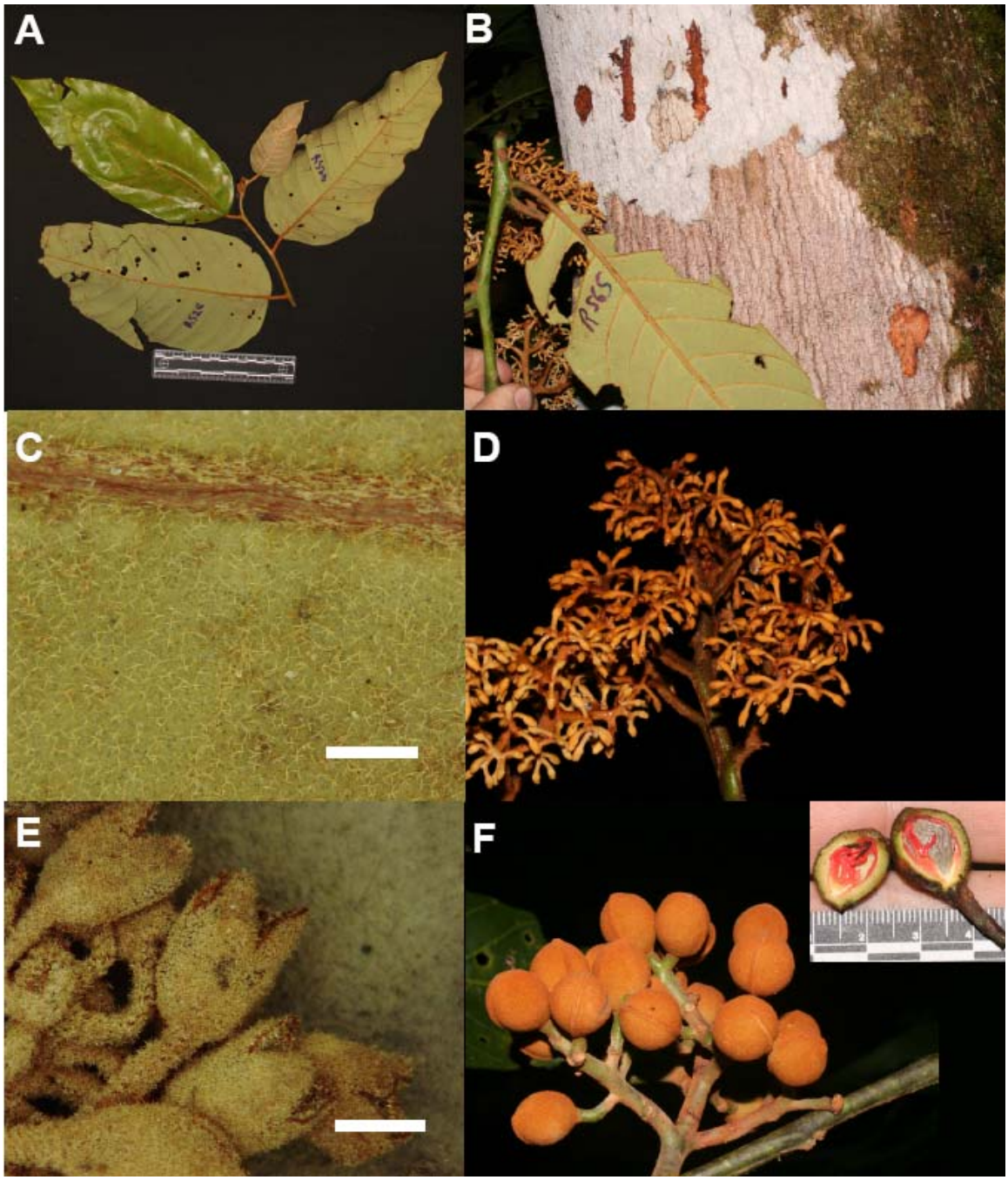

Figure 3.12 *Virola sebifera-SL* (Small Leaf morphotype): A) Leaves with orange-red pubescence (15 cm ruler) B) Bark, leaves and inflorescence of mature tree. C) Abaxial leaf pubescence (scale bar= 1 mm). D) Inflorescence. E) Flowers (scale bar= 1 mm). F) Nearly mature fruits with dissected fruit showing seed testa and aril (inset).

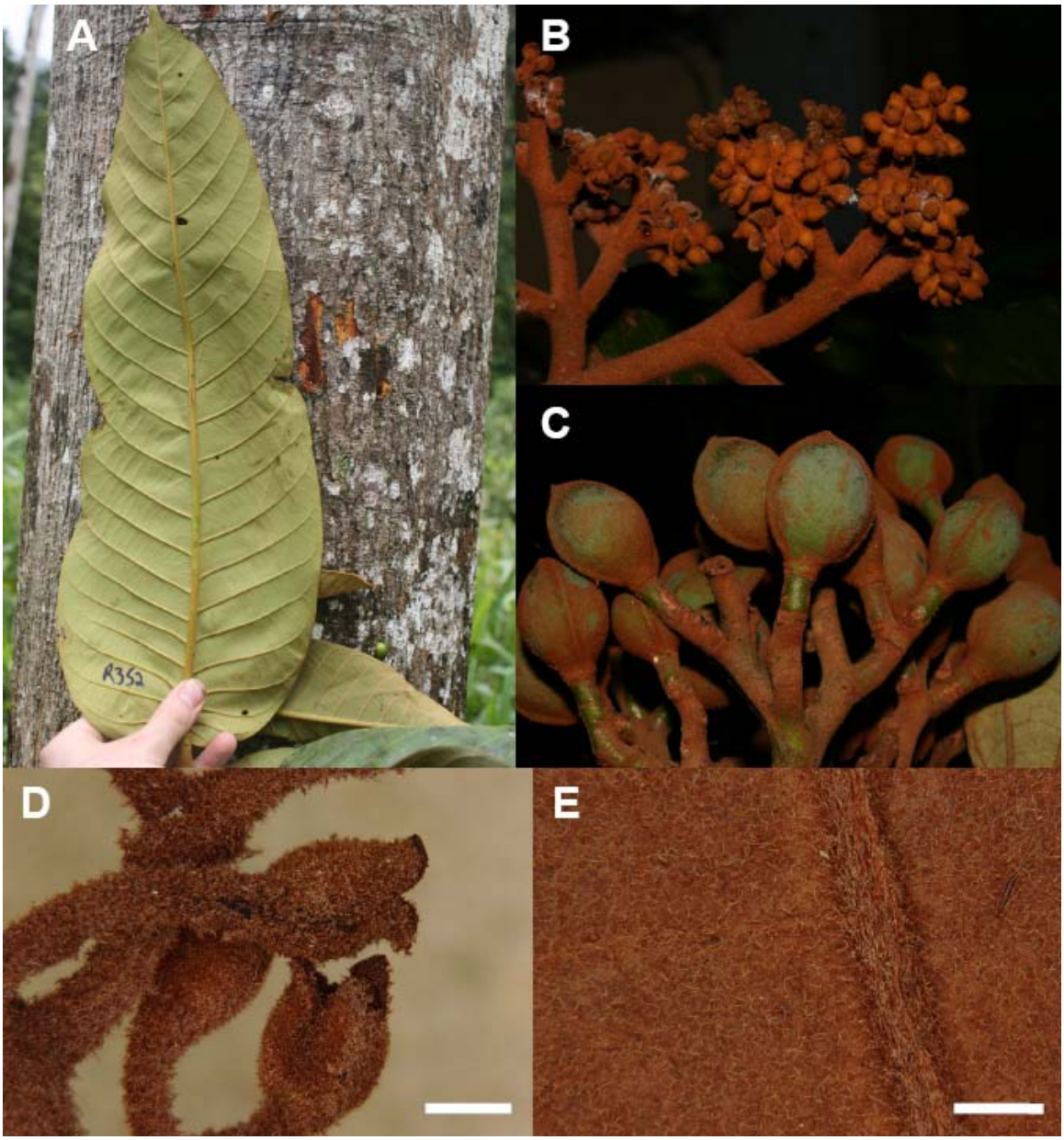

Figure 3.13 *Virola sebifera-BL* (Big Leaf morphotype) : A) Bark with leaf showing abaxial surface. B) Mature pistillate inflorescence. C) Nearly mature infructescence showing deciduos red pubescence. D) Close-up of staminate flowers (scale bar= 1 mm). E) Abaxial leaf pubescence and secondary vein (scale bar= 1 mm).

Figure 3.14 *Virola dixonii* branch showing adaxial and abaxial leaf surfaces (15 cm ruler).

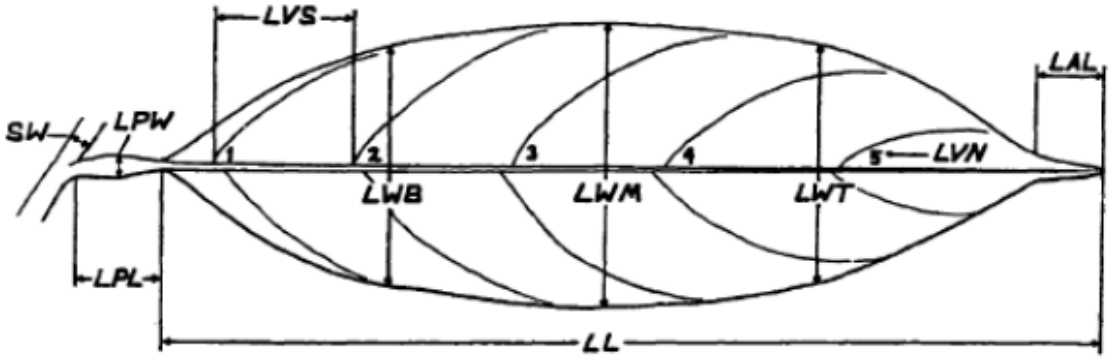

Figure 3.15 Illustration of morphological metrics measured on provisional new species: SW=Stem Width, LPL=Petiole Length, LPW=Petiole Width, LVS=space between secondary veins, LVN=Vein Number, LL=Lamina Length, LWB= Leaf Width at ¼ length, LWM=Leaf Width at ½ length, LWT=Leaf Width at ¾ length, LVN=Leaf Vein Number, LAL=Leaf Acumen Length.

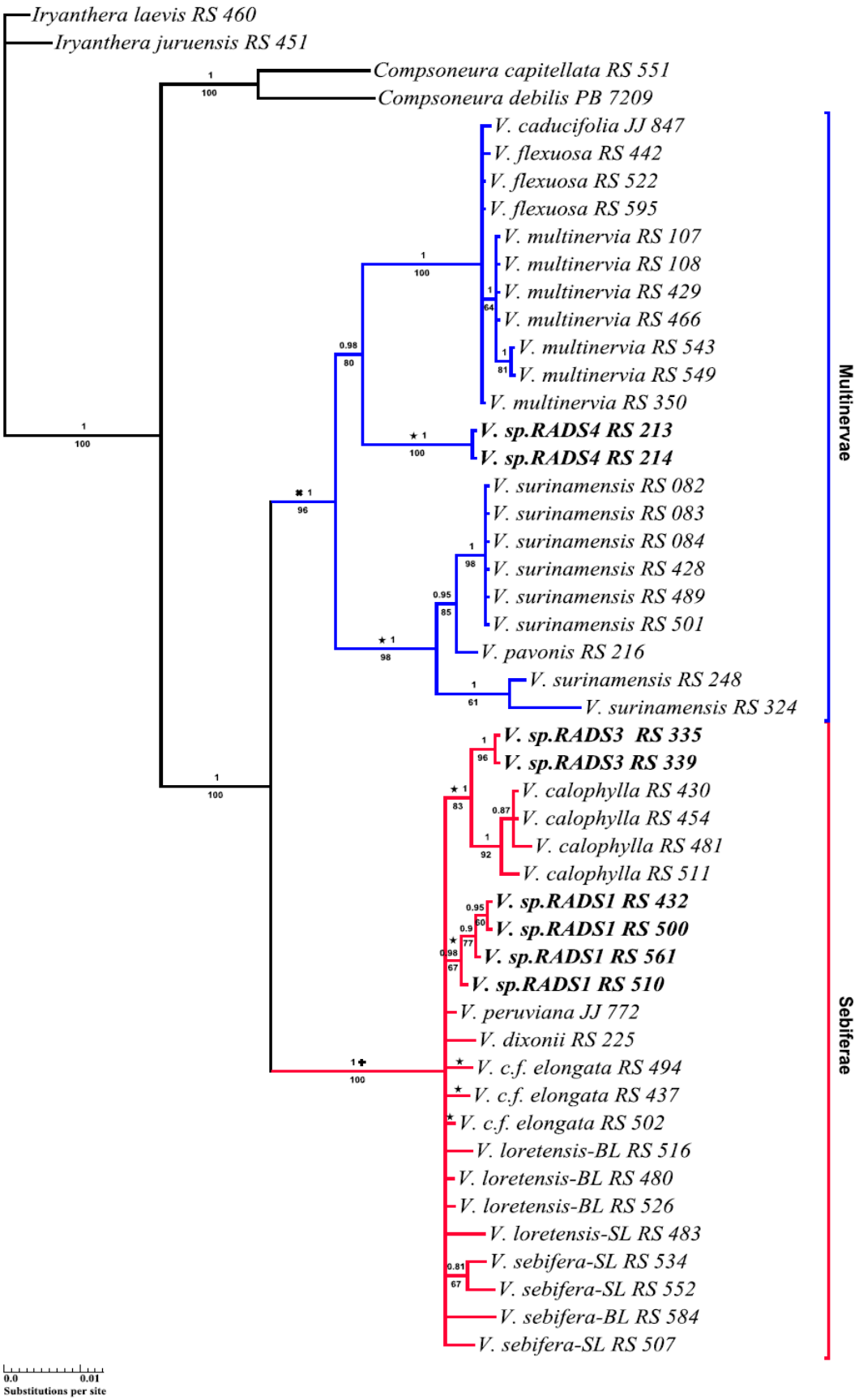
Iryanthera laevis RS 460
Iryanthera juruensis RS 451
Compsoneura capitellata RS 551
Compsoneura debilis PB 7209
V. caducifolia JJ 847
V. flexuosa RS 442
V. flexuosa RS 522
V. flexuosa RS 595
V. multinervia RS 107
V. multinervia RS 108
V. multinervia RS 429
V. multinervia RS 466
V. multinervia RS 543
V. multinervia RS 549
V. multinervia RS 350
V. sp.RADS4 RS 213
V. sp.RADS4 RS 214
V. surinamensis RS 082
V. surinamensis RS 083
V. surinamensis RS 084
V. surinamensis RS 428
V. surinamensis RS 489
V. surinamensis RS 501
V. pavonis RS 216
V. surinamensis RS 248
V. surinamensis RS 324
V. sp.RADS3 RS 335
V. sp.RADS3 RS 339
V. calophylla RS 430
V. calophylla RS 454
V. calophylla RS 481
V. calophylla RS 511
V. sp.RADS1 RS 432
V. sp.RADS1 RS 500
V. sp.RADS1 RS 561
V. sp.RADS1 RS 510
V. peruviana JJ 772
V. dixonii RS 225
V. c.f. elongata RS 494
V. c.f. elongata RS 437
V. c.f. elongata RS 502
V. loretensis-BL RS 516
V. loretensis-BL RS 480
V. loretensis-BL RS 526
V. loretensis-SL RS 483
V. sebifera-SL RS 534
V. sebifera-SL RS 552
V. sebifera-BL RS 584
V. sebifera-SL RS 507
Multinervae
Sebiferae
0.0
0.01
Substitutions per site

Figure 3.16 Bayesian majority rule consensus tree of 45 *Virola* taxa with proportional branch lengths and Bayesian posterior probabilities (PP) above, and bootstrap support measures below branches. Stars (★) indicate clades with sessile leaf hairs (compared to stalked leaf hairs), the "✖" indicates a clade with type II pollen and the "✚" indicates type I pollen (*sensu* Walker and Walker 1979). "SL" and "BL" following *Virola sebifera* and *Virola loretensis* collections indicate small-leaf and big-leaf morphotypes respectively. Additional letters following taxa names indicate collectors names and numbers (RS=Royce Steeves, JJ=John Janovec, and PB=Paul Berry). Additional information can be found about these specimens in table 3.1 as well as the barcode of life database. Taxa in bold are provisional new species.

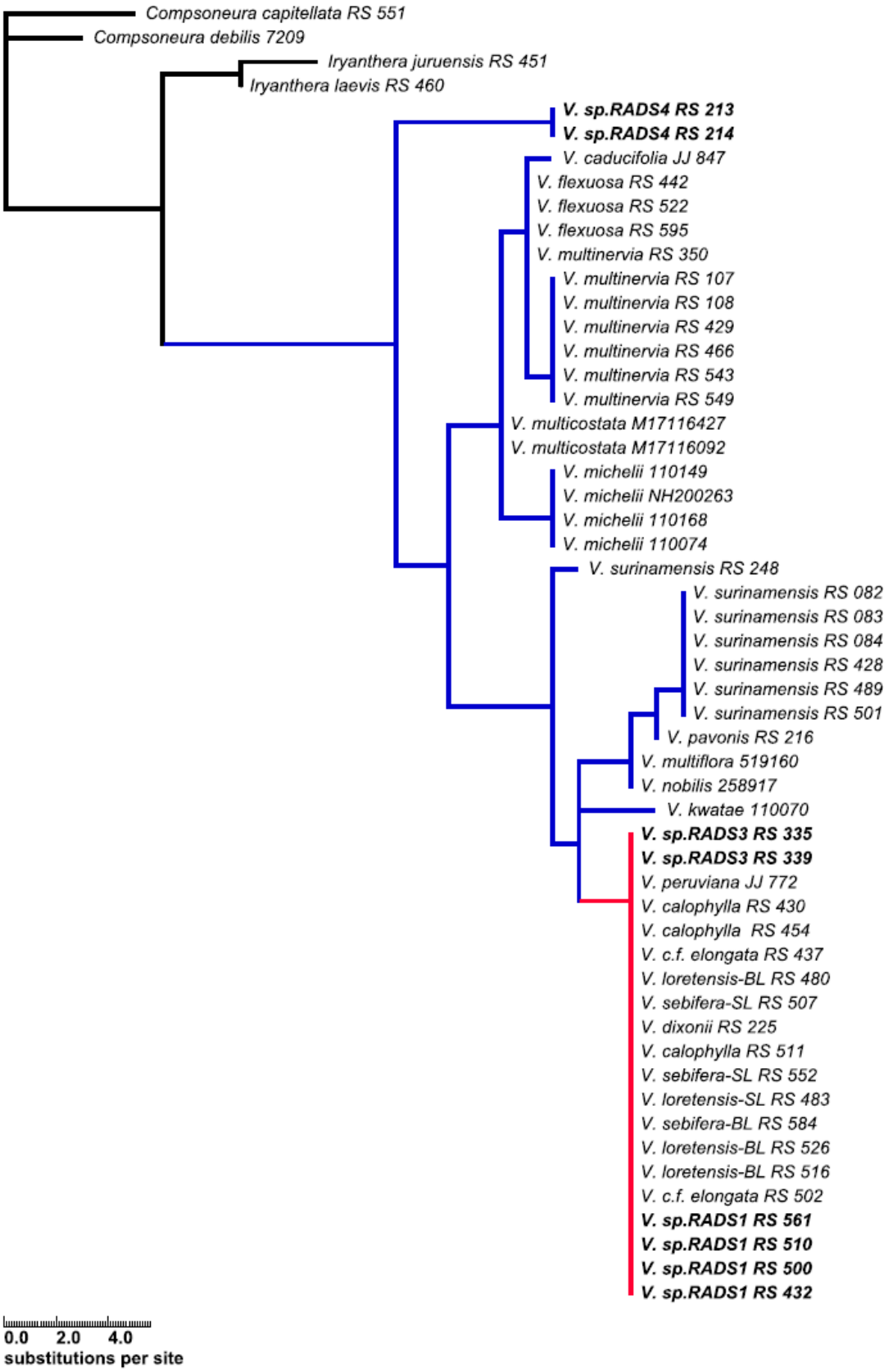
Compsoneura capitellata RS 551
Compsoneura debilis 7209
Iryanthera juruensis RS 451
Iryanthera laevis RS 460
V. sp.RADS4 RS 213
V. sp.RADS4 RS 214
V. caducifolia JJ 847
V. flexuosa RS 442
V. flexuosa RS 522
V. flexuosa RS 595
V. multinervia RS 350
V. multinervia RS 107
V. multinervia RS 108
V. multinervia RS 429
V. multinervia RS 466
V. multinervia RS 543
V. multinervia RS 549
V. multicostata M17116427
V. multicostata M17116092
V. michelii 110149
V. michelii NH200263
V. michelii 110168
V. michelii 110074
V. surinamensis RS 248
V. surinamensis RS 082
V. surinamensis RS 083
V. surinamensis RS 084
V. surinamensis RS 428
V. surinamensis RS 489
V. surinamensis RS 501
V. pavonis RS 216
V. multiflora 519160
V. nobilis 258917
V. kwatae 110070
V. sp.RADS3 RS 335
V. sp.RADS3 RS 339
V. peruviana JJ 772
V. calophylla RS 430
V. calophylla RS 454
V. c.f. elongata RS 437
V. loretensis-BL RS 480
V. sebifera-SL RS 507
V. dixonii RS 225
V. calophylla RS 511
V. sebifera-SL RS 552
V. loretensis-SL RS 483
V. sebifera-BL RS 584
V. loretensis-BL RS 526
V. loretensis-BL RS 516
V. c.f. elongata RS 502
V. sp.RADS1 RS 561
V. sp.RADS1 RS 510
V. sp.RADS1 RS 500
V. sp.RADS1 RS 432
0.0 2.0 4.0
substitutions per site

Figure 3.17 Neighbour-joining uncorrected p-distance gene tree of 17 *Virola* species. Taxa in bold are proposed new species. Blue branches indicate taxa of the Multinervae clade and red those of the Sebiferae clade. “SL” and “BL” following *Virola sebifera* taxa indicate whether they are of the small-leaf or big-leaf morphotypes. Letters and numbers proceeding species names indicate the collector and collection number.

Literature Cited

Ackerly, D. D., Rankin-De-Merona, J. M., & Rodrigues, W. A. 1990. Tree densities and sex ratios in breeding populations of dioecious central Amazonian Myristicaceae. *Journal of Tropical Ecology* **6**: 239-348.

Alfaro, M. E., & Holder, M. T. 2006. The posterior and the prior in Bayesian phylogenetics. *Annual review of ecology, evolution, and systematics* **37**: 19-42.

Armstrong, J. E., & Irvine, A. K. 1989. Floral biology of *Myristica insipida* (Myristicaceae), a distinctive beetle pollination syndrome. *American Journal of Botany* **76**: 86–94.

Asif, M. J., & Cannon, C. H. 2005. DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*). *Plant Molecular Biology Reporter* **23**: 185–192.

Beloz, A. 1992. Brine shrimp bioassay screening of two medicinal plants used by the Warao: *Solanum straminifolium* and *Virola surinamensis*. *Journal of Ethnopharmacology* **37**: 225-227.

Bennett, B. C., & Alarcón, R. 1994. *Osteophloeum platyspermum* and *Virola duckei* (Myristicaceae): newly reported as hallucinogens from Amazonian Ecuador. *Economic Botany* **48**: 152–158.

Candolle, A. D. 1856. Myristicaceae. *Prodromus Systematis Naturali Vegetablis* **14**: 187-208.

Chagnon, N. A., Le Quesne, P., & Cook, J. M. 1971. Yanomamo¨ Hallucinogens: Anthropological, Botanical, and Chemical Findings. *Current Anthropology* **12**: 72–74.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y. L., & others. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. *Annals of the Missouri Botanical Garden* **80**: 528–580.

Clark, J. L., Neill, D. A., Weber, A., Gruhn, J. A., & Katan, T. 2010. *Shuaria* (Gesneriaceae), an Arborescent New Genus from the Cordillera del Cóndor and Amazonian Ecuador. *Systematic Botany* **35**: 662–674.

Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In *Dynamics of populations*. Proceedings of the Advanced Study Institute on Dynamics of numbers in populations. Wageningen: Oosterbeek (pp. 298-310).

Davis, E. W., & Yost, J. A. 1983. The ethnomedicine of the Waorani of Amazonian Ecuador. *Journal of Ethnopharmacology* **9**: 273–297.

Davis, J. I., Stevenson, D. W., Petersen, G., Seberg, O., Campbell, L. M., Freudenstein, J. V., Goldman, D. H., Hardy, C. R., Michelangeli, F. A., Simmons, M. P., & others. 2004. A phylogeny of the monocots, as inferred from rbcL and atpA sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. *Systematic Botany* **29**: 467–510.

Degen, B., Bandou, E., & Caron, H. 2004. Limited pollen dispersal and biparental inbreeding in *Symphonia globulifera* in French Guiana. *Heredity* **93**: 585–591.

Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F., & Douzery, E. J. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution* **20**: 248-254.

Draheim, H., Cui, M., & Dick, C. W. 2009. Characterization of 14 microsatellite DNA markers for the tropical forest tree *Virola surinamensis* (Rol.) Warb.(Myristicaceae). *Molecular Ecology Resources* **9**: 1386–1388.

Farris, J. S., Kallersjo, M., Kluge, A. G., & Bult, C. 1995. Testing significance of incongruence. *Cladistics* **10**: 315-319.

Fazekas, A. J., Steeves, R., & Newmaster, S. G. 2010. Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques* **48**: 277–285.

Forget, P. M., & Cuijpers, L. 2008. Survival and scatterhoarding of frugivores-dispersed seeds as a function of forest disturbance. *Biotropica* **40**: 380–385.

Forget, P. M., Milleron, T., Feer, F., Henry, O., & Dubost, G. 2000. Effects of dispersal pattern and mammalian herbivores on seedling recruitment for *Virola michelii* (Myristicaceae) in French Guiana. *Biotropica* **32**: 452–462.

Gentry, A. H. 1988. Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 156-159.

Golden, J. L., & Bain, J. F. 2000. Phylogeographic patterns and high levels of chloroplast DNA diversity in four *Packera* (Asteraceae) species in southwestern Alberta. *Evolution* **54**: 1566–1579.

Gonzalez, M. A., Baraloto, C., Engel, J., Mori, S. A., Pétronelli, P., Riéra, B., Roger, A., Thébaud, C., & Chave, J. 2009. Identification of Amazonian trees with DNA barcodes. *PLoS One* **4**: e7483.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series*. (pp. 95–98).

Holbrook, K. M., & Loiselle, B. A. 2009. Dispersal in a Neotropical tree, *Virola flexuosa* (Myristicaceae): Does hunting of large vertebrates limit seed removal? *Ecology* **90**: 1449–1455.

Holbrook, K. M., Loiselle, B. A., & Clark, A. M. 2007. Eight polymorphic microsatellite loci for a Neotropical nutmeg, *Virola flexuosa* (Myristicaceae). *Molecular ecology notes* **7**: 248–250.

Hollingsworth, P. M., Graham, S. W., & Little, D. P. 2011. Choosing and using a plant DNA barcode. *PloS one* **6**: e19254.

Howe, H. F. 1993. Aspects of variation in a neotropical seed dispersal system. *Plant Ecology* **107**: 149–162.

Howe, H. F. 1981. Dispersal of a neotropical nutmeg (*Virola sebifera*) by birds. *The Auk* **98**: 88–98.

Howe, H. F., & Vande Kerckhove, G. A. 1981. Removal of wild nutmeg (*Virola surinamensis*) crops by birds. *Ecology* **62**: 1093–1106.

Howe, H. F., Schupp, E. W., & Westley, L. C. 1985. Early consequences of seed dispersal for a neotropical tree (*Virola surinamensis*). *Ecology* **66**: 781–791.

Janovec, J. P. 2000. A systematic study of *Compsoneura* (A. DC.) Warb., A Neotropical member of the nutmeg family. Texas A&M University Dissertation: 1-359.

Janovec, J. P., & Harrison, J. S. 2002. A morphological analysis of the Compsoneura sprucei complex (Myristicaceae), with a new combination for the Central American species Compsoneura mexicana. *Systematic Botany* **27**: 662–673.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. *The American Naturalist* **104**: 501-528.

Jaramillo, T. S., Muriel, P., Rodrigues, W. A., & Balslev, H. 2000. Myristicaceae novelties from Ecuador. *Nordic Journal of Botany* **20**: 443–447.

Julliot, C. 1996. Fruit choice by red howler monkeys (*Alouatta seniculus*) in a tropical rain forest. *American Journal of Primatology* **40**: 261–282.

Julliot, C., & Sabatier, D. 1993. Diet of the red howler monkey (*Alouatta seniculus*) in French Guiana. *International Journal of Primatology* **14**: 527–550.

Kreader, C. A. 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology* **62**: 1102-1106.

Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjur, O., & Bermingham, E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences* **106**: 18621-18626.

Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H., & Varotto, C. 2008. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* **24**: 727–745.

Lopes, N. P., Kato, M. J., & others. 1999. Antimalarial use of volatile oil from leaves of *Virola surinamensis* (Rol.) Warb. by Waiãpi Amazon Indians. *Journal of ethnopharmacology* **67**: 313–319.

Macedo, D. S., & Anderson, A. B. 1993. Early ecological changes associated with logging in an Amazon floodplain. *Biotropica* **25**: 151-163.

Macrae, W. D., & Towers, G. H. 1984a. *Justicia pectoralis*: a study of the basis for its use as a hallucinogenic snuff ingredient. *Journal of ethnopharmacology* **12**: 93–111.

Macrae, W., & Towers, G. 1984b. An ethnopharmacological examination of *Virola elongata* bark: A South American arrow poison. *Journal of ethnopharmacology* **12**: 75–92.

McKenna, D. J., Towers, G. H. N., & Abbott, F. S. 1984. Monoamine oxidase inhibitors in South American hallucinogenic plants part 2: Constituents of orally-active Myristicaceous hallucinogens. *Journal of Ethnopharmacology* **12**: 179–211.

Müller, K. 2006. Incorporating information from length-mutational events into phylogenetic analysis. *Molecular phylogenetics and evolution* **38**: 667–676.

Nylander, J. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Opler, P. A., & Bawa, K. S. 1978. Sex ratios in tropical forest trees. *Evolution* **32**: 812–821.

Palma-Silva, C., Wendt, T., Pinheiro, F., Barbará, T., FAY, M. F., Cozzolino, S., & Lexer, C. Sympatric bromeliad species (*Pitcairnia spp.)* facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* 20: 3185–3201

Palmé, A. E., Semerikov, V., & Lascoux, M. 2003. Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity* **91**: 465–474.

Pascal, J. P., & Pelissier, R. 1996. Structure and floristic composition of a tropical evergreen forest in south-west India. *Journal of Tropical Ecology* **12**: 191–214.

Pitman, N. C., Mogollón, H., Dávila, N., Ríos, M., García-Villacorta, R., Guevara, J., Baker, T. R., Monteagudo, A., Phillips, O. L., Vásquez-Martínez, R., & others. 2008. Tree community change across 700 km of lowland Amazonian forest from the Andean foothills to Brazil. *Biotropica* **40**: 525–535.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2002. A comparison of tree species diversity in two upper Amazonian forests. *Ecology* **83**: 3210–3224.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* **82**: 2101–2117.

Poinar, G. J., & Poinar, R. 1999. *The amber forest: a reconstruction of a vanished world*. Princeton, N.J.: Princeton University Press.

Poinar, H. N. 1998. Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281**: 402-406.

Prance, G. T. 1972. Ethnobotanical notes from Amazonian Brazil. *Economic Botany* **26**: 221–237.

Qiu, Y. L., Chase, M. W., Les, D. H., & Parks, C. R. 1993. Molecular phylogenetics of the Magnoliidae: cladistic analyses of nucleotide sequences of the plastid gene rbcL. *Annals of the Missouri Botanical Garden* **80**: 587–606.

Rieseberg, L.H., & Soltis, D.E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**: 65-84.

Rodrigues, W. A. 1980. Revisão taxonômica das espécies de *Virola* Aublet (Myristicaceae) do Brasil. *Acta Amazonica* **(suplemento)**: 1-127.

Rodrigues, W. A. 1989. A new Venezuelan *Virola* (Myristicaceae). *Annals of the Missouri Botanical Garden* **76**: 1163-1164.

Rodrigues, W. A. 2002. Notas taxonômicas sobre Myristicaceae neotropicais. *Acta Biológica Paranaense* **31**: 71-77.

Rogers, Z. S. 2002. A new species of *Weinmannia* (Cunoniaceae: Cunonieae) from southern Ecuador. *Novon* **12**: 249–252.

Rohlf, F. J. 2006. *tpsDig version 2.10, Department of Ecology and Evoluation*. State University of New York at Stony Brook, New York.

Ronquist, F., & Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574.

Roumy, V., Garcia-Pizango, G., Gutierrez-Choquevilca, A. L., Ruiz, L., Jullian, V., Winterton, P., Fabre, N., Moulis, C., & Valentin, A. 2007. Amazonian plants from Peru used by Quechua and Mestizo to treat malaria with evaluation of their activity. *Journal of ethnopharmacology* **112**: 482–489.

Russo, S.E. 1992. Linking spatial patterns of seed dispersal and plant recruitment in a neotropical tree, *Virola calophylla* (Myristicaceae). Harvard University. PhD dissertation: 1-148.

Russo, S. E. 2003. Responses of dispersal agents to tree and fruit traits in *Virola calophylla* (Myristicaceae): implications for selection. *Oecologia* **136**: 80-87.

Sabatier, D. 1997. Description et biologie d'une nouvelle espèce de *Virola* (Myristicaceae) de Guyane. *Adansonia, Série* **3**: 273–278.

Sang, T., Crawford, D. J., & Stuessy, T. F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**: 1120-1136.

Sasaki, Y., Miyoshi, D., & Sugimoto, N. 2006. Effect of molecular crowding on DNA polymerase activity. *Biotechnology Journal* **1**: 440–446.

Sauquet, H. 2004. Systematic revision of Myristicaceae (Magnoliales) in Madagascar, with four new species of Mauloutchia. *Botanical Journal of the Linnean Society* **146**: 351–368.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W., & Le Thomas, A. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* **142**: 125–186.

Sauquet, H., & Le Thomas, A. 2003. Pollen diversity and evolution in Myristicaceae (Magnoliales). *International Journal of Plant Sciences* **164**: 613–628.

Sawadjoon, S., Kittakoop, P., Kirtikara, K., Vichai, V., Tanticharoen, M., & Thebtaranonth, Y. 2002. Atropisomeric myristinins: Selective COX-2 inhibitors and antifungal agents from *Myristica cinnamomea*. *The Journal of Organic Chemistry* **67**: 5470-5475.

Schultes, R. E., & Holmstedt, B. 1968. The vegetal ingredients of the Myristicaceous snuffs of the Northwest Amazon. *Rhodora* **70**: 113-160.

Schultes, R. E., & Raffauf, R. F. 1990. *The Healing Forest*. Portland, Oregon: Dioscorides Press.

Schultes, R. 1981. Iconography of the New World Hallucinogens. *Arnoldia* **41**: 80-125.

Simmons, M., & Ochoterena, H. 2000. Gaps as characters in sequence-based phylogenetic analysis. *Systematic Botany* **49**: 369-381.

Smith, A. C. 1937. The American species of Myristicaceae. *Brittonia* **2**: 393-510.

Soltis, D. E., Smith, S. A., Cellinese, N., Wurdack, K. J., Tank, D. C., Brockington, S. F., Refulio-Rodriguez, N. F., Walker, J. B., Moore, M. J., Carlsward, B. S., & others. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* **98**: 704.

Speiss, A., T. Mueller, & Ivell, R. 2004. Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of Taq polymerase by the disaccharide trehalose. *Clinical chemistry* **50**: 1256-1259.

Stöver, B. C., & Müller, K. F. 2010. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC bioinformatics* **11**: 1-9.

Suzuki, Y., Glazko, G. V., & Nei, M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 16138-16143.

Swafford, D. 1999. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4 . Sunderland : Sinauer Associates.

Tate, J. A., & Simpson, B. B. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* **28**: 723–737.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**: 4673-4680.

Ulloa Ulloa, C., & Neill, D. A. 2006. *Phainantha shuariorum* (Melastomataceae), una especie nueva de la Cordillera del Condor, Ecuador, Disyunta de un genero Guyanes. *Novon* **16**: 281-285.

Warburg, O. 1897. Monographie der Myristicaceen. *Nova Acta Acad. Caes. Leop.-Carol* **68**: 1-680.

Walker, J.W. & A.G. Walker. 1979. Comparative pollen morphology of the American Myristicaceous genera *Compsoneura* and *Virola*. *Annals of the Missouri Botanical Garden* **66**: 731-755.

Williams, L. O. 1960. Ucuuba and related wax-like vegetable tallows. *Plant Foods for Human Nutrition (Formerly Qualitas Plantarum)* **7**: 295–296.

Yang, Z., & Rannala, B. 2010. Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences* **107**: 9264-9269.

Zimmerman, S. B., & Harrison, B. 1987. Macromolecular crowding increases binding of DNA polymerase to DNA: an adaptive effect. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 1871-1875.

Zipparro, V. B., & Morellato, L. P. 2005. Seed predation of *Virola bicuhyba* (Schott) Warb.(Myristicaceae) in the Atlantic forest of south-eastern Brazil. *Revista Brasileira de Botânica* **28**: 515–522.

Chapter 4

A MORPHOLOGICAL AND MOLECULAR INVESTIGATION OF *VIROLA SEBIFERA* Aublet AND *VIROLA LORETENSIS* A. C. Sm.

Abstract

Virola is a widespread and species-rich genus of neotropical Myristicaceae. Although *Virola* is an ecologically dominant and ethnobotanically significant group of trees, they have received little contemporary taxonomic study. In this study I test whether ecologically separate and polymorphic morphotypes of *V. sebifera* and *V. loretensis* represent undescribed cryptic species using morphological and molecular data. Species hypotheses were tested using a multivariate analysis of morphological characters as well as a statistical parsimony analysis using a low-copy nuclear locus. Morphological analyses show discrete variation in morphotypes of both *V. sebifera* and *V. loretensis* that corresponds to upland and floodplain forest habitats. Haplotype networks constructed from molecular data reveal some distinct clades as well as a high degree of apparent incomplete lineage sorting in these taxa. It is concluded that these data suggest that these ecologically separated morphotypes of each species represent distinct undescribed species.

Introduction

There are an estimated 235,000-420,000 (Govaerts 2003, Scotland and Wortley 2003) seed plant species in the world. The vast majority (about 200,000) of these plant species reside in tropical areas, with the neotropical region contributing about 90,000 species (Thomas 1999). It is now widely recognized that tropical moist forests of the world are being rapidly destroyed by human activity (Dodson and Gentry 1991, Gentry 1988, 1992). The sheer organismal diversity and short history of intensive study in tropical regions are demonstrated by the meager understanding we have of these fragile ecosystems and their astonishing number of species. Many species remain to be discovered, described, and classified; many more remain to be understood from ecological and evolutionarily perspectives. Examinations of the composition and diversity of tropical rainforests have been the focus of many contemporary plant ecologists, however, they are faced with a flora that is exceedingly diverse and difficult to identify.

Myristicaceae is composed of 21 genera and approximately 500 species of trees that have significant ecological and ethnobotanical importance worldwide in wet lowland tropical forests. Floristic and ecological studies have revealed that Myristicaceae ranks among the ten, and often five, most diverse and prominent tree families in these ecosystems (Gentry 1982, 1988, Pascal and Pelissier 1996, Pitman et al. 2001, 2002, 2008). Worldwide, the most important species economically is the Asiatic *Myristica fragrans*, which is widely cultivated in tropical areas for the spices nutmeg and mace. In the American tropics, numerous species are valued as sources of food, medicine, narcotics, and timber (Gottlieb 1973, Prance 1972, Schultes and Raffaf 1990). The most

widely employed Myristicaceae genus in the Neotropics for both commercial and traditional ethnobotanical uses is *Virola*.

Virola is a genus of canopy to sub-canopy trees endemic to South and Central America. They are the most speciose nutmeg genus in the new world tropics and often rank amongst the most abundant trees in this region (Pitman et al. 2001, 2002, 2008). Despite being the most species-rich genus of neotropical Myristicaceae, *Virola* has received relatively little recent taxonomic attention. *Virola* was first described by Aublet (1775) as a genus endemic to Central and South America. The genus was divided into 6 groups (*Mollissimae*, *Sebiferae*, *Calophyllae*, *Rugulosae*, *Surinamenses*, and *Subsessilis*) by Smith's (1937) monograph of the neotropical Myristicaceae where he recognized 38 species. The last comprehensive treatment of *Virola* was performed by Rodrigues (1980) and was largely restricted to the species residing within the political boundaries of Brazil. Since 1980, taxonomic investigations of *Virola* have been primarily alpha-taxonomic in nature and about 60 species are currently recognized in the genus (Rodrigues 1989, Sabatier 1997, Jaramillo et al. 2000, Rodrigues 2002).

Virola sebifera is a common and widespread canopy tree that is found in tropical rainforests from 1700 m a.s.l. in the Andes to near sea level in Eastern Brazil. This species also ranges across a wide latitudinal gradient as it can be found from Costa Rica to Southern parts of Brazil (Provinces of São Paulo and Matto Grosso del Sul). Across its range, and even within populations, members of *V. sebifera* exhibit a great deal of morphological diversity. Smith (1937) described *V. sebifera* leaves as being ovate, oblong, deltoid, or elliptic in shape with cordate, rounded, or truncate bases and ranging from 15-47 cm in length and 6-15 cm in width. In spite of this great variation in leaf

morphologies, Smith (1937) united 9 species previously described spcies (*V. macoa*, *V. boliviensis*, *V. panamensis*, *V. cordifolia*, *V. fulva*, *V. warburgii*, *V. peruviana*, *V. mycetis*, and *V. venezuelensis*) into *V. sebifera*, citing numerous connecting morphological forms for this monospecific concept.

Virola loretensis is a small tree (3-10 m according to Smith 1937) of floodplain forests of the Northwestern Amazon basin (Ecuador, Peru and Colombia). It was first described by Smith (1937) from 12 collections from the Peruvian province of Loreto and Acre, Brazil. The leaves of *V. loretensis* are distinguished by having a ferruginous-tomentose pubescence which is particularly prominent on abaxial veins, large size (15-37 cm long and 4-10 cm broad) and deciduous habit during pronounced dry seasons. The inflorescences of this species are also easily distinguished by a dense ferruginous-tomentose pubescence which is also present on the fruits (Smith 1937).

In the course of field and herbarium investigations, it became apparent that both *V. sebifera* and *V. loretensis* possess a great deal of morphological variation within each species, often within the same locality. The objective of this investigation was to test whether this morphological diversity is indicative of within-species phenotypic plasticity or cryptic speciation. This was tested by multivariate analysis of morphological data as well as the construction of a haplotype network from a low-copy nuclear gene.

Materials and Methods

Collection sites

Since members of *Virola* are particularly recalcitrant with respect to DNA extraction and sequencing, previously collected herbarium specimens were of no utility to

this investigation. Consequently, only recent collections made where leaf tissue was promptly preserved in silica dessicant could be employed in this study, thereby greatly limiting the geographic and morphological breadth of sampling. In attempts to capture the inherent morphological and genetic variability encountered in these species, five populations were sampled from Southern Peru (n=4) and Brazil (n=1) spanning 2560 km in distance and 1600 m of elevation (Figure 4.1). *Virola sebifera* samples were collected in Peru from two cloud forests (Manu Paradise Lodge and Quincemil) and two lowland rainforests (Centro de Investigación y Capacitación Río Los Amigos [CICRA] and Puerto Maldonado). *Virola sebifera* samples were also collected from a single Brazilian population by Mark Leithead and Lucas Silva at the Reserva Ecologica do IBGE near the city of Brasilia. *Virola loretensis* samples were collected from CICRA and Puerto Maldonado, Peru. All Peruvian samples were collected by the author and associated collectors. Herbarium vouchers were made for all collections and have been deposited at the OAC herbarium (Table 4.1). Leaf tissue was placed in silica gel immediately upon collection to minimize oxidation as members of this genus typically do not yield DNA amenable to PCR unless tissue is desiccated immediately.

Study species

Two relatively common morphotypes of *Virola sebifera* were collected. A big-leafed form (hereafter *V. sebifera*-BL) was found growing at all locations in either seasonal flooded forests of the Madre de Dios river or in well drained upland habitats in cloud forests (750-1700 m asl). A second and comparatively smaller-leafed form (*V.*

sebifera-SL) was found to grow exclusively in terra firme forests (upland, non-flooded, forests) of the Madre de Dios watershed and near streams in cloud forest habitats.

Large and small-leafed forms of *V. loretensis* were also collected from the Madre de Dios watershed. The big-leafed variety (*V. loretensis*-BL) was relatively rare (only 3 trees encountered in the course of 2 months) and was collected from floodplain forests of two locations. The small-leafed variety (*V. loretensis*-SL) was rarely collected (only 3 trees encountered) in terra firme forests at the Los Amigos Biological Station. Pictures were taken of vegetative and reproductive material (when available) to document the 4 different morphotypes (Figures 4.2-4.5).

Morphological analysis

After sterile material was photographed, pressed, and dried, measurements of 10 quantitative vegetative metrics (Figure 4.6) were made from digital photographs using tpsDig 2.0 (Rohlf 2006). A principal component analysis (PCA) was performed on leaf morphometric data using Canoco 4.5 (ter Braak, 1998) to identify the length of the ordination axis and unimodal ordination model was applied (Correspondence Analysis, CA). The relationship between quantitative leaf characters was analysed via nonmetric multi-dimensional scaling (NMS; Kruskal 1964, Primer 2002). In NMS, the Bray-Curtis distance measure was used because of its robustness for both large and small scales on the axes (Minchin 1987). Data were standardized by species maxima and two-dimensional solutions were appropriately chosen based on plotting a measure of fit ('stress') to the number of dimensions. Stress represents distortion in the data and a stress value over 0.2 is high enough that the results are invalidated (Primer 2002). One thousand

iterations were used for each NMS run, using random start coordinates. The first two ordination axes were rotated to enhance interpretability with the different axes. As an independent check, detrended correspondence analysis (DCA; ter Braak 1998) was used to evaluate the NMS classification. A Pearson-Correlation analysis was performed to investigate which metrics contributed the most to the differentiation of taxa.

DNA extraction, amplification, molecular cloning, sequencing and alignment

A total of 16 samples including 5 *V. sebifera*-BL, 5 *V. sebifera*-SL, 3 *V. loretensis*-BL, and 3 *V. loretensis*-SL were used in genetic analyses. Collection information for samples used in genetic analyses can be found in Table 4.1. Total genomic DNA was extracted from leaf tissue of silica-dried specimens using the Macherey-Nagel Nucleospin II plant Kit. Lysis buffer 1 was used according to the manufactures' instructions with the exception of the elongation of the post homogenization incubation period to 1hr (from 10 minutes) and the addition of 20mM N-Phenacylthiazolium Bromide which has been found to improve amplification of recalcitrant samples (Poinar et al. 1998, Asif and Cannon 2005).

A phylogenetic investigation of *Virola* (chapter 3 of this thesis) found the AGT1 gene and associated exons (Li et al. 2008) to be the most variable loci in *Virola* and was therefore employed in this study. This locus was amplified using the primers AGT1-MYR-F (5'-GGGCATTGACGTAGCTTTGACAGG-3'; this thesis), and AGT1-MYR-R (5'-GTGCAGTTCTTCAAGCCCCAAGC-3'; this thesis). PCR was performed with 0.5U of Platinum*Taq*® (Invotrogen, Burlington, Ontario) DNA polymerase in a 20 µl reaction containing 1X reaction buffer, 2.5 mM MgCl, 8% W/V Polyethylene glycol

(Zimmerman and Harrison 1987, Sasaki et al. 2006) , 0.2 M trehalose (Speiss et al. 2004), 2 µg BSA, 0.2 mM each DNTP, and 0.2 µM of each primer. Cycling conditions entailed an initial denaturation step of 4 min at 95°; 30 cycles of 95° for 15 s, 67° for 10 s, 72° for 30 s: and a final elongation step of 72° for 2 min followed by a 4° hold. Platinum*Taq*® was used as it was found to work well with the AGT1 primer set and be more robust than two high fidelity polymerases (Kapa HIFI and Finnzymes Phusion) to the inhibiting substances present in these samples.

Since there exists the possibility of multiple alleles with low-copy nuclear genes, PCR amplicons were cloned to sequence individual alleles. PCR product was diluted 10X and then cloned using the StrataClone™ PCR Cloning Kit (VWR, Mississauga, Ontario). PCR fragments were ligated to the StrataClone™ PCR cloning vector in a reaction mixture containing 1.5 µl of StrataClone™ Cloning Buffer, 1 µl of diluted PCR product and 0.5 µl of StrataClone™ Vector Mix amp/kan. The reaction mixture was then incubated at room temperature for 10 minutes and transformed into SoloPack competent *E.coli* cells according to the manufacturer's instructions. The transformed cells were incubated in Luria-Bertani (LB) Medium with agitation at 37^{o}C for 2 hours to allow the cells to recover. Transformed cell were then spread onto 1% agar LB plates containing 100mg/L ampicillin (Sigma, Oakville, Ontario) and 64mg/L X-Gal (Fisher Scientific, Ottawa, Ontario, Canada). Plates were then incubated at 37^{o}C for 20 hours.

Colonies are white in color if they fail to express the β-Galactosidase gene, which is interrupted with the successfully insertion of a sequence into the vector. These white colonies were selected and transferred into a buffer containing low TE and 2% V/V tween-20 and then subjected to 95 ^{o}C for 10 min to lyse the cells. This lysate was then

used as template DNA for a subsequent PCR reaction using the same primers and cycling conditions employed for the initial amplification reaction.

In case multiple copies of this locus were present in any individuals, 16 colonies were sequenced in one direction from each individual tree. This number of colonies was selected as it gives a 0.96 probability of detecting all alleles present in a sample with 4 alleles according to the formula:

$$P= [1-(t-1/t)^{n})]^{t}$$

where "t" is the number of alleles in an individual and "n" is the number of colonies sequenced (Joly et al. 2006).

Amplification products were sequenced directly using the AGT1-MYR-F primer used for PCR. Cycle sequencing reactions were performed in a 10.5 μL reaction volume containing 0.5 μL of BigDye terminator mix v3.1, 1.88 μL of 5x sequencing buffer (Applied Biosystems), 1.0 μM of primer and 0.5 μL of PCR product. Thermal cycling parameters were 96° for 2 min; 30 cycles of 96° for 30s, and 60° for 4 min; and a 4° hold. Cycle sequencing reactions were cleaned using sephadex columns (Cat. no. S5897; Sigma-Aldrich, St. Louis, MO, USA) and the samples were run on an ABI 3730 sequencer (Applied Biosystems).

Sequences were edited using Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA). DNA sequences were aligned using the default settings of the ClustalW algorithm (Thompson et al. 1994) in Bioedit (Hall 1999) and adjusted manually. Alignments can be found in appendix 4. Single nucleotide polymorphisms (SNP's) found in only one of the cloned sequences of a specimen were assumed to be due to polymerase error and were therefore omitted from the alignment (i.e. the allele was

retained but suspected erroneous base pairs were changed to the nucleotide used at that position by the other alleles of that individual). If two or more clones of an individual specimen contained the same alleles they were retained but it was kept in mind that these may be artifacts of PCR. All unique alleles from each individual were then aligned for haplotype network analysis.

Haplotype networks were constructed by statistical parsimony (Templeton et al. 1992, Templeton 1998) using the program TCS (Clement et al. 2000) which calculates the number of differences between haplotypes that are due to a sequence of single mutations at each site. Networks were calculated with TCS as implemented by ANeCA (Panchal 2007) by treating gaps as a 5th character state and with a 95% confidence interval. Haplotype networks may be preferable to traditional phylogenetic methods at low taxonomic levels because population level data often violate many of the assumptions made by traditional tree-building methods, such as Maximum Parsimony, Maximum Likelihood and Bayesian Analyses (Posada and Crandall 2001). In comparison, networks are able to utilize haplotypic data that contains low levels of genetic divergence, ancestral haplotypes, multiple mutational variants from a given haplotype within the population, and reticulation that may be caused by recombination or hybridization (Templeton et al. 1992, Posada and Crandall 2001).

Results

Morphological Analyses

Morphological measurements were recorded for 3 specimens each of *V. loretensis*-BL and *V. loretensis*-SL, 11 *V. sebifera*-BL and 9 *V. sebifera*-SL. Leaf length of *V. sebifera*-BL ranged from 24-49 cm while *V. sebifera*-SL leaves were 13-32 cm

long; *V. loretensis*-BL was 24.5-42 cm and *V. loretensis*-SL ranged from 16-22 cm in length. The ordination analyses of quantitative leaf characters showed a pronounced distinction between big-leaf and small-leaf varieties of the two species but both *V. sebifera* and *V. loretensis* clustered together with respect to big or small leaf-type (Figure 4.7). One sample of *V. loretensis*-BL (RS 480) was positioned approximately mid-way between big-leaf and small-leaf varieties. *Virola loretensis*-BL is deciduous during the dry season in South-Eastern Peru and this sample was in the process of regenerating its leaves and therefore this sample may not represent its fully developed size. Additionally, one sample of *V. sebifera*-SL was grouped with *V. sebifera*-BL and this may be due to the fact that this sample was the only juvenile included in the analysis and it was observed that juveniles growing under the canopy tend to have larger leaves than adult trees that have emerged from the sub-canopy. Pearson correlation analyses revealed that the X-axis is strongly correlated with numerous measures of leaf length and width (Figure 4.7; Table 4.2) and the y-axis was relatively weakly correlated to leaf acumen length and stem width (Fig 4.7; Table 4.2).

Although the fruits of these species were not available for quantitative analysis they were qualitatively observed in the field (Figures 4.2-4.5). The fruits of *V. sebifera*-BL are generally elliptic in shape, strongly carinate at suture of the pericarps, and ferruginous-tomentose but quickly deciduous. The fruits of *V. sebifera*-SL are more globose in shape, weakly carinate at pericarp suture, and persistently ferruginous-tomentose. Individuals of *V. sebifera*-BL were typically found growing in the floodplain forests or cloud forests; *V. sebifera*-SL were typically found inhabiting terra firme forests.

The fruits of *V. loretensis*-BL are globose in shape but are readily distinguished from other taxa in having long (6-9 mm) persistent ferruginous-tomentose pubescence on the pericarp surface. The fruits of *V. loretensis*-SL are covered in a comparatively shorter indument (<1 mm) and their size, shape and color are is similar to the fruits of *V. sebifera*-SL. In addition to fruit characters, these *V. loretensis* taxa can be distinguished based on their habitats (flood plain versus terra firme forest) and height (*V. loretensis*-BL = 3-5 m, *V. loretensis*-SL= 20-40 m).

Molecular Analyses

In total, 233 *AGT1* clones were successfully sequenced for the 16 individuals used in these analyses. The alignment consistent of 686 bp, 20 of which were parsimony informative. There were 30 unique alleles found and only four alleles were shared by multiple individuals. Two alleles were shared by members of *V. sebifera*-SL, two amongst *V. loretensis*-BL, and one was shared amongst *V. loretensis*-BL and *V. loretensis*-SL. The number of alleles per individual varied from 1-4 but most individuals (n=8) contained only 2 alleles. Of the 233 AGT1 sequences recovered from clones, 90 (38.6%) contained one or more single nucleotide polymorphisms not found in any other cloned sequence of their respective individual specimens. These SNP's were assumed to be due to PCR error and the nucleotide(s) was/were changed in the alignment to the same nucleotide used in equivalent alleles of that individual.

The haplotype network (Figure 4.8) shows 30 unique alleles and contains only a single loop. The node representing *V. sebifera*-BL allele "RS 636-AL-R" connects primarily to other *V. sebifera*-BL alleles and one *V. loretensis*-SL allele (RS 465-AL-X), which itself connects solely to other *V. loretensis* individuals. The one loop recovered in

the haplotype network (between AL-N and AL-O) analysis may be due to an ancient recombination event or may be an artifact of PCR mediated recombination (Jumpponen and Johnson 2005, Lahr and Katz 2009).

Discussion

This investigation represents one of the first assessments of infraspecific genetic variation in Myristicaceae (Chapter 1 of this thesis, Degen et al. 2001). To date, investigations of infraspecific genetic variation in neotropical trees have mainly been restricted to questions concerning population genetics (Hamrick et al. 1993, Dick et al. 2003), with few investigations testing taxonomic hypotheses with molecular methods (Duminil et al. 2006). The Myristicaceae, like many other diverse tropical plant families, have received relatively little taxonomic study since the advent of molecular methods. The following discourse outlines the implications of the morphological and molecular investigations with respect to the taxonomic status of *V. sebifera* and *V. loretensis*.

Morphological Analyses

Observations of *V. sebifera* and *V. loretensis* taxa in the field suggested that both of these species were either polymorphic, morphologically plastic depending on which habitat they were found in, or contain undescribed species. The multivariate analysis indicates that both *V. sebifera* small and big-leaf morphotypes form practically discrete clusters on each side of the y-axis of the ordination. This suggests that leaf size variation is not continuous between these two morphotypes. Although sample sizes were relatively small, the two morphotypes of *V. loretensis* also form two discrete clusters on each side of the y-axis. Big-leaf and small-leaf morphotypes of *V. sebifera* and *V. loretensis* also

grouped together in the ordination of leaf characters. This clustering is of interest since both small-leafed varieties were found growing in terra firme forests and big-leafed varieties were found growing primarily in floodplain forests. The two small-leafed morphotypes were both elliptic-lanceolate in shape with rounded bases and acuminate tips, while the big-leafed morphotypes were oblong in shape and had cordate bases and acute tips. Increased morphological sampling from additional trees and geographic regions would be beneficial to explore whether these morphological patterns between the two habitats are maintained with increased taxon and geographic sampling.

Molecular Analyses

In spite of relatively clear vegetative and reproductive morphological differences among *V. sebifera* and *V. loretensis* and their respective morophotypes, there was little structure in terms of species and morphotypes in the haplotype network. Six alleles of *V. sebifera*-BL (alleles R-W, Figure 4.8) formed a clade as did 9 *V. loretensis*-BL and SL alleles (alleles X-AC, Figure 4.8). Aside from these two clades, alleles from both species and leaf types were distributed throughout the remaining network. This may indicate that contemporary gene flow and/or incomplete lineage sorting may be responsible for the lack of resolution between *V. sebifera* and *V. loretensis* and their respective morphotypes. Gene flow resulting in introgression of alleles among these taxa is probably as interspecific hybridization is thought to be a commonplace among closely related angiosperms (Golden and Bain 2000, Palme et al. 2003 Palma-Silva et al. 2001) and is likely an important mechanism in plant speciation (Whitney et al. 2010, Southcott et al. 2011). Incomplete lineage sorting and/or low nucleotide substitution levels could also be responsible for the lack of resolution in this haplotype network as a previous

phylogenetic analysis of *Virola spp.* found little to no resolution among *V. sebifera*, *V. loretensis* and closely related species (see Sebiferae clade, Figure 3.16 in Chapter 3 of this thesis).

Hybridization in tropical trees is presumed to be rare (Ashton 1969) but has been observed with morphological and molecular data in paleotropical Dipterocarpaceae (Kamiya et al. 2010) and neotropical Meliaceae (Duminil et al. 2006). In the course of 4 months of field work no apparent big-leaf x small-leaf hybrids were observed in either *V. sebifera* or *V. loretensis*. It was noted in the field that *Virola loretensis*-SL had morphological affinities to both *V. sebifera* and *V. loretensis*. *Virola loretensis*-SL has leaf size and shape (rounded leaf bases) comparable to *V. sebifera*-SL in addition to similar inflorescences, fruit indument, and tree height, however, *V. loretensis*-SL has a rich reddish leaf indument and orange aril typical of *V. loretensis*-BL. Furthermore, both *V. loretensis* varieties and *V. sebifera*-SL flower at the same time where they were found in South-Eastern Peru so it is conceivable that hybridization could occur. However, of 7 *V. loretensis*-SL alleles sampled only one (AL-Q) was apparently derived from a *V. sebifera*-SL allele. Although it is possible that *V. loretensis*-SL is a contemporary hybrid derived from *V. loretensis*-BL and *V. sebifera*-SL, this may be unlikely given the fact that no sampled alleles were shared between them and few alleles were similar to *V. sebifera*-SL. These potential hybrid collections may merit further investigation as hybridization in tropical trees may be underestimated due to the difficulty in collecting fertile specimens, few morphological differences between species, and general paucity of rigorous taxonomic investigations employing both morphological and molecular data.

It also appears that AGT1 may potentially exist in multiple copies in these *Virola* taxa. Of the 16 trees sampled for molecular analyses, 5 individuals had clones with 3 or 4 alleles (Table 4.1). If alleles differing by only a single nucleotide polymorphism from another allele within the same individual are deleted, then only 3 individuals potentially have multiple copies (RS 480, RS 503, and ML 013). Two of these collections (RS 480 and RS 503) contain alleles that differ by only two polymorphisms from other alleles of the same individual and therefore may be due to PCR error as well. This leaves the single collection from Brazil (ML 013) as an individual with 3 or more alleles that differ by more than two polymorphisms from other alleles in the same individual. It is well known that DNA polymerase error can introduce erroneous bases into PCR amplicons (Ashelford et al. 2005, Dickie 2010, Fazekas et al 2010) and 38.6% of sequenced clones in this study contained suspected erroneous single nucleotide polymorphisms. Since the error rate of Platinum*Taq*® is not advertised, it cannot be ascertained whether this percentage of errors is to be expected. However, a PCR reaction performed using comparable parameters to this study (i.e. 30 cycles, 700 bp fragment, and a *Taq*-based polymerase with an error rate of 2.28×10^{-5}) is expected to produce errors in about 48% of PCR-generated amplicons (Finnzymes PCR fidelity calculator-http://finnzymes.com/pcr/fidelity_calc.php). Although attemps made in this study to use more accurate DNA polymerases were largely ineffectual, it would be preferable to perform additional PCR reactions with a higher fidelity DNA polymerase for subsequent cloning reactions to determine whether the multiple alleles detected are an artifact of polymerase error or due to the existence of multiple copies.

Conclusions

This investigation endeavored to elucidate taxonomic hypotheses in *V. sebifera* and *V. loretensis* by integrating morphological and molecular data. Although the haplotype analysis was somewhat ambiguous in terms of delimiting the 4 morphotypes studied, the combination of morphological discontinuity among conspecific morphotypes and separation of morphotypes among different edaphic habitats is indicative of multiple unrecognized species. These species are not provisionally described in this study as increased sample sizes, more variable molecular markers, and the study of more type specimens are desired prior to the recognition of putative new species.

The discovery of two novel taxa would be highly significant as both were found at one of the most active research stations in the lowland Amazon (Los Amigos Biological Research Station). These two cryptic taxa differed morphologically and ecologically from their larger-leafed conspecifics yet they were not recognized in ecological plots as distinct taxa. A recent study of *Inga spp*. in South-Eastern Peru found that error rates in the identification of *Inga* taxa were around 7% and significantly impacted the accuracy of some, but not all, ecological conclusions from such data (Dexter et al. 2010). After observing permanent ecological plots in Ecuador and Peru, I estimated that mis-identification rates often exceed 20% for Myristicaceae species if they are identified as the correct family, although high, this is still lower than the estimated 50% mis-identification rate estimated for herbarium material (J. Janovec, personal communication).

This study demonstrates the utility of field observation and integrative taxonomic approaches, both of which will be essential to help discover new species in the highly diverse neotropical region. Identification of Myristicaceae taxa solely by molecular

means will likely prove exceedingly difficult as their plastid DNA exhibits low rates of molecular evolution (Sauquet et al. 2003, chapter 1,2, and 3 of this thesis), more variable nuclear loci may be prone to incomplete lineage sorting, and these species may have rapidly diversified relatively recently as has been found in the speciose genus *Inga* (Richardson et al. 2001). The need for comprehensive studies of tropical flowering plants precludes knowledgeable conservation decisions and sustainable utilization of tropical forests. This study shows that our taxonomic understanding of one of the most widespread and abundant genera of the neotropics is rudimentary at best and that much work remains to be done if we wish to explore the ecological and evolutionary history of this important floristic region.

Tables

Table 4.1 Species collected, collection numbers (Coll. #), Morphotype (type)[BL=Big Leaf, SL=Small Leaf], location (CICRA=Los Amigos Biological Station), decimal degrees latitude (Lat.) and longitude (Long.), herbarium accession numbers (Herb. Acc.), and data set sample was used in (M=molecular, m=morphology, # indicates how many alleles were recovered) for 29 collections employed in this study.

Species	Coll. #	type	Location	Lat.	Long.	Herb. Acc.	Data Set
V. loretensis	RS 480	BL	Peru-CICRA	-12.56	-70.09	OAC 94718	Mm-3
V. loretensis	RS 516	BL	Peru-Puerto Maldonado	-12.72	-69.24	OAC 94754	Mm-2
V. loretensis	RS 526	BL	Peru-Puerto Maldonado	-12.72	-69.24	OAC 94764	Mm-2
V. loretensis	RS 452	SL	Peru-CICRA	-12.56	-70.09	OAC 94690	Mm-3
V. loretensis	RS 465	SL	Peru-CICRA	-12.56	-70.09	OAC 94703	Mm-4
V. loretensis	RS 483	SL	Peru-CICRA	-12.56	-70.09	OAC 94721	Mm-2
V. sebifera	ML 013	BL	Brazil-Brasilia	-15.89	-47.86		M-3
V. sebifera	RS 435	BL	Peru-CICRA	-12.56	-70.09	OAC 94673	M-2
V. sebifera	RS 472	BL	Peru-CICRA	-12.56	-70.09	OAC 94710	m
V. sebifera	RS 496	BL	Peru-CICRA	-12.56	-70.09	OAC 94734	m
V. sebifera	RS 503	BL	Peru-Puerto Maldonado	-12.72	-69.24	OAC 94741	Mm-3
V. sebifera	RS 506	BL	Peru-Puerto Maldonado	-12.72	-69.24	OAC 94744	m
V. sebifera	RS 578	BL	Peru-Quincemil	-13.24	-70.78	OAC 94816	m
V. sebifera	RS 582	BL	Peru-Quincemil	-13.24	-70.78	OAC 94820	Mm-2
V. sebifera	RS 587	BL	Peru-Quincemil	-13.24	-70.78	OAC 94825	m
V. sebifera	RS 611	BL	Peru-Manu	-13.05	-71.53	OAC 94849	m
V. sebifera	RS 618	BL	Peru-Manu	-13.05	-71.53	OAC 94856	m
V. sebifera	RS 624	BL	Peru-Manu	-13.05	-71.53	OAC 94862	m
V. sebifera	RS 636	BL	Peru-Manu	-13.05	-71.53		Mm-2
V. sebifera	RS 434	SL	Peru-CICRA	-12.56	-70.09	OAC 94672	m
V. sebifera	RS 443	SL	Peru-CICRA	-12.56	-70.09	OAC 94681	Mm-2
V. sebifera	RS 444	SL	Peru-CICRA	-12.56	-70.09	OAC 94682	Mm-1
V. sebifera	RS 464	SL	Peru-CICRA	-12.56	-70.09	OAC 94702	Mm-1
V. sebifera	RS 513	SL	Peru-Puerto Maldonado	-12.72	-69.24	OAC 94751	m
V. sebifera	RS 529	SL	Peru-Quincemil	-13.24	-70.78	OAC 94767	Mm-2
V. sebifera	RS 533	SL	Peru-Quincemil	-13.24	-70.78	OAC 94771	M-2
V. sebifera	RS 553	SL	Peru-Quincemil	-13.24	-70.78	OAC 94791	m
V. sebifera	RS 565	SL	Peru-Quincemil	-13.24	-70.78	OAC 94803	m
V. sebifera	RS 602	SL	Peru-Quincemil	-13.24	-70.78	OAC 94840	m

Table 4.2 Correspondence analysis of 10 morphological metrics (minimum and maximum) measured for *V. sebifera* and *V. loretensis* taxa. Bolded Pearson correlations (P Corr.) indicate the 4 metrics most significant to the differentiation of taxa (p value < 0.01).

Metric	X-axis		Y-axis	
	P Corr.	Sig. (2-tailed)	P Corr.	Sig.(2-tailed)
Stem width-min.	-0.577	0.002	**0.496****	0.010
Stem width-max	-0.494	0.010	**0.565****	0.003
Petiole length-min	-0.574	0.002	0.126	0.540
Petiole length-max	-0.683	0.000	0.238	0.242
Leaf Petiole width-min	-0.792	0.000	0.334	0.096
Leaf Petiole width-max	-0.792	0.000	0.354	0.076
Lamina Length-min	**-0.966****	0.000	-0.074	0.718
Lamina Length-max	**-0.959****	0.000	-0.098	0.635
Leaf width ¼ length-min	-0.841	0.000	0.020	0.922
Leaf width ¼ length-max	-0.871	0.000	0.197	0.336
Leaf width ½ length-min	-0.914	0.000	-0.134	0.515
Leaf width ½ length-max	**-0.940****	0.000	-0.064	0.756
Leaf width ¾ length-min	-0.873	0.000	-0.287	0.155
Leaf width ¾ length-max	-0.921	0.000	-0.264	0.193
Leaf vein number-min	-0.917	0.000	0.282	0.163
Leaf vein number-max	**-0.927****	0.000	0.220	0.280
Leaf vein space-min	-0.458	0.019	0.067	0.745
Leaf vein space-max	-0.360	0.071	-0.090	0.661
Leaf acumen length-min	-0.133	0.517	**-0.574****	0.002
Leaf acumen length-max	-0.134	0.515	**-0.703****	0.000

Figures

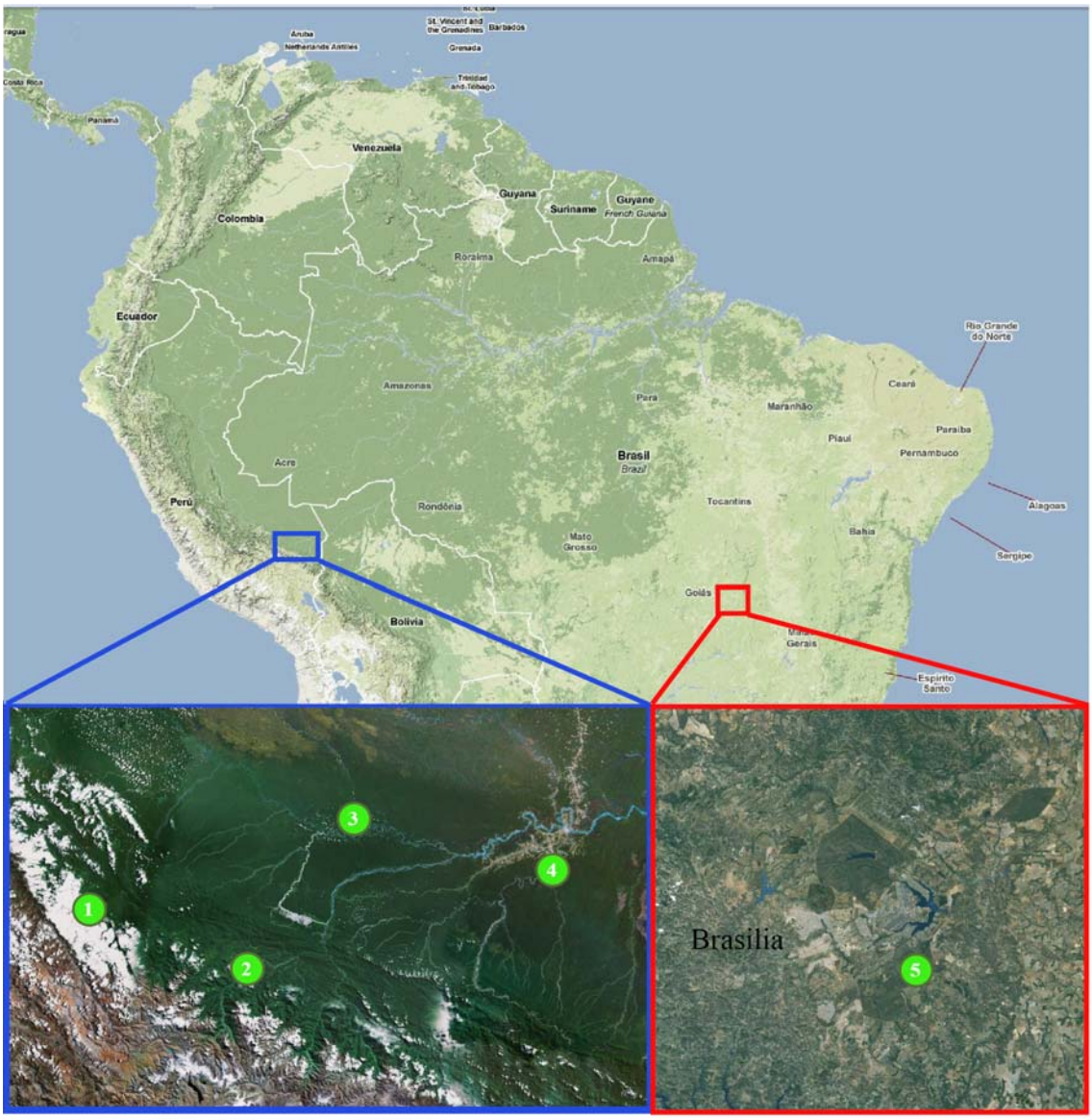

Figure 4.1 Collection locations in Peru (n=4) and Brazil (n=1): 1) Manu Paradise Lodge, 2) Quincemil, 3) CICRA-Los Amigos Biological Reserve, 4) Puerto Maldonado, 5) Brasilia-Reserva Ecologia do IBGE.

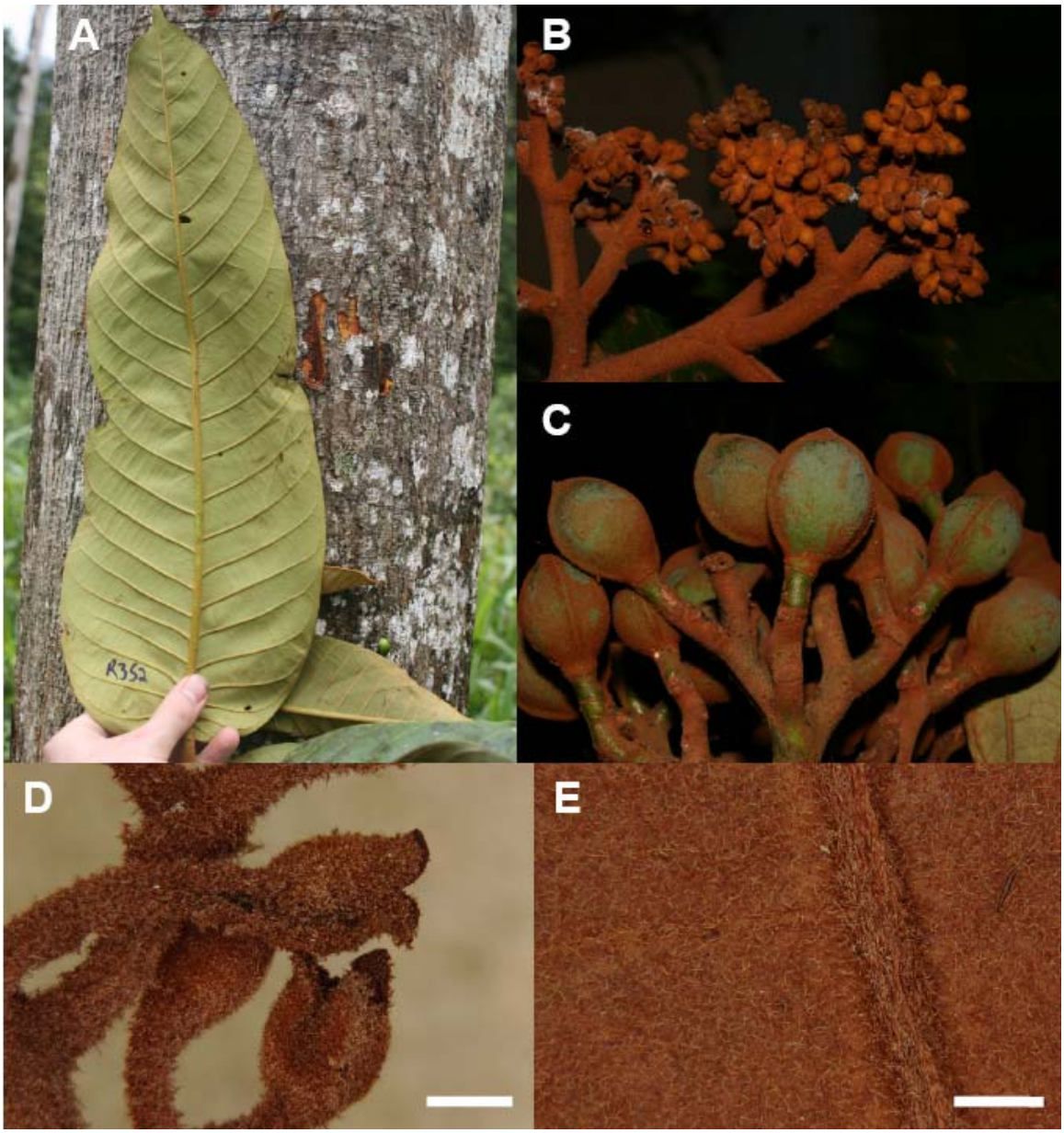

Figure 4.2 *Virola sebifera*-BL (Big Leaf morphotype): A) Bark with leaf showing abaxial surface. B) Mature pistillate inflorescence. C) Nearly mature infructescence showing deciduous red pubescence. D) Close-up of staminate flowers (scale bar= 1 mm). E) Abaxial leaf pubescence and secondary vein (scale bar= 1 mm).

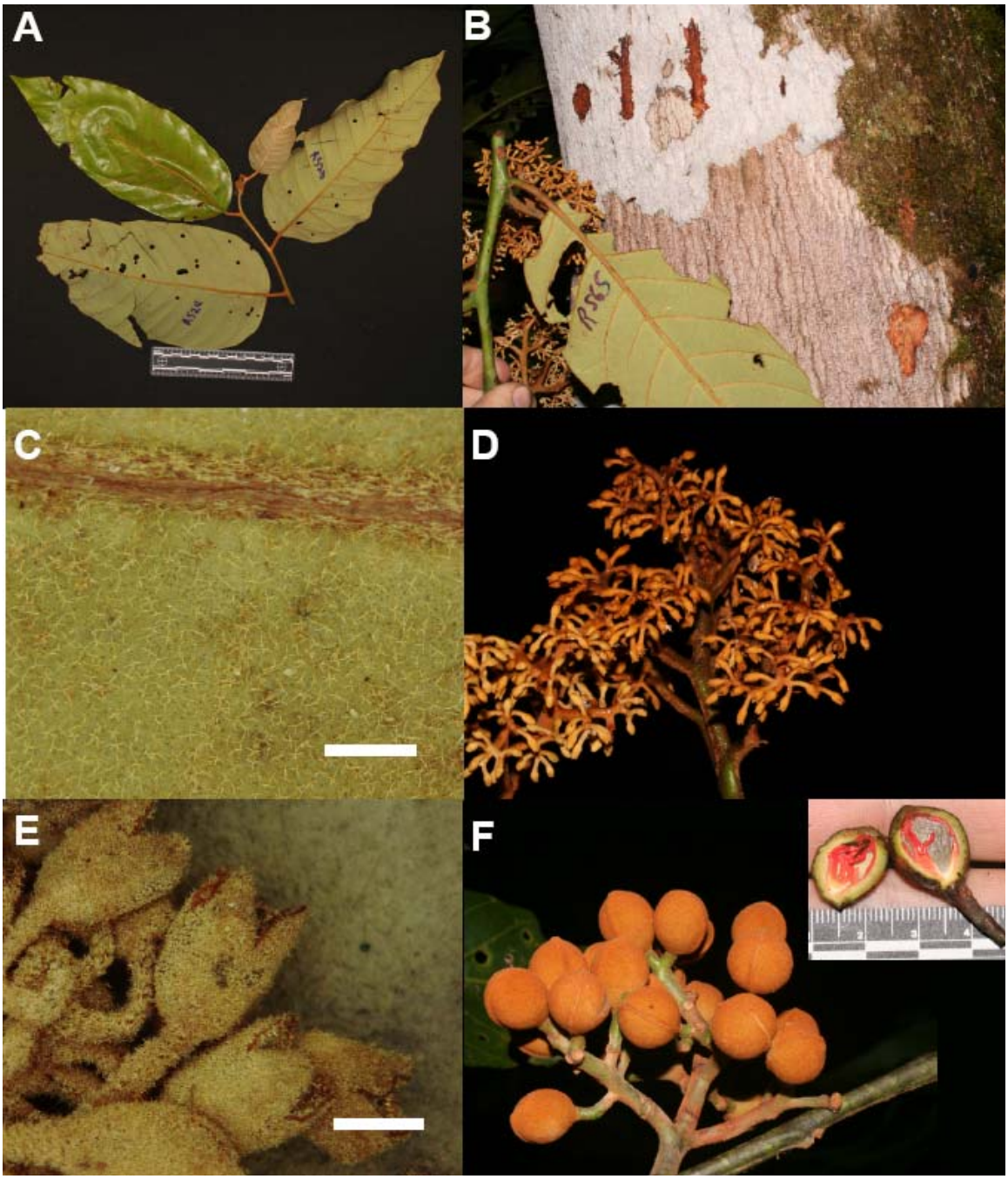

Figure 4.3 *Virola sebifera*-SL (Small Leaf morphotype): A) Leaves with orange-red pubescence (15 cm ruler) B) Bark, leaves and inflorescence of mature tree. C) Abaxial leaf pubescence (scale bar= 1 mm). D) Inflorescence. E) Flowers (scale bar= 1 mm). F) Nearly mature fruits with dissected fruit showing seed testa and aril (inset).

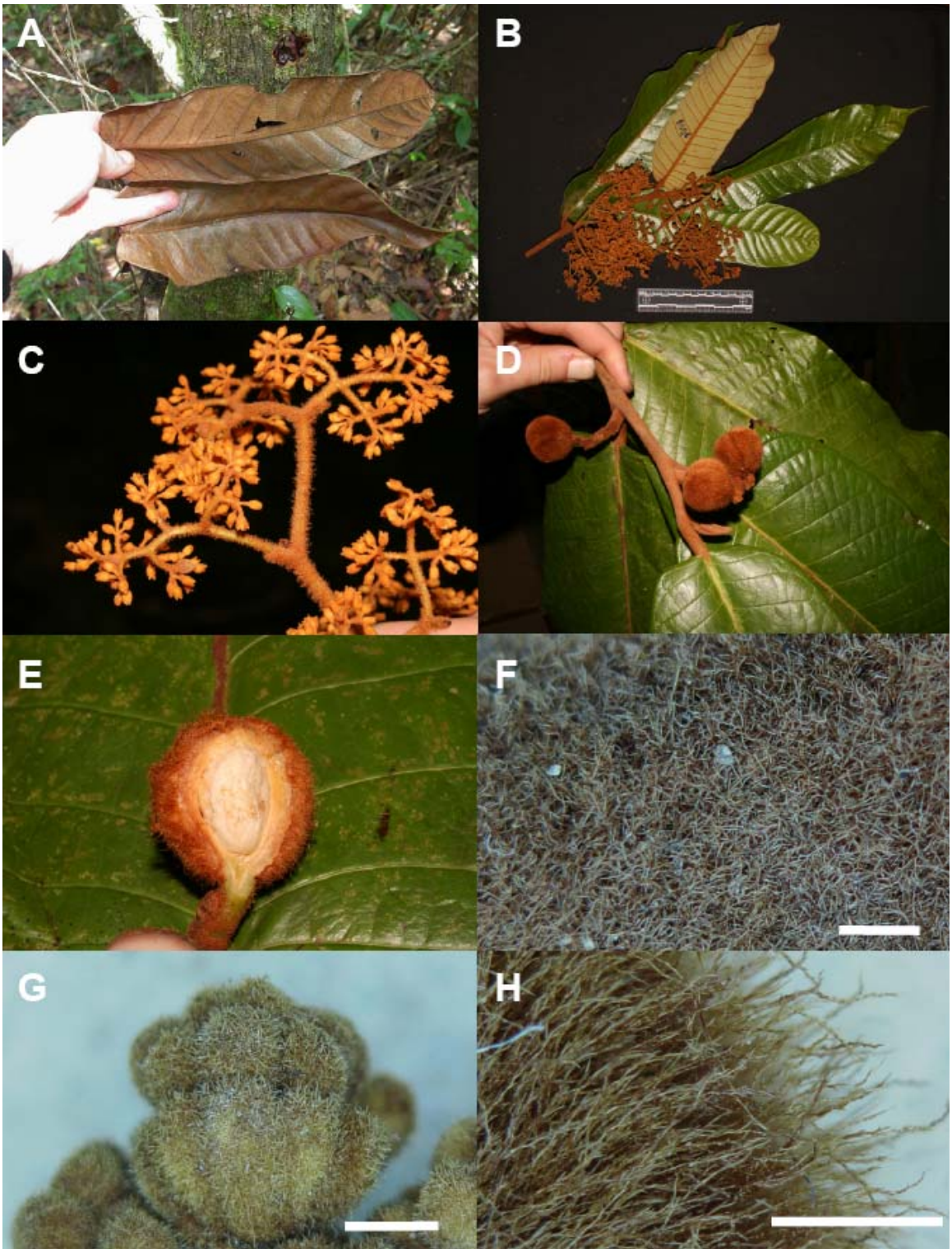

Figure 4.4 *Virola loretensis*-BL (Big Leaf morphotype): A) Bark, deep red latex and dried (dropped) leaves. B) Abaxial, adaxial leaf surfaces of sample with inflorescences (15 cm ruler). C) Close-up of inflorescence. D) Branch with mature fruits. E) Dissected

immature fruit showing dense pubescence. F) Adaxial leaf pubescence. G) Developing flower buds (scale bar=1 mm). H) Close-up of fruit indument (scale bar= 1 mm).

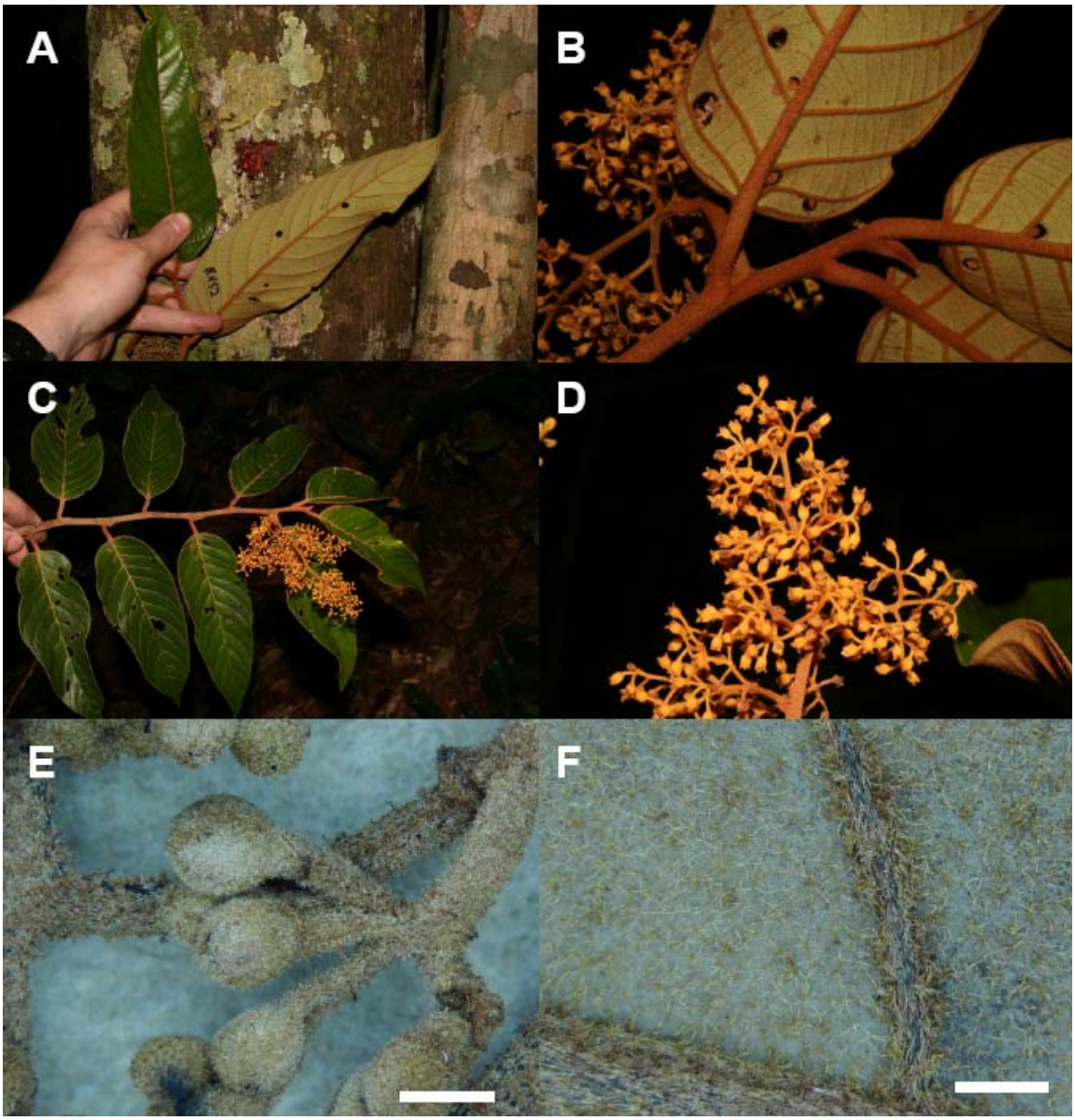

Figure 4.5 *Virola loretensis*-SL (Small Leaf morphotype): A) Bark, leaves and red latex. B) Adaxial leaf surfaces with inflorescence. C) Branch with inflorescence. D) Close-up of mature inflorescence. E) Flower pubescence (scale bar= 1 mm). F) Adaxial leaf pubescence (scale bar= 1 mm).

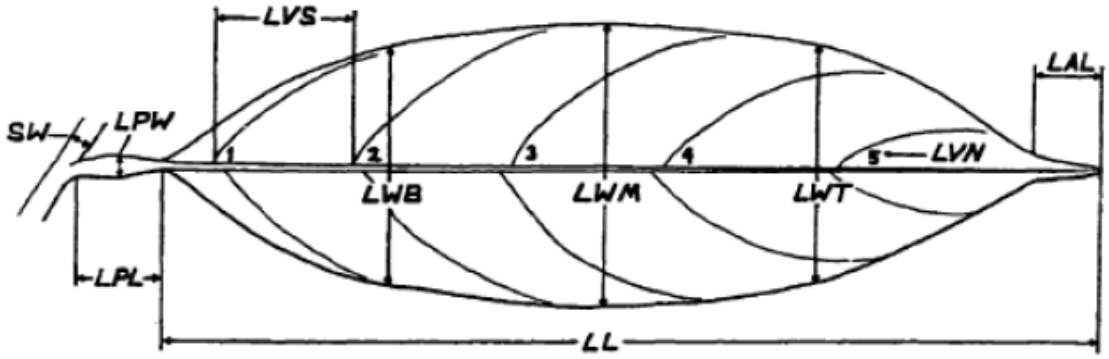

Figure 4.6 Illustration of 10 vegetative morphological metrics measured: SW=Stem Width, LPL=Petiole Length, LPW=Petiole Width, LVS=space between secondary veins, LVN=Vein Number, LL=Lamina Length, LWB= Leaf Width at ¼ length, LWM=Leaf Width at ½ length, LWT=Leaf Width at ¾ length, LVN=Leaf Vein Number, LAL=Leaf Acumen Length. A maximum and minimum value of each metric was recorded for every specimen measured.

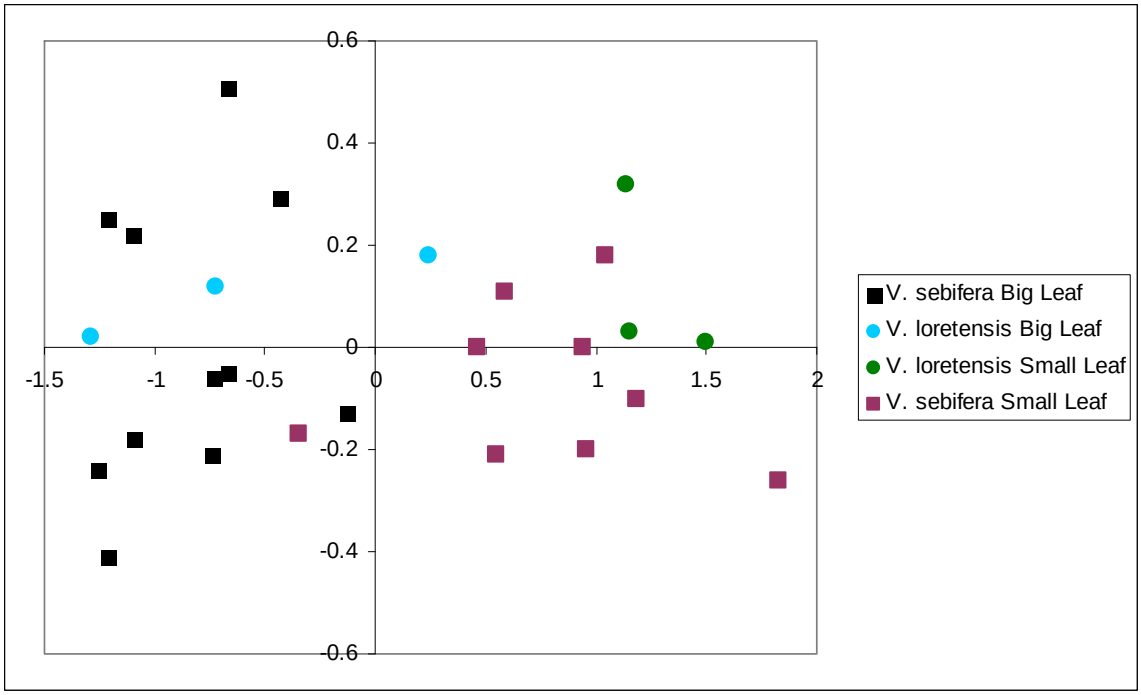

Figure 4.7 NMS ordination of quantitative leaf metrics of *V. sebifera* and *V. loretensis* morphotypes.

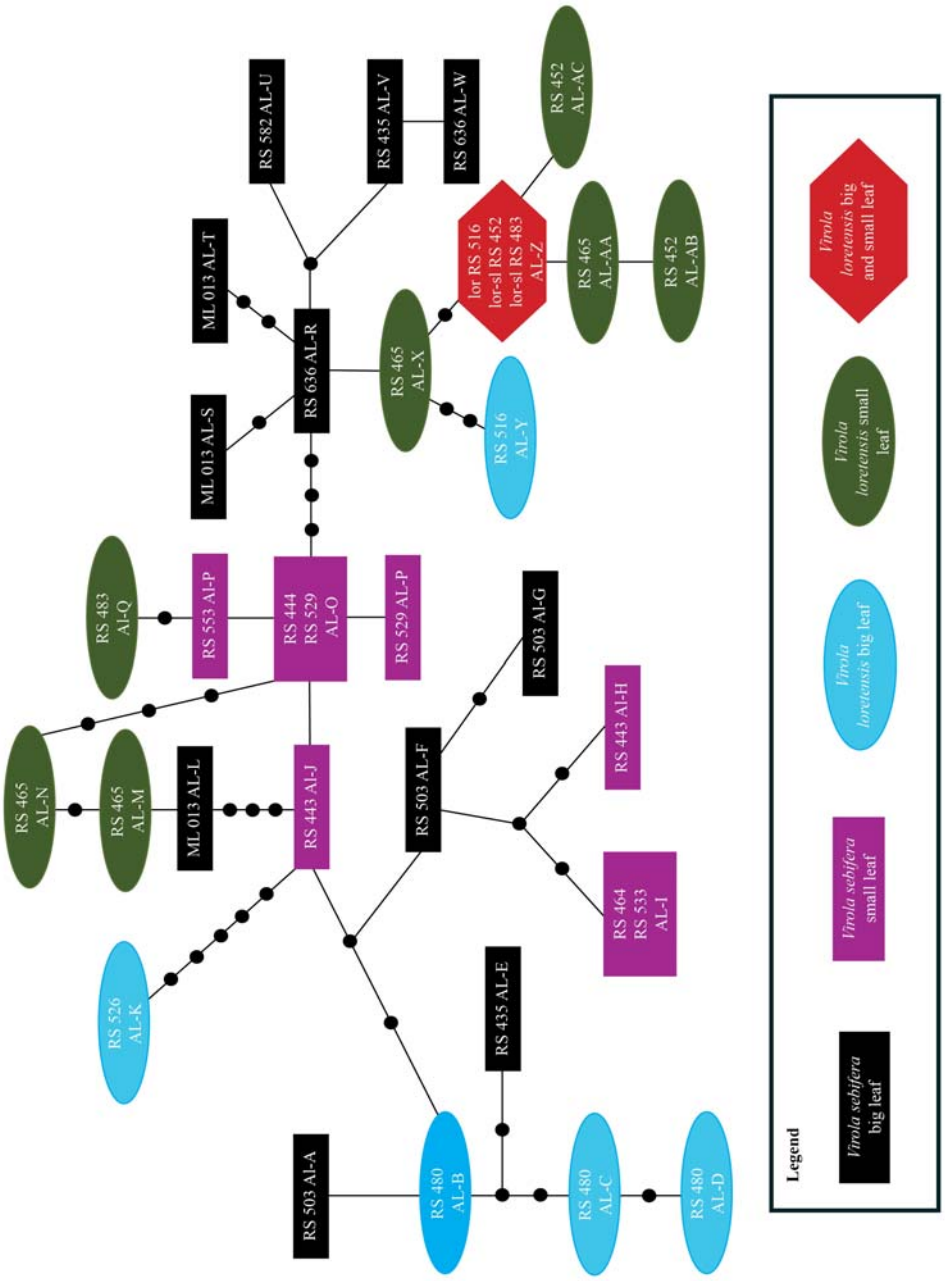

Figure 4.8 Haplotype network of 35 sequences from *V. sebifera* and *V. loretensis* clones. Legend indicates the shapes and colours used for each of the morphotypes and the one shared allele.

Literature Cited

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., & Weightman, A. J. 2005. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Applied and Environmental Microbiology* **71**: 7724-7736.

Ashton, P. S. 1969. Speciation among tropical forest trees: some deductions in the light of recent evidence. *Biological Journal of the Linnean Society* **1**: 155–196.

Asif, M. J., & Cannon, C. H. 2005. DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*). *Plant Molecular Biology Reporter* **23**: 185–192.

Clement, M., Posada, D., & Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1659.

Degen, B., Caron, H., Bandou, E., Maggia, L., Chevallier, M. H., Leveau, A., & Kremer, A. 2001. Fine-scale spatial genetic structure of eight tropical tree species as analysed by RAPDs. *Heredity* **87**: 497–507.

Dexter, K. G., Pennington, T. D., & Cunningham, C. W. 2010. Using DNA to assess errors in tropical tree identifications: How often are ecologists wrong and when does it matter? *Ecological Monographs* **80**: 267–286.

Dick, C. W., Abdul-Salim, K., & Bermingham, E. 2003. Molecular systematic analysis reveals cryptic tertiary diversification of a widespread tropical rain forest tree. *American Naturalist* **162**: 691–703.

Dickie, I. A. 2010. Insidious effects of sequencing errors on perceived diversity in molecular surveys. *New Phytologist* **188**: 916-918.

Dodson, C. H., & Gentry, A. H. 1991. Biological extinction in Western Ecuador. *Annals of the Missouri Botanical Garden* **78**: 273-295.

Duminil, J., Caron, H., Scotti, I., CAZAL, S. O., & Petit, R. J. 2006. Blind population genetics survey of tropical rainforest trees. *Molecular Ecology* **15**: 3505–3513.

Gentry, A. H. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* **69**: 557–593.

Gentry, A. H. 1988. Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 156-159.

Gentry, A. H. 1992. Tropical forest biodiversity: distributional patterns and their conservational significance. *Oikos* **63**: 19–28.

Gottlieb, O., Loureiro, A., Dos Santo Carneiro, M., & Da Rocha, A. 1973. Distribution of diarylpropanoids in Amazonian *Virola* species. *Phytochemistry* **12**: 1830.

Golden, J. L., & Bain, J. F. 2000. Phylogeographic patterns and high levels of chloroplast DNA diversity in four *Packera* (Asteraceae) species in southwestern Alberta. *Evolution* **54**: 1566–1579.

Govaerts, R. 2003. How many species of seed plants are there?-a response. *Taxon* **52**: 583-584.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series*. (pp. 95–98).

Hamrick, J. L., Murawski, D. A., & Nason, J. D. 1993. The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Plant Ecology* **107**: 281–297.

Jaramillo, T. S., Muriel, P., Rodrigues, W. A., & Balslev, H. 2000. Myristicaceae novelties from Ecuador. *Nordic Journal of Botany* **20**: 443–447.

Joly, S., Starr, J. R., Lewis, W. H., & Bruneau, A. 2006. Polyploid and hybrid evolution in roses east of the Rocky Mountains. *American Journal of Botany* **93**: 412-425.

Jumpponen, A., & Johnson, L. C. 2005. Can rDNA analyses of diverse fungal communities in soil and roots detect effects of environmental manipulations–a case study from tallgrass prairie. *Mycologia* **97**: 1177-1194.

Kamiya, K., Gan, Y. Y., Lum, S. K. Y., Khoo, M. S., Chua, S. C., & Faizu, N. N. H. 2010. Morphological and molecular evidence of natural hybridization in *Shorea* (Dipterocarpaceae). *Tree Genetics & Genomes*. Retrieved from http://www.springerlink.com/index/10.1007/s11295-010-0332-8

Kruskal, J. B. 1964. Nonmetric multidimensional scaling: a numerical method. *Psychometrika* **29**: 115–129.

Lahr, D. J., & Katz, L. A. 2009. Reducing the impact of PCR-mediated recombination in molecular evolution and environmental studies using a new-generation high-fidelity DNA polymerase. *BioTechniques* **47**: 351-357.

Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H., & Varotto, C. 2008. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* **24**: 727–745.

Minchin, P. R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. *Plant Ecology* **69**: 89–107.

Panchal, M. 2007. The automation of nested clade phylogeographic analysis. *Bioinformatics* **23**: 509-510.

Pascal, J. P., & Pelissier, R. 1996. Structure and floristic composition of a tropical evergreen forest in south-west India. *Journal of Tropical Ecology* **12**: 191–214.

Pitman, N. C., Mogollón, H., Dávila, N., Ríos, M., García-Villacorta, R., Guevara, J., Baker, T. R., Monteagudo, A., Phillips, O. L., Vásquez-Martínez, R., & others. 2008. Tree community change across 700 km of lowland Amazonian forest from the Andean foothills to Brazil. *Biotropica* **40**: 525–535.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2002. A comparison of tree species diversity in two upper Amazonian forests. *Ecology* **83**: 3210–3224.

Pitman, N. C., Terborgh, J. W., Silman, M. R., Núñez V, P., Neill, D. A., Cerón, C. E., Palacios, W. A., & Aulestia, M. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* **82**: 2101–2117.

Poinar, H. N. 1998. Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281**: 402-406.

Posada, D., & Crandall, K. A. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**: 37–45.

Prance, G. T. 1972. Ethnobotanical notes from Amazonian Brazil. *Economic Botany* **26**: 221–237.

ter Braak, C. 1998. *Canoco 4 Centre for Biometry.* Wageningenm, The Netherlands

Richardson, J. E. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* **293**: 2242-2245.

Rodrigues, W. A. 2002. Notas taxonômicas sobre Myristicaceae neotropicais. *Acta Biológica Paranaense* **31**: 71-77.

Rodrigues, W. A. 1980. Revisão taxonômica das espécies de *Virola* Aublet (Myristicaceae) do Brasil. *Acta Amazonica* **(suplemento)**: 1-127.

Rodrigues, W. A. 1989. Two new neotropical species of *Compsoneura* (Myristicaceae). *Brittionia* **4**: 160-163.

Rohlf, F. J. 2006. *tpsDig version 2.10, Department of Ecology and Evoluation*. State University of New York at Stony Brook, New York.

Sabatier, D. 1997. Description et biologie d'une nouvelle espèce de *Virola* (Myristicaceae) de Guyane. *Adansonia, Série* **3**: 273–278.

Sasaki, Y., Miyoshi, D., & Sugimoto, N. 2006. Effect of molecular crowding on DNA polymerase activity. *Biotechnology Journal* **1**: 440–446.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W., & Le Thomas, A. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* **142**: 125–186.

Schultes, R. E., & Raffauf, R. F. 1990. *The Healing Forest*. Portland, Oregon: Dioscorides Press.

Scotland, R. W., & Wortley, A. H. 2003. How many species of seed plants are there? *Taxon* **52**: 101–104.

Southcott, L., & K.L. Ostevic. 2011. Bromeliad population genetics reveals species cohesion agaist the odds. *Molecular Ecology* **20**: 3081-3083

Smith, A. C. 1937. The American species of Myristicaceae. *Brittonia* **2**: 393-510.

Speiss, A., T. Mueller, & Ivell, R. 2004. Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of Taq polymerase by the disaccharide trehalose. *Clinical chemistry* **50**: 1256-1259.

Templeton, A. R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* **7**: 381–397.

Templeton, A. R., Crandall, K. A., & Sing, C. F. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619-633.

ter Braak, C. 1998. *Canoco 4 Centre for Biometry.* Wageningenm, The Netherlands.

Thomas, W. W. 1999. Conservation and monographic research on the flora of Tropical America. *Biodiversity and Conservation* **8**: 1007–1015.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.

Whitney, K.D., Randell, R.A., & L.H. Rieseberg. 2010. Adaptive introgression of abioatic tolerance traits in the sunflower Helianthus annuus. *New Phytologist* **187**: 230-239.

Zimmerman, S. B., & Harrison, B. 1987. Macromolecular crowding increases binding of DNA polymerase to DNA: an adaptive effect. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 1871-1875.

CONCLUDING STATEMENTS

This research represents a long overdue molecular and morphological revision of two of the most elemental genera of neotropical forests. The following remarks summarize the implications of my findings as they relate to future research and conservation efforts in the tropics.

This work is also among the first tests of DNA barcoding in a diverse group of plants (Spooner 2009). Most plant DNA barcoding projects to date have focused on local/regional floras and have included few congeneric species and have rarely included multiple collections from across a species' range (Lahaye et al. 2008, Kress et al. 2008). This approach could be likened to that of assembling a flora, or regional checklist, as compared to that of a monograph that focuses on taxonomy. There are a number of reasons why a "flora-like" approach has been applied with plants. Firstly, plant DNA barcoders have been somewhat hindered by low mitochondrial and chloroplast DNA sequence divergence and have spent much time converging on a set of loci that are to be employed as the core plant DNA barcodes (Hollingsworth et al 2011). Furthermore, sister-species are often not sympatric so little emphasis has been put on thorough taxon sampling.

Although most members of *Compsoneura* could be indentified by the trnH-psbA region (Chapter 2 of this thesis), the low levels of plastid sequence variability in *Virola* currently makes molecular identification difficult (Chapter 3 of this thesis) and plant DNA barcoding efforts are likely to be hindered by haplotype sharing among species (Hollingsworth et al. 2011). Consequently, species rich groups such as *Virola* may require multiple rapidly evolving loci for molecular discrimination. Similar levels of

variation within and among described species in Myristicaceae also confounds molecular species delimitation via barcodes. In my opinion, future plant DNA barcoding efforts should take on a molecular-monographic approach where described taxa are tested with dense taxon sampling across a species' known range and rigorously combined with additional data sources (morphological, edaphic, environmental, ethnobotanical, etc) for an integrative taxonomic approach.

Thomas (1999) found that 29% of species included in recent monographs were newly described, causing him to estimate that the South American flora contains approximately 90,000 species of seed plants. A total of 4 provision new species and one resurrection are presented within this thesis entirely from collections made by myself. I find this highly significant as I collected in only 7 locations, most of which were amongst the most active research stations in the Amazon. Furthermore, these new species were often among the most abundant trees in these forests (R. Steeves and J. Janovec, unpublished data). In my investigations of the infraspecific morphological and molecular variation in *Compsoneura capitellata*, *Virola sebifera*, and *V. loretensis* two provisional new species and possibly four or more, were found within these widespread species. It is likely that a great deal of additional species remain to be documented in these and other genera of Myristicaceae, as well as other tropical plant families. If future investigations of infraspecific molecular diversity discover a similar rate of cryptic species we may have to recalculate our estimates of neotropical plant diversity. It is likely that we may find significantly more than Thomas' (1999) estimate as a study of herbarium collection patterns in Peru done by Tobler et al.(2007) found that the vast majority of collections were made near very few localities that were close to roads and cities, leaving much of

the Peruvian rainforest unexplored by botanists.

According to most estimates, well over half of our medicinal drugs come either directly or indirectly from plants. With roughly a quarter to a third of the world's plant diversity, the neotropics offers enormous prospects for the discovery of new drugs for implementation in Western medicinal practice. However, this great diversity can also be a hindrance to bio-prospecting in the region. Species, genera and even families are often mis-identified and taxon identification are often inconsistent even with the same collector (R. Steeves, personal observations). This means that even if desirable properties are found in a particular plant extract, subsequent expeditions often collect a different species with correspondingly dissimilar, and potentially ineffective, chemical constitution. When DNA identification is possible in tropical plants it will undoubtedly help collectors ensure their identifications are consistent.

My research forms a much needed foundation for future investigations aimed at understanding the neotropical Myristicaceae and the Amazon as a whole. This work represents only the beginning of a modern systematic revision of the genera *Compsoneura* and *Virola*. To these ends, this investigation generated about 600 novel collections of Myristicaceae from the Northwest Amazon and produced about as many DNA nucleotide sequences from coding and non-coding regions. Future research should attempt to resolve evolutionary relationships amongst the genera of the Myristicaceae, which has remained largely unresolved due to low levels of nucleotide variation and the difficulty of extracting and amplifying DNA from samples. Although low levels of sequence divergence were found in this study, the low copy nuclear and non-coding chloroplast regions employed in this study show increased sequence divergence

compared to those used in pervious investigations (Sauquet et al 2003). Additionally, the difficulty of extracting DNA from both fresh and archived (Herbarium) specimens was found to be greatly alleviated by the use of PTB (Phenacylthiazolium bromide) and the use of PCR enhancers such as trehalose and polypropylene glycol. Future investigations should seek to determine the causes of chloroplast haplotype sharing among species and illuminate phylogeographic patterns in medicinal and edible members of the Myristicaceae as well as other families as the possibility exists that humans played a pivotal role in the determination of tree distributions in the Amazon basin.

Literature Cited

Hollingsworth, P. M., Graham, S. W., & Little, D. P. 2011. Choosing and using a plant DNA barcode. *PloS one* **6**: e19254.

Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjur, O., & Bermingham, E. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences* **106**: 18621-18626.

Lahaye, R., Van Der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T. G., & Savolainen, V. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences* **105**: 2923-2928.

Spooner, D. M. 2009. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany* **96**: 1177-1189.

Thomas, W. W. 1999. Conservation and monographic research on the flora of Tropical America. *Biodiversity and Conservation* **8**: 1007–1015.

Tobler, M., Honorio, E., Janovec, J., & Reynel, C. 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: an example of two neotropical plant families (Moraceae and Myristicaceae) in Peru. *Biodiversity and Conservation* **16**: 659-677.

APPENDICES

Appendix 1

Alignment 1.1

Concatenated DNA nucleotide sequence alignment 3-loci,1464 characters, and 17 taxa used in Bayesian and Maximum parsimony analyses in chapter 1. 50 base pairs per line.

trnH-psba 1-342, AGT1 343-1052, AT103 1053-1464

```
C_atopa_1374   TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_cap_RS_527   TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTCGATTACTAGT
C_cap_RS_551   TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTCGATTACTAGT
C_cap_835      TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_debilis_7209 TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_excelsa_671  TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_mexicana_362 TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_mexicana_720 TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_mutisii_914  TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_mutisii_1295 TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_sprucei_817  TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_sprucei_887  TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTATAGATTACTAGT
C_ulei_88      TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
C_ulei_6192    TCCGCCCCTTGTCCTTTCT-----AAAGAAAAAAATTTTAGATTACTAGT
I_juru_451     TCCGCCCCTTGTCTTTTCTTTTCTAAAGACAAAAATTTTAGATTACTAGT
I_laev_460     TCCGCCCCTTGTCTTTTCTTTTCTAAAGACAAAAATTTTAGATTACTAGT
O_parvi_RS_598 TCCGCCCCTTGTCTTTTCG-----AAAGACTAAAATTTTAGATTACTAGT

C_atopa_1374   CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGACAATA
C_cap_RS_527   CTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATCGACAATA
C_cap_RS_551   CTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATCGACAATA
C_cap_835      TTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGACAATA
C_debilis_7209 CTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA----CAATA
C_excelsa_671  CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----CAATA
C_mexicana_362 CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----CAATA
C_mexicana_720 CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----CAATA
C_mutisii_914  CTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----CAATA
C_mutisii_1295 CTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----CAATA
C_sprucei_817  CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTTGAAAA----CAATA
C_sprucei_887  CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----CAATA
C_ulei_88      CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----CAATA
C_ulei_6192    CTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----CAATA
I_juru_451     CTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGACAATA
I_laev_460     CTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGACAATA
O_parvi_RS_598 CTTTCTTATTTTTTTCATACTAATTTCTACCCTTTCTAAAATTTACAATA

C_atopa_1374   GGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCAAATC
C_cap_RS_527   GGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCAAATC
C_cap_RS_551   GGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCAAATC
C_cap_835      GGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGATCAGTTCAAATC
C_debilis_7209 GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGTTAAAATC
C_excelsa_671  GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_mexicana_362 GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
```

```
C_mexicana_720 GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_mutisii_914  GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_mutisii_1295 GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_sprucei_817  GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_sprucei_887  GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_ulei_88      GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
C_ulei_6192    GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
I_juru_451     GGAAAAAATGCATTTTTGGAATGTACATAAACTGAAGATCAGTTAAAATA
I_laev_460     GGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAAAATA
O_parvi_RS_598 GGAAAAAAT----TTTTGGAATGTACATAAACAGAAGATCAGTTAAAAAC

C_atopa_1374   AAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATTCATA
C_cap_RS_527   AAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATTCATA
C_cap_RS_551   AAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATTCATA
C_cap_835      AAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATTCATA
C_debilis_7209 AAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_excelsa_671  AAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_mexicana_362 AAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_mexicana_720 AAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_mutisii_914  AAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATTCATA
C_mutisii_1295 AAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_sprucei_817  AAAAAAAAAAA----GGTATGATGTTCGATCATGAACCAAATAATGAATA
C_sprucei_887  AAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_ulei_88      AAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATTAATA
C_ulei_6192    AAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATTAATA
I_juru_451     AAAATAAAAAAAAAAGGTATGATGTTCGATCATGAA-CAAA-AATGAATA
I_laev_460     AAAAAAAA-------GGTATGATGTTCGATCATGAA-CAAA-AATGAATA
O_parvi_RS_598 AAAAAAA--------GGTATGATGTTCGATCCTGAACCAACTAATTAATA

C_atopa_1374   TTTTCTGAAATTGAAAAAAAAA-TCTTATGTGAGTAAACCACTACTGAAC
C_cap_RS_527   TTTTCTGAAATTGAAAAAACAAATCTTATGTGAGTAAACCACTACTGAAC
C_cap_RS_551   TTTTCTGAAATTGAAAAAACAAATCTTATGTGAGTAAACCACTACTGAAC
C_cap_835      TTTTCTGAAATTGAAAAAAAAA-TCTTATGTGAGTAAACCACTACTGAAC
C_debilis_7209 TTTTCTGAAATTGAAAAAAAAA-TCTTATGTGAGTAAACCACTACTGAAC
C_excelsa_671  TTTTCTGAAATGGAAAAAAAA--TCTTATGTGAGTAAACCACTACTGAAC
C_mexicana_362 TTTTCTGAAATGGAAAAAAAAA-TCTTATGTGAGTAAACCACTACTGAAC
C_mexicana_720 TTTTCTGAAATGGAAAAAAAA--TCTTATGTGAGTAAACCACTACTGAAC
C_mutisii_914  TTTTCTTAAATGGAAAAAAAA--TCTTATGTGAGTAAACCACTACTGAAC
C_mutisii_1295 TTTTCTTAAATGGAAAAAAAA--TCTTATGTGAGTAAACCACTACTGAAC
C_sprucei_817  TTTTCTTAAATGGAAAAAAAA--TCTTATGTGAGTAAACCACTACTGAAC
C_sprucei_887  TTTTCTTAAATGGAAAAAAAA--TCTTATGTGAGTAAACCACTACTGAAC
C_ulei_88      TTTTCTTAAATTGAAAAAAAAA-TCTTATGTGAGTAAACCACTACTGAAC
C_ulei_6192    TTTTCTTAAATTGAAAAAAAAA-TCTTATGTGAGTAAACCACTACTGAAC
I_juru_451     TTTTCTTAAATTGAAAAAAAAAAATCTTATGTGAGTAAACCACTACTGAAC
I_laev_460     TTTTCTTAAATTGAAAAACAAAATCTTATGTGAGTAAACCACTACTGAAC
O_parvi_RS_598 TTTTTTAAAATTGAAAAAAAAAAATCTTATGTGAGTAAACCACTACTGAAA

C_atopa_1374   CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_cap_RS_527   CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_cap_RS_551   CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_cap_835      CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_debilis_7209 CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_excelsa_671  CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_mexicana_362 CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_mexicana_720 CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_mutisii_914  CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_mutisii_1295 CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
```

```
C_sprucei_817  CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_sprucei_887  CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_ulei_88      CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
C_ulei_6192    CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
I_juru_451     CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
I_laev_460     CAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCGTATACA
O_parvi_RS_598 CAGATCAATACCCATGGGTATTGATCTGATCTTTCAATGACTCGTATACA

C_atopa_1374   CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACANNNNNNNN
C_cap_RS_527   CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_cap_RS_551   CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_cap_835      CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACANNNNNNNN
C_debilis_7209 CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_excelsa_671  CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACANNNNNNNN
C_mexicana_362 CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_mexicana_720 CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_mutisii_914  CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_mutisii_1295 CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_sprucei_817  CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_sprucei_887  CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACATCTTTCTA
C_ulei_88      CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACANNNNNNNN
C_ulei_6192    CTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACANNNNNNNN
I_juru_451     CTAATACCGAAGTATTAGCCATTTGTTGATAGAGCTTCAACATCTTTCTA
I_laev_460     CTAATACCGAAGTATTAGCCATTTGTTGATAGAGCTTCAACATCTTTCTA
O_parvi_RS_598 CTAATACCGAAGCATTNGCCATTTGTTGATAGAGCTTCAACATCTTTCTA

C_atopa_1374   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_cap_RS_527   TGCCTACTGGCATGGGAATCGTATGCGCAAGTCCAAAAGCAATAGAAGCA
C_cap_RS_551   TGCCTACTGGCATGGGANTCGTATGCGCAAGTCCAAAAGCAATAGAAGCA
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 TGCCTACTGGNATGGGAATCGTATGCGCAAGTCCAAAAGCAATGGAAGCA
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 TGCCTACTGGTATGGGAATCGTATGCGCAAGTCCAAAAGCAATGGAAGCA
C_mexicana_720 TGCCTACTGGTATGGGAATCGTATGCGCAAGTCCAAAAGCAATGGAAGCA
C_mutisii_914  TGCCTACTGGTATGGGAATCGTATGCGCAAGTCCAAAAGCAATGGAAGCA
C_mutisii_1295 TGCCTACTGGTATGGGAATCGTATGCGCAAGTCCAAAAGCAATGGAAGCA
C_sprucei_817  TGCCTACTGGTATGGGAATCGTATGCGCAAGTCCAAAAGCAATGGAAGCA
C_sprucei_887  TGCCTACTGGTATGGGAATCGTATGCGCAAGTNCAAAAGCAATGGAAGCA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     TGCCTACTGGTATGGGAATCGTATGCGCAAGTCCAAAAGCAATAGAAGCA
I_laev_460     TGCCTACTGGTATGGGANTCGTATGCGCAAGTCCAAAAGCAATAGAAGCA
O_parvi_RS_598 TGCCTACTGGCATGGGAATCGTATGCGCAAGTCCAAAAGCAATAGAAGCA

C_atopa_1374   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_cap_RS_527   TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_cap_RS_551   TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACNT
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_mexicana_720 TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_mutisii_914  TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGNCTGGANAGACTACCT
C_mutisii_1295 TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_sprucei_817  TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_sprucei_887  TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
```

```
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
I_laev_460     TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT
O_parvi_RS_598 TCCAAGACTGCTAAATCAGTCAGAGTATTCTTTGACTGGAAAGACTACCT

C_atopa_1374   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_cap_RS_527   GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCTTCTATACAAC
C_cap_RS_551   GAAGTTCTATAANTTGGGAACATATTGGCCCTACACACCTTCTATACAAC
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 GAAGTTCTATAAGCTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_mexicana_720 GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_mutisii_914  GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_mutisii_1295 GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_sprucei_817  GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_sprucei_887  GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCCTCTATACAAC
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCTTCTATACAAC
I_laev_460     GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCTTCTATACAAC
O_parvi_RS_598 GAAGTTCTATAAGTTGGGAACATATTGGCCCTACACACCTTCTATACAAC

C_atopa_1374   NNNNGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTGTGGAAGGACTT
C_cap_RS_527   TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTGTGGAAGGACTT
C_cap_RS_551   TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTGTGGAAGGACTT
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_mexicana_720 TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_mutisii_914  TNTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_mutisii_1295 TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_sprucei_817  TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_sprucei_887  TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTTTGGAAGGACTT
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTGTGGAAGGACTT
I_laev_460     TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTGTGGAAGGACTT
O_parvi_RS_598 TTTTGTATGGCCTGAGAGCAGCTTTGGATCTCATATTTGTGGAAGGACTT

C_atopa_1374   GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_cap_RS_527   GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_cap_RS_551   GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 GAAAATGTGTTTGNCAGACACAAACGTTTGGGCAAAGCAACAAGGTAATN
C_mexicana_720 GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_mutisii_914  GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_mutisii_1295 GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_sprucei_817  GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_sprucei_887  GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
I_laev_460     GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATA
```

```
O_parvi_RS_598 GAAAATGTGTTTGACAGACACAAACGTTTGGGCAAAGCAACAAGGTAA--

C_atopa_1374   GATTTTGGAAAGATGTAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_cap_RS_527   GATTTTGGAAAGATGTAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_cap_RS_551   GATTTTGGAAAGATGTAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 GATTTTGGAAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 GATTTTGGAAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_mexicana_720 GATTTTGGAAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_mutisii_914  GATTTTGGGAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_mutisii_1295 GATTTTGGGAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_sprucei_817  GATTTTGGAAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_sprucei_887  GATTTTGGAAAGATGCAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     GATTTTGGAAATATGTAGTGCTGAATGCTAATATTGAGCACTTGATGTCC
I_laev_460     GATTTTGGAAAGATGTAGTGCTGAATGCTAATATTGAGCACTCGATGTCC
O_parvi_RS_598 ----------AGATGTAGTACTGAATGCTAATATTGAGCACTCGATGTCC

C_atopa_1374   ATAGTAAGAGAGGNACACCATTTATTTATTGGGAACTACCCTGGATGGGN
C_cap_RS_527   ATAGTAAGAGAGGNACACCATTTATTTATTGGGAACTACCCTGGATGGGA
C_cap_RS_551   ATAGTAAGAGAGGGACACCATTTATTTATTGGGAACTACCCTGGATGGGA
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 ATAGTAAGAGAGGCACATCATTTATTTATTGGGAACTACCCTGGATGGGA
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTAACCTGGATGGGA
C_mexicana_720 ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTAACCTGGATGGGA
C_mutisii_914  ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTAACCTGGATGGGA
C_mutisii_1295 ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTAACCTGGATGGGA
C_sprucei_817  ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTAACCTGGATGGGA
C_sprucei_887  ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTAACCTGGATGGGA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTGCCCTGGATGGGT
I_laev_460     ATAGTAAGAAAGGCACACCATTTATTTATTGGGAACTGCCCTGGATGGGT
O_parvi_RS_598 ATAGTAAGAGAGGCACACCATTTATTTATTGGGAACTACACTGGATGAGC

C_atopa_1374   GGGAAATATT-CCAAAAATTTCAGCAGTGGGACAGTCCTAGACACTGATA
C_cap_RS_527   GGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_cap_RS_551   GGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 GGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATG
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 GGGGAATATC-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_mexicana_720 GGGGCATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_mutisii_914  GGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_mutisii_1295 GGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_sprucei_817  GGGGAATATT-CCAAAGATTTCAGCAGTGGGACAATCCTAGACACTGATA
C_sprucei_887  GGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTNGACACTGATA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     GGGGAATATT-CAAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATT
I_laev_460     GGGGAATATT-CAAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATT
O_parvi_RS_598 AGGGAATATT-CCAAAAATTTCAGCAGTGGGACAATCCTAGACACTGATA

C_atopa_1374   TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATGCTTTTTGAACTN
```

```
C_cap_RS_527    TTGNGAACTTTTTCCCCAGTGAATGCTGACCATATATGCTTTTTGAACTT
C_cap_RS_551    TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATGCTTTTTGAACTT
C_cap_835       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209  TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATGTTTTTTGAGCTT
C_excelsa_671   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362  TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATCTTTTTTGAACTT
C_mexicana_720  TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATCTTTTTTGAACTT
C_mutisii_914   TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATCTTTTTTGAACTT
C_mutisii_1295  TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATCTTTTTTGAACTT
C_sprucei_817   TTGAGAACTTTTTCCCCAGTGAATGCTGACCATATATGTTTTTTGAACTT
C_sprucei_887   TTGAGAACTTTTTCCCCAGTGANTGCTGACCNTATATNTTTTTTGAACTT
C_ulei_88       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451      TTGAGAACTTTTGCCCCAGTGACTGTTGACCATATATGTTTTTTGAACTT
I_laev_460      TTGAGAACTTTTGCCCCAGTGACTGTTGACCATATATGTTTTTTGAACTT
O_parvi_RS_598  TTGAGAACTTTTGACCCGGTGAATGTTGACCATATATGTTTTTTGAGCTT

C_atopa_1374    CCTATTTATAATCCTCANAA-NTCCTCCAAACATCAGCTAAAATCATCAT
C_cap_RS_527    CCTATTTATAATCCTCAAAA-CTCCTCCAAACATCAGCTAAAATCATCAT
C_cap_RS_551    CCTATTTATAATCCTCAAAA-CTCCTCCAAACATCAGCTAAAATCATCAT
C_cap_835       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209  TCTATTTATAATCCTCAAAA-TTCCTCCAAACAGCAGCTAAAATCATTGT
C_excelsa_671   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362  TCTATTTATAATCCTCAAAAATTCCTCCAAACAGCAGCTAAAATCACCTT
C_mexicana_720  TCTATTTATAATCCTCAAAAATTCCTCCAAACAGCAGCTAAAATCACCTT
C_mutisii_914   TCNATTTATAATCCTCAAAAATTCCTCCAAACAGCAGCTAAAATCACCTT
C_mutisii_1295  TCTATTTATAATCCTCAAAAATTCCTCCAAACAGCAGCTAAAATCACCTT
C_sprucei_817   TCTATTTATAATCCTCAAAAATTCCTCCAAACAGCATCTAAAATCACCTT
C_sprucei_887   TCTATTTATAATCCTCAAAAATTCCTCCAAACAGCANCTAAAATCACCTT
C_ulei_88       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451      CCTATTTATAATCCTCAAAA-TTCCTCCATACAGCTGCTAAAATCATCAT
I_laev_460      CCTATTTATAATCCTCAAAA-TTCCTCCATACAGCTGCTAAAATCATCAT
O_parvi_RS_598  TCTATTTATAATCCTCAAAA-TTCCTCCAAACAGCAGCTAAAATCATCAT

C_atopa_1374    ATTCATTTACTTTTAGGCTGTGACCCGCCCAATGGATGANTGGTTCAATT
C_cap_RS_527    ATTCATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_cap_RS_551    ATTCATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_cap_835       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209  ATTCATTTACTTTTAGGCTGTGACCTGCCCAATGGATGAATGGTTCAATT
C_excelsa_671   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362  ATTTATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_mexicana_720  ATTTATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_mutisii_914   ATTTATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_mutisii_1295  ATTTATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_sprucei_817   ATTTATTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
C_sprucei_887   ATTTATTTACTTTTAGGCTGTGACCCNCCCAATGGATGAATGGTTCAATT
C_ulei_88       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451      ATTCGTTTACTTTTTGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
I_laev_460      ATTCGTTTACTTTTCGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT
O_parvi_RS_598  ATTCGTTTACTTTTAGGCTGTGACCCGCCCAATGGATGAATGGTTCAATT

C_atopa_1374    CTCATGAAGGNGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCNNN
C_cap_RS_527    CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCNGAGATCTAA
C_cap_RS_551    CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCNGAGATCTAA
C_cap_835       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
```

```
C_debilis_7209 CTCATGAAGGTGTGACATT-GTAAGATGATAA-GAGTGCCNGAGATCTAA
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA
C_mexicana_720 CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA
C_mutisii_914  CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA
C_mutisii_1295 CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA
C_sprucei_817  CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA
C_sprucei_887  CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     CTCATGAAGGTGTGACATTTGTAAGATGATAAAGAGTGCCTGAGATCTAA
I_laev_460     CTCATGAAGGTGTGACATTTGTAAGATGATAAAGAGTGCCTGAGATCTAA
O_parvi_RS_598 CTCATGAAGGTGTGACATTTGTAAGATGATAA-GAGTGCCTGAGATCTAA

C_atopa_1374   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_cap_RS_527   CAAATAAGGACTGCAGGAAAATTTTCTCATATAAAT----GACTTGGAAT
C_cap_RS_551   CAAATAAGGACTGCAGGAAAATTTTCTCATATAAAT----GACTTGGAAT
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 CAAATAAGGACTGCAGGAAAATTTTCTCATATACAT----GACTTGGAAT
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 CAAATAAGGACTGCAGGAAAATTTTCTCATATA----------------
C_mexicana_720 CAAATAAGGACTGCAGGAAAATTTTCTCATATA----------------
C_mutisii_914  CAAATGAGGACTGCAGGAAAATTTTCTCATATATATATATGACTTGGAAT
C_mutisii_1295 CAAATGAGGACTGCAGGAAAATTTTCTCATATATATATATGACTTGGAAT
C_sprucei_817  CAAATAAGGACTGCAGGAAAATTTTCTCATATATATAT--GAATTGGAAT
C_sprucei_887  CAAATAAGGACTGCAGGAAAATTTTCTCATNTATATAT--GACTTGGAAT
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     CAAATAAGGACTGCAGGAAAAATTTCTCATATATATAT--GATTTGAAAT
I_laev_460     CAAATAAGGACTGCAGGAAAAATTTCTCATATATATAT--GATTTGAAAT
O_parvi_RS_598 CAAATAAGGACTGCAGGAAAAATTTCTCATATATAT----GATTTGAAAT

C_atopa_1374   NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_cap_RS_527   GTCTATAGACAGAGAAATGGCCTTTCACTCATGTTCTTTGTCCTTTTTAA
C_cap_RS_551   GTCTATAGACAGAGAAATGGCCTTTCACTCATGTTCTTTGTCCTTTTTAA
C_cap_835      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_debilis_7209 GTCTATAGACAGAGAAATGGCCTTTCACTAATGTTCTTTGTCCTTTTTAA
C_excelsa_671  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_mexicana_362 -------GACAGAGAAATGGCCTTTCACTAATGCTCTTTGTCCTTTTTAA
C_mexicana_720 -------GACAGAGAAATGGCCTTTCACTAATGCTCTTTGTCCTTTTTAA
C_mutisii_914  GTCTATAGACAGAGAAATGGCCTTTCACTAATGCTCTTTGTCCTTTTTAA
C_mutisii_1295 GTCTATAGACAGAGAAATGGCCTTTCACTAATGCTCTTTGTCCTTTTTAA
C_sprucei_817  GTCTATAGACAGAGAAATGGCCTTTCACTAATGCTCTTTGT---------
C_sprucei_887  GTCTATAGACAGAGAAATGGCCTTTCACTAATGCTCTTTGTCCTTGTTAA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
I_juru_451     GTCTATCGACAGA--AATGGCCTTTCACTAATGTTCTTTGTCCTTTTTAA
I_laev_460     GTCTATCGACAGA--AATGGCCTTTCACTAATGTTCTTTGTCCTTTTTAA
O_parvi_RS_598 GTCTATAGACAGAGAAATGGCCTTTCACTAATGTTCTCTGTCCTTTTTTA

C_atopa_1374   NNATCTTCTACGCCACATACTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_cap_RS_527   ATATNTTCTNCGCCACATACTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_cap_RS_551   ATATCTTCTACGCCACATACTTGTCTGNGAAAATTGGGTATTGGAGGTAC
C_cap_835      NNATCTTCTACGCCACATACTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_debilis_7209 ATATCTTCTATGCCACATANTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_excelsa_671  NNATCTTCTATGCCACATATTTGTCNGAGAAAATTGGGTATTGGAGGTAC
C_mexicana_362 ATATCTTCTATGCCACATATTTGTCTGAGAAAATTGGGTATTGGAGGTAC
```

```
C_mexicana_720 ATATCTTCTATGCCACATATTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_mutisii_914  ATATCTTCTATGCCACATATTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_mutisii_1295 ATATCTTCTATGCCACATATTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_sprucei_817  --ATCTTCTATGCCACATATTTNTCTGAGAAAATTGGGTATTGGAGGTAC
C_sprucei_887  ATATCTTCTATGCCACATATTTGTCTGAGAAAATTGGGTATTGGAGGTAC
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    NNATCTTCTATGCCACATATTTGTCTGAGAAAATTGGGTATTGGAGGTAC
I_juru_451     ATATCTTCTATGCCACATACCTGTCTGAGAAAATTGGGTATTGGAGGTAC
I_laev_460     ATATCTTCTATGCCACGTACTTGTCTGAGAAAATTGGGTATTGGAGGTAC
O_parvi_RS_598 AAATCTTCTACGCCACATACCTGTCTGAGAAAATTGGGTATTGGAGGTAC

C_atopa_1374   ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_cap_RS_527   ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_cap_RS_551   ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_cap_835      ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_debilis_7209 ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_excelsa_671  ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_mexicana_362 ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_mexicana_720 ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_mutisii_914  ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_mutisii_1295 ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_sprucei_817  ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_sprucei_887  ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
I_juru_451     ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
I_laev_460     ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC
O_parvi_RS_598 ATAACTATTTACAGGCATTTGAAGGCCAACCCAGAGTACCAACTTTACCC

C_atopa_1374   AATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_cap_RS_527   AATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_cap_RS_551   AATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_cap_835      AATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_debilis_7209 AATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_excelsa_671  GATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_mexicana_362 GATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_mexicana_720 GATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_mutisii_914  GATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_mutisii_1295 GATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_sprucei_817  AATCTTCAAGTATTTTGAAAACTGGTGCCAGGATGAGAATAGGCANGGAG
C_sprucei_887  AATCTTCAAGTATTTTGAAAACTGGTGCCAGGATGAGAATAGGCATGGAG
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    GATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAATAGGCATGGAG
I_juru_451     AATCTTCAAGTATTTTGAGAACTGGTGCCAGGATGAGAACAGGCATGGAG
I_laev_460     AATCTTCAAGTATTTTGAGAACTGGTGCCAGGATGAGAACAGGCATGGAG
O_parvi_RS_598 AATCTTCAAGTATTTCGAGAACTGGTGCCAGGATGAGAACAGGCATGGAG

C_atopa_1374   ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_cap_RS_527   ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_cap_RS_551   ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_cap_835      ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_debilis_7209 ANTTCTTCTCTGCATTGATGAAGGCNCAGCCTCAATTCCTCAATGACTGG
C_excelsa_671  ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_mexicana_362 ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_mexicana_720 ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_mutisii_914  ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_mutisii_1295 ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
```

```
C_sprucei_817  ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_sprucei_887  ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG
I_juru_451     ATTTCTTCTCTGCATTGATGAAGGCACAGCCTCAATTCCTCAATGACTGG
I_laev_460     ATTTCTTCTCTGCATTGATGAAGGCACAGCCTCAATTCCTCAATGACTGG
O_parvi_RS_598 ATTTCTTCTCTGCATTGATGAAGGCGCAGCCTCAATTCCTCAATGACTGG

C_atopa_1374   AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
C_cap_RS_527   AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
C_cap_RS_551   AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
C_cap_835      AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTNA
C_debilis_7209 AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
C_excelsa_671  AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTTA
C_mexicana_362 AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTTA
C_mexicana_720 AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTTA
C_mutisii_914  AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTTA
C_mutisii_1295 AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTTA
C_sprucei_817  AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
C_sprucei_887  AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
I_juru_451     AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
I_laev_460     AAAGCGAAGCTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATGTTATTCA
O_parvi_RS_598 AAAGCGAAGTTGTGGGCTCGTTTCTTCTGTCTATCGGTAAATTTTATTCA

C_atopa_1374   ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTATTC
C_cap_RS_527   ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTATTC
C_cap_RS_551   ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTATTC
C_cap_835      ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTATTC
C_debilis_7209 ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_excelsa_671  ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_mexicana_362 ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_mexicana_720 ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCGTGATTCTTC
C_mutisii_914  ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_mutisii_1295 ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_sprucei_817  ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_sprucei_887  ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTC
I_juru_451     ATTTTATTATATTTTTTCAGCCTTTTCATATTTCACTGGGCATGATTCTTA
I_laev_460     ATTTTATTATATTTTTTAGCCTTTTCATATTTCACTGGGCATGATTCTTA
O_parvi_RS_598 ATTTTATTATATTTTTTAGCCTTTTCATATTTAACTGGGAATGATTCTTC

C_atopa_1374   TGATTATTGGAGACTCACAATCTGTTACTTCCAGCTTTTGATTTTTCATG
C_cap_RS_527   TGATTATTGGAGACTCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATG
C_cap_RS_551   TGATTATTGGAGACTCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATG
C_cap_835      TGATTATTGGAGACTCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATG
C_debilis_7209 TGATTATTGTAGACTCACAATCTGTTGCTTCAAGNTTTTGATTTTTCATG
C_excelsa_671  TGATTATTGTAGACTCACAATCTGTTGCTTCAAGCTTTTGATTTTTCATG
C_mexicana_362 TGATTATTGTAGACTCACAATCTGTTGCTTCAAGGTTTTGATTTTTCATG
C_mexicana_720 TGATTATTGTAGACTCACAATCTGTTGCTTCAAGGTTTTGATTTTTCATG
C_mutisii_914  TGATTATTGTAGACTCACAATCTGTTGCTTCAAGCTTTTGATTTTTCATG
C_mutisii_1295 TGATTATTGTAGACTCACAATCTGTTGCTTCAAGCTTTTGATTTTTCATG
C_sprucei_817  TGATTATTGTAGACTCACAATCTGTTGCTTCNAGCTTTTGATTTTTCATG
C_sprucei_887  TGATTATTGTAGACTCACAATCTGTTGCTTCGAGCTTTTGATTTTTCATG
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
```

```
C_ulei_6192    TGATTATTGTAGACTCACAATCTGTTGCTTCAAGCTTTTGATTTTTCATG
I_juru_451     TGATTATTGGATACTCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATG
I_laev_460     TGATTATTGGATACTCACAATCTGTNGCTTCCAGCTTTTGATTTTTCATG
O_parvi_RS_598 TGATTATTGGAGACTCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATG

C_atopa_1374   GATGAAGAGGTTTCAAAAGTGGATTTTGTTGGTCATCTGTTCTCTTTTTA
C_cap_RS_527   GATGAAGTGGTTTCAAAAGTGGATTTTGTTGGTCATCTGTTCTCTTTTTA
C_cap_RS_551   GATGAAGTGGTTTCAAAAGTGGATTTTGTTGGTCATCTGTTCTCTTTTTA
C_cap_835      GATGAAGAGGTTTCAAAAGTGGATTTTGTTGGTCATCTGTTCTCTTTTTA
C_debilis_7209 GATGAAGNNGTTTCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_excelsa_671  GATGAAGTGGTTTCAAGAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_mexicana_362 GATGAAGTGGTTTCAAGAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_mexicana_720 GATGAAGTGGTTTCAAGAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_mutisii_914  GATGAAGTGGTTTCAAGAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_mutisii_1295 GATGAAGTGGTTTCAAGAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_sprucei_817  GATGAAGTGGTTTCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_sprucei_887  GATGAAGTAGTTTCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
C_ulei_88      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
C_ulei_6192    GATGAAGTGGTTTCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA
I_juru_451     GATGAAGTAGTTTCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTCA
I_laev_460     GATGAAGTAGTTTCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTCA
O_parvi_RS_598 GATGAAGTGGTTNCAAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA

C_atopa_1374   CATTGCTTATGCCT
C_cap_RS_527   CGTTGCTTATGCCT
C_cap_RS_551   CGTTGCTTATGCCT
C_cap_835      CGTTGCTTATGCCT
C_debilis_7209 TGTNGCTTATGCTT
C_excelsa_671  TGTTGCTNANGCTT
C_mexicana_362 TGTTGCTTATGCTT
C_mexicana_720 TGTNGCTTATGCTT
C_mutisii_914  TGTTGCTTATGCTT
C_mutisii_1295 TGTNGCTNANGCTT
C_sprucei_817  TGTNGCTTATGCTT
C_sprucei_887  TGTTGCTTATGCTT
C_ulei_88      NNNNNNNNNNNNN
C_ulei_6192    TNNNGNTNANNNTT
I_juru_451     TNTNNCTNANNNTT
I_laev_460     TGTNGCTNATGCTT
O_parvi_RS_598 TGTTGCTTATGCTT
```

Alignment 1.2

DNA nucleotide sequence alignment of trnH-psbA for 50 taxa and 348 characters used to construct NJ tree in chapter 1. 50 base pairs per line.

```
C_atopa_1374     TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_cap_835        TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_cap_855        TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_cap_872        TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_cap_875        TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_cap_RS_551     TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTCGATTAC
C_cap_RS_527     TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTCGATTAC
C_diaz_7644A     TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_cap_889        TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_sp1_A848       TCCGCCCCTTGTCCTTTCTAAACTTTCTAAAGAAAATAATTTTAGATTAC
C_cuatre_40682   TCCGCCCCTTGTCCTTTCTAAACTTTCTAAAGAAAATAATTTTAGATTAC
C_rigid_10008    TCCGCCCCTTGTCCTTTCTAAACTTTCTAAAGAAAATAATTTTAGATTAC
C_sp2_23826      TCCGCCCCTTGTCCTTTCTAAACTTTCTAAAGAAAATAATTTTAGATTAC
C_debilis_190    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_debilis_6172   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_debilis_7209   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_debilis_22972  TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_excelsa_636    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_excelsa_666    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_excelsa_668    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_excelsa_669    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_excelsa_671    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_07    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_354   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_362   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_696   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_701   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_719   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_720   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_757   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mexicana_1283  TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mutisii_911    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mutisii_913    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mutisii_914    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mutisii_1290   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_mutisii_1295   TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_sprucei_812    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_sprucei_817    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_sprucei_821    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_sprucei_884    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTATAGATTAC
C_sprucei_887    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTATAGATTAC
C_sprucei_903    TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTATAGATTAC
C_ulei_88        TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_ulei_6192      TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
C_ulei_42644     TCCGCCCCTTGTCCTTTCT---------AAAGAAAAAAATTTTAGATTAC
O_gly_RS_546     TCCGCCCCTTGTCTTTTCG---------AAAGACTAAAATTTTAGATTAC
O_parvi_RS_598   TCCGCCCCTTGTCTTTTCG---------AAAGACTAAAATTTTAGATTAC
V_sur_RS_078     ?CCG?CCCTTGTCTTTTCT---TTTTCTAAAGAAAAAAATTTTAGATTAC
I_juru_RS_451    TCCGCCCCTTGTCTTTTCT----TTTCTAAAGACAAAAATTTTAGATTAC
I_laev_RS_460    TCCGCCCCTTGTCTTTTCT----TTTCTAAAGACAAAAATTTTAGATTAC
```

```
C_atopa_1374    TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_cap_835       TAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_cap_855       TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCTAC
C_cap_872       TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCTAC
C_cap_875       TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCTAC
C_cap_RS_551    TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATCGAC
C_cap_RS_527    TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATCGAC
C_diaz_7644A    TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_cap_889       TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_sp1_A848      TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_cuatre_40682  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_rigid_10008   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_sp2_23826     TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAATCGAC
C_debilis_190   TAGTCTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA----C
C_debilis_6172  TAGTCTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA----C
C_debilis_7209  TAGTCTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA----C
C_debilis_22972 TAGTCTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA----C
C_excelsa_636   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_excelsa_666   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_excelsa_668   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_excelsa_669   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_excelsa_671   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_07   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_354  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_362  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_696  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_701  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_719  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_720  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_757  TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mexicana_1283 TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_mutisii_911   TAGTCTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----C
C_mutisii_913   TAGTCTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----C
C_mutisii_914   TAGTCTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----C
C_mutisii_1290  TAGTCTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----C
C_mutisii_1295  TAGTCTTTCTTATTTTTTTCATACAAATTTATACCCTTTAGAAAA----C
C_sprucei_812   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTTGAAAA----C
C_sprucei_817   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTTGAAAA----C
C_sprucei_821   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTTGAAAA----C
C_sprucei_884   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_sprucei_887   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_sprucei_903   TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_ulei_88       TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_ulei_6192     TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
C_ulei_42644    TAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAAAA----C
O_gly_RS_546    TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTCTAAAATTTAC
O_parvi_RS_598  TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTCTAAAATTTAC
V_sur_RS_078    TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
I_juru_RS_451   TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
I_laev_RS_460   TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC

C_atopa_1374    AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cap_835       AATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cap_855       AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cap_872       AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cap_875       AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
```

```
C_cap_RS_551     AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cap_RS_527     AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_diaz_7644A     AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cap_889        AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_sp1_A848       AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_cuatre_40682   AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_rigid_10008    AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_sp2_23826      AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_debilis_190    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGTTAA
C_debilis_6172   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGTTAA
C_debilis_7209   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGTTAA
C_debilis_22972  AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGTTAA
C_excelsa_636    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_excelsa_666    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_excelsa_668    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_excelsa_669    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_excelsa_671    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_07    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_354   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_362   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_696   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_701   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_719   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_720   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_757   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mexicana_1283  AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mutisii_911    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mutisii_913    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mutisii_914    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mutisii_1290   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_mutisii_1295   AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_sprucei_812    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_sprucei_817    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_sprucei_821    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_sprucei_884    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_sprucei_887    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_sprucei_903    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_ulei_88        AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_ulei_6192      AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
C_ulei_42644     AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
O_gly_RS_546     AATAGGAAAAAAT----TTTTGGAATGTACATAAACAGAAGATCAGTTAA
O_parvi_RS_598   AATAGGAAAAAAT----TTTTGGAATGTACATAAACAGAAGATCAGTTAA
V_sur_RS_078     AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
I_juru_RS_451    AATAGGAAAAAATGCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
I_laev_RS_460    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA

C_atopa_1374     AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_835        AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_855        AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_872        AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_875        AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_RS_551     AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_RS_527     AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_diaz_7644A     AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_cap_889        AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_sp1_A848       AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAA?G
C_cuatre_40682   AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATG
```

```
C_rigid_10008   AATCAAAAAAA--------GGTATGATGTTCGATCATGAACCAAATAATG
C_sp2_23826     AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_debilis_190   AATCAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_debilis_6172  AATCAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_debilis_7209  AATCAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_debilis_22972 AATCAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_excelsa_636   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_excelsa_666   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_excelsa_668   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_excelsa_669   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_excelsa_671   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_07   AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_354  AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_362  AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_696  AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_701  AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_719  AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_720  AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_757  AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_mexicana_1283 AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mutisii_911   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mutisii_913   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mutisii_914   AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mutisii_1290  AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_mutisii_1295  AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_sprucei_812   AATAAAAAAAAAAAA----GGTATGATGTTCGATCATGAACCAAATAATG
C_sprucei_817   AATAAAAAAAAAAAA----GGTATGATGTTCGATCATGAACCAAATAATG
C_sprucei_821   AATAAAAAAAAAAAA----GGTATGATGTTCGATCATGAACCAAATAATG
C_sprucei_884   AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_sprucei_887   AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_sprucei_903   AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_ulei_88       AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
C_ulei_6192     AATAAAAAAAAA-------GGTATGATGTTCGATCATGAACCAAATAATT
C_ulei_42644    AATAAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
O_gly_RS_546    AAACAAAAAAA--------GGTATGATGTTCGATCCTGAACCAACTAATT
O_parvi_RS_598  AAACAAAAAAA--------GGTATGATGTTCGATCCTGAACCAACTAATT
V_sur_RS_078    AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG
I_juru_RS_451   AATAAAAATAAAAAAAAAAGGTATGATGTTCGATCATGAA-CAAA-AATG
I_laev_RS_460   AATAAAAAAAAA-------GGTATGATGTTCGATCATGAA-CAAA-AATG

C_atopa_1374    CATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_cap_835       CATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_cap_855       CATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTA
C_cap_872       CATATTTTCTGAAATTGAAAAAAAAAAAATCTTATGTGAGTAAACCACTA
C_cap_875       CATATTTTCTGAAATTGAAAAAAAAAAAATCTTATGTGAGTAAACCACTA
C_cap_RS_551    CATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTA
C_cap_RS_527    CATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTA
C_diaz_7644A    CATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTA
C_cap_889       CATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTA
C_sp1_A848      AATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTA
C_cuatre_40682  AATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTA
C_rigid_10008   AATATTTTCTGAAATTGAAAAAAAAAA--??????????????????????
C_sp2_23826     CATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTA
C_debilis_190   AATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_debilis_6172  AATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_debilis_7209  AATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_debilis_22972 AATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
```

```
C_excelsa_636   AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_excelsa_666   AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_excelsa_668   AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_excelsa_669   AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_excelsa_671   AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_07   AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_354  AATATTTTCTGAAATGGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_mexicana_362  AATATTTTCTGAAATGGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_mexicana_696  AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_701  AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_719  AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_720  AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_757  AATATTTTCTGAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mexicana_1283 AATATTTTCTGAAATGGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_mutisii_911   CATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mutisii_913   CATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mutisii_914   CATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mutisii_1290  AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_mutisii_1295  AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_sprucei_812   AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_sprucei_817   AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_sprucei_821   AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_sprucei_884   AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_sprucei_887   AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_sprucei_903   AATATTTTCTTAAATGGAAAAAAAA----TCTTATGTGAGTAAACCACTA
C_ulei_88       AATATTTTCTTAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_ulei_6192     AATATTTTCTTAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
C_ulei_42644    CATATTTTCTTAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTA
O_gly_RS_546    AATATTTTTTAAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTA
O_parvi_RS_598  AATATTTTTTAAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTA
V_sur_RS_078    AATATTTTCTTAAAAAAAAAAAGAAA--TCTTATGTGAGTAAACCACTA
I_juru_RS_451   AATATTTTCTTAAATTGAAAAAAAAA--TCTTATGTGAGTAAACCACTA
I_laev_RS_460   AATATTTTCTTAAATTGAAAAACAAAA--TCTTATGTGAGTAAACCACTA

C_atopa_1374    CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_835       CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_855       CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_872       CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_875       CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_551    CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_527    CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_diaz_7644A    CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_889       CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sp1_A848      CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cuatre_40682  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_rigid_10008   ??????????????????????????????????????????????????
C_sp2_23826     CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_debilis_190   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_debilis_6172  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_debilis_7209  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_debilis_22972 CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_excelsa_636   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_excelsa_666   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_excelsa_668   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_excelsa_669   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_excelsa_671   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_07   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
```

```
C_mexicana_354  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_362  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_696  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_701  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_719  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_720  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_757  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mexicana_1283 CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mutisii_911   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mutisii_913   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mutisii_914   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mutisii_1290  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_mutisii_1295  CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sprucei_812   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sprucei_817   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sprucei_821   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sprucei_884   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sprucei_887   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_sprucei_903   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_ulei_88       CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_ulei_6192     CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
C_ulei_42644    CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
O_gly_RS_546    CTGAAACAGATCAATACCCATGGGTATTGATCTGATCTTTCAATGACTCG
O_parvi_RS_598  CTGAAACAGATCAATACCCATGGGTATTGATCTGATCTTTCAATGACTCG
V_sur_RS_078    CTGAACCAGATCAATACCCATGGGTATTGATCTGATCCTTCAATGACTCG
I_juru_RS_451   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG
I_laev_RS_460   CTGAACCAGATCAATACCCATGGGTATTGATCTGGTCCTTCAATGACTCG

C_atopa_1374    TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_cap_835       TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_cap_855       TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACA
C_cap_872       TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACA
C_cap_875       TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACA
C_cap_RS_551    TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_cap_RS_527    TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_diaz_7644A    TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_cap_889       TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACA
C_sp1_A848      TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_cuatre_40682  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_rigid_10008   ??????????????????????????????????????????????
C_sp2_23826     TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_debilis_190   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_debilis_6172  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_debilis_7209  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_debilis_22972 TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_excelsa_636   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_excelsa_666   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_excelsa_668   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_excelsa_669   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_excelsa_671   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_07   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_354  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_362  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_696  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_701  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_719  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_720  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
```

```
C_mexicana_757  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mexicana_1283 TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mutisii_911   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mutisii_913   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mutisii_914   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mutisii_1290  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_mutisii_1295  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_sprucei_812   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_sprucei_817   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_sprucei_821   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_sprucei_884   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_sprucei_887   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_sprucei_903   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_ulei_88       TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_ulei_6192     TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
C_ulei_42644    TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACA
O_gly_RS_546    TATACACTAATACCGAAGCATTAGCCATTTGTTGATAGA?CTTCA???
O_parvi_RS_598  TATACACTAATACCGAAGCATT?GCCATTTGTTGATAGAGCTTCAACA
V_sur_RS_078    TATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGCTTCAACA
I_juru_RS_451   TATACACTAATACCGAAGTATTAGCCATTTGTTGATAGAGCTTCAACA
I_laev_RS_460   TATACACTAATACCGAAGTATTAGCCATTTGTTGATAGAGCTTCAACA
```

Appendix 2

Alignment 2.1
The trnH-psbA DNA nucleotide sequence alignment of 61 taxa and 302 characters used in haplotype network analysis in chapter 2. 50 base pairs per line.

```
C_cap_JJ_1542 AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_1543 AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_1544 AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_1545 AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_827  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_829  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_830  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_831  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_832  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_833  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_834  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_835  AGATTACTAGTTTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_843  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_844  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_855  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_860  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_862  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_863  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_872  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_873  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_874  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_875  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_893  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_894  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_899  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_900  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_JJ_902  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_271  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_274  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_278  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_280  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_284  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_293  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_294  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_295  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_332  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_334  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_336  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_338  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_346  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_364  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_370  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_406  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_409  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_418  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_427  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_cap_RS_527  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_531  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_535  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
```

```
C_cap_RS_538  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_550  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_551  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_562  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_566  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_568  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_569  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_570  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_571  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_cap_RS_572  CGATTACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAA
C_atopa_1374  AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA
C_diaz_7644   AGATTACTAGTCTTTCTTATTTTTTTCATACAAATTTCTACCCTTTAGAA

C_cap_JJ_1542 AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_1543 AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_1544 AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_1545 AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_827  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_829  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_830  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_831  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_832  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_833  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_834  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_835  AATCGACAATAGGAAAAAATTCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_843  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_844  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_855  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_860  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_862  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_863  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_872  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_873  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_874  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_875  AATCTACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_893  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_894  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_899  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_900  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_JJ_902  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_271  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_274  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_278  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_280  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_284  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_293  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_294  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_295  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_332  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_334  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_336  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_338  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_346  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_364  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_370  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_406  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_409  AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
```

```
C_cap_RS_418   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_427   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_527   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_531   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_535   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_538   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_550   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_551   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_562   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_566   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_568   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_569   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_570   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_571   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_cap_RS_572   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_atopa_1374   AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT
C_diaz_7644    AATCGACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGAT

C_cap_JJ_1542 CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_1543 CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_1544 CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_1545 CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_827   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_829   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_830   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_831   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_832   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_833   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_834   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_835   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_843   CAGTTCAAATCAAAAAAAGGTATGAGGTTCGATTATGAACCAAATAATTC
C_cap_JJ_844   CAGTTCAAATCAAAAAAAGGTATGAGGTTCGATTATGAACCAAATAATTC
C_cap_JJ_855   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_860   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_862   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_863   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_872   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_873   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_874   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_875   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_893   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_894   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_899   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_900   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_JJ_902   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_271   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_274   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_278   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_280   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_284   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_293   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_294   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_295   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_332   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_334   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_336   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_338   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
```

```
C_cap_RS_346  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_364  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_370  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_406  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_409  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_418  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_427  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_527  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_531  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_535  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_538  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_550  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_551  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_562  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_566  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_568  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_569  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_570  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_571  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_cap_RS_572  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_atopa_1374  CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC
C_diaz_7644   CAGTTCAAATCAAAAAAAGGTATGATGTTCGATTATGAACCAAATAATTC

C_cap_JJ_1542 ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_1543 ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_1544 ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_1545 ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_827  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_829  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_830  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_831  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_832  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_833  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_834  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_835  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_JJ_843  ATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTAC
C_cap_JJ_844  ATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTAC
C_cap_JJ_855  ATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTAC
C_cap_JJ_860  ATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTAC
C_cap_JJ_862  ATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTAC
C_cap_JJ_863  ATATTTTCTGAAATTGAAAAAAAAAAA-TCTTATGTGAGTAAACCACTAC
C_cap_JJ_872  ATATTTTCTGAAATTGAAAAAAAAAAAAATCTTATGTGAGTAAACCACTAC
C_cap_JJ_873  ATATTTTCTGAAATTGAAAAAAAAAAAAATCTTATGTGAGTAAACCACTAC
C_cap_JJ_874  ATATTTTCTGAAATTGAAAAAAAAAAAAATCTTATGTGAGTAAACCACTAC
C_cap_JJ_875  ATATTTTCTGAAATTGAAAAAAAAAAAAATCTTATGTGAGTAAACCACTAC
C_cap_JJ_893  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_JJ_894  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_JJ_899  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_JJ_900  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_JJ_902  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_271  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_274  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_278  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_280  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_284  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_293  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_294  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
```

```
C_cap_RS_295  ATATTTTTTGAAATTGAAAAAAAAA---TCTTATGTGGGTAAACCACTAC
C_cap_RS_332  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_RS_334  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_RS_336  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_RS_338  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_RS_346  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_cap_RS_364  ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_370  ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_406  ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_409  ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_418  ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_427  ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_527  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_531  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_535  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_538  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_550  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_551  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_562  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_566  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_568  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_569  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_570  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_571  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_cap_RS_572  ATATTTTCTGAAATTGAAAAAACAAA--TCTTATGTGAGTAAACCACTAC
C_atopa_1374  ATATTTTCTGAAATTGAAAAAAAAA---TCTTATGTGAGTAAACCACTAC
C_diaz_7644   ATATTTTCTGAAATTGAAAAAAAAAA--TCTTATGTGAGTAAACCACTAC

C_cap_JJ_1542 TGAACCAGATCAATACCCAATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_1543 TGAACCAGATCAATACCCAATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_1544 TGAACCAGATCAATACCCAATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_1545 TGAACCAGATCAATACCCAATGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_827  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_829  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_830  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_831  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_832  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_833  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_834  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_835  TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_843  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_844  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_855  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_860  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_862  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_863  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_872  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_873  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_874  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_875  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_893  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_894  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_899  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_900  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_JJ_902  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_271  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_274  TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
```

```
C_cap_RS_278   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_280   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_284   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_293   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_294   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_295   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_332   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_334   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_336   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_338   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_346   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_364   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_370   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_406   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_409   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_418   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_427   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_527   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_531   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_535   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_538   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_550   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_551   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_562   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_566   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_568   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_569   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_570   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_571   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_cap_RS_572   TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG
C_atopa_1374   TGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTCAATGACTCG
C_diaz_7644    TGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTCAATGACTCG

C_cap_JJ_1542 TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_1543 TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_1544 TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_1545 TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_827   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_829   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_830   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_831   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_832   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_833   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_834   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_835   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_843   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_844   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_855   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_860   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_862   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_863   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_872   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_873   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_874   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_875   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_893   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_894   TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
```

```
C_cap_JJ_899  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_900  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_JJ_902  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_271  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_274  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_278  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_280  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_284  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_293  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_294  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_295  TATACACTAATACCGAAATATTAACCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_332  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_334  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_336  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_338  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_346  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_364  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_370  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_406  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_409  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_418  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_427  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_527  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_531  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_535  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_538  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_550  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_551  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_562  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_566  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_568  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_569  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_570  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_571  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_cap_RS_572  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_atopa_1374  TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
C_diaz_7644   TATACACTAATACCGAAATATTAGCCATTTATTGATAGAGCTTCAACAGC
```

Appendix 3

Alignment 3.1

Concatenated DNA nucleotide sequence alignment of 3-loci, 1517 characters, and 49 taxa used in Bayesian and Maximum parsimony analyses in chapter 3. 50 base pairs per line.

loci order and positions in alignment:
trnH-psba 1-357, AGT1 358-1078, AT103 1079-1517

```
C_cap_RS_551      TCCGCCCCTTGTCCTTTCT-------AAAGAAAAAAA--TTTTCGATTAC
C_debilis_7209    TCCGCCCCTTGTCCTTTCT-------AAAGAAAAAAA--TTTTAGATTAC
I_juru_RS_451     TCCGCCCCTTGTCTTTTCTTTT--CTAAAGACAAAAA--TTTTAGATTAC
I_laev_RS_460     TCCGCCCCTTGTCTTTTCTTTT--CTAAAGACAAAAA--TTTTAGATTAC
V_cadu_JJ_847     TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAAAAAATTTAGATTAC
V_spRADS3_RS_335  TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_spRADS3_RS_339  TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_peru_JJ_772     TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_calo_RS_430     TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_calo_RS_454     ??????CCTTGTCTTTTCTTTTT-CT-----------------------
V_calo_RS_481     ????????????????????????-----------------------
V_calo_RS_511     TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_spRADS1_RS_432  TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_spRADS1_RS_500  TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_spRADS1_R1_510  TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_spRADS1_RS_561  TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_dix_RS_225      TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_elon_RS_437     ???????CTTGTCTTTTCTTTTT-CT-----------------------
V_elon_RS_494     ????????????????????????-----------------------
V_elon_RS_502     TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_flex_RS_442     TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAAAAAATTTAGATTAC
V_flex_RS_522     TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAAAAAATTTAGATTAC
V_flex_RS_595     TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAAAAAATTTAGATTAC
V_multin_RS_107   TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAGAAAATTTAGATTAC
V_multin_RS_108   TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAGAAAATTTAGATTAC
V_loret_RS_480    TCC?CCCCTTGTCTTTTCTTTTT-CT-----------------------
V_loret_RS_516    TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_loret_RS_526    TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_multi_RS_350    TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAAAAAATTTAGATTAC
V_multi_RS_429    TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAGAAAATTTAGATTAC
V_multi_RS_466    TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAGAAAATTTAGATTAC
V_multi_RS_543    TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAGAAAATTTAGATTAC
V_multi_RS_549    TCCGCCCCTTGTCTTTT-----T-CTAAAGAAAAAGAAAATTTAGATTAC
V_sebBL_RS_584    TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_lorSL_RS_483    TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_lorSL_RS_507    TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_sebSL_RS_534    ????????????????????????-----------------------
V_sebSL_RS_552    TCCGCCCCTTGTCTTTTCTTTTT-CT-----------------------
V_spRADS4_RS_213  TCCGCCCCTTGTCTTTTCCTTTTTGTACAGAAAAAAA--TTTTGGATTTC
V_spRADS4_RS_214  TCCGCCCCTTGTCTTTTCCTTTTTGTACAGAAAAAAA--TTTTGGATTTC
V_sur_RS_082      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC
V_sur_RS_083      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC
```

```
V_sur_RS_084      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC
V_sur_RS_216      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC
V_sur_RS_248      ??????????????????????????????GAAAAAAA--TTTTAGATTAC
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC
V_sur_RS_489      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC
V_sur_RS_501      TCCGCCCCTTGTCTTTTCTTTTT-CTAAAGAAAAAAA--TTTTAGATTAC

C_cap_RS_551      TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATCGAC
C_debilis_7209    TAGTCTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA----C
I_juru_RS_451     TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
I_laev_RS_460     TAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_cadu_JJ_847     TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_spRADS3_RS_335  --------------------------------------------------
V_spRADS3_RS_339  --------------------------------------------------
V_peru_JJ_772     --------------------------------------------------
V_calo_RS_430     --------------------------------------------------
V_calo_RS_454     --------------------------------------------------
V_calo_RS_481     --------------------------------------------------
V_calo_RS_511     --------------------------------------------------
V_spRADS1_RS_432  --------------------------------------------------
V_spRADS1_RS_500  --------------------------------------------------
V_spRADS1_R1_510  --------------------------------------------------
V_spRADS1_RS_561  --------------------------------------------------
V_dix_RS_225      --------------------------------------------------
V_elon_RS_437     --------------------------------------------------
V_elon_RS_494     --------------------------------------------------
V_elon_RS_502     --------------------------------------------------
V_flex_RS_442     TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_flex_RS_522     TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_flex_RS_595     TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_multin_RS_107   TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_multin_RS_108   TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_loret_RS_480    --------------------------------------------------
V_loret_RS_516    --------------------------------------------------
V_loret_RS_526    --------------------------------------------------
V_multi_RS_350    TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_multi_RS_429    TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_multi_RS_466    TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_multi_RS_543    TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_multi_RS_549    TAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATTGAC
V_sebBL_RS_584    --------------------------------------------------
V_lorSL_RS_483    --------------------------------------------------
V_lorSL_RS_507    --------------------------------------------------
V_sebSL_RS_534    --------------------------------------------------
V_sebSL_RS_552    --------------------------------------------------
V_spRADS4_RS_213  TAGTCTTTCTTATTTTTTTCATACTAA----------------------
V_spRADS4_RS_214  TAGTCTTTCTTATTTTTTTCATACTAA----------------------
V_sur_RS_082      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
V_sur_RS_083      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
V_sur_RS_084      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
V_sur_RS_216      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
V_sur_RS_248      TAGTCTTTCTTATTTTTT-CATACTAATTCCTACCCTTTAGAAAATTGAC
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
V_sur_RS_489      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
V_sur_RS_501      TAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATTGAC
```

```
C_cap_RS_551      AATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGTTCA
C_debilis_7209    AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGTTAA
I_juru_RS_451     AATAGGAAAAAATGCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
I_laev_RS_460     AATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
V_cadu_JJ_847     AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_spRADS3_RS_335  --------------------------------------------------
V_spRADS3_RS_339  --------------------------------------------------
V_peru_JJ_772     --------------------------------------------------
V_calo_RS_430     --------------------------------------------------
V_calo_RS_454     --------------------------------------------------
V_calo_RS_481     --------------------------------------------------
V_calo_RS_511     --------------------------------------------------
V_spRADS1_RS_432  --------------------------------------------------
V_spRADS1_RS_500  --------------------------------------------------
V_spRADS1_R1_510  --------------------------------------------------
V_spRADS1_RS_561  --------------------------------------------------
V_dix_RS_225      --------------------------------------------------
V_elon_RS_437     --------------------------------------------------
V_elon_RS_494     --------------------------------------------------
V_elon_RS_502     --------------------------------------------------
V_flex_RS_442     AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_flex_RS_522     AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_flex_RS_595     AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_multin_RS_107   AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_multin_RS_108   AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_loret_RS_480    --------------------------------------------------
V_loret_RS_516    --------------------------------------------------
V_loret_RS_526    --------------------------------------------------
V_multi_RS_350    AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_multi_RS_429    AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_multi_RS_466    AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_multi_RS_543    AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_multi_RS_549    AATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGTTAA
V_sebBL_RS_584    --------------------------------------------------
V_lorSL_RS_483    --------------------------------------------------
V_lorSL_RS_507    --------------------------------------------------
V_sebSL_RS_534    --------------------------------------------------
V_sebSL_RS_552    --------------------------------------------------
V_spRADS4_RS_213  ---AGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
V_spRADS4_RS_214  ---AGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_082      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_083      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_084      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_216      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_248      AATAGGAAAAAATTCATTTTTGGAATGTACAGAAACTGAAGATCAGTTAA
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_489      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA
V_sur_RS_501      AATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGTTAA

C_cap_RS_551      AATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAATAATT
C_debilis_7209    AATCAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAATAATT
I_juru_RS_451     AATAAAAATAAAAAAAAAAGGTATGATGTTCGATCATGAA-CAAA-AATG
I_laev_RS_460     AATAAAAAAAAA-------GGTATGATGTTCGATCATGAA-CAAA-AATG
V_cadu_JJ_847     AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_spRADS3_RS_335  --------------------------------------------------
```

```
V_spRADS3_RS_339 --------------------------------------------------
V_peru_JJ_772    --------------------------------------------------
V_calo_RS_430    --------------------------------------------------
V_calo_RS_454    --------------------------------------------------
V_calo_RS_481    --------------------------------------------------
V_calo_RS_511    --------------------------------------------------
V_spRADS1_RS_432 --------------------------------------------------
V_spRADS1_RS_500 --------------------------------------------------
V_spRADS1_R1_510 --------------------------------------------------
V_spRADS1_RS_561 --------------------------------------------------
V_dix_RS_225     --------------------------------------------------
V_elon_RS_437    --------------------------------------------------
V_elon_RS_494    --------------------------------------------------
V_elon_RS_502    --------------------------------------------------
V_flex_RS_442    AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_flex_RS_522    AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_flex_RS_595    AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_multin_RS_107  AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_multin_RS_108  AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_loret_RS_480   --------------------------------------------------
V_loret_RS_516   --------------------------------------------------
V_loret_RS_526   --------------------------------------------------
V_multi_RS_350   AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_multi_RS_429   AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_multi_RS_466   AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_multi_RS_543   AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_multi_RS_549   AATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_sebBL_RS_584   --------------------------------------------------
V_lorSL_RS_483   --------------------------------------------------
V_lorSL_RS_507   --------------------------------------------------
V_sebSL_RS_534   --------------------------------------------------
V_sebSL_RS_552   --------------------------------------------------
V_spRADS4_RS_213 AATAAAAAAGAAAAA----GGTATAATGTTCGATCATTAAACAAATAATT
V_spRADS4_RS_214 AATAAAAAAGAAAAA----GGTATAATGTTCGATCATTAAACAAATAATT
V_sur_RS_082     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG
V_sur_RS_083     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG
V_sur_RS_084     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG
V_sur_RS_216     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_sur_RS_248     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATT
V_sur_RS_324     ????????????????----????????????????????????????
V_sur_RS_428     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG
V_sur_RS_489     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG
V_sur_RS_501     AATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAATAATG

C_cap_RS_551     CATATTTTCTGAAATTGAAAAAACAAA-----TCTTATG-----TGAGTA
C_debilis_7209   AATATTTTCTGAAATTGAAAAAAAAA------TCTTATG-----TGAGTA
I_juru_RS_451    AATATTTTCTTAAATTGAAAAAAAAA-----TCTTATG-----TGAGTA
I_laev_RS_460    AATATTTTCTTAAATTGAAAAACAAAA-----TCTTATG-----TGAGTA
V_cadu_JJ_847    AATATTTTAT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_spRADS3_RS_335 -----------AAAAAAAAAAA------GAAATCTTATGTTATGTGAGTA
V_spRADS3_RS_339 -----------AAAAAAAAAAA------GAAATCTTATGTTATGTGAGTA
V_peru_JJ_772    -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_calo_RS_430    -----------AAAAAAAAAAAAA--------TCTTATGTTATGTGAGTA
V_calo_RS_454    -----------AAAAAAAAAAAAA--------TCTTATGTTATGTGAGTA
V_calo_RS_481    -----------??????????????????????????????????????
V_calo_RS_511    -----------AAAAAAAAAAAAA--------TCTTATGTTATGTGAGTA
V_spRADS1_RS_432 -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
```

```
V_spRADS1_RS_500 -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_spRADS1_R1_510 -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_spRADS1_RS_561 -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_dix_RS_225     -----------AAAAAAAAAAAA---------TCTTATGTTATGTGAGTA
V_elon_RS_437    -----------AAAAAAAAAAAAAAA------TCTTATGTTATGTGAGTA
V_elon_RS_494    -----------????????????????????????????????????
V_elon_RS_502    -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_flex_RS_442    AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_flex_RS_522    AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_flex_RS_595    AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_multin_RS_107  AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_multin_RS_108  AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_loret_RS_480   -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_loret_RS_516   -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_loret_RS_526   -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_multi_RS_350   AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_multi_RS_429   AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_multi_RS_466   AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_multi_RS_543   AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_multi_RS_549   AATATTTTCT-AAAAAAAAA--------GAAATCTTATG-----TGAGTA
V_sebBL_RS_584   -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_lorSL_RS_483   -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_lorSL_RS_507   -----------AAAAAAAAAAAA---------TCTTATGTTATGTGAGTA
V_sebSL_RS_534   -----------????????????????????????????????????
V_sebSL_RS_552   -----------AAAAAAAAAAAAAA-------TCTTATGTTATGTGAGTA
V_spRADS4_RS_213 AATATTTTCTTAAAAAAAAAAA------GAAATCTTATG-----TGAGTA
V_spRADS4_RS_214 AATATTTTCTTAAAAAAAAAAA------GAAATCTTATG-----TGAGTA
V_sur_RS_082     AATATTTTCTTAAAAAAAAAAAA-----GAAATCTTATG-----TGAGTA
V_sur_RS_083     AATATTTTCTTAAAAAAAAAAAA-----GAAATCTTATG-----TGAGTA
V_sur_RS_084     AATATTTTCTTAAAAAAAAAAAA-----GAAATCTTATG-----TGAGTA
V_sur_RS_216     AATATTTTCTTAAAAAAAAAAAA----GAAATCTTATG-----TGAGTA
V_sur_RS_248     AATATTTTCTTAAAAAAAAAAAAAAAAGAAATCTTATG-----TGAGTA
V_sur_RS_324     ????????????????????????????????????????????
V_sur_RS_428     AATATTTTCTTAAAAAAAAAAA-----GAAATCTTATG-----TGAGTA
V_sur_RS_489     AATATTTTCTTAAAAAAAAAAA-----GAAATCTTATG-----TGAGTA
V_sur_RS_501     AATATTTTCTTAAAAAAAAAAA-----GAAATCTTATG-----TGAGTA

C_cap_RS_551     AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTC
C_debilis_7209   AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTC
I_juru_RS_451    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGGTCCTTC
I_laev_RS_460    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGGTCCTTC
V_cadu_JJ_847    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_spRADS3_RS_335 AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_spRADS3_RS_339 AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_peru_JJ_772    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_calo_RS_430    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_calo_RS_454    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_calo_RS_481    ????????????????????????????????????????????????
V_calo_RS_511    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_spRADS1_RS_432 AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_spRADS1_RS_500 AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_spRADS1_R1_510 AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_spRADS1_RS_561 AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_dix_RS_225     AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_elon_RS_437    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_elon_RS_494    ????????????????????????????????????????????????
V_elon_RS_502    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
```

```
V_flex_RS_442     AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_flex_RS_522     AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_flex_RS_595     AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_multin_RS_107   AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_multin_RS_108   AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_loret_RS_480    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_loret_RS_516    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_loret_RS_526    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_multi_RS_350    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_multi_RS_429    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_multi_RS_466    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_multi_RS_543    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_multi_RS_549    AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_sebBL_RS_584    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_lorSL_RS_483    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_lorSL_RS_507    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sebSL_RS_534    ??????????????????????????????????????????????????
V_sebSL_RS_552    AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_spRADS4_RS_213  AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_spRADS4_RS_214  AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_sur_RS_082      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sur_RS_083      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sur_RS_084      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sur_RS_216      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sur_RS_248      AACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGATCCTTC
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sur_RS_489      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC
V_sur_RS_501      AACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGATCCTTC

C_cap_RS_551      AATGACTCGTATACACTAATACCGAAATATTAGCCATTTATTGATAGAGC
C_debilis_7209    AATGACTCGTATACACTAATACCGAAATATTAGCCATTTATTGATAGAGC
I_juru_RS_451     AATGACTCGTATACACTAATACCGAAGTATTAGCCATTTGTTGATAGAGC
I_laev_RS_460     AATGACTCGTATACACTAATACCGAAGTATTAGCCATTTGTTGATAGAGC
V_cadu_JJ_847     AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_spRADS3_RS_335  AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_spRADS3_RS_339  AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_peru_JJ_772     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_calo_RS_430     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_calo_RS_454     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_calo_RS_481     ??????????????????????????????????????????????????
V_calo_RS_511     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_spRADS1_RS_432  AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_spRADS1_RS_500  AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_spRADS1_R1_510  AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_spRADS1_RS_561  AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_dix_RS_225      AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_elon_RS_437     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_elon_RS_494     ??????????????????????????????????????????????????
V_elon_RS_502     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_flex_RS_442     AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_flex_RS_522     AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_flex_RS_595     AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_multin_RS_107   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_multin_RS_108   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_loret_RS_480    AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_loret_RS_516    AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
```

```
V_loret_RS_526   AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_multi_RS_350   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_multi_RS_429   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_multi_RS_466   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_multi_RS_543   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_multi_RS_549   AATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sebBL_RS_584   AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_lorSL_RS_483   AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_lorSL_RS_507   AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_sebSL_RS_534   ??????????????????????????????????????????????????
V_sebSL_RS_552   AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATAGAGC
V_spRADS4_RS_213 AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_spRADS4_RS_214 AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_082     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_083     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_084     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_216     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_248     AATGACTCGTATACACTAATACC?????????????????????????????
V_sur_RS_324     ??????????????????????????????????????????????????
V_sur_RS_428     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_489     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC
V_sur_RS_501     AATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGATGGAGC

C_cap_RS_551     TTCAACA??????????CTCTTTCTATGCCTACTGGCATGGGA?TCGTAT
C_debilis_7209   TTCAACA??????????CTCTTTCTATGCCTACTGG?ATGGGAATCGTAT
I_juru_RS_451    TTCAACA??????????CTCTTTCTATGCCTACTGGTATGGGAATCGTAT
I_laev_RS_460    TTCAACA??????????CTCTTTCTATGCCTACTGGTATGGGA?TCGTAT
V_cadu_JJ_847    TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_spRADS3_RS_335 TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCTTGGGAATCGTAT
V_spRADS3_RS_339 TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCTTGGGAATCGTAT
V_peru_JJ_772    TTCAACT????????????????????GCCTAC?GGCATGGGAATCGTAT
V_calo_RS_430    TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_calo_RS_454    TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_calo_RS_481    ???????TCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_calo_RS_511    TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_spRADS1_RS_432 TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
V_spRADS1_RS_500 TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
V_spRADS1_R1_510 TTCAACT??????????????????????????????GGGAATCGTAT
V_spRADS1_RS_561 TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
V_dix_RS_225     TTCAACT??????????????TCTATGCCTACTGGCATGGGAATCGTAT
V_elon_RS_437    TTCAACT?????????????????????TACCGGCATGGGAATCGTAT
V_elon_RS_494    ???????TCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_elon_RS_502    TTCAACT??????????????????????????????????????AT
V_flex_RS_442    TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_flex_RS_522    TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_flex_RS_595    TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multin_RS_107  TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multin_RS_108  TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_loret_RS_480   TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_loret_RS_516   TTCAACT?????????????????????????????????ATCGTAT
V_loret_RS_526   TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multi_RS_350   TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multi_RS_429   TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multi_RS_466   TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multi_RS_543   TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_multi_RS_549   TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sebBL_RS_584   TTCAACTTTGACAGGGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
```

```
V_lorSL_RS_483   TTCAACT??????????????????????????????????????????????
V_lorSL_RS_507   TTCAACTTCTCAAAAGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
V_sebSL_RS_534   ???????TCTCAAAAGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
V_sebSL_RS_552   TTCAACTTTTGACAGGGCTCTTTCTATGCCTACCGGCATGGGAATCGTAT
V_spRADS4_RS_213 TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_spRADS4_RS_214 TTCAACA????????GGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sur_RS_082     TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGA?TCGTAT
V_sur_RS_083     TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sur_RS_084     TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sur_RS_216     TTCAACA??CcAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sur_RS_248     ????????????????????????????????????????????????
V_sur_RS_324     ????????????????????????????????????????????????
V_sur_RS_428     TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sur_RS_489     TTCAACATCCCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT
V_sur_RS_501     TTCAACA??CCAAAAGGCTCTTTCTATGCCTACTGGCATGGGAATCGTAT

C_cap_RS_551     GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
C_debilis_7209   GCGCAAGTCCAAAAGCAATGGAAGCATCCAAGACTGCTAAATCAGTCAGA
I_juru_RS_451    GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
I_laev_RS_460    GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_cadu_JJ_847    GYGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS3_RS_335 GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS3_RS_339 GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_peru_JJ_772    GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_calo_RS_430    GCGCAAGTCCAAAAGCATTKGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_calo_RS_454    GCGCAAGTCCAAAAGCATTKGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_calo_RS_481    GCGCAAGTCCAAAAGCATTWGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_calo_RS_511    GCGCAAGTCCAAAAGCATTGGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS1_RS_432 GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS1_RS_500 GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS1_R1_510 G?GCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS1_RS_561 GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_dix_RS_225     GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_elon_RS_437    GCGCMAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCACTCAGA
V_elon_RS_494    GTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_elon_RS_502    GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_flex_RS_442    GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_flex_RS_522    GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_flex_RS_595    GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multin_RS_107  GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multin_RS_108  GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_loret_RS_480   GYGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_loret_RS_516   GTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_loret_RS_526   GTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multi_RS_350   GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multi_RS_429   GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multi_RS_466   GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multi_RS_543   GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_multi_RS_549   GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sebBL_RS_584   GCGCMAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_lorSL_RS_483   ????????????????????????????????????????????????
V_lorSL_RS_507   GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sebSL_RS_534   GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sebSL_RS_552   GCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS4_RS_213 GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_spRADS4_RS_214 GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sur_RS_082     GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
```

```
V_sur_RS_083      GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sur_RS_084      GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sur_RS_216      GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sur_RS_248      ??????????????????????????????????????????????????
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sur_RS_489      GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA
V_sur_RS_501      GCGCAAGTCCAAAAGCAATAGAAGCATCCAAGACTGCTAAATCAGTCAGA

C_cap_RS_551      GTATTCTTTGACTGGAAAGACTACCTGAAGTTCTATAA?TTGGGAACATA
C_debilis_7209    GTATTCTTTGACTGGAAAGACTAC?TGAAGTTCTATAAGCTGGGAACATA
I_juru_RS_451     GTATTCTTTGACTGGAAAGACTACCTGAAGTTCTATAAGTTGGGAACATA
I_laev_RS_460     GTATTCTTTGACTGGAAAGACTACCTGAAGTTCTATAAGTTGGGAACATA
V_cadu_JJ_847     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_spRADS3_RS_335  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_spRADS3_RS_339  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_peru_JJ_772     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_calo_RS_430     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_calo_RS_454     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_calo_RS_481     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_calo_RS_511     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_spRADS1_RS_432  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_spRADS1_RS_500  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_spRADS1_R1_510  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_spRADS1_RS_561  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_dix_RS_225      GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_elon_RS_437     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_elon_RS_494     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_elon_RS_502     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_flex_RS_442     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_flex_RS_522     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_flex_RS_595     GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_multin_RS_107   GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_multin_RS_108   GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_loret_RS_480    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_loret_RS_516    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_loret_RS_526    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_multi_RS_350    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_multi_RS_429    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_multi_RS_466    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_multi_RS_543    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_multi_RS_549    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sebBL_RS_584    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_lorSL_RS_483    ??????????????????????????????????????????????????
V_lorSL_RS_507    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_sebSL_RS_534    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_sebSL_RS_552    GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGGGAACATA
V_spRADS4_RS_213  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_spRADS4_RS_214  GTATTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sur_RS_082      GTGTTCTTYGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sur_RS_083      GTGTTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sur_RS_084      GTGTTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sur_RS_216      GTGTTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sur_RS_248      ??????????????????????????????????????????????????
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      GTGTTCTTYGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
V_sur_RS_489      GTGTTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA
```

```
V_sur_RS_501     GTGTTCTTTGACTGGAAAGACTACCTCAAGTTCTATAAGTTGGGTACATA

C_cap_RS_551     TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
C_debilis_7209   TTGGCCCTACACACCCTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
I_juru_RS_451    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
I_laev_RS_460    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_cadu_JJ_847    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_spRADS3_RS_335 TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS3_RS_339 TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_peru_JJ_772    TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_calo_RS_430    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_calo_RS_454    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_calo_RS_481    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_calo_RS_511    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS1_RS_432 TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS1_RS_500 TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS1_R1_510 TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS1_RS_561 TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_dix_RS_225     TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_elon_RS_437    TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_elon_RS_494    TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_elon_RS_502    TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_flex_RS_442    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_flex_RS_522    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_flex_RS_595    TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_multin_RS_107  TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_multin_RS_108  TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_loret_RS_480   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_loret_RS_516   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_loret_RS_526   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_multi_RS_350   TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_multi_RS_429   TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_multi_RS_466   TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_multi_RS_543   TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_multi_RS_549   TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCATCTT
V_sebBL_RS_584   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_lorSL_RS_483   ??????????????????????????????????????????????????
V_lorSL_RS_507   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_sebSL_RS_534   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_sebSL_RS_552   TTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS4_RS_213 TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_spRADS4_RS_214 TTGGCCCTACACACCTTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_082     TTGGCCCTACACACCKTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_083     TTGGCCCTACACACCGTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_084     TTGGCCCTACACACCGTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_216     TTGGCCCTACACACCGTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_248     ??????????????????????????????????????????????????
V_sur_RS_324     ??????????????????????????????????????????????????
V_sur_RS_428     TTGGCCCTACACACCKTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_489     TTGGCCCTACACACCGTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT
V_sur_RS_501     TTGGCCCTACACACCGTCTATACAACTTTTGTATGGCCTGAGAGCAGCTT

C_cap_RS_551     TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
C_debilis_7209   TGGATCTCATATTTTTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
I_juru_RS_451    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
I_laev_RS_460    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_cadu_JJ_847    TGGATCTCATATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
```

```
V_spRADS3_RS_335 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_spRADS3_RS_339 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_peru_JJ_772    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_calo_RS_430    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_calo_RS_454    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_calo_RS_481    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_calo_RS_511    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_spRADS1_RS_432 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_spRADS1_RS_500 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_spRADS1_R1_510 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_spRADS1_RS_561 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_dix_RS_225     TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_elon_RS_437    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_elon_RS_494    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_elon_RS_502    TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_flex_RS_442    TGGATCTCATATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_flex_RS_522    TGGATCTCATATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_flex_RS_595    TGGATCTCATATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_multin_RS_107  TGGATCTCATATTTGAGGAAGGACTCGAAAATGTGTTTGACAGACACAAA
V_multin_RS_108  TGGATCTCATATTTGAGGAAGGACTYGAAAATGTGTTTGACAGACACAAA
V_loret_RS_480   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_loret_RS_516   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_loret_RS_526   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_multi_RS_350   TGGATCTCATATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_multi_RS_429   TGGATCTCATATTTGAGGAAGGACTCGAAAATGTGTTTGACAGACACAAA
V_multi_RS_466   TGGATCTCATATTTGAGGAAGGACTYGAAAATGTGTTTGACAGACACAAA
V_multi_RS_543   TGGATCTCATATTTGAGGAAGGACTCGAAAATGTGTTTGACAGACACAAA
V_multi_RS_549   TGGATCTCATATTTGAGGAAGGACTYGAAAATGTGTTTGACAGACACAAA
V_sebBL_RS_584   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_lorSL_RS_483   ??????????????????????????????????????????????????
V_lorSL_RS_507   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sebSL_RS_534   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sebSL_RS_552   TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_spRADS4_RS_213 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTCACAGACACAAA
V_spRADS4_RS_214 TGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTCACAGACACAAA
V_sur_RS_082     TGGATCTCTTATTTGWGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sur_RS_083     TGGATCTCTTATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sur_RS_084     TGGATCTCTTATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sur_RS_216     TGGATCTCTTATTTGTGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sur_RS_248     ??????????????????????????????????????????????????
V_sur_RS_324     ??????????????????????????????????????????????????
V_sur_RS_428     TGGATCTCTTATTTGWGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sur_RS_489     TGGATCTCTTATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA
V_sur_RS_501     TGGATCTCTTATTTGAGGAAGGACTTGAAAATGTGTTTGACAGACACAAA

C_cap_RS_551     CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGCTGA
C_debilis_7209   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGCAGTGCTGA
I_juru_RS_451    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAATATGTAGTGCTGA
I_laev_RS_460    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGCTGA
V_cadu_JJ_847    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGAWGTAGTTATGA
V_spRADS3_RS_335 CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_spRADS3_RS_339 CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_peru_JJ_772    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_calo_RS_430    CGTTTGGGCAAGGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_calo_RS_454    CGTTTGGGCAAGGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_calo_RS_481    CGTTTGGGCAAGGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_calo_RS_511    CGTTTGGGCAAGGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
```

```
V_spRADS1_RS_432 CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_spRADS1_RS_500 CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_spRADS1_R1_510 CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_spRADS1_RS_561 CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_dix_RS_225     CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_elon_RS_437    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_elon_RS_494    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_elon_RS_502    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_flex_RS_442    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_flex_RS_522    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_flex_RS_595    CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_multin_RS_107  CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_multin_RS_108  CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_loret_RS_480   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_loret_RS_516   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_loret_RS_526   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_multi_RS_350   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_multi_RS_429   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_multi_RS_466   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_multi_RS_543   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_multi_RS_549   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTTATGA
V_sebBL_RS_584   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_lorSL_RS_483   ??????????????????????????????????????????????????
V_lorSL_RS_507   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sebSL_RS_534   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sebSL_RS_552   CGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_spRADS4_RS_213 CGTTTGGGCAAAGCAACAAGGTAATAGAGTTTGGAAAGATGTAGTGATGA
V_spRADS4_RS_214 CGTTTGGGCAAAGCAACAAGGTAATAGAGTTTGGAAAGATGTAGTGATGA
V_sur_RS_082     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sur_RS_083     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sur_RS_084     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sur_RS_216     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sur_RS_248     ??????????????????????????????????????????????????
V_sur_RS_324     ??????????????????????????????????????????????????
V_sur_RS_428     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sur_RS_489     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA
V_sur_RS_501     CGTCTGGGCAAAGCAACAAGGTAATAGATTTTGGAAAGATGTAGTGATGA

C_cap_RS_551     ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGGACACCATTTA
C_debilis_7209   ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACATCATTTA
I_juru_RS_451    ATGCTAATATTGAGCACTTGATGTCCATAGTAAGAGAGGCACACCATTTA
I_laev_RS_460    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAAAGGCACACCATTTA
V_cadu_JJ_847    ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_spRADS3_RS_335 ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_spRADS3_RS_339 ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_peru_JJ_772    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_calo_RS_430    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_calo_RS_454    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_calo_RS_481    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_calo_RS_511    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_spRADS1_RS_432 ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_spRADS1_RS_500 ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_spRADS1_R1_510 ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_spRADS1_RS_561 ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_dix_RS_225     ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_elon_RS_437    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_elon_RS_494    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
```

```
V_elon_RS_502     ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_flex_RS_442     ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_flex_RS_522     ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCAYACCATTTA
V_flex_RS_595     ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_multin_RS_107   ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_multin_RS_108   ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_loret_RS_480    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_loret_RS_516    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_loret_RS_526    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_multi_RS_350    ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_multi_RS_429    ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_multi_RS_466    ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_multi_RS_543    ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_multi_RS_549    ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_sebBL_RS_584    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_lorSL_RS_483    ????????????????????????????????????????????????
V_lorSL_RS_507    CTGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_sebSL_RS_534    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_sebSL_RS_552    ATGCTAATATTGAGCACTCGATGTCCATAGTAAGAGAGGCACACCATTTA
V_spRADS4_RS_213 ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGGGAAGCACACCATTTA
V_spRADS4_RS_214 ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGGGAAGCACACCATTTA
V_sur_RS_082      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAKCACACCATTTA
V_sur_RS_083      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_sur_RS_084      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_sur_RS_216      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAATCACACCATTTA
V_sur_RS_248      ????????????????????????????????????????????????
V_sur_RS_324      ????????????????????????????????????????????????
V_sur_RS_428      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAKCACACCATTTA
V_sur_RS_489      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA
V_sur_RS_501      ATGCTAATATTGAGCACTCGATGTCCGTAGTAAGAGAAGCACACCATTTA

C_cap_RS_551      TTTATTGGGAACTACCCTGGATGGGAGGGGAATATT-CCAAAAATTTCAG
C_debilis_7209    TTTATTGGGAACTACCCTGGATGGGAGGGGAATATT-CCAAAAATTTCAG
I_juru_RS_451     TTTATTGGGAACTGCCCTGGATGGGTGGGGAATATT-CAAAAAATTTCAG
I_laev_RS_460     TTTATTGGGAACTGCCCTGGATGGGTGGGGAATATT-CAAAAAATTTCAG
V_cadu_JJ_847     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_spRADS3_RS_335 TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_spRADS3_RS_339 TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_peru_JJ_772     TTTATTGGGAACTMCCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_calo_RS_430     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAAWTTCAG
V_calo_RS_454     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAAWTTCAG
V_calo_RS_481     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAAWTTCAG
V_calo_RS_511     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_spRADS1_RS_432 TTTATTGGGAACT-CCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_spRADS1_RS_500 TTTATTGGGAACT-CCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_spRADS1_R1_510 TTTATTGGGAACT??????????????????????????????????????
V_spRADS1_RS_561 TTTATTGGGAACT-CCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_dix_RS_225      TTTATTGGGAACTACMCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_elon_RS_437     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_elon_RS_494     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_elon_RS_502     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_flex_RS_442     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_flex_RS_522     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_flex_RS_595     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_multin_RS_107   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_multin_RS_108   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_loret_RS_480    TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
```

```
V_loret_RS_516   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_loret_RS_526   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_multi_RS_350   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_multi_RS_429   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_multi_RS_466   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_multi_RS_543   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_multi_RS_549   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAA
V_sebBL_RS_584   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_lorSL_RS_483   ??????????????????????????????????????????????????
V_lorSL_RS_507   TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTGCAAAAAATTTCAG
V_sebSL_RS_534   TTTATTGGGAACTACCCTGGATCGGCGGGGAATATTGCAAAAAATTTCAG
V_sebSL_RS_552   TTTATTGGGAACTACCCTGGATCGGCGGGGAATATTGCAAAAAATTTCAG
V_spRADS4_RS_213 TTTATTGGGAACTACCCTGGATGGGCGGGGACTATTCCAAAAA-TTTCAG
V_spRADS4_RS_214 TTTATTGGGAACTACCCTGGATGGGCGGGGACTATTCCAAAAA-TTTCAG
V_sur_RS_082     TTTATTGGGAACTACCYTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG
V_sur_RS_083     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG
V_sur_RS_084     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG
V_sur_RS_216     TTTATTGGGAACTACCTTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG
V_sur_RS_248     ??????????????????????????????????????????????????
V_sur_RS_324     ??????????????????????????????????????????????????
V_sur_RS_428     TTTATTGGGAACTACCYTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG
V_sur_RS_489     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG
V_sur_RS_501     TTTATTGGGAACTACCCTGGATGGGCGGGGAATATTCCAAAAA-TTTCAG

C_cap_RS_551     CAGTGGGACAATCCTAGACACTGATATTGAGAACTTTTTCCCCAGTGAAT
C_debilis_7209   CAGTGGGACAATCCTAGACACTGATGTTGAGAACTTTTTCCCCAGTGAAT
I_juru_RS_451    CAGTGGGACAATCCTAGACACTGATTTTGAGAACTTTTGCCCCAGTGACT
I_laev_RS_460    CAGTGGGACAATCCTAGACACTGATTTTGAGAACTTTTGCCCCAGTGACT
V_cadu_JJ_847    CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_spRADS3_RS_335 CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_spRADS3_RS_339 CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_peru_JJ_772    CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_calo_RS_430    CAGTGGGACTATCCTMGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_calo_RS_454    CAGTGGGACTATCCTMGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_calo_RS_481    CAGTGGGACTATCCTMGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_calo_RS_511    CAGTGGTACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_spRADS1_RS_432 CAGTGGGGCTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_spRADS1_RS_500 CAGTGGGGCTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_spRADS1_R1_510 ??????????????????????????????????????????????????
V_spRADS1_RS_561 CAGTGGGGCTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_dix_RS_225     CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_elon_RS_437    CAGTGGGACTATCCTAGRCACTGATATTGACAACTTTTGCCCCAGTGAAT
V_elon_RS_494    CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_elon_RS_502    CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_flex_RS_442    CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_flex_RS_522    CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_flex_RS_595    CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multin_RS_107  CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multin_RS_108  CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_loret_RS_480   CAGTGGGACTRTCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_loret_RS_516   CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_loret_RS_526   CAGTSGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multi_RS_350   CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multi_RS_429   CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multi_RS_466   CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multi_RS_543   CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_multi_RS_549   CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
```

```
V_sebBL_RS_584   CAGTGGGACTATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_lorSL_RS_483   ??????????????????????????????????????????????????
V_lorSL_RS_507   CAGTGGGACTGTCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_sebSL_RS_534   CAGTGGGACTATCCTAGACACTGATATTGACTACTTTTGCCCCAGTGAAT
V_sebSL_RS_552   CAGTGGGACTATCCTAGACACTGATATTGACTACTTTTGCCCCAGTGAAT
V_spRADS4_RS_213 CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_spRADS4_RS_214 CAGTGGGACAATCCTAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_082     CAGTGGGACAATCATAGACACTKATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_083     CAGTGGGACAATCATAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_084     CAGTGGGACAATCATAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_216     CAGTGGGACAATCATAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_248     ??????????????????????????????????????????????????
V_sur_RS_324     ??????????????????????????????????????????????????
V_sur_RS_428     CAGTGGGACAATCATAGACACTKATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_489     CAGTGGGACAATCATAGACACTGATATTGACAACTTTTGCCCCAGTGAAT
V_sur_RS_501     CAGTGGGACAATCATAGACACTGATATTGACAACTTTTGCCCCAGTGAAT

C_cap_RS_551     GCTGACCATATATGCTTTTT-G--AACTTCCTATTTATAATCCTCAAAAC
C_debilis_7209   GCTGACCATATATGTTTTTT-G--AGCTTTCTATTTATAATCCTCAAAAT
I_juru_RS_451    GTTGACCATATATGTTTTTT-G--AACTTCCTATTTATAATCCTCAAAAT
I_laev_RS_460    GTTGACCATATATGTTTTTT-G--AACTTCCTATTTATAATCCTCAAAAT
V_cadu_JJ_847    GTTGACCATATRTATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_spRADS3_RS_335 GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_spRADS3_RS_339 GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_peru_JJ_772    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_calo_RS_430    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_calo_RS_454    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_calo_RS_481    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_calo_RS_511    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_spRADS1_RS_432 GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_spRADS1_RS_500 GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_spRADS1_R1_510 ????????????KTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_spRADS1_RS_561 GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCAAAAT
V_dix_RS_225     GTTGACCATATGTGTTTTTTTTT-AACTTTCCATTTATAATCCTCAAAAT
V_elon_RS_437    GTTGACCATATGTGTTTTTTT??????????????????????????????
V_elon_RS_494    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTTAAAAT
V_elon_RS_502    GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCT?AAAAT
V_flex_RS_442    GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_flex_RS_522    GTTGACCATATRTATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_flex_RS_595    GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_multin_RS_107  GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_multin_RS_108  GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_loret_RS_480   GTTGACCATATGTGTTTTTTTG--AACTTTCCATKTATAATCCTCAAAAT
V_loret_RS_516   GTTGACCATATATGTGTTTTTTTGAACTTTCCATKTATAATCCTCAAAAT
V_loret_RS_526   GTTGACCATATGTGTTTTTTTTTTAACTTTCCATTTATAATCCTCAAAAT
V_multi_RS_350   GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_multi_RS_429   GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_multi_RS_466   GTTGACCATATATATTTTTCTR--AACTTTCCATTTATAATCCTCAAAAT
V_multi_RS_543   GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_multi_RS_549   GTTGACCATATATATTTTTCTG--AACTTTCCATTTATAATCCTCAAAAT
V_sebBL_RS_584   GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTYAAAAT
V_lorSL_RS_483   ??????????????????????????????????????????????????
V_lorSL_RS_507   GTTGACCATATGTGTTTTTTTG--AACTTTCCATTTATAATCCTCCAAAT
V_sebSL_RS_534   GTTGACCATATGTGTTTTTTTCTTAACTTTCCATTTATAATCCTCAAAAT
V_sebSL_RS_552   GTTGACCATATGTGTTTTTTTCTTAACTTTCCATTTATAATCCTCAAAAT
V_spRADS4_RS_213 GTTGACCACATATGTTTTTTTG--AACTT-----------TCCTCAAAAT
V_spRADS4_RS_214 GTTGACCACATATGTTTTTTTG--AACTT-----------TCCTCAAAAT
```

```
V_sur_RS_082      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT
V_sur_RS_083      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT
V_sur_RS_084      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT
V_sur_RS_216      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT
V_sur_RS_248      ??????????????????????????????----------??????????
V_sur_RS_324      ??????????????????????????????----------??????????
V_sur_RS_428      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT
V_sur_RS_489      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT
V_sur_RS_501      GTTGATCATATATGTTTTTTTG--AACTT----------TCCTCAAAAT

C_cap_RS_551      TCCTCCAAACATCAGCTAAAATCATCATATTCATTTACTTTTAGGCTGTG
C_debilis_7209    TCCTCCAAACAGCAGCTAAAATCATTGTATTCATTTACTTTTAGGCTGTG
I_juru_RS_451     TCCTCCATACAGCTGCTAAAATCATCATATTCGTTTACTTTTTGGCTGTG
I_laev_RS_460     TCCTCCATACAGCTGCTAAAATCATCATATTCGTTTACTTTTCGGCTGTG
V_cadu_JJ_847     TCCTCCAAACAGCAGCTAAARTCATCATATTCGTTTACTTTTAGGCTGTG
V_spRADS3_RS_335  TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_spRADS3_RS_339  TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_peru_JJ_772     TCCTCCAAACAGTASCTAAAATCATTATAKTCTTTTACTTTAAGGCTGTG
V_calo_RS_430     TCCTCCAAACAGTAGCTAAAATCATTATATTATTTTACTTTAAGGCTGTG
V_calo_RS_454     TCCTCCAAACAGTAGCTAAAATCATTATATTATTTTACTTTAAGGCTGTG
V_calo_RS_481     TCCTCCAAACAGTAGCTAAAATCATTATATTATTTTACTTTAAGGCTGTG
V_calo_RS_511     TCCTCCAAACAGTAGCTAAAATCATTATATTATTTTACTTTAAGGCTGTG
V_spRADS1_RS_432  TCCTCCAAATAGTAGCTAAAATCGTTATATTCTTTTACTTTAAGGCTGTG
V_spRADS1_RS_500  TCCTCCAAATAGTAGCTAAAATCGTTATATTCTTTTACTTTAAGGCTGTG
V_spRADS1_R1_510  TCCTCCAAA??GTAGCTAAAATCGTTATATTCTTTTACTTTAAGGCTGTG
V_spRADS1_RS_561  TCCTCCAAACAGTAGCTAAAATCGTTATATTCTTTTACTTTAAGGCTGTG
V_dix_RS_225      TCCTCCAAACAGTAGCTAAAATCRTTATATTCTTTTACTTTAAGGCTGTG
V_elon_RS_437     ??????????????????????????????????????????????????
V_elon_RS_494     TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_elon_RS_502     TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_flex_RS_442     TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_flex_RS_522     TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_flex_RS_595     TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_multin_RS_107   TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_multin_RS_108   TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_loret_RS_480    TCCTCCAAACAGTAGCTAAAATCRTTATATTCTTTTACTTTAAGGCTGTG
V_loret_RS_516    TCCTCCAAACAGTAGCTAAAATCATTATAKTCTTTTACTTTAAGGCTGTG
V_loret_RS_526    TCCTCCAAACAGTAGCTAAAATCATTATAGTCTTTTACTTTAAGGCTGTG
V_multi_RS_350    TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_multi_RS_429    TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_multi_RS_466    TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_multi_RS_543    TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_multi_RS_549    TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sebBL_RS_584    TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_lorSL_RS_483    ??????????????????????????????????????????????????
V_lorSL_RS_507    TCCTCCAAACAGTAGCTAAAATCGTTATATTCTTTTACTTTAAGGCTGTG
V_sebSL_RS_534    TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_sebSL_RS_552    TCCTCCAAACAGTAGCTAAAATCATTATATTCTTTTACTTTAAGGCTGTG
V_spRADS4_RS_213  TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_spRADS4_RS_214  TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sur_RS_082      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sur_RS_083      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sur_RS_084      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sur_RS_216      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sur_RS_248      ??????????????????????????????????????????????????
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
```

```
V_sur_RS_489      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG
V_sur_RS_501      TCCTCCAAACAGCAGCTAAAATCATCATATTCGTTTACTTTTAGGCTGTG

C_cap_RS_551      ACCCGCCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
C_debilis_7209    ACCTGCCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATT-GT
I_juru_RS_451     ACCCGCCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
I_laev_RS_460     ACCCGCCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_cadu_JJ_847     ACYCACCCAATGGATGAATGGTTCAATTCTCACRAAGGTGTGACATTTGT
V_spRADS3_RS_335  ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_spRADS3_RS_339  ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_peru_JJ_772     ACCCACCCAATGGATRAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_calo_RS_430     ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_calo_RS_454     ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_calo_RS_481     ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_calo_RS_511     ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_spRADS1_RS_432  ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_spRADS1_RS_500  ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_spRADS1_R1_510  ACCCACCCAATGGAT?AATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_spRADS1_RS_561  ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_dix_RS_225      ACCCACCCAATGGATAWATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_elon_RS_437     ??????????????????????????????????????????ACATTTGT
V_elon_RS_494     ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_elon_RS_502     ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_flex_RS_442     ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_flex_RS_522     ACYCACCCAATGGATGAATGGTTCAATTCTCACRAAGGTGTGACATTTGT
V_flex_RS_595     ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_multin_RS_107   ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_multin_RS_108   ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_loret_RS_480    ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_loret_RS_516    ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_loret_RS_526    ACCCRCCCAATGGATAAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_multi_RS_350    ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_multi_RS_429    ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_multi_RS_466    ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_multi_RS_543    ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_multi_RS_549    ACTCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_sebBL_RS_584    ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_lorSL_RS_483    ??????????????????????????????????????????????????
V_lorSL_RS_507    ACCCACCCAATGGATGAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_sebSL_RS_534    ACCCACCCAATGGATAAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_sebSL_RS_552    ACCCACCCAATGGATAAATGGTTCAATTCTCATGAAGGTGTGACATTTGT
V_spRADS4_RS_213  ATCCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_spRADS4_RS_214  ATCCACCCAATGGATGAATGGTTCAATTCTCACGAAGGTGTGACATTTGT
V_sur_RS_082      ACCCACCCAATGGATGAATAGTTCRGTTCTCATGAAGGTGTGACATTTGT
V_sur_RS_083      ACCCACCCAATGGATGAATAGTTCAGTTCTCATGAAGGTGTGACATTTGT
V_sur_RS_084      ACCCACCCAATGGATGAATAGTTCAGTTCTCATGAAGGTGTGACATTTGT
V_sur_RS_216      ACCCACCCAATGGATGAATAGTTCAGTTCTCATGAAGGTGTGACATTTGT
V_sur_RS_248      ??????????????????????????????????????????????????
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      ACCCACCCAATGGATGAATAGTTCRGTTCTCATGAAGGTGTGACATTTGT
V_sur_RS_489      ACCCACCCAATGGATGAATAGTTCAGTTCTCATGAAGGTGTGACATTTGT
V_sur_RS_501      ACCCACCCAATGGATGAATAGTTCAGTTCTCATGAAGGTGTGACATTTGT

C_cap_RS_551      AAGATGATAAGA-GTGCC?GAGATCTAACAAATAAGGACTGCAGGAAAAT
C_debilis_7209    AAGATGATAAGA-GTGCC?GAGATCTAACAAATAAGGACTGCAGGAAAAT
I_juru_RS_451     AAGATGATAAAGAGTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
I_laev_RS_460     AAGATGATAAAGAGTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
```

```
V_cadu_JJ_847     AAGATGATAAGA-GTRCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_spRADS3_RS_335 AAGATGATATGA-GTGCCTGAGACGTAACAAGTAAGGAGTGCAGGAAAAA
V_spRADS3_RS_339 AAGATGATATGA-GTGCCTGAGACGTAACAAGTAAGGAGTGCAGGAAAAA
V_peru_JJ_772     AAGATGATATGA-GTGCCTGAGACGTAACAMATAAGGAGTGCRGGAAAAA
V_calo_RS_430     AAGATGATATGA-GTGCCTGAGACGTAACAAGTAAGGAGTGCAGGAAAAA
V_calo_RS_454     AAGATGATATGA-GTGCCTGAGACGTAACAAGTAAGGAGTGCAGGAAAAA
V_calo_RS_481     AAGATGATATGA-GTGCCTGAGACGTAACAAGTAAGGAGTGCAGGAAAAA
V_calo_RS_511     AAGATGATATGA-GTGCCTGAGACGTAACAAGTAAGGAGTGCAGGAAAAA
V_spRADS1_RS_432 AAGATGATATGA-GTGCCTGAGACATAACAAATAAGGAGTGCAGGAAAGA
V_spRADS1_RS_500 AAGATGATATGA-GTGCCTGAGACATAACAAATAAGGAGTGCAGGAAAGA
V_spRADS1_R1_510 AAGATGATATGA-GTGCCTGAGACATAACAAATAAGGAGTGCAGGAAA?A
V_spRADS1_RS_561 AAGATGATATGA-GTGCCTGAGACATAACAAATAAGGAGTGCAGGAAAGA
V_dix_RS_225      AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAARA
V_elon_RS_437     AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_elon_RS_494     AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_elon_RS_502     AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_flex_RS_442     AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_flex_RS_522     AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_flex_RS_595     AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_multin_RS_107   AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_multin_RS_108   AAGATGATAAGA-GTRCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_loret_RS_480    AAGATGATAWGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAGA
V_loret_RS_516    AAGATGATAKGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAARA
V_loret_RS_526    AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_multi_RS_350    AAGATGATAAGA-GTGYCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_multi_RS_429    AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_multi_RS_466    AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_multi_RS_543    AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_multi_RS_549    AAGATGATAAGA-GTRCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_sebBL_RS_584    AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_lorSL_RS_483    ??????????????????????????????????????????????????
V_lorSL_RS_507    AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAGA
V_sebSL_RS_534    AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_sebSL_RS_552    AAGATGATATGA-GTGCCTGAGACGTAACAAATAAGGAGTGCAGGAAAAA
V_spRADS4_RS_213 AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_spRADS4_RS_214 AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAT
V_sur_RS_082      AAGATGATAARA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
V_sur_RS_083      AAGATGATAAAA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
V_sur_RS_084      AAGATGATAAAA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
V_sur_RS_216      AAGATGATAAGA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
V_sur_RS_248      ??????????????????????????????????????????????????
V_sur_RS_324      ??????????????????????????????????????????????????
V_sur_RS_428      AAGATGATAARA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
V_sur_RS_489      AAGATGATAAAA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA
V_sur_RS_501      AAGATGATAAAA-GTGCCTGAGATCTAACAAATAAGGACTGCAGGAAAAA

C_cap_RS_551      TTTCTCATATAAAT--GACTTGGAATGTCTATAGACAGAGAAATGGCCTT
C_debilis_7209    TTTCTCATATACAT--GACTTGGAATGTCTATAGACAGAGAAATGGCCTT
I_juru_RS_451     TTTCTCATATATATATGATTTGAAATGTCTATCGACAGA--AATGGCCTT
I_laev_RS_460     TTTCTCATATATATATGATTTGAAATGTCTATCGACAGA--AATGGCCTT
V_cadu_JJ_847     TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_spRADS3_RS_335 TTTCTCATATTTAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS3_RS_339 TTTCTCATATTTAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_peru_JJ_772     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_calo_RS_430     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_calo_RS_454     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_calo_RS_481     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
```

```
V_calo_RS_511     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS1_RS_432  TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS1_RS_500  TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS1_R1_510  T?TCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS1_RS_561  TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_dix_RS_225      TTTCTCATATATAT--GATTTGWGATGTCTATAKACAGAGAAATGGCCTT
V_elon_RS_437     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_elon_RS_494     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_elon_RS_502     TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_flex_RS_442     TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_flex_RS_522     TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_flex_RS_595     TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_multin_RS_107   TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_multin_RS_108   TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_loret_RS_480    TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_loret_RS_516    TYTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_loret_RS_526    TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_multi_RS_350    TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_multi_RS_429    TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_multi_RS_466    TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_multi_RS_543    TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_multi_RS_549    TTTCTCTTATATAT--GATTTGAGATGTCCATAGACAAAGAAATGGCCTT
V_sebBL_RS_584    TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_lorSL_RS_483    ????????????????--??????????????????????????????????
V_lorSL_RS_507    TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sebSL_RS_534    TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sebSL_RS_552    TTTCTCATATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS4_RS_213  TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_spRADS4_RS_214  TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_082      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_083      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_084      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_216      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_248      ????????????????--??????????????????????????????????
V_sur_RS_324      ????????????????--??????????????????????????????????
V_sur_RS_428      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_489      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT
V_sur_RS_501      TTTCTCTTATATAT--GATTTGAGATGTCTATAGACAGAGAAATGGCCTT

C_cap_RS_551      TCACTCATGTTCTTTGTCCTTTTTAAATATCTTCTACGCCACATACTTGT
C_debilis_7209    TCACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATA?TTGT
I_juru_RS_451     TCACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACCTGT
I_laev_RS_460     TCACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACGTACTTGT
V_cadu_JJ_847     TCACTAATGTTCTTTGTCCTTTATACAT????????????ATACTTGT
V_spRADS3_RS_335  -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_spRADS3_RS_339  -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_peru_JJ_772     -CACTA??????????????????????ATCTTCTATGCCACATACTTGT
V_calo_RS_430     -CACTAATGTTCTTTGTCCTTATTAAATATCTTCTATGCCACATACTTGT
V_calo_RS_454     -CACTAATGTTCTTTGTCCTTA?TAA?TATCTTCTATGCCACATACTTGT
V_calo_RS_481     -CACTAATGTTCTTTGTCCTTATTAAATATCTTCTATGCCACATACTTGT
V_calo_RS_511     -CACTAATGTTCTTTGTCCTTT??????ATCTTCTATGCCACATACTTGT
V_spRADS1_RS_432  -CACTAATGTTCT????????????????ATCTTCTATGCCACATACTTGT
V_spRADS1_RS_500  -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_spRADS1_R1_510  -CACTAATGTTCTT????????????????ATCTTCTATGCCACATACTTGT
V_spRADS1_RS_561  -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_dix_RS_225      -CACTAATGTTCTT????????????????ATCTTCTATGCCACATACTTGT
V_elon_RS_437     -CAC????????????????????????ATCTTCTATGCCACATACTTGT
```

```
V_elon_RS_494     -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_elon_RS_502     -CACTAACGTTCTT????????????????ATCTTCTATGCCACATACTTGT
V_flex_RS_442     TCACTAATGTTCTTTGTCCTTTATACATATCTTCTACGCCACATACTTGT
V_flex_RS_522     TCACTAATGTTCTTTGTCCTTTATACAT?????????????????????
V_flex_RS_595     TCACTAATGTTCTTTGTCCTTTATACAT??????????????????????
V_multin_RS_107   TCACTAATGTTCTTTGTCCTTTATACATAT?TTCTACGCCACATACTTGT
V_multin_RS_108   TCACTAATGTTCTTTGYCCTTTATACATATSTTCTACGCCACATACTTGT
V_loret_RS_480    -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_loret_RS_516    -CACTAA??????????????????????ATCTTCTATGCCACATACTTGT
V_loret_RS_526    -CACTA??GTTCTTTGYCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_multi_RS_350    TCACTAATGTTCTTTGTCCTTTATACATATCTTCTACGCCACATACTTGT
V_multi_RS_429    TCACTAATGTTCTTTGTCCTTTATACATATCTTCT?YGCCACATACTTGT
V_multi_RS_466    TCACTAATGTTCTTTGTCCTTTATACATATCTTCTACGCCACATACTTGT
V_multi_RS_543    TCACTAATGTTCTTTGTCCTTTATACATATCTTCTATGCCACATACTTGT
V_multi_RS_549    TCACTAATGTTCTTTGTCCTTTATACATATCTTCTATGCCACATACTTGT
V_sebBL_RS_584    -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_lorSL_RS_483    ????????????????????????????ATCTTCTATGCCACATACTTGT
V_lorSL_RS_507    -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_sebSL_RS_534    -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_sebSL_RS_552    -CACTAATGTTCTTTGTCCTTTTTAAATATCTTCTATGCCACATACTTGT
V_spRADS4_RS_213  TCACTAATGTTCTTTGTCCTTTATAAATATCTTCTACGCCACATACTTGT
V_spRADS4_RS_214  TCACTAATGTTCTTTGTCCTTTATAAATATCTTCTACGCCACATACTTGT
V_sur_RS_082      TCACTAATGTTCTTTGTCCTTTAT?AATATCTTCTACGCCACATACTTGT
V_sur_RS_083      TCACTAATGTTCTTTGTCCTTTATGAATATCTTCTACGCCACATACTTGT
V_sur_RS_084      TCACTAATGTTCTTTGTCCTTTATGAATAT?TTCTACGCCACATACTTGT
V_sur_RS_216      TCACTAATGTTCTTTGTCCTTTATAAATATCTTCTACGCCACATACTTGT
V_sur_RS_248      ????????????????????????????ATCTTCTACGCCACATACTTGT
V_sur_RS_324      ????????????????????????????ATCTTCTACGCCACATACTTGT
V_sur_RS_428      TCACTAATGTTCTTTGTCCTTTATGAATATCTTCTACGCCACATACTTGT
V_sur_RS_489      TCACTAATGTTCTTTGTCCTTTATGAATATCTTCTACGCCACATACTTGT
V_sur_RS_501      TCACTAATGTTCTTTGTCCTTTATGAATATCTTCTACGCCACATACTTGT

C_cap_RS_551      CTG?GAAAATTGGGTATTGGAGGTACATAACTATTTACAGGCATTTGAAG
C_debilis_7209    CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACAGGCATTTGAAG
I_juru_RS_451     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACAGGCATTTGAAG
I_laev_RS_460     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACAGGCATTTGAAG
V_cadu_JJ_847     CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_spRADS3_RS_335  CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_spRADS3_RS_339  CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_peru_JJ_772     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_calo_RS_430     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_calo_RS_454     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_calo_RS_481     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_calo_RS_511     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_spRADS1_RS_432  CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_spRADS1_RS_500  CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_spRADS1_R1_510  CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_spRADS1_RS_561  CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_dix_RS_225      CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_elon_RS_437     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_elon_RS_494     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_elon_RS_502     CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_flex_RS_442     CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_flex_RS_522     ???????????????TGGAGGTACATAACCATTTACAGGCATTTGAAG
V_flex_RS_595     ???????????????????GAGGTACATAACCATTTACAGGCATTTGAAG
V_multin_RS_107   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_multin_RS_108   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
```

```
V_loret_RS_480   CTGAGAAAATTGGGTATTGGMGGTACATAACTATTTACCGGCATTTGAAG
V_loret_RS_516   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_loret_RS_526   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_multi_RS_350   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_multi_RS_429   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_multi_RS_466   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_multi_RS_543   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_multi_RS_549   CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_sebBL_RS_584   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_lorSL_RS_483   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_lorSL_RS_507   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_sebSL_RS_534   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_sebSL_RS_552   CTGAGAAAATTGGGTATTGGAGGTACATAACTATTTACCGGCATTTGAAG
V_spRADS4_RS_213 CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_spRADS4_RS_214 CTGAGAAAATTGGGTATTGGAGGTACATAACCATTTACAGGCATTTGAAG
V_sur_RS_082     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_083     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_084     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_216     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_248     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_324     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_428     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_489     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG
V_sur_RS_501     CTGAGAAAATTGGGTATTGGAGGTACATCACTATTTACAGGCATTTGAAG

C_cap_RS_551     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
C_debilis_7209   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
I_juru_RS_451    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTTGAGAACTG
I_laev_RS_460    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTTGAGAACTG
V_cadu_JJ_847    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS3_RS_335 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS3_RS_339 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_peru_JJ_772    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_calo_RS_430    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_calo_RS_454    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_calo_RS_481    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_calo_RS_511    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS1_RS_432 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS1_RS_500 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS1_R1_510 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS1_RS_561 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_dix_RS_225     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_elon_RS_437    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_elon_RS_494    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_elon_RS_502    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_flex_RS_442    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_flex_RS_522    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_flex_RS_595    GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_multin_RS_107  GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_multin_RS_108  GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_loret_RS_480   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_loret_RS_516   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_loret_RS_526   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_multi_RS_350   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_multi_RS_429   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_multi_RS_466   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_multi_RS_543   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
```

```
V_multi_RS_549   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sebBL_RS_584   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_lorSL_RS_483   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_lorSL_RS_507   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sebSL_RS_534   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sebSL_RS_552   GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS4_RS_213 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_spRADS4_RS_214 GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_082     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_083     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_084     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_216     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_248     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_324     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_428     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_489     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG
V_sur_RS_501     GCCAACCCAGAGTACCAACTTTACCCAATCTTCAAGTATTTCGAGAACTG

C_cap_RS_551     GTGCCAGGATGAGAATAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
C_debilis_7209   GTGCCAGGATGAGAATAGGCATGGAGA?TTCTTCTCTGCATTGATGAAGG
I_juru_RS_451    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
I_laev_RS_460    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_cadu_JJ_847    GTGCCAGGATGAGAACAGGCAYGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS3_RS_335 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS3_RS_339 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_peru_JJ_772    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_calo_RS_430    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_calo_RS_454    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_calo_RS_481    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_calo_RS_511    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS1_RS_432 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS1_RS_500 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS1_R1_510 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS1_RS_561 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_dix_RS_225     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_elon_RS_437    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_elon_RS_494    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_elon_RS_502    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_flex_RS_442    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_flex_RS_522    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_flex_RS_595    GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multin_RS_107  GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multin_RS_108  GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_loret_RS_480   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_loret_RS_516   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_loret_RS_526   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multi_RS_350   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multi_RS_429   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multi_RS_466   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multi_RS_543   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_multi_RS_549   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sebBL_RS_584   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_lorSL_RS_483   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_lorSL_RS_507   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sebSL_RS_534   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sebSL_RS_552   GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_spRADS4_RS_213 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
```

```
V_spRADS4_RS_214 GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_082     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_083     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_084     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_216     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_248     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_324     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_428     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_489     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG
V_sur_RS_501     GTGCCAGGATGAGAACAGGCATGGAGATTTCTTCTCTGCATTGATGAAGG

C_cap_RS_551     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
C_debilis_7209   C?CAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
I_juru_RS_451    CACAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
I_laev_RS_460    CACAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_cadu_JJ_847    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS3_RS_335 CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS3_RS_339 CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_peru_JJ_772    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_calo_RS_430    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_calo_RS_454    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_calo_RS_481    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_calo_RS_511    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS1_RS_432 CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS1_RS_500 CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS1_R1_510 CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS1_RS_561 CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_dix_RS_225     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_elon_RS_437    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_elon_RS_494    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_elon_RS_502    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_flex_RS_442    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_flex_RS_522    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_flex_RS_595    CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_multin_RS_107  CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_multin_RS_108  CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_loret_RS_480   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGYTGTGGGCTCGTTTC
V_loret_RS_516   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_loret_RS_526   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_multi_RS_350   CGCAGCCTCAATTCCTCAATGACTGGAAAGCRAAGCTGTGGGCTCGTTTC
V_multi_RS_429   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_multi_RS_466   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_multi_RS_543   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_multi_RS_549   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sebBL_RS_584   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGTTGTGGGCTCGTTTC
V_lorSL_RS_483   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_lorSL_RS_507   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sebSL_RS_534   CRCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sebSL_RS_552   CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_spRADS4_RS_213 CGCAGCCTCAATTCCTCAATGACTGGAAAGCAAAGCTGTGGGCTCGTTTC
V_spRADS4_RS_214 CGCAGCCTCAATTCCTCAATGACTGGAAAGCAAAGCTGTGGGCTCGTTTC
V_sur_RS_082     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_083     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_084     CRCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_216     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_248     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_324     CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
```

```
V_sur_RS_428      CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_489      CGCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC
V_sur_RS_501      CRCAGCCTCAATTCCTCAATGACTGGAAAGCGAAGCTGTGGGCTCGTTTC

C_cap_RS_551      TTCTGTCTATCGGTAAATGTTATTCAATTTTATTATA----------TTT
C_debilis_7209    TTCTGTCTATCGGTAAATGTTATTCAATTTTATTATA----------TTT
I_juru_RS_451     TTCTGTCTATCGGTAAATGTTATTCAATTTTATTATA----------TTT
I_laev_RS_460     TTCTGTCTATCGGTAAATGTTATTCAATTTTATTATA----------TTT
V_cadu_JJ_847     TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_spRADS3_RS_335  TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_spRADS3_RS_339  TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_peru_JJ_772     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_calo_RS_430     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_calo_RS_454     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_calo_RS_481     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_calo_RS_511     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_spRADS1_RS_432  TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_spRADS1_RS_500  TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_spRADS1_R1_510  TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_spRADS1_RS_561  TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_dix_RS_225      TTCTGTCTATCAGTAAATGCTATTCAATTTCCTTAT----------ATTT
V_elon_RS_437     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_elon_RS_494     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_elon_RS_502     TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_flex_RS_442     TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_flex_RS_522     TTCTGTCTATCGGTAAATGTTATTCAATTTTCTWAT----------ATTT
V_flex_RS_595     TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_multin_RS_107   TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_multin_RS_108   TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_loret_RS_480    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_loret_RS_516    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_loret_RS_526    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_multi_RS_350    TTCTGTCTATCGGTAAATKTTATTCAATTTTCTTAT----------ATTT
V_multi_RS_429    TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_multi_RS_466    TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_multi_RS_543    TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_multi_RS_549    TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_sebBL_RS_584    TTCTGTCTATCAGTAAATGCTATTAAATTTTCTTAT----------ATTT
V_lorSL_RS_483    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_lorSL_RS_507    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_sebSL_RS_534    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_sebSL_RS_552    TTCTGTCTATCAGTAAATGCTATTCAATTTTCTTAT----------ATTT
V_spRADS4_RS_213  TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_spRADS4_RS_214  TTCTGTCTATCGGTAAATGTTATTCAATTTTCTTAT----------ATTT
V_sur_RS_082      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_083      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_084      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_216      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_248      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_324      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_428      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_489      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT
V_sur_RS_501      TTCTGTCTATCGGTAAATATTATTCAATTTTATTATTACTATTATTATTT

C_cap_RS_551      TTTAGCCTTTTCATATTTCACTGGGCATGATTATTCTGATTATTGGAGAC
C_debilis_7209    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGTAGAC
I_juru_RS_451     TTCAGCCTTTTCATATTTCACTGGGCATGATTCTTATGATTATTGGATAC
```

```
I_laev_RS_460    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTATGATTATTGGATAC
V_cadu_JJ_847    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGYTGGAGAC
V_spRADS3_RS_335 ATTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS3_RS_339 ATTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_peru_JJ_772    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_calo_RS_430    ATTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_calo_RS_454    ATTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_calo_RS_481    ATTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_calo_RS_511    ATTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS1_RS_432 TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS1_RS_500 TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS1_R1_510 TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS1_RS_561 TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_dix_RS_225     TTTAGCCTTTTCGTATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_elon_RS_437    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_elon_RS_494    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_elon_RS_502    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_flex_RS_442    TTTAGCCTTTTCATATTTCACTGGGCTTGATTCTTCTGATTGTTGGAGAC
V_flex_RS_522    TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGTTGGAGAC
V_flex_RS_595    TTTAGCCTTTTCATATTTCACTGGGCWTGATTCTTCTGATTGTTGGAGAC
V_multin_RS_107  TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGTTGGAGAC
V_multin_RS_108  TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGTTGGAGAC
V_loret_RS_480   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_loret_RS_516   TTTAGCCTTTTCATATTTSACTGGGCATGATTCTTCTGATTATTGGAGAC
V_loret_RS_526   TTTAGCCTTTTCATATTTSACTGGGCATGATTCTTCTGATTATTGGAGAC
V_multi_RS_350   TTTAGCCTTTTCATATWTCACTGGGCRWGATTCTTCTGATTGTTGGAGAC
V_multi_RS_429   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTRTTGGAGAC
V_multi_RS_466   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGTTGGAGAC
V_multi_RS_543   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGTTGGAGAC
V_multi_RS_549   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTGTTGGAGAC
V_sebBL_RS_584   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_lorSL_RS_483   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_lorSL_RS_507   TTTAGCCTTTTCA?ATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_sebSL_RS_534   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_sebSL_RS_552   TTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS4_RS_213 CTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_spRADS4_RS_214 CTTAGCCTTTTCATATTTCACTGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_082     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_083     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_084     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_216     TTTTGCCTTWWCATATTTCACCGGGGWTGATTCTTCTGATTATTGGAGAC
V_sur_RS_248     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_324     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_428     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_489     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC
V_sur_RS_501     TTTTGCCTTTTCATATTTCACCGGGCATGATTCTTCTGATTATTGGAGAC

C_cap_RS_551     TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
C_debilis_7209   TCACAATCTGTTGCTTCAAG?TTTTGATTTTTCATGGATGAAG??GTTTC
I_juru_RS_451    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTAGTTTC
I_laev_RS_460    TCACAATCTGT?GCTTCCAGCTTTTGATTTTTCATGGATGAAGTAGTTTC
V_cadu_JJ_847    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_spRADS3_RS_335 TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_spRADS3_RS_339 TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_peru_JJ_772    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_calo_RS_430    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_calo_RS_454    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
```

```
V_calo_RS_481     TCACAATCTGTTGSTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_calo_RS_511     TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_spRADS1_RS_432  TYACAATCTGTTGCTTCCAGCTTTTGAYTTTTCATGGATGAAGTGGTTTC
V_spRADS1_RS_500  TYACAATCTGTTGCTTCCAGCTTTTGAYTTTTCATGGATGAAGTGGTTTC
V_spRADS1_R1_510  TTACAATCTGTTGCTTCCAGCTTTTGAYTTTTCATGGATGAAGTGGTTTC
V_spRADS1_RS_561  TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_dix_RS_225      TTACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_elon_RS_437     TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_elon_RS_494     TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_elon_RS_502     TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_flex_RS_442     TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_flex_RS_522     TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTY
V_flex_RS_595     TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTT?
V_multin_RS_107   TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_multin_RS_108   TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_loret_RS_480    TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_loret_RS_516    TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_loret_RS_526    TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_multi_RS_350    TCACAATCTGTTGCTTCCAGYTTTTGATTTTTCATGGATGAAGTGGTTTC
V_multi_RS_429    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAARTGGTTTC
V_multi_RS_466    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_multi_RS_543    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_multi_RS_549    TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_sebBL_RS_584    TTACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_lorSL_RS_483    TTACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_lorSL_RS_507    TTACAATCTGTTGCTTCCAGCTTTTGATTTTTCATG?ATGAAGTGGTTTC
V_sebSL_RS_534    TYACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_sebSL_RS_552    TTACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_spRADS4_RS_213  TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_spRADS4_RS_214  TCACAATCTGTTGCTTCCAGCTTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_082      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_083      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_084      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_216      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGA?GRAGTGGTTTC
V_sur_RS_248      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_324      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_428      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_489      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC
V_sur_RS_501      TCACAATCTGTTGCTTCCAGATTTTGATTTTTCATGGATGAAGTGGTTTC

C_cap_RS_551      AAAAGTGGATTTTGTTGGTCATCTGTTCTCTTTTTACGTTGCTTATGCCT
C_debilis_7209    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGT?GCTTATGCTT
I_juru_RS_451     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTCAT?T??CT?A???TT
I_laev_RS_460     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTCATGT?GCT?ATGCTT
V_cadu_JJ_847     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_spRADS3_RS_335  AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGTTT
V_spRADS3_RS_339  AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGTTT
V_peru_JJ_772     AAAAGTGGATTTTGTAGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_calo_RS_430     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGTTT
V_calo_RS_454     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGTTT
V_calo_RS_481     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGTTT
V_calo_RS_511     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGTTT
V_spRADS1_RS_432  AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_spRADS1_RS_500  AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_spRADS1_R1_510  AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_spRADS1_RS_561  AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_dix_RS_225      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
```

```
V_elon_RS_437     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_elon_RS_494     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_elon_RS_502     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_flex_RS_442     AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_flex_RS_522     AA??????????????????????????????????????????????
V_flex_RS_595     ??????????????????????????????????????????????
V_multin_RS_107   AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_multin_RS_108   AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_loret_RS_480    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_loret_RS_516    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_loret_RS_526    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_multi_RS_350    AAAAGTGGATTTTGTGGGTCATCTGTTCWCTTTTTATGTTGCTTATGCTT
V_multi_RS_429    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTYGCTTATGCTT
V_multi_RS_466    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_multi_RS_543    AAAAGTGGATTCTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_multi_RS_549    AAAAGTGGATTCTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sebBL_RS_584    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_lorSL_RS_483    AAAAGTGGATTTTGTGGGTCATCTGTTMTCTTTTTATGTTGCTTATGCTT
V_lorSL_RS_507    AAAAGTGGATTTTGTGG????????????????????????????????
V_sebSL_RS_534    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sebSL_RS_552    AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_spRADS4_RS_213  AAAAGTGGATTTTGTGG-TCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_spRADS4_RS_214  AAAAGTGGATTTTGTGG-TCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_082      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_083      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_084      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_216      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_248      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_324      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTA?????????????
V_sur_RS_428      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_489      AAAAGTGGATTTTGTGGGTCATCTGTTCTCTTTTTATGTTGCTTATGCTT
V_sur_RS_501      AAAAGTGGATTTTGTGGGTCATCTGTTCTMTTTTTATGTTGCTTATGCTT

C_cap_RS_551      ???????????????
C_debilis_7209    ???????????????
I_juru_RS_451     ???????????????
I_laev_RS_460     ???????????????
V_cadu_JJ_847     CATTTTCTTGCAGGTAT
V_spRADS3_RS_335  CATTCTCTTGCAGGTCT
V_spRADS3_RS_339  CATTCTCTTGCAGGTCT
V_peru_JJ_772     CATTCTCTTGCAGGTCT
V_calo_RS_430     CATTCTCTTGCAGGTCT
V_calo_RS_454     CATTCTCTTGCAGGTCT
V_calo_RS_481     CATTCTCTTGCAGGTCT
V_calo_RS_511     CATTCTCTTGCAGGTCT
V_spRADS1_RS_432  CATTCTCTTGCAGGTCT
V_spRADS1_RS_500  CATTCTCTTGCAGGTCT
V_spRADS1_R1_510  CATTCTCTTGCAGGTCT
V_spRADS1_RS_561  CATTCTCTTGCAGGTCT
V_dix_RS_225      CATTCTCTTGCAGGTCT
V_elon_RS_437     CATTCTCTTGCAGGTCT
V_elon_RS_494     CATTCTCTTGCAGGTCT
V_elon_RS_502     CATTCTCTTGCAGGTCT
V_flex_RS_442     CATTTTCTTGCAGGTAT
V_flex_RS_522     ???????????????
V_flex_RS_595     ???????????????
V_multin_RS_107   CATTTTCTTGCAGGTAT
```

```
V_multin_RS_108  CATTTTCTTGCAGGTAT
V_loret_RS_480   CATTCTCTTGCAGGTCT
V_loret_RS_516   CATTCTCTTGCAGGTCT
V_loret_RS_526   CATTCTCTTGCAGGTCT
V_multi_RS_350   CATTTTCTTGCAGGTAT
V_multi_RS_429   CATTTTCTTGCAGGTAT
V_multi_RS_466   CATTTTCTTGCAGGTAT
V_multi_RS_543   CATTTTCTTGCAGGTAT
V_multi_RS_549   CATTTTCTTGCAGGTAT
V_sebBL_RS_584   CATTCTCTTGCAGGTCT
V_lorSL_RS_483   CATTCTCTTGCAGGTCT
V_lorSL_RS_507   ?????????????????
V_sebSL_RS_534   CATTCTCTTGCAGGTCT
V_sebSL_RS_552   CATTCTCTTGCAGGTCT
V_spRADS4_RS_213 CATTCTCTTGCAGGTAT
V_spRADS4_RS_214 CATTCTCTTGCAGGTAT
V_sur_RS_082     CATTCTCTTGCAGGTAT
V_sur_RS_083     CATTCTCTTGCAGGTAT
V_sur_RS_084     CATTCTCTTGCAGGTAT
V_sur_RS_216     C?????????????????
V_sur_RS_248     CATTCTCTTGCAGGTAT
V_sur_RS_324     ?????????????????
V_sur_RS_428     CATTCTCTTGCAGGTAT
V_sur_RS_489     CATTCTCTTGCAGGTAT
V_sur_RS_501     CATTCTCTTGCAGGTAT
```

Alignment 3.2

The trnH-psbA DNA nucleotide sequence alignment of 54 taxa and 363 characters used in NJ analyses in chapter 3. 50 base pairs per line.

```
C_cap_RS_551       TCCGCCCCTTGTCCTTT-CT-------AAAGAAAAAAA----TTTTCGAT
C_debilis_7209     TCCGCCCCTTGTCCTTT-CT-------AAAGAAAAAAA----TTTTAGAT
I_juru_RS_451      TCCGCCCCTTGTCTTTT-CTTTTCT--AAAGACAAAAA----TTTTAGAT
I_laev_RS_460      TCCGCCCCTTGTCTTTT-CTTTTCT--AAAGACAAAAA----TTTTAGAT
V_cadu_JJ_847      TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA--TTTAGAT
V_flex_RS_442      TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA--TTTAGAT
V_flex_RS_522      TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA--TTTAGAT
V_flex_RS_595      TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA--TTTAGAT
V_multi_RS_107     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAGAAAA--TTTAGAT
V_multi_RS_108     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAGAAAA--TTTAGAT
V_multi_RS_350     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA--TTTAGAT
V_multi_RS_429     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAGAAAA--TTTAGAT
V_multi_RS_466     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAGAAAA--TTTAGAT
V_multi_RS_543     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAGAAAA--TTTAGAT
V_multi_RS_549     TCCGCCCCTTGTCTTTTTCT-------AAAGAAAAAGAAAA--TTTAGAT
V_spRADS4_RS_213   TCCGCCCCTTGTCTTTT-CCTTTTTGTACAGAAAAAAA----TTTTGGAT
V_spRADS4_RS_214   TCCGCCCCTTGTCTTTT-CCTTTTTGTACAGAAAAAAA----TTTTGGAT
V_surRS_082        TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_sur_RS_083       TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_sur_RS_084       TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_pav_RS_216       TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_sur_RS_248       ????????????????-????????--??GAAAAAAA----TTTTAGAT
V_sur_RS_428       TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_sur_RS_489       TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_sur_RS_501       TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_multicost_1      ????????????????????????T-AAAGAAAAAAAAAAATTTTAGAT
V_multicost_2      ????????????????????TTTTTTT-AAAGAAAAAAAAAAATTTTAGAT
V_nobilis          TCCGCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_multiflora       ???GCCCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_michelii_110149  ?????CCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA-TTTTAGAT
V_michelii_200263  ?????????????????????????-?AAGAAAAAAAAAA-TTTTAGAT
V_michelii_110168  ?????CCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA-TTTTAGAT
V_michelii_110074  ?????CCCTTGTCTTTTTCT-------AAAGAAAAAAAAAA-TTTTAGAT
V_kwatae_110070    ?????CCCTTGTCTTTT-CTTTTTCT-AAAGAAAAAAA----TTTTAGAT
V_spRADS3_RS_335   TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_RADS3_RS_339     TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_peru_JJ_772      TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_calo_RS_430      TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_calo_RS_454      ??????CCTTGTCTTTT-CTTTTTCT-----------------------
V_calo_RS_511      TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_spRADS1_RS_432   TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_spRADS1_RS_500   TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_spRADS1_RS_510   TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_spRADS1_RS_561   TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_dix_RS_225       TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_elon_RS_437      ???????CTTGTCTTTT-CTTTTTCT-----------------------
V_elonRS_502       TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
V_lorBL_RS_480     TCC?CCCCTTGTCTTTT-CTTTTTCT-----------------------
V_lorBL_RS_526     TCCGCCCCTTGTCTTTT-CTTTTTCT-----------------------
```

```
V_sebBL_RS_584    TCCGCCCCTTGTCTTTT-CTTTTTCT----------------------
V_lorBL_RS_516    TCCGCCCCTTGTCTTTT-CTTTTTCT----------------------
V_lorSL_RS_483    TCCGCCCCTTGTCTTTT-CTTTTTCT----------------------
V_sebSL_RS_507    TCCGCCCCTTGTCTTTT-CTTTTTCT----------------------
V_sebSL_RS_552    TCCGCCCCTTGTCTTTT-CTTTTTCT----------------------

C_cap_RS_551      TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATC
C_debilis_7209    TACTAGTCTTTCTTATTTTTTTCATACAAATTTCTATCCTTTAGAAAA--
I_juru_RS_451     TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
I_laev_RS_460     TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_cadu_JJ_847     TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_flex_RS_442     TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_flex_RS_522     TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_flex_RS_595     TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_107    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_108    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_350    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_429    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_466    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_543    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multi_RS_549    TACTAGTCTTTCTTCTTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_spRADS4_RS_213  TTCTAGTCTTTCTTATTTTTTTCATACTAA--------------------
V_spRADS4_RS_214  TTCTAGTCTTTCTTATTTTTTTCATACTAA--------------------
V_surRS_082       TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_sur_RS_083      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_sur_RS_084      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_pav_RS_216      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_sur_RS_248      TACTAGTCTTTCTTATTTTTT-CATACTAATTCCTACCCTTTAGAAAATT
V_sur_RS_428      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_sur_RS_489      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_sur_RS_501      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_multicost_1     TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_multicost_2     TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_nobilis         TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_multiflora      TACTAGTCTTTCTTATTTTTT-CATACTAATTTCTACCCTTTAGAAAATT
V_michelii_110149 TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_michelii_200263 TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_michelii_110168 TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_michelii_110074 TACTAGTCTTTCTTATTTTTTTCATACTAATTTCTACCCTTTAGAAAATT
V_kwatae_110070   TACTAGTCTTTCTTATTTTTT-CATACTAATTCCTACCCTTTAGAAAATT
V_spRADS3_RS_335  --------------------------------------------------
V_RADS3_RS_339    --------------------------------------------------
V_peru_JJ_772     --------------------------------------------------
V_calo_RS_430     --------------------------------------------------
V_calo_RS_454     --------------------------------------------------
V_calo_RS_511     --------------------------------------------------
V_spRADS1_RS_432  --------------------------------------------------
V_spRADS1_RS_500  --------------------------------------------------
V_spRADS1_RS_510  --------------------------------------------------
V_spRADS1_RS_561  --------------------------------------------------
V_dix_RS_225      --------------------------------------------------
V_elon_RS_437     --------------------------------------------------
V_elonRS_502      --------------------------------------------------
V_lorBL_RS_480    --------------------------------------------------
V_lorBL_RS_526    --------------------------------------------------
V_sebBL_RS_584    --------------------------------------------------
V_lorBL_RS_516    --------------------------------------------------
```

```
V_lorSL_RS_483    --------------------------------------------------
V_sebSL_RS_507    --------------------------------------------------
V_sebSL_RS_552    --------------------------------------------------

C_cap_RS_551      GACAATAGGAAAAAATGCATTTTAGGAATGTACATGAACTGAAGATCAGT
C_debilis_7209    --CAATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCCGT
I_juru_RS_451     GACAATAGGAAAAAATGCATTTTTGGAATGTACATAAACTGAAGATCAGT
I_laev_RS_460     GACAATAGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGT
V_cadu_JJ_847     GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_flex_RS_442     GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_flex_RS_522     GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_flex_RS_595     GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_107    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_108    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_350    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_429    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_466    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_543    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multi_RS_549    GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_spRADS4_RS_213  ------AGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGT
V_spRADS4_RS_214  ------AGGAAAAAATGCATTTTAGGAATGTACATAAACTGAAGATCAGT
V_surRS_082       GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_sur_RS_083      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_sur_RS_084      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_pav_RS_216      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_sur_RS_248      GACAATAGGAAAAAATTCATTTTTGGAATGTACAGAAACTGAAGATCAGT
V_sur_RS_428      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_sur_RS_489      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_sur_RS_501      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_multicost_1     GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_multicost_2     GACAATAGGAAAAAATTCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_nobilis         GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_multiflora      GACAATAGGAAAAAATTCATTTTTGGAATGTACATAAACTGAAGATCAGT
V_michelii_110149 GACAATAGGAAAAAATGCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_michelii_200263 GACAATAGGAAAAAATGCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_michelii_110168 GACAATAGGAAAAAATGCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_michelii_110074 GACAATAGGAAAAAATGCATTTTAGGAATGAACATAAACTGAAGATCAGT
V_kwatae_110070   GACAATAGGAAAAAATTCTTTTTTGGAATGTACATAAACTGAAGATCAGT
V_spRADS3_RS_335  --------------------------------------------------
V_RADS3_RS_339    --------------------------------------------------
V_peru_JJ_772     --------------------------------------------------
V_calo_RS_430     --------------------------------------------------
V_calo_RS_454     --------------------------------------------------
V_calo_RS_511     --------------------------------------------------
V_spRADS1_RS_432  --------------------------------------------------
V_spRADS1_RS_500  --------------------------------------------------
V_spRADS1_RS_510  --------------------------------------------------
V_spRADS1_RS_561  --------------------------------------------------
V_dix_RS_225      --------------------------------------------------
V_elon_RS_437     --------------------------------------------------
V_elonRS_502      --------------------------------------------------
V_lorBL_RS_480    --------------------------------------------------
V_lorBL_RS_526    --------------------------------------------------
V_sebBL_RS_584    --------------------------------------------------
V_lorBL_RS_516    --------------------------------------------------
V_lorSL_RS_483    --------------------------------------------------
V_sebSL_RS_507    --------------------------------------------------
```

```
V_sebSL_RS_552    --------------------------------------------------

C_cap_RS_551      T-CAAATCAAAAAAA--------GGTATGATGTTCGATTATGAACCAAAT
C_debilis_7209    T-AAAATCAAAAAAAAA------GGTATGATGTTCGATCATGAACCAAAT
I_juru_RS_451     T-AAAATAAAAATAAAAAAAAAAGGTATGATGTTCGATCATGAA-CAAA-
I_laev_RS_460     T-AAAATAAAAAAAAA-------GGTATGATGTTCGATCATGAA-CAAA-
V_cadu_JJ_847     T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_flex_RS_442     T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_flex_RS_522     T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_flex_RS_595     T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_107    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_108    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_350    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_429    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_466    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_543    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multi_RS_549    T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_spRADS4_RS_213  T-AAAATAAAAAAGAAAAA----GGTATAATGTTCGATCATTAAACAAAT
V_spRADS4_RS_214  T-AAAATAAAAAAGAAAAA----GGTATAATGTTCGATCATTAAACAAAT
V_surRS_082       T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_sur_RS_083      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_sur_RS_084      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_pav_RS_216      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_sur_RS_248      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_sur_RS_428      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_sur_RS_489      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_sur_RS_501      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multicost_1     T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multicost_2     T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_nobilis         T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_multiflora      T-AAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_michelii_110149 T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_michelii_200263 T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_michelii_110168 T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_michelii_110074 T-AAAATAAAAAAGAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_kwatae_110070   TAAAAATAAAAAATAAAAA----GGTATGATGTTCGATCATGAAACAAAT
V_spRADS3_RS_335  --------------------------------------------------
V_RADS3_RS_339    --------------------------------------------------
V_peru_JJ_772     --------------------------------------------------
V_calo_RS_430     --------------------------------------------------
V_calo_RS_454     --------------------------------------------------
V_calo_RS_511     --------------------------------------------------
V_spRADS1_RS_432  --------------------------------------------------
V_spRADS1_RS_500  --------------------------------------------------
V_spRADS1_RS_510  --------------------------------------------------
V_spRADS1_RS_561  --------------------------------------------------
V_dix_RS_225      --------------------------------------------------
V_elon_RS_437     --------------------------------------------------
V_elonRS_502      --------------------------------------------------
V_lorBL_RS_480    --------------------------------------------------
V_lorBL_RS_526    --------------------------------------------------
V_sebBL_RS_584    --------------------------------------------------
V_lorBL_RS_516    --------------------------------------------------
V_lorSL_RS_483    --------------------------------------------------
V_sebSL_RS_507    --------------------------------------------------
V_sebSL_RS_552    --------------------------------------------------
```

```
C_cap_RS_551       AATTCATATTTTCTGAAATTGAAAAAA------CAAA-TCTTATG-----
C_debilis_7209     AATTAATATTTTCTGAAATTGAAAAAAAAA--------TCTTATG-----
I_juru_RS_451      AATGAATATTTTCTTAAATTGAAAAAAAAAA-------TCTTATG-----
I_laev_RS_460      AATGAATATTTTCTTAAATTGAAAAA-------CAAAATCTTATG-----
V_cadu_JJ_847      AATTAATATTTTAT-AAAAAAAAA---------GAAA-TCTTATG-----
V_flex_RS_442      AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_flex_RS_522      AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_flex_RS_595      AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_107     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_108     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_350     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_429     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_466     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_543     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_multi_RS_549     AATTAATATTTTCT-AAAAAAAAA---------GAAA-TCTTATG-----
V_spRADS4_RS_213   AATTAATATTTTCTTAAAAAAAAAAA-------GAAA-TCTTATG-----
V_spRADS4_RS_214   AATTAATATTTTCTTAAAAAAAAAAA-------GAAA-TCTTATG-----
V_surRS_082        AATGAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_sur_RS_083       AATGAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_sur_RS_084       AATGAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_pav_RS_216       AATTAATATTTTCTTAAAAAAAAAAAAA-----GAAA-TCTTATG-----
V_sur_RS_248       AATTAATATTTTCTTAAAAAAAAAAAAAAAAA-GAAA-TCTTATG-----
V_sur_RS_428       AATGAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_sur_RS_489       AATGAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_sur_RS_501       AATGAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_multicost_1      AATTAATATTTTCT-AAAAAAAAAA--------GAAA-TCTTATG-----
V_multicost_2      AATTAATATTTTCT-AAAAAAAAAA--------GAAA-TCTTATG-----
V_nobilis          AATTAATATTTTCTTAAAAAAAAAAAA------GAAA-TCTTATG-----
V_multiflora       AATTAATATTTTCTTAAAAAAAAAAAAA-----GAAA-TCTTATG-----
V_michelii_110149  AATGAATATTTTCT-AAAAAAAAAAA-------GAAA-TCTTATG-----
V_michelii_200263  AATGAATATTTTCT-AAAAAAAAAAA-------GAAA-TCTTATG-----
V_michelii_110168  AATGAATATTTTCT-AAAAAAAAAAA-------GAAA-TCTTATG-----
V_michelii_110074  AATGAATATTTTCT-AAAAAAAAAAA-------GAAA-TCTTATG-----
V_kwatae_110070    AATTAATATTTTCTTAAAAAAAAAAAAAAAAAA----CTTTTAGG-----
V_spRADS3_RS_335   ---------------AAAAAAAAAA-------GAAA-TCTTATGTTATG
V_RADS3_RS_339     ---------------AAAAAAAAAA-------GAAA-TCTTATGTTATG
V_peru_JJ_772      ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_calo_RS_430      ---------------AAAAAAAAAAAAA----------TCTTATGTTATG
V_calo_RS_454      ---------------AAAAAAAAAAAAA----------TCTTATGTTATG
V_calo_RS_511      ---------------AAAAAAAAAAAAA----------TCTTATGTTATG
V_spRADS1_RS_432   ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_spRADS1_RS_500   ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_spRADS1_RS_510   ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_spRADS1_RS_561   ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_dix_RS_225       ---------------AAAAAAAAAAAA-----------TCTTATGTTATG
V_elon_RS_437      ---------------AAAAAAAAAAAAAAA--------TCTTATGTTATG
V_elonRS_502       ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_lorBL_RS_480     ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_lorBL_RS_526     ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_sebBL_RS_584     ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_lorBL_RS_516     ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_lorSL_RS_483     ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG
V_sebSL_RS_507     ---------------AAAAAAAAAAAA-----------TCTTATGTTATG
V_sebSL_RS_552     ---------------AAAAAAAAAAAAAA---------TCTTATGTTATG

C_cap_RS_551       TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGG
C_debilis_7209     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGG
```

```
I_juru_RS_451      TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGG
I_laev_RS_460      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGG
V_cadu_JJ_847      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_flex_RS_442      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_flex_RS_522      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_flex_RS_595      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_107     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_108     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_350     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_429     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_466     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_543     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multi_RS_549     TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_spRADS4_RS_213   TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_spRADS4_RS_214   TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_surRS_082        TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sur_RS_083       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sur_RS_084       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_pav_RS_216       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sur_RS_248       TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_sur_RS_428       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sur_RS_489       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sur_RS_501       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_multicost_1      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_multicost_2      TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_nobilis          TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_multiflora       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_michelii_110149  TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_michelii_200263  TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_michelii_110168  TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_michelii_110074  TGAGTAAACCACTACTGAACCAGATCAATACCCAGTGGGTATTGATCTGA
V_kwatae_110070    TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_spRADS3_RS_335   TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_RADS3_RS_339     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_peru_JJ_772      TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_calo_RS_430      TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_calo_RS_454      TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_calo_RS_511      TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_spRADS1_RS_432   TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_spRADS1_RS_500   TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_spRADS1_RS_510   TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_spRADS1_RS_561   TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_dix_RS_225       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_elon_RS_437      TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_elonRS_502       TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_lorBL_RS_480     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_lorBL_RS_526     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sebBL_RS_584     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_lorBL_RS_516     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_lorSL_RS_483     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sebSL_RS_507     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA
V_sebSL_RS_552     TGAGTAAACCACTACTGAACCAGATCAATACCCACTGGGTATTGATCTGA

C_cap_RS_551       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTATTGA
C_debilis_7209     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTATTGA
I_juru_RS_451      TCCTTCAATGACTCGTATACACTAATACCGAAGTATTAGCCATTTGTTGA
I_laev_RS_460      TCCTTCAATGACTCGTATACACTAATACCGAAGTATTAGCCATTTGTTGA
```

```
V_cadu_JJ_847      TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_flex_RS_442      TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_flex_RS_522      TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_flex_RS_595      TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_107     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_108     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_350     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_429     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_466     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_543     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multi_RS_549     TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS4_RS_213   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS4_RS_214   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_surRS_082        TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sur_RS_083       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sur_RS_084       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_pav_RS_216       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sur_RS_248       TCCTTCAATGACTCGTATACACTAATACC??????????????????
V_sur_RS_428       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sur_RS_489       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sur_RS_501       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multicost_1      TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_multicost_2      TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGT???
V_nobilis          TCCTTCAATGACTCGTATACCCTAATACCGAAATATTAG??????????
V_multiflora       TCCTTCAATGACTCGTATACCCTAATACCGAAATATTAG??????????
V_michelii_110149  TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_michelii_200263  TCCTTCAATTACTCGTATACACTAATACCGAAATATTA???????????
V_michelii_110168  TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCA????????
V_michelii_110074  TCCTTCAATTACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_kwatae_110070    TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS3_RS_335   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_RADS3_RS_339     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_peru_JJ_772      TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_calo_RS_430      TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_calo_RS_454      TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_calo_RS_511      TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS1_RS_432   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS1_RS_500   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS1_RS_510   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_spRADS1_RS_561   TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_dix_RS_225       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_elon_RS_437      TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_elonRS_502       TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_lorBL_RS_480     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_lorBL_RS_526     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sebBL_RS_584     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_lorBL_RS_516     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_lorSL_RS_483     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sebSL_RS_507     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA
V_sebSL_RS_552     TCCTTCAATGACTCGTATACACTAATACCGAAATATTAGCCATTTGTTGA

C_cap_RS_551       TAGAGCTTCAACA
C_debilis_7209     TAGAGCTTCAACA
I_juru_RS_451      TAGAGCTTCAACA
I_laev_RS_460      TAGAGCTTCAACA
V_cadu_JJ_847      TGGAGCTTCAACA
V_flex_RS_442      TGGAGCTTCAACA
```

```
V_flex_RS_522     TGGAGCTTCAACA
V_flex_RS_595     TGGAGCTTCAACA
V_multi_RS_107    TGGAGCTTCAACA
V_multi_RS_108    TGGAGCTTCAACA
V_multi_RS_350    TGGAGCTTCAACA
V_multi_RS_429    TGGAGCTTCAACA
V_multi_RS_466    TGGAGCTTCAACA
V_multi_RS_543    TGGAGCTTCAACA
V_multi_RS_549    TGGAGCTTCAACA
V_spRADS4_RS_213  TGGAGCTTCAACA
V_spRADS4_RS_214  TGGAGCTTCAACA
V_surRS_082       TGGAGCTTCAACA
V_sur_RS_083      TGGAGCTTCAACA
V_sur_RS_084      TGGAGCTTCAACA
V_pav_RS_216      TGGAGCTTCAACA
V_sur_RS_248      ?????????????
V_sur_RS_428      TGGAGCTTCAACA
V_sur_RS_489      TGGAGCTTCAACA
V_sur_RS_501      TGGAGCTTCAACA
V_multicost_1     TGGAGC???????
V_multicost_2     ?????????????
V_nobilis         ?????????????
V_multiflora      ?????????????
V_michelii_110149 ?????????????
V_michelii_200263 ?????????????
V_michelii_110168 ?????????????
V_michelii_110074 ?????????????
V_kwatae_110070   ?????????????
V_spRADS3_RS_335  TAGAGCTTCAACT
V_RADS3_RS_339    TAGAGCTTCAACT
V_peru_JJ_772     TAGAGCTTCAACT
V_calo_RS_430     TAGAGCTTCAACT
V_calo_RS_454     TAGAGCTTCAACT
V_calo_RS_511     TAGAGCTTCAACT
V_spRADS1_RS_432  TAGAGCTTCAACT
V_spRADS1_RS_500  TAGAGCTTCAACT
V_spRADS1_RS_510  TAGAGCTTCAACT
V_spRADS1_RS_561  TAGAGCTTCAACT
V_dix_RS_225      TAGAGCTTCAACT
V_elon_RS_437     TAGAGCTTCAACT
V_elonRS_502      TAGAGCTTCAACT
V_lorBL_RS_480    TAGAGCTTCAACT
V_lorBL_RS_526    TAGAGCTTCAACT
V_sebBL_RS_584    TAGAGCTTCAACT
V_lorBL_RS_516    TAGAGCTTCAACT
V_lorSL_RS_483    TAGAGCTTCAACT
V_sebSL_RS_507    TAGAGCTTCAACT
V_sebSL_RS_552    TAGAGCTTCAACT
```

Appendix 4

Alignment 4.1
DNA nucleotide sequence alignment for AGT1 containing 683 characters and 36 unique alleles from cloned sequences employed in haplotype analyses in chapter 4. 50 base pairs per line.

```
V_lorBL_RS516    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorBL_RS516_2  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorBL_RS526    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorBL_RS526_2  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorBL_RS480    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorBL_RS480_2  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorBL_RS480_2_ TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS452    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS452_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS452_3  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS465    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS465_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS465_3  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS465_4  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_ML13     TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_ML13_2   TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_ML13_3   TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS435    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS435_2  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS503    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS503_2  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS503_3  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS582    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS582_2  TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS636    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebBL_RS636_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS483    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_lorSL_RS483_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS443    TATGTGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS443_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS444    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS464    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS529    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS529_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS533    TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG
V_sebSL_RS533_2  TATGCGCAAGTCCAAAAGCATTTGAAGCATCCAAGACTGCTAAATCAG

V_lorBL_RS516    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorBL_RS516_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorBL_RS526    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorBL_RS526_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorBL_RS480    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorBL_RS480_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorBL_RS480_2_ TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS452    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS452_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS452_3  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS465    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
```

```
V_lorSL_RS465_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS465_3  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS465_4  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_ML13     TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_ML13_2   TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_ML13_3   TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS435    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS435_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS503    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS503_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS503_3  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS582    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS582_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS636    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebBL_RS636_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS483    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_lorSL_RS483_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS443    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS443_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS444    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS464    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS529    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS529_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS533    TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG
V_sebSL_RS533_2  TCAGAGTATTCTTTGACTGGAAAGACTACCTCAAGTTCTACAAGTTGG

V_lorBL_RS516    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorBL_RS516_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorBL_RS526    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorBL_RS526_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorBL_RS480    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorBL_RS480_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorBL_RS480_2_ GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS452    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS452_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS452_3  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS465    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS465_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS465_3  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS465_4  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_ML13     GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_ML13_2   GAACATATTGGCCCTACACACCTTCTATACATCTTTTATATGGCCTGA
V_sebBL_ML13_3   GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS435    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS435_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS503    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS503_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS503_3  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS582    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS582_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS636    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebBL_RS636_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS483    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_lorSL_RS483_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS443    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS443_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS444    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
```

```
V_sebSL_RS464    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS529    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS529_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS533    GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA
V_sebSL_RS533_2  GAACATATTGGCCCTACACACCTTCTATACATCTTTTGTATGGCCTGA

V_lorBL_RS516    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorBL_RS516_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorBL_RS526    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorBL_RS526_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorBL_RS480    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorBL_RS480_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorBL_RS480_2_ GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS452    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS452_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS452_3  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS465    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS465_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS465_3  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS465_4  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_ML13     GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_ML13_2   GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_ML13_3   GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS435    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS435_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS503    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS503_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS503_3  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS582    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS582_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS636    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebBL_RS636_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS483    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_lorSL_RS483_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS443    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS443_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS444    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS464    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS529    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS529_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS533    GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG
V_sebSL_RS533_2  GAGCAGCTTTGGATCTCATATTTGTGGAAGGACTTGAAAATGTGTTTG

V_lorBL_RS516    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorBL_RS516_2  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorBL_RS526    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorBL_RS526_2  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorBL_RS480    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorBL_RS480_2  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorBL_RS480_2_ ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS452    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS452_2  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS452_3  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS465    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS465_2  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS465_3  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS465_4  ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
```

```
V_sebBL_ML13      ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_ML13_2    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_ML13_3    ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS435     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS435_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS503     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS503_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS503_3   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS582     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS582_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS636     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebBL_RS636_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS483     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_lorSL_RS483_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS443     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS443_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS444     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS464     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS529     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS529_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS533     ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA
V_sebSL_RS533_2   ACAGACACAAACGTTTGGGCAAAGCAACAAGGTAATAGATTTTGGAAA

V_lorBL_RS516     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorBL_RS516_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorBL_RS526     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorBL_RS526_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorBL_RS480     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorBL_RS480_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorBL_RS480_2_  GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS452     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS452_2   GATGTAGTGATGACTGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS452_3   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS465     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS465_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS465_3   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS465_4   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_ML13      GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_ML13_2    GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGCAAGA
V_sebBL_ML13_3    GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS435     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS435_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS503     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS503_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS503_3   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS582     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS582_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS636     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebBL_RS636_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS483     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_lorSL_RS483_2   GATGTAGTGATAAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS443     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS443_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS444     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS464     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS529     GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS529_2   GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
```

```
V_sebSL_RS533    GATGTAGTGATAAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA
V_sebSL_RS533_2  GATGTAGTGATGAATGCTAATATTGAGCACTCGATGTCCATAGTAAGA

V_lorBL_RS516    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorBL_RS516_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorBL_RS526    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorBL_RS526_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorBL_RS480    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorBL_RS480_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorBL_RS480_2_ GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS452    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS452_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS452_3  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS465    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS465_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS465_3  GAGGCACACCATTTATTTATTGGGAACTACACTGGATGGGCGGGGAAT
V_lorSL_RS465_4  GAGGCACACCATTTATTTATTGGGAACTACACTGGATGGGCGGGGAAT
V_sebBL_ML13     GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_ML13_2   GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_ML13_3   GAGGCACACCATTTATTTATTGGGAACTACACTGGATGGGCGGGGAAT
V_sebBL_RS435    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_RS435_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCCGGGAAT
V_sebBL_RS503    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_RS503_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_RS503_3  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_RS582    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_RS582_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebBL_RS636    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCCGGGAAT
V_sebBL_RS636_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS483    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_lorSL_RS483_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS443    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS443_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS444    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS464    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATCGGCGGGGAAT
V_sebSL_RS529    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS529_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS533    GAGGCACACCATTTATTTATTGGGAACTACCCTGGATGGGCGGGGAAT
V_sebSL_RS533_2  GAGGCACACCATTTATTTATTGGGAACTACCCTGGATCGGCGGGGAAT

V_lorBL_RS516    ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_lorBL_RS516_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorBL_RS526    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorBL_RS526_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorBL_RS480    ATTGCAAAAAATTTCAGCAGTCGGACTATCCTAGACACTGATATTGAC
V_lorBL_RS480_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorBL_RS480_2_ ATTGCAAAAAATTTCAGCAGTCGGACTATCCTAGACACTGATATTGAC
V_lorSL_RS452    ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_lorSL_RS452_2  ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_lorSL_RS452_3  ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_lorSL_RS465    ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_lorSL_RS465_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorSL_RS465_3  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorSL_RS465_4  ATTGCAAAAA-TTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_ML13     ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_sebBL_ML13_2   ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_ML13_3   ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
```

```
V_sebBL_RS435    ATTGCAAAAAATTTCAGCAGTCGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS435_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS503    ATTGCAAAAAATTTCAGCAGTCGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS503_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS503_3  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS582    ATTGCAAAAAATTTCAGCAGTCGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS582_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS636    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebBL_RS636_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_lorSL_RS483    ATTGCAAAAAATTTCAGCAGTGGGACTGTCCTAGACACTGATATTGAC
V_lorSL_RS483_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS443    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS443_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS444    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS464    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS529    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS529_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS533    ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC
V_sebSL_RS533_2  ATTGCAAAAAATTTCAGCAGTGGGACTATCCTAGACACTGATATTGAC

V_lorBL_RS516    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorBL_RS516_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorBL_RS526    AACTTTTGCCCCAGTGAATGTTGACCATATATGTGTTTTTTT-GAACT
V_lorBL_RS526_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorBL_RS480    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTTT-AACT
V_lorBL_RS480_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTTT-AACT
V_lorBL_RS480_2_ AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTTT-AACT
V_lorSL_RS452    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS452_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS452_3  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS465    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS465_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS465_3  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_lorSL_RS465_4  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebBL_ML13     AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_sebBL_ML13_2   AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_sebBL_ML13_3   AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebBL_RS435    AACTTTTGCCCCAGTGAATGTTGACCACATGTGTTTTTTTTTT-AACT
V_sebBL_RS435_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_sebBL_RS503    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTTT-AAAT
V_sebBL_RS503_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTCTT-AACT
V_sebBL_RS503_3  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTCTT-AACT
V_sebBL_RS582    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTTT-AACT
V_sebBL_RS582_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_sebBL_RS636    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_sebBL_RS636_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS483    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTT---GAACT
V_lorSL_RS483_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTTT-AACT
V_sebSL_RS443    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebSL_RS443_2  AACTTTTGCCCCAGTGAATGTTGACCATATGCGTTTTTTTCTT-AACT
V_sebSL_RS444    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebSL_RS464    TACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTCTT-AACT
V_sebSL_RS529    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebSL_RS529_2  AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebSL_RS533    AACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTTT--AACT
V_sebSL_RS533_2  TACTTTTGCCCCAGTGAATGTTGACCATATGTGTTTTTTTCTT-AACT
```

```
V_lorBL_RS516    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCGTT
V_lorBL_RS516_2  TTCCATGTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorBL_RS526    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorBL_RS526_2  TTCCATGTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorBL_RS480    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorBL_RS480_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorBL_RS480_2_ TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorSL_RS452    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCGTT
V_lorSL_RS452_2  TTCCATTTATAATCCTCCAAATTCCTCCAAACAGTAGCTAAAATCGTT
V_lorSL_RS452_3  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCGTT
V_lorSL_RS465    TTCCATTTATAATCCTCCAAATTCCTCCAAACAGTAGCTAAAATCGTT
V_lorSL_RS465_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorSL_RS465_3  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorSL_RS465_4  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_ML13     TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_ML13_2   TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_ML13_3   TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS435    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTACCTAAAATCATT
V_sebBL_RS435_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS503    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS503_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS503_3  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS582    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS582_2  TTCCATTTATAATCCTTAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS636    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebBL_RS636_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_lorSL_RS483    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCGTT
V_lorSL_RS483_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS443    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS443_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS444    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS464    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS529    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS529_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS533    TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT
V_sebSL_RS533_2  TTCCATTTATAATCCTCAAAATTCCTCCAAACAGTAGCTAAAATCATT

V_lorBL_RS516    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorBL_RS516_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorBL_RS526    ATAGTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorBL_RS526_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorBL_RS480    ATAGTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_lorBL_RS480_2  ATAGTCTTTTACTTTAAGGCTGTGACCCGCCCAATGGATAAATGGTTC
V_lorBL_RS480_2_ ATAGTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_lorSL_RS452    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS452_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS452_3  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS465    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS465_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS465_3  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATATATGGTTC
V_lorSL_RS465_4  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATATATGGTTC
V_sebBL_ML13     ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_sebBL_ML13_2   ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_sebBL_ML13_3   ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATATATGGTTC
V_sebBL_RS435    ATAGTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebBL_RS435_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_sebBL_RS503    ATAGTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
```

```
V_sebBL_RS503_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGCTC
V_sebBL_RS503_3  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebBL_RS582    ATAGTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebBL_RS582_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_sebBL_RS636    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_sebBL_RS636_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS483    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATGAATGGTTC
V_lorSL_RS483_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS443    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS443_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS444    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS464    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS529    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS529_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS533    ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC
V_sebSL_RS533_2  ATATTCTTTTACTTTAAGGCTGTGACCCACCCAATGGATAAATGGTTC

V_lorBL_RS516    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorBL_RS516_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATAAGAGTGCCTGAGA
V_lorBL_RS526    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorBL_RS526_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATAAGAGTGCCTGAGA
V_lorBL_RS480    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorBL_RS480_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorBL_RS480_2_ AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS452    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS452_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS452_3  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGG
V_lorSL_RS465    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS465_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS465_3  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS465_4  AATTCTCGTGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_ML13     AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_ML13_2   AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGTGTGCCTGAGA
V_sebBL_ML13_3   AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS435    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS435_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS503    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS503_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS503_3  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS582    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS582_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS636    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebBL_RS636_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS483    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_lorSL_RS483_2  AATTCTCATGAAGGTGTGACATTTGTAAGATTATATGAGTGCCTGAGA
V_sebSL_RS443    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS443_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS444    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS464    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS529    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS529_2  AATTCTCTTGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS533    AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA
V_sebSL_RS533_2  AATTCTCATGAAGGTGTGACATTTGTAAGATGATATGAGTGCCTGAGA

V_lorBL_RS516    CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorBL_RS516_2  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorBL_RS526    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
```

```
V_lorBL_RS526_2  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorBL_RS480    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_lorBL_RS480_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_lorBL_RS480_2_ CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_lorSL_RS452    CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorSL_RS452_2  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorSL_RS452_3  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorSL_RS465    CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorSL_RS465_2  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorSL_RS465_3  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGT
V_lorSL_RS465_4  CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGT
V_sebBL_ML13     CGTAACAAATAAGGAGTGCAGGAAAAAATTCTCATATATATGATTTGA
V_sebBL_ML13_2   CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_ML13_3   CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGT
V_sebBL_RS435    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS435_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS503    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS503_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS503_3  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS582    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS582_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS636    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebBL_RS636_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_lorSL_RS483    CGTAACAAATAAGGAGTGCAGGAAAGATTTCTCATATATATGATTTGA
V_lorSL_RS483_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS443    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS443_2  TGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS444    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS464    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS529    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS529_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS533    CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA
V_sebSL_RS533_2  CGTAACAAATAAGGAGTGCAGGAAAAATTTCTCATATATATGATTTGA

V_lorBL_RS516    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorBL_RS516_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorBL_RS526    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTCTGTCCTTT
V_lorBL_RS526_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorBL_RS480    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorBL_RS480_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAACGTTCTTTGTCCTTT
V_lorBL_RS480_2_ GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS452    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS452_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS452_3  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS465    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS465_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS465_3  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS465_4  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_ML13     GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_ML13_2   GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_ML13_3   GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_RS435    GATGTCTATAGACAGAGAAATGGCCTTCACTAACGTTCTTTGTCCTTT
V_sebBL_RS435_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_RS503    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_RS503_2  GATGCCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_RS503_3  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_RS582    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
```

```
V_sebBL_RS582_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebBL_RS636    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCGTTT
V_sebBL_RS636_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS483    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_lorSL_RS483_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS443    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS443_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS444    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS464    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS529    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS529_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS533    GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT
V_sebSL_RS533_2  GATGTCTATAGACAGAGAAATGGCCTTCACTAATGTTCTTTGTCCTTT

V_lorBL_RS516    TTAAATGTAGG
V_lorBL_RS516_2  TTAAATGTAGG
V_lorBL_RS526    TTAAATGTAGG
V_lorBL_RS526_2  TTAAATGTAGG
V_lorBL_RS480    TTAAATGTAGG
V_lorBL_RS480_2  TTAAATGTAGG
V_lorBL_RS480_2_ TTAAATGTAGG
V_lorSL_RS452    TTAAATGTAGG
V_lorSL_RS452_2  TTAAATGTAGG
V_lorSL_RS452_3  TTAAATGTAGG
V_lorSL_RS465    TTAAATGTAGG
V_lorSL_RS465_2  TTAAATGTAGG
V_lorSL_RS465_3  TTAAATGTAGG
V_lorSL_RS465_4  TTAAATGTAGG
V_sebBL_ML13     TTAAATGTAGG
V_sebBL_ML13_2   TTAAATGTAGG
V_sebBL_ML13_3   TTAAATGTAGG
V_sebBL_RS435    TTAAATGTAGG
V_sebBL_RS435_2  TTAAATGTAGG
V_sebBL_RS503    TTAAATGTAGG
V_sebBL_RS503_2  TTAAATGTAGG
V_sebBL_RS503_3  TTAAATGTAGG
V_sebBL_RS582    TTAAATGTAGG
V_sebBL_RS582_2  TTAAATGTAGG
V_sebBL_RS636    TTAAATGTAGG
V_sebBL_RS636_2  TTAAATGTAGG
V_lorSL_RS483    TTAAATGTAGG
V_lorSL_RS483_2  TTAAATGTAGG
V_sebSL_RS443    TTAAATGTAGG
V_sebSL_RS443_2  TTAAATGTAGG
V_sebSL_RS444    TTAAATGTAGG
V_sebSL_RS464    TTAAATGTAGG
V_sebSL_RS529    TTAAATGTAGG
V_sebSL_RS529_2  TTAAATGTAGG
V_sebSL_RS533    TTAAATGTAGG
V_sebSL_RS533_2  TTAAATGTAGG
```