2016-5

The phylogeography and cytogenetics of Rhipsalis baccifera: the cactus that made it to the Old World

Elizabeth Ann Tan

Follow this and additional works at: https://scholarworks.smith.edu/theses

Part of the Biology Commons

Recommended Citation

Tan, Elizabeth Ann, "The phylogeography and cytogenetics of Rhipsalis baccifera: the cactus that made it to the Old World" (2016). Theses, Dissertations, and Projects. 1679.
https://scholarworks.smith.edu/theses/1679

This Honors Project has been accepted for inclusion in Theses, Dissertations, and Projects by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu.
The phylogeography and cytogenetics of *Rhipsalis baccifera*: the cactus that made it to the Old World

Elizabeth Tan

Submitted to the Department of Biological Sciences of Smith College in partial fulfillment of the requirements for the degree of Bachelor of Arts

Robert Merritt, Honors Project Adviser
Laura Katz, Honors Project Adviser
May 2016
Table of Contents:

Introduction 4
  An overview of plant phylogenetics 4
  Molecular approaches for understanding cactus evolution 5
  The unusual distribution of Rhipsalis baccifera 6
  Cactus phylogenetics 9
  Speciation by polyploidy 11
  Polyploidy in cacti 15
  Relating ploidy and biogeography 20
  Cactaceae phylogenomics 23
  Selection of markers 25
  Nuclear gene: phosphoglucose isomerase 26

Research questions 27

Materials 29
  Database assembly 29

Methods 33
  DNA extraction and amplification 33
  Molecular markers for species identification 36
  Histology and microscopy 42

Results 45
  Species identification 45
  Phosphoglucose isomerase 49
  Microscopy 50
  Morphology of nuclei 50
  Nuclear area as a proxy for ploidy levels 51

Discussion 55
  Molecular phylogeny of Rhipsalis baccifera 55
  Microscopy: cytotyping individual plants 59
  Broader significance 63

Acknowledgements 64

References 65

Appendix 71

Figures & Tables

Figure 1. Polyploidy in Plants 14
Figure 2. Morphological characteristics of Rhipsalis baccifera subspecies 18
Figure 3. Geographic distribution of samples 33
Figure 4. Distance tree for all samples 48
Figure 5. Distance tree for Brazilian samples 49
Figure 6. Feulgen nuclear staining of Rhipsalis baccifera and garlic 52
Figure 7. Feulgen nuclear staining of Rhipsalis baccifera Puerto Rico #2 53
Figure 8. Nuclear area measurements 54

Table 1. Collection location and morphology of Rhipsalis baccifera 30
Table 2. Chloroplast PCR reaction 38
Table 3. Primer sequences 39
Table 4. Phosphoglucose isomerase PCR reaction 41
Table 5. Summary of results: cytology & molecular 61
OVERVIEW

*Rhipsalis baccifera* is the only species of cactus that grows endemically in both the Americas and the Old World. Furthermore, genome duplication events within the species resulted in diploid (2n), tetraploid (4n), and octoploid (8n) populations. Polyploidy contributes to species diversification resulting from instant speciation events. The subspecies of *Rhipsalis baccifera* are defined by their morphological characteristics. These subspecies delineations generally correlate with different polyploidy levels and geographic distributions. I conducted a population genetics study to assess the DNA sequence divergence of individual plants collected across a global distribution. I also estimated ploidy counts of individuals by measuring nuclear area to compare my data set with the putative ploidy levels associated with bio-geographic ranges. Results from chloroplast DNA suggest that the species dispersed rapidly and genetic divergence has yet to occur between the New World and the Old World populations. New World individuals morphologically identified as subspecies *R. baccifera shaferi* and *R. baccifera hileiabaiana*, which is found exclusively in the Bahia region of Brazil, have low sequence identity with other *Rhipsalis baccifera* subspecies. I used nuclear area as a proxy for ploidy counts and nuclear area to show that variation in polyploidy existed in my dataset. Pilot findings from the cytology suggest that polyploid populations do correspond with subspecies categorizations and biogeography. I found that forming connection between polyploidy and genetic divergence is necessary to determine speciation in plants.
Introduction

An overview of plant phylogenetics:

Phylogenetic studies determining species delimitation in plants must account for biological, morphological, and phylogenetic species definitions. Evolutionary patterns driven by climate change, genome duplication, and long-distance dispersal provide a framework for understanding tropical plant communities and related taxa (Clayton et al. 2009). Moreover, studying recent radiation events of tropical species is necessary to understand neotropical species adaptation in modern ecosystems (Clayton et al. 2009). Plant phylogenetics relies on molecular biology for understanding relationships between genera and for the molecular dating of dispersal events (Bonatelli et al. 2013). Fossil records can provide dates for identifying the emergence of a species or clade. However, the absence of a fossil record for succulents leaves much of the early evolutionary history of cacti a mystery (Arakaki et al. 2011). Additionally, in comparison to other succulents, cacti radiated and diversified quickly, but the processes that allow for these genetic changes and adaptability to a range of environments are not well understood (Arakaki et al. 2011).

This study investigates long-range dispersal and evolution in cacti by focusing on *Rhipsalis baccifera*, the cactus species with the largest geographic distribution. All cacti are native exclusively to the New World, with the sole exception of *Rhipsalis baccifera* - the only species native to both the New World and the Old World (Korotkova et al. 2011). Studying *Rhipsalis baccifera* as an exceptional species gives insight into dispersal
because developmental adaptations can be compared to genotypic variations, or alleles. Using existing samples of *Rhipsalis baccifera* to understand long-range dispersal and genome duplication, also known as polyploidy, elucidates the agents of genetic divergence in cacti. The evolution of *Rhipsalis baccifera* can be better understood by analyzing living specimens to determine the phylogenetic divergence of populations (Hilu *et al.* 2003). The geographically isolated New World and Old World populations serve as models for the diversification of other cactus species. I sampled individuals from a large geographic breadth to include individuals with different ploidy levels and genetic diversity to support the cytological and molecular work of this study.

**Molecular approaches for understanding cactus evolution**

Molecular methods can be used to roughly estimate dispersal and speciation events for plants. I looked at sequencing data to determine when *Rhipsalis baccifera* dispersed to the Old World. Although Bayesian inferences can be used to estimate the emergence of clades at both the interspecies and intraspecies taxonomic level, molecular data showing the highly similar sequences also suggests a recent introduction of *Rhipsalis baccifera* to the Old World. Bayesian inferences can also be applied to molecular dating in Cactaceae and *Rhipsalis* samples (Calvente *et al.* 2011). For example, one Bayesian inference estimated a most recent common ancestor of Cactaceae emerging between 3.1 to 19.1 million years ago (OCampo *et al.* 2010). Alternatively, Arakaki *et al.* (2011) propose that Cactoideae, the subfamily that includes the genus *Rhipsalis*, diverged from other Cactaceae around 21 million years ago in the Andes. These Bayesian estimates
also illustrate the ranges presented in the phylogenies used to recreate common ancestors and relationships for the early Cactineae (Caryophyllales) (OCampo et al. 2010, and Arakaki et al. 2011). Intrataxonomic studies can investigate the human influence on plant hybridizations in cacti that identify recent phylogenetic changes (Realini et al. 2014). Bayesian inferences can also estimate the dispersal of species of cacti that travelled from both human and bird vectors (Griffith, 2004). One study on the origins and importance of the agriculturally important Opuntiae cactus family used Bayesian inferences to estimate the instance of domestication of the desert crop (Griffith, 2004). These examples from both Griffith (2004) and Realini (2014) show that recent human cultivation of species of Cactaceae can be traced using molecular techniques. The efficacy of molecular dating to understand *Rhipsalis baccifera* at the species level is viable, provided that markers exist that show sufficient genetic variation to track intraspecific evolution (Arakaki et al. 2011). These Bayesian estimates confirm that *Rhipsalis baccifera* was introduced to the Old World well after the Gondwanan split, but the mechanism of the dispersal from the New World remains unclear (OCampo et al. 2010, and Arakaki et al. 2011). I reason that *Rhipsalis baccifera* likely dispersed to the Old World via a bird vector rather than a human introduction.

**The unusual distribution of *Rhipsalis baccifera***

Historically, *Rhipsalis baccifera* has presented a riddle to botanists and plant phylogeneticists due to its unusual distribution. Three models exist to explain the unusual distribution of *Rhipsalis baccifera* in the Old World. These include bird, human,
and an illogical vicariance model that suggests that *Rhipsalis baccifera* existed on both the African and South American continents and the tectonic separation resulted in the evolution of different species. The Gondwanan split model, which is one hypothesis presented by Backeberge (1942) and Croizat (1952), cannot temporally accommodate the dispersal of *Rhipsalis baccifera* into the Old World because cacti and the genus *Rhipsalis* emerged after South America separated from the African continent, which is estimated to have occurred in the late Cretaceous period between 65-90 million years ago (Nyffeler, 2002). The Gondwanan split date confirms that cacti and *Rhipsalis* are indeed neotropical monophyletic clades (OCampo et al. 2010, Arakaki et al. 2011). Vicariance refers to a previously large effective population size that on Gondwanaland that split with the tectonic plates leading to genetic divergence between the separated populations. This vicariance model suggests that geographic barriers, such as the Atlantic Ocean, divided a previously large effective population size resulting in genetically divergent populations (Christenhusz and Chase, 2013).

Long-range dispersal of *Rhipsalis baccifera* accounts for the establishment of subspecies as distinct populations that are geographically isolated by the Atlantic and Indian oceans (Christenhusz and Chase, 2013). Similarly, other morphologically distinct subspecies are observed in other neotropical species (Christenhusz and Chase, 2013). *Rhipsalis baccifera* morphological differences across subspecies suggests that this cactus is a neotropical species distinct populations. However the mechanisms for dispersal are unknown for *Rhipsalis baccifera* (Korotkova et al. 2011). Seed distribution from a bird
vector and global trade vectors in a post-Colombian human introduction are the two most common hypotheses for the dispersal of *Rhipsalis baccifera* from the New World to the Old World. Of these two proposals, trade wind patterns and fruit morphology provide conditions that explain a likely bird dispersal from South America to Africa for the dispersal of *R. baccifera* (Christenhusz and Chase, 2013; Korotkova *et al.* 2011).

The bird-vector dispersal scenario, in which *Rhipsalis baccifera* was introduced into different ecologies and continents, explains how geographic reproductive isolation of populations resulted in morphological divergence across subspecies associated with the New World and Old World populations (Korotkova *et al.* 2011). This hypothesis is supported by the seed and fruit morphology of Cactaceae because *Rhipsalis baccifera* has small seeds in mucilaginous fruits that are often eaten by birds (Calvente *et al.* 2011; Echevarría-Machado *et al.* 2005). Additionally, bird dispersal is known to encourage gene flow in low effective population sizes (Ornelas, et al. 2015). This seed morphology is advantageous for isolated individuals that remain confined to specific areas of growth (Calvente *et al.* 2011).

Furthermore, *Rhipsalis baccifera* self-fertilizes and establishes clones from cutting propagations, allowing a single plant to establish a full population in subsequent generations (Barthlott, 1983). Alternatively, the introduction of *R. baccifera* by man seems unlikely because observable morphological differences from genetic divergence
would have to have arisen within the last 500 years under this post-Columbian hypothesis (Barthlott, 1983).

**Cactus phylogenetics**

Understanding cactus phylogeny with molecular markers can provide details about the wide diversity of succulents in both desert and tropical environments. Many of the synapomorphic features of Cactaceae emerged multiple times (Korotkova et al. 2011). Similarly, morphology can be misleading due to convergent evolution within Cactaceae (Korotkova et al. 2011). Studying *Rhipsalis baccifera* with phylogenetics has elucidate evolutionary trends in epiphytes and polyploids to understand the genetically and morphologically diverse Cactaceae (Nyffeler, 2002). Cactaceae, a monophyletic clade of angiosperms that emerged around 35 million years ago, evolved fairly recently compared to other clades of angiosperms (Arakaki et al. 2011). Cacti radiated quickly across a large geographic area at the end of an Eocene era of intense aridification and increased CO₂ in the American desert (Majure et al. 2012). Cactaceae radiated and diversified rapidly to include plants with photosynthetic stems and reduced leaves that suit a wide range of arid environments (Edwards et al. 2005).

The subfamily *Cactoideae* originated around 30 to 20 million years ago in the central Andes on the western side of South America. Unlike the *Cactoideae*, that grow in highly arid environments, the *Rhipsalideae* clade likely evolved and originated in humid rainforests in coastal Brazil (Calvente et al. 2011). The *Rhipsalideae* is a recently
emerged and diverse clade of cacti that live in wide range of environments as epiphytes. *Rhipsalis baccifera* is no exception to the morphological diversity present within the *Rhipsalideae* clade. Similarly, the diversification of the tribe *Rhipsalideae* was likely driven by adaptations to grow in humid shaded environments such as cloud forests (Ornelas *et al.* 2015). *Rhipsalis baccifera* can be found in arid Old World environments, cloud forests, and rainforests with subspecies specificity to each climate. Species in the genus *Rhipsalis* oftentimes do not have spines, or only have spines at the younger portions of their stems to account for sun radiation in especially dry and hot environments (Manzano *et al.* 2010).

The *Rhipsalideae*, the focus on this thesis, compose one of the nine tribes of the subfamily *Cactoideae*, which includes the genus *Rhipsalis* (Korotkova *et al.* 2011). *Rhipsalideae* includes 52 accepted species morphologically defined by hanging stems that are cylindrical or flattened like a leaf (Korotkova *et al.* 2011). *Rhipsalis*, a genus composed of epiphytes and lithophytes, includes cacti that grow in subtropical and tropical environments (Korotkova *et al.* 2011). *Rhipsalis* are limited to growing in select microenvironments within a forest, which results in the isolated establishment of individuals (Andrade *et al.* 1997). Species within the genus *Rhipsalis* are usually rare and occur in endemically-restricted distributions, such as the rainforests of Central and South America (Calvente *et al.* 2011). Additionally, *Rhipsalis baccifera* has a putative origin in a humid, tropical environment, but an expanded bio-geographical range to more arid environments like Madagascar (Almeida *et al.* 2013).
*Rhipsalis baccifera* has small flowers, branched stems, and epiphytic growth that place it morphologically in the *Rhipsalis* genus (Cota-Sanchez et al. 2010). *Rhipsalis baccifera* has skinny, pendent, cylindrical stems with adventitious roots that allow individuals to grow epiphytically in tropical forest systems (Calvente et al. 2011). Adventitious roots grow directly from the stems, usually towards the base, to allow for growth on a tree trunk or branch (North and Nobel, 1994). *R. baccifera* will germinate and grow in the lower branches where there is more shade and humidity (Manzano et al. 2010). I observed that the young shoot tips of *Rhipsalis baccifera* also have spines, even if the stems do not, which is similar to the other species of Cactaceae (Manzano et al. 2010). Slight morphological differences in spine density, stem thickness, and color distinguish the various subspecies of *R. baccifera*, although these variations do not necessarily warrant species delineation (Cota-Sanchez et al. 2010).

**Speciation by polyploidy**

Genome duplication, or polyploidy, accounts for about 2-4% of speciation events in angiosperms, and is an important mechanism for plant diversification (Negron-Ortiz et al. 2007; Otto and Whitton, 1990; Soltis et al. 2004). Cactaceae are not an exception, and are known to have autopolyploid species and populations (Arakaki et al. 2007). However, the continuum of classifications of polyploids presents a problem in negotiating the polyploid cytotypes as potential candidate populations that merit species ranking (Soltis and Soltis, 2009). Cytotypes refers to different chromosome
numbers present within the same species. Although botanists and phylogeneticists regard polyploidy as a mechanism that drives speciation, it remains difficult to determine when reproductive isolation occurs from genetic divergence rather than a biological species concept (Soltis and Soltis, 2009). Autopolyploidy refers to genome duplication within species or across individuals that are highly similar within a population (Figure 1; Soltis and Soltis, 2000). Allopolyploidy refers to genome duplication resulting from hybridization across species or genetically divergent individuals, and is more prevalent than autopolyploidy (Soltis and Soltis, 2000). In addition to cytological count differences, polyploidy contributes to morphological differences noted by increased cell sizes and decreased stomatal density in angiosperms (Beaulieu et al. 2008). Phylogenetic data combined with cytological counts provide information for distinguishing between allopolyploids and autopolyploids (Majure et al. 2012). For example, sequencing data may suggest autopolyploidy in a plant with homozygous genotypes and synteny, where genes are located in the same positions on the chromosomes (Soltis and Soltis, 2009; Judd et al. 2007). Allopolyploidy results in increased heterozygosity compared to diploid progenitors because alleles from different species are introduced into the hybridized individual (Soltis et al. 2009). Intraclade studies investigating morphologically similar autopolyploid species are necessary to disentangle the species delineation debate because genetically similar yet biogeographically distinct populations with separate evolutionary paths warrant species ranking (Soltis et al. 2007). Autopolyploidy allows for increased heterozygosity and
genetic diversity within a population because of polysomic inheritance (Soltis and Soltis, 2004). The increase in gene copy number allows for more allelic combinations.

Allopolyploid populations have a more allelic variation and occur commonly to account for the majority of instant speciation events associated with polyploidy. Also, allopolyploidy presents the most common polyploidization events that result in increased species diversity in angiosperms (Soltis and Soltis, 2009). Similarly, a homoploid individual is also sterile because it lacks pairs of sufficiently homologous chromosomes (Figure 1). However, allopolyploids are genetically viable with other allopolyploid hybrids. Although allopolyploidy is well understood and easily detectable with genetic sequencing data, autopolyploidy may be more commonplace than previously suspected (Soltis et al. 2000). Autopolyploidy should be considered as a driver of speciation despite resulting in genetically identical populations because reproductive isolation occurs.
Figure 1. Hybridization and polyploidy play a role in plant evolution in angiosperms. While a homoploid hybrid species may be diploid like the progenitor parentals, the heterologous chromosomes cannot participate in meiosis because they lack genetic complementarity. Tetraploid allopolyploids and autopolyploids are displayed with heterologous and homologous pairs of chromosomes, respectively. This figure is from Soltis, PS, Soltis DE. 2009. Annual Review of Plant Biology. 60:561-88.

Speciation events can occur in autopolyploid populations because of an instant biological species is generated (Soltis and Soltis, 2009). For example, a tetraploid resulting from autopolyploidization cannot produce fertile triploid offspring with a diploid making polypoids a biological species with identical genetic information. However, autopolyploids are effectively genetically identical to their diploid progenitors. Polyploids without other polyploids within pollination range are reproductively isolated.
Polyploidy may drive the evolution of New World and Old World populations of *Rhipsalis baccifera* because it results in reproductively isolated populations. However, polyploidy also often results in morphological and physical differences, which allow individuals to adapt to different environments (Soltis and Soltis, 2009). The geographic endemism of populations combined with an independent evolutionary path from the diploid species factor into the species merit of autopolyploids (Soltis and Soltis, 2009). For example, the plant *Tolmiea menziesii* has autotetraploids that are endemic from central Oregon to southeastern Alaska, compared to the diploid cytotype that remains south of central Oregon (Soltis et al. 2007). The chemical, DNA, and isozyme data support that these autotetraploids warrant species ranking as they are molecularly distinct from the diploid population (Soltis et al. 2007). Additionally, polyploids may exhibit increased tolerance to drought, or temperature extremes to account for population segregation and decreased gene flow (Soltis and Soltis, 2009). Diploid, tetraploid, and octoploid individuals exist within *R. baccifera* populations. In this study, I correlate polyploidy and genetic divergence to determine species delineation between the New World and Old World populations of *Rhipsalis baccifera* (Cota-Sanchez et al. 2010).

**Polyploidy in cacti**

Cacti species vary considerably in morphology and ploidy levels, which has affected their genetic divergence and morphological diversity as a family of angiosperms (Majure et al. 2012). The subfamilies *Opuntiae* and *Cactoideae*, which includes *Rhipsalis baccifera*,
more commonly include polyploidy cytotypes compared to other subfamilies of cacti (Cota-Sanchez and Philbrick, 1994.) Polyploidy results from the fusion of two unreduced gametes. Although most new polyploidy population go extinct from being reproductively isolated, polyploidy allows for more allelic diversity (Soltis and Soltis, 2000). Genetic diversity in Cactaceae is advantageous for dispersal and allowing for adaptations to new environments (Majure et al. 2012).

Reproductively isolated individuals are more likely to self-fertilize, which is observed in the terrestrial cacti in the Sonoran desert (Murawski et al. 1994). Furthermore, populations of the columnar tetraploid cactus *Pachycereus pringlei* have been known to change their reproductive system from cross-pollinating to selfing after a polyploidization event (Murawski et al. 1994; Moraes et al. 2005). Triploid hybrid populations in cacti are limited to contact zones between diploid and tetraploid populations and consist of individuals produced from stem fragmentations (Ferriol et al. 2014; Arakaki et al. 2012). This model illustrates that polyploids can only cross with individuals of the same cytotype to produce fertile offspring in usually limited geographies for populations of *Rhipsalis baccifera*. Autopolyploid species delineation can result from biogeographic isolation or population segregation from reproductive isolation (Plume et al. 2013). In addition, autopolyploidy may occur more frequently in cacti and other lineages of angiosperms with small chromosomes, although there is no significant causal relationship between the two characteristics (Soltis and Soltis, 2009).
Agamospermy, or production of fertile seeds without the fertilization of gametes, can also play a role in producing long-standing clonal polyploid populations (Arakaki et al. 2012). Consequently, the effective population size for new polyploids, or neopolyploids, is smaller and therefore the effects of genetic drift are greater resulting in allele fixation or loss (Ferriol et al. 2014). Therefore, reproductively isolated populations with small effective population sizes can lead to speciation events. The maintenance of a reproductively isolated population over a long period of time serves as an example of the advantages of clonal propagation in polyploid populations of cacti that share morphologically similar succulent stems (Arakaki et al. 2012).

**Rhipsalis baccifera: morphological variation among subspecies**

*Rhipsalis baccifera* has the largest geographic distribution of any species of cactus, and can be found in South America, Central America, the Caribbean, Florida, mainland Africa, Madagascar, Sri Lanka, and India (Cota-Sánchez et al. 2010). In conjunction with the biogeography of the populations of *Rhipsalis baccifera*, observable morphological differences define the subspecies. *Rhipsalis baccifera* has been categorized into six subspecies: *baccifera, hileiabaiana,* and *shaferi* in the New World; and, *erythrocarpa, horrida,* and *mauritiania* in the Old World (Cota-Sánchez et al. 2010).
Figure 2. Morphological characteristics of *Rhipsalis baccifera* subspecies

Figure 2. Morphological characteristics, such as stem width, presence and density of spines, and color, delineate subspecies of *Rhipsalis baccifera*. (A) Vera Cruz, Mexico ssp. *baccifera* shows the pendant branching structure indicative of all populations in this species. New World populations present thin, light-green stems with little to no spines. (B) Madagascar ssp. *horrida* shows a thicker stem width, dense spine coverage, and putative octoploid populations. (C) Zimbabwe, Chiande Forest ssp. *mauritiania* shows dark green, thick stem widths with tetraploid populations. (D) Individual from panel A shows the full body of *Rhipsalis baccifera*, which shows its hanging growth as an epiphyte.

The morphology of the various subspecies correlates with biogeography in *Rhipsalis baccifera*. Figure 2 displays the range of morphologies in the subspecies of *Rhipsalis baccifera*. Subspecies *baccifera*, shown in picture D of Figure 2, is the divergent subspecies to the other subspecies, and is endemic throughout the New World. *Rhipsalis baccifera* ssp. *hileiabaiana* grows exclusively in the Bahia region of Brazil, which contains tropical forests with increased humidity and rainfall compared to the environments of endemic growth of other New World subspecies (Korotkova et al. 2011).
Old World subspecies have morphologies, that may be associated with increased ploidy level, that allow for increased fitness in the more arid, hotter environments on the African continent. Old World subspecies typically have thicker stems covered in spines in comparison to their New World counterparts (Cota-Sanchez et al. 2010). The Old World populations and subspecies display morphological features that are advantageous in arid, hot environments. The subspecies *horrida*, which grows endemically in Madagascar and southeastern Africa, has considerably thicker and shorter stems densely covered with soft spines, which may be an advantageous characteristic to deflect solar radiation (Cota-Sanchez et al. 2010). Subspecies *horrida* is depicted in picture C of Figure 2, as an individual from Madagascar, shows this morphology.

Subspecies *horrida* also displays vivipary during germination (Cota-Sanchez, 2004). In vivipary, seeds are germinated before they are separated from the fruit or the mother plant, which provides an advantage during periods of intermittent drought during the initial stages of dicotyledonous growth (Andrade et al. 1997).

The *R. baccifera* subspecies *horrida* and *baccífera* both have light-green colored fruits and rotate flowers with radial symmetry that grow at the apex of the long hanging stems (Almeida et al. 2013; Figure2). The Old World subspecies *mauritiania*, (picture C of Figure 2), grows endemically across the African continent, Sri Lanka, and Seychelles, an archipelago located in the middle of the Indian ocean (Cota-Sanchez et al. 2010). Subspecies *mauritiania* individuals are covered sparsely in spines, and have thicker
stems. These morphological differences across subspecies may be explained by examining developmental differences associated with ploidy level, genetic divergence, or nutrient availability in the environment.

The ranges in seed size of subspecies of *R. baccifera* also suggest dramatic morphological differences across individuals. These morphological differences correlate with increased ploidy levels, with polyploids exhibiting larger cell sizes and fruit sizes (Cota-Sanchez et al. 2010). For example, *R. baccifera* ssp. *horrida* has the largest and widest seeds (around 1.37 mm) compared to other subspecies (Cota-Sanchez et al. 2010). Comparatively, the New World ssp. *shaferi* notably has smaller seed sizes consisting of only 0.38 mm in the Caribbean and 0.94 mm in samples from Argentina. These data were collected using fruits and seeds obtained from indoor botanical collections with controlled environmental conditions, which suggests that these morphological differences resulted from genetic differences (Cota-Sanchez et al. 2010). Also, these differences in seed and cell sizes attributed to varying ploidy levels affect morphological categorizations of subspecies (Cota-Sanchez et al. 2010).

**Relating ploidy and biogeography in *Rhipsalis baccifera***

*Rhipsalis baccifera* polyploid populations are no exception to the influential role polyploidy plays in plant speciation (Soltis and Soltis, 2009). However, *Rhipsalis baccifera* and its subspecies were defined taxonomically based on morphological characteristics without considering ploidy levels (Korotkova et al. 2011). Polyploidy is
specific to biogeographic populations of *R. baccifera*, with the presence of tetraploid *R. baccifera* reported to exist in the tropics of the Old World and the Caribbean, and octoploids are reportedly present in Madagascar (Cota-Sanchez *et al.* 2010; Spencer, 1995). Higher levels of ploidy correlate with increased geographic distance from Brazil, which is the origin of *Rhipsalis baccifera* and other species in the *Rhipsalis* genus (Ornelas *et al.* 2015; Calvente *et al.* 2011). Unusual distributions across geographic barriers occurred expanding from coastal Brazil into southern Brazil, the Andes, Central America, and the Caribbean, and Florida, which makes *Rhipsalis baccifera* unique from other *Rhipsalis* epiphytes (Calvente *et al.* 2011).

The subspecies *baccifera*, *mauritiania*, and *horrida* have polyploid populations (Cota-Sanchez *et al.* 2010). The base number of chromosomes for all cacti is 11. Consequently, *Rhipsalis baccifera* diploids have 22 chromosomes, tetraploids - 44, and octoploids - 88. Individual plants characterized under the subspecies of *mauritiania* and *baccifera* have been recorded as tetraploids in the Old World and parts of the Caribbean, respectively (Cota-Sanchez *et al.* 2010). *Horrida* individuals have also been recorded as tetraploids and octoploids (Cota-Sanchez *et al.* 2010).

The subspecies of *Rhipsalis baccifera* are defined by morphology that may be associated with ploidy level differences that result in larger cell and fruit sizes than their parental diploid counterparts (Cota-Sanchez *et al.* 2010). This correlation suggests mutualistic relationships with seed dispersal vectors specific to each region may lead to the genetic
isolation of *Rhipsalis* populations (Cota-Sánchez et al. 2010). For example, the presence of spines and increased stem width may deflect sun radiation and reduce the surface area to volume ratio in the octoploid *horrida* individuals (Cota-Sanchez et al. 2004). In *R. baccifera*, the tetraploid subspecies found in the Old World have larger fruit sizes compared to the New World diploid subspecies (Cota-Sanchez et al. 2010).

Although some studies have used *R. baccifera* karyotypes to determine ploidy in a few samples for intra-clade studies, there is no study that determines the *R. baccifera* ploidy counts of more than a dozen plants. The *Cactaceae* have been studied as a family full of autopolyploids, yet there is no study that relates ploidy in *Rhipsalis baccifera* to its geographic isolation and potential for genetic variation (Soltis et al. 2007; Moreno et al. 2015). Wilhelm Barthlott, a *Rhipsalis*-specialized botanist has noted the varying ploidy levels in *Rhipsalis baccifera* using the karyotyping data published in the journal *Taxon* that covered the ploidy counts of many cacti (Barthlott, 1983 and Love, 1976). Similarly, Cota-Sánchez et al. (2010) support that there are polyploids present in the Old World throughout Africa and in the New World in regions north of Brazil, but there have been no recent publications recording ploidy levels in *Rhipsalis baccifera* (Cota-Sánchez et al. 2010). Recent phylogenetic analyses of nuclear genetic data do not involve sampling individuals from different populations with different ploidy levels, but rather focus on a single subspecies without accounting for reproductive isolation from polyploidization (Ornelas et al. 2015; Calvente et al. 2011).
**Cactaceae phylogenomics**

Intrataxonomic studies for rapidly diverging clades such as the Cactaceae use highly variable regions to molecularly understand population divergence. J. Shaw developed primers that amplify the most variable intergenic chloroplast regions of angiosperms to aid botanists with low taxonomic categorizations and intra-clade species delimitations (Shaw et al. 2007). These regions included the \textit{rpl32F-trnL}^{(UAG)} intergenic region of the chloroplast genome (Calvente et al., 2011). Ornelas et al (2015) and Korotkova et al (2011) have used this region in \textit{Rhipsalis baccifera} samples to understand population divergence and clade delimitations, respectively.

Two recent publications featured the phylogenetic analysis of samples classified under the morphospecies of \textit{Rhipsalis baccifera}. One study analyzed DNA sequences of \textit{Rhipsalis baccifera} samples from Mexican and Guatemalan cloud forests to understand gene flow in species after the last glacial maximum (Ornelas et al. 2015). Ornelas sampled 154 individuals of \textit{Rhipsalis baccifera} to direct sequence the intergenic chloroplast region \textit{rpl32-trnL}, and the data suggest decreased gene flow during the historical period of the last glacial maximum (Ornelas et al. 2015). Speciation and genetic divergence could have occurred during the ice age period of isolation (Ornelas et al. 2015). The Central American isthmus decreased gene flow and separated populations, which led to restricted microenvironments for successful growth because of the patchy establishment of epiphytes in cloud forest trees (Ornelas et al. 2015).

Cloud forests are defined by their frequent or seasonal low-level cloud cover, which
usually occurs at the canopy level. This study also showed that cpDNA intergenic regions had more genetic divergence than the internal transcribed spacer ribotypes of the nuclear genome (Ornelas et al. 2015). This population genetics study by Ornelas et al. (2015) used a large dataset to understand gene flow by looking at pollen flow and seed flow using different markers, but this study did not analyze speciation events or the current reproductive isolation of the extant *Rhipsalis baccifera* populations in Mexican and Guatemalan cloud forests.

Korotkova et al. (2011) investigated species delimitation in *Rhipsalis baccifera* to understand speciation after polyploidization and geographic isolation. Previous research looking at the phylogenetic relationships of the entire *Rhipsalis* clade suggests that the subspecies *shaferi* and *hileiabaiana* of *Rhipsalis baccifera* should be classified as their own species ranking (Korotkova et al. 2011). *Rhipsalis baccifera* ssp. *shaferi* grows in the Old World and it has divergence represented in SNPs and indels in the sequenced data in the cpDNA that aligns more closely with *Rhipsalis teres*, a sister outgroup. *Rhipsalis baccifera* ssp. *hileiabaiana* grows endemically in the New World in the Bahia region of Brazil, which contains tropical forests with increased humidity and rainfall. Morphological characters can identify most species within a phylogenetic subgenera in the Rhipsalideae clade (Korotkova et al. 2011; Calvente et al. 2011). With the increase in accessibility to molecular techniques and direct sequencing, phylogenetic trees have been produced to model the evolution of the *Rhipsalis* clade, as seen in both the Korotkova and Calvente studies. These studies also use genetic molecular markers
rather than the physiological characteristics that previously defined species (Korotkova et al. 2011; Calvente et al. 2011).

**Selection of molecular markers**

A survey of the molecular markers used to classify cacti was necessary for intraspecies differentiations. The barcoding region *matK*, which is a maturase coding gene found in the intron of the chloroplast genome, has historically served as the baseline phylogenetic marker in plant studies (Korotkova et al. 2011). However, a molecular marker, found in regions of DNA with low conservation and high variability, was needed for an intraspecific phylogenetic analysis where the sequence identity across individuals was high. Two cpDNA markers used on the *Rhipsalis* clade with the most amount of variation are psbA-trnH and rps3-rpl16 (Korotkova et al. 2011). However, the available data fail to incorporate many samples of *Rhipsalis baccifera*. The intergenic region rpl32-trnL(UAG) of the chloroplast genome, which evolves more rapidly compared to the nuclear genome, has been used for species identification of *Rhipsalis* and to find genetic variation amongst populations of *Rhipsalis baccifera* (Ornelas et al. 2015). Similarly, microsatellites or variable number of tandem repeats have been used in other phylogenetic studies on cacti for species delimitations, but are difficult to locate using polymerase chain reaction amplification and direct sequencing (Arakaki et al. 2011). High throughput sequencing, provides a costly, yet effective alternative method for locating these repeats. Allozyme analysis using proteomic biochemistry also reveals population divergence. Phosphoglucose isomerase, in particular, is an enzyme with a lot
of allelic variation at the population level (Terauchi et al. 1997). Looking at allelic variation by direct sequencing and observing nucleotide differences provides an alternative molecular marker for low taxonomic phylogenies.

**Nuclear gene analysis- Phosphoglucone isomerase**

Chloroplast sequencing data only provides information for maternal inheritance because plasmid DNA is transferred through the egg cytoplasm. Therefore, nuclear DNA sequences are necessary to provide information representative of pollen flow and seed flow to suggest phylogenetic differences across both maternal and paternal inheritance (Ornelas et al. 2015). The intronic regions of nuclear genes are not functionally conserved, and as a result have faster evolution rates because genetic variation in these regions results in neutral fitness. The introns of phosphoglucone isomerase, in particular, have high sequence divergence that can be used to determine population differences within a single species (Toulson Wimmer and Merritt, 2008). The metabolic enzyme phosphoglucone isomerase catalyzes the reversible isomerization of glucose-6-phosphate and fructose-6-phosphate during the second step of the glycolytic pathway (Terauchi et al. 1997). Phosphoglucone isomerase has been used to understand population divergence in different organisms including the Japanese yam (*Dioscorea tokoro*) and the Joshua tree (*Yucca brevifolia*) (Toulson Wimmer and Merritt, 2008). Terauchi et al. (1997) designed degenerate primers to amplify the partial coding region of phosphoglucone isomerase in the Japanese yam (*Dioscorea tokoro*), which had a product length of approximately 750 base pairs (Terauchi et al. 1997). These degenerate
primers could be used to initially amplify and obtain the sequence of this partial coding region of PGI in other plants.

Starch gels have been historically used to determine the electrophoretic mobility of allozymes. Allozymes include all allelic differences in enzymes that may or may not be structurally or functionally conserved. Starch gels can be used to determine allozyme differences across populations because the allelic products have charge differences from nonsynonymous mutations that affect the way the proteins move across an electrophoresis gel. Although these methods account for varying sizes and shapes of proteins that have been expressed differently, their actual genetic sequences cannot be determined (Martinez-Palacios et al. 1999). Direct sequencing to understand variations in the intronic regions of phosphoglucose isomerase provides an alternative method to observe genetic divergence in non-functionally conserved regions with fast evolution rates. The nuclear gene phosphoglucose isomerase was selected to determine how many alleles are included in populations of *Rhipsalis baccifera* (Terauchi et al. 1997).

**Research questions**

I hypothesize that the various subspecies of *Rhipsalis baccifera* include diploids, tetraploids and octoploids that correspond with the putative polyploidy populations reported in the literature (Cota-Sanchez et al. 2010). The polyploid populations likely resulted from autopolyploidization events, which would be evident if a species identified as a polyploid was homozygous for a gene that had fixed alleles in related
species. *Rhipsalis baccifera* samples are often collected in isolation from other plants of the same species, which is a characteristic of autoployploids. This makes a within species polyploidization event more likely because only unreduced gametes from a selfing plant can result in an autoployploid.

I hypothesize that molecular data will correlate genetic divergence with individuals with different ploidy levels because reproductive isolation from polyploidization and geographic separation leads to genetic isolation. I expect that genetic divergence between the New World and the Old World populations occurred resulting in dissimilar sequences. Looking at a larger sampling size of ploidy levels in *Rhipsalis baccifera* provides a framework for the biogeographic ranges of subspecies. In particular, I hypothesize that sequence divergence will be present in samples with further geographic locations from Brazil, the putative origin of all *Rhipsalis* diversity and *Rhipsalis baccifera*. The *Rhipsalis baccifera* samples in Madagascar, and the Caribbean will notably have higher ploidy levels because these populations have had longer periods of geographic isolation to have independent autopolyploidization events. Therefore, I hypothesize that the dispersal of *Rhipsalis baccifera* to the Old World occurred as an ancient event. Understanding both the phylogenetic and cytologic components of plant speciation is needed to categorize *Rhipsalis baccifera* and other species of cacti because of its wide distribution and different ploidy levels.
In addition, I predict that if future studies investigate the rapidly evolving nuclear gene phosphoglucone isomerase, allopolyploids will have more variation in the alleles in both intronic regions and allozyme variants because these differences can occur independently in two different species and become included during the hybridization event that results in polyploidy. Autopolyploids will show less variation in the allozymes because there is no hybridization event that introduces genetic variation in a homozygous population.

Materials

Database assembly

This dataset included 49 *Rhipsalis baccifera* individuals. The morphology and the putative locations of collection for these individuals are listed (Table 1). This dataset does not include individuals morphologically determined as *Rhipsalis baccifera*, but rather more genetically similar to another species in the *Rhipsalis* genus. The living botanical collection in the Lyman Plant House at Smith College provides a diverse sampling of plants in the *Rhipsalis* genus and many individuals identified as *Rhipsalis baccifera* in the hot genetics room (approximately 25° C). These samples also provided the materials needed for microscopy and DNA extraction for cytology and molecular analysis, respectively. The collection sites of the individual samples span a wide geographic breadth with locations in the Caribbean, Central America, South America, Eastern Africa, Western Africa, Madagascar, Sri Lanka, and India. Morphology varies regarding stem color, stem width, spine density that corresponds to the identification of
subspecies (Figure 3). The geographic breadth of the samples used in this study accounted for genetic divergence associated with geographic isolation (Figure 3).

Most individuals were procured from collections at other botanical institutions including Huntington Botanical Gardens, University of California Botanical Garden at Berkeley, the New York Botanical Garden, a collection from South Africa, and the Kew Royal Botanic Gardens. The accession numbers specific to these botanical gardens’ inventory are listed in the samples used in this study. The five samples from Puerto Rico were acquired from the wild for this project from a single neighborhood. Similarly, Rob Nicholson, the manager of the Lyman Conservatory, collected two samples from Brazil from the wild. Mr. Nicholson anecdotally noted that *Rhipsalis baccifera* individuals are collected in isolation in tropical rainforests.

**Table 1.** Collection location and morphology of all samples confirmed as *Rhipsalis baccifera*. New World subspecies include *hileiabaiana* in Northeastern Brazil, *shaferi*, and *baccifera*. Old World subspecies include *horrida* in Madagascar, *mauritiania*, and *erythrocarpa*. The morphological identifiers of each subspecies correspond to the data set. Subspecies were sampled multiple times. Similarly, the dataset reflects the large geographic distribution of the individuals in the Lyman Plant House collection.

<table>
<thead>
<tr>
<th>NEW WORLD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geographic location of sample</strong></td>
<td><strong>Morphology (stem, color, spine density)</strong></td>
</tr>
<tr>
<td>Everglades, Florida</td>
<td>Thin, dark color, no spines</td>
</tr>
<tr>
<td>Puerto Rico #1</td>
<td>Thin, light green, no spines</td>
</tr>
<tr>
<td>Puerto Rico #2</td>
<td>Thin, light green, no spines</td>
</tr>
<tr>
<td>Puerto Rico #3</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Location Description</td>
<td>Morphology</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Puerto Rico #4</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Puerto Rico #5</td>
<td>Thin, dark color, no spines</td>
</tr>
<tr>
<td>Chiapas, Mexico</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Monaco, Mexico</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Oaxaca, Mexico</td>
<td>thin, green, no spines</td>
</tr>
<tr>
<td>Mexico (wild collected)</td>
<td>thin, green, no spines</td>
</tr>
<tr>
<td>Vera Cruz, Mexico (Berkeley BG)</td>
<td>thin, light green-yellow, no spines</td>
</tr>
<tr>
<td>Vera Cruz, Mt. Blanco, Mexico</td>
<td>thin, light-green, no spines</td>
</tr>
<tr>
<td>Vera Cruz (South African collection)</td>
<td>thin, green, no spines</td>
</tr>
<tr>
<td>Botswana or Vera Cruz (location unknown)</td>
<td>thin, green, no spines</td>
</tr>
<tr>
<td>Colombia 42082</td>
<td>thin, light green-yellow, no spines</td>
</tr>
<tr>
<td>Ecuador 408-84</td>
<td>thin, green, no spines</td>
</tr>
<tr>
<td>Peru (wild collected; Berkeley BG)</td>
<td>thick, green, no spines</td>
</tr>
<tr>
<td>Brazil #1 (wild collected by Rob Nicholson)</td>
<td>medium thickness, green, no spines</td>
</tr>
<tr>
<td>Brazil #2 (wild collected by Rob Nicholson) (<em>Rhipsalis teres</em>)</td>
<td>medium thickness, green, no spines</td>
</tr>
<tr>
<td>Subspecies shaferi</td>
<td>thin, light green, no spines</td>
</tr>
<tr>
<td>70-97 (South African collection)</td>
<td>thin, light green, no spines</td>
</tr>
<tr>
<td>NYBG 587//80</td>
<td>medium thickness, green, no spines</td>
</tr>
<tr>
<td>San Luis Potosi, Mexico (Kew RBG)</td>
<td>(No morphology data)</td>
</tr>
<tr>
<td>2000-420 (unknown location) (Kew RBG)</td>
<td>(no morphology data)</td>
</tr>
<tr>
<td>Rupununi District, Guyana (Kew RBG)</td>
<td>Thick, green, no spines</td>
</tr>
<tr>
<td>Sierra Cubitas, Cuba 1996-2735 (Kew RBG)</td>
<td>(no morphology data)</td>
</tr>
<tr>
<td>Floresta Azul, Bahia, Brazil 1994-1144 (Kew RBG)</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Georgetown, Guyana 1979-3177 (Kew RBG)</td>
<td>Thick, green, no spines</td>
</tr>
<tr>
<td>Bahia, Brazil 1977-652 (Kew RBG)</td>
<td>Thick, green, no spines</td>
</tr>
<tr>
<td>Location</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Floresta Azul, Bahia, Brazil (Kew RBG)</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Floresta Azul, Bahia, Brazil subsp. hileiabaiana (Kew RBG)</td>
<td>Thin, green, no spines</td>
</tr>
<tr>
<td>Estado Falcon, Venezuela (Kew RBG)</td>
<td>(seedling material)</td>
</tr>
<tr>
<td>Subspecies shaferi (Kyoto BG)</td>
<td>Thick, green, no spines</td>
</tr>
<tr>
<td>Peru Huntington (Rhipsalis neves-armondii)</td>
<td>thick, green, no spines, ribbed stem</td>
</tr>
<tr>
<td>Colombia 97166 (Rhipsalis neves-armondii)</td>
<td>thick, green, no spines, ribbed stem</td>
</tr>
</tbody>
</table>

**OLD WORLD**

<table>
<thead>
<tr>
<th>Location</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Madagascar 85646</td>
<td>thick, light-green/green, no spines</td>
</tr>
<tr>
<td>Madagascar 46-09</td>
<td>thick and clubby, dark green, spines</td>
</tr>
<tr>
<td>Madagascar New York BG 407/84</td>
<td>thin, green, no spines</td>
</tr>
<tr>
<td>subspecies horrida (found in Madagascar)</td>
<td>thick, green, spines</td>
</tr>
<tr>
<td>Tanzania</td>
<td>thin, dark green, no spines</td>
</tr>
<tr>
<td>Cameroon Huntington 35802</td>
<td>medium thickness, green, no spines</td>
</tr>
<tr>
<td>Zimbabwe, Chiande Forest #1</td>
<td>medium thickness, dark green, sparse number of spines</td>
</tr>
<tr>
<td>Zimbabwe, Chiande Forest #2</td>
<td>medium thickness, dark green, sparse number of spines</td>
</tr>
<tr>
<td>Zimbabwe 68-98 (South African collection)</td>
<td>thin to medium thickness, dark green, sparse spines</td>
</tr>
<tr>
<td>India (putatively collected)</td>
<td>thin, dark-green, no spines</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>thin, green, sparse number of spines</td>
</tr>
<tr>
<td>Sierra Leone (Kew RBG)</td>
<td>(no morphology data) (epiphytic)</td>
</tr>
<tr>
<td>Nwarakaya, Kenya (Kew RBG)</td>
<td>thick, green, red tips at stem shoots, no spines (lithophytic)</td>
</tr>
<tr>
<td>Madagascar 1982-2498 (Kew RBG)</td>
<td>Thick, green, spines,(octoploid)</td>
</tr>
<tr>
<td>Subspecies mauritiana (Kew RBG)</td>
<td>(seedling material)</td>
</tr>
<tr>
<td>Natal, South Africa Subspecies mauritiana</td>
<td>(no morphology data)</td>
</tr>
<tr>
<td>Liberia subspecies mauritiana</td>
<td>Thick, green, no spines (epiphytic)</td>
</tr>
</tbody>
</table>
Figure 3. Geographic distribution of samples used in this study

Figure 3. Geographic distribution of *Rhipsalis baccifera* based on putative collection sites. The number depicted in the country-based stars represents the number of individual plants collected from a named location. This study produces a large and widely distributed set of samples to understand long-range recent dispersal.

Methods

DNA Extraction and Amplification

Direct sequencing results can be used to determine phylogenetic relationships between individuals but require that a specific region have many replicates. Polymerase chain reaction (PCR) amplifies specific sequences using oligonucleotides with sequence complementarity to DNA regions that flank the target region. The Phire Plant Direct PCR
kit with Phire Hot Start II DNA Polymerase was used for all polymerase chain reactions (ThermoFisher Scientific). Purified template DNA was not necessary under the protocol of the direct PCR kit. Instead, macerated plant material from the stems of *Rhipsalis baccifera* was added to dilution buffer, which was provided in the kit, and vortexed for three seconds to prepare the 1x template solution. 1:10 dilutions of the template DNA solution were made in 1x TE buffer (Tris pH 8.0 buffer combined with 1 mM EDTA) and used in the PCR reaction. This template DNA worked well for PCR reactions involving chloroplast DNA, with higher genome copy numbers. However, the template DNA preparation using the dilution buffer may have introduced impurities and confounding variables that inhibited the amplification of nuclear genes like phosphoglucone isomerase. I decided that an alternative DNA extraction method would be necessary to yield a repeatable PCR product.

Qiagen DNEasy Plant Minikit (Valencia, California, USA, cat. nos. 69104 and 69106) was used to obtain a purified DNA template. In short, the lytic release of DNA from *Rhipsalis baccifera* was performed using a detergent solution with RNase in a silica-membrane column purification process. The DNA binds to the spin columns and is removed using an elution buffer. DNA extractions were prepared using the kit’s provided instructions. The protocol for DNA extraction was slightly altered to account for the highly mucilaginous tissue of succulent stems (appendix: protocol 1). The eluted DNA was used directly in PCR reactions, and 1:10 dilutions of the template DNA from the extraction were also prepared in 10mM Tris (pH 8), for use in PCR to standardize
variable DNA extraction concentrations. Variations in DNA concentration resulted from different stem tissue quality and the amount of tissue incorporated into the lysis buffer during the maceration step.

Polymerase chain reaction (PCR) was used to amplify the rpl32-trnL\textsuperscript{UAG} chloroplast intergenic region and the nuclear coding region of the gene phosphoglucone isomerase. The concentration and volumes of reagents are listed (Table 2). The cycling conditions of the chloroplast PCR were designed according to Joey Shaw’s conditions specific to the intergenic chloroplast primers (Shaw \textit{et al.} 2007). PCR reactions amplifying chloroplast DNA (20 uL) were run under the following thermocycling conditions (98 °C for 5:00, 98°C for 1:00, 56 °C for 4:00, 56°C - 65°C temperature ramp at 0.3°C /second, 65°C for 4:00, repeat from step two 27 times, 65°C for 5:00, and 4°C for forever). The stoichiometry of reagents used came from the suggested quantities in the Phire Plant Direct PCR Kit. The sequences of the primers used to amplify the chloroplast DNA and direct sequencing reactions are listed in Table 3. The PCR products were cleaned using ExoSAP-IT® (1 uL per 2.5 uL of PCR product), which eliminates unincorporated dNTPs and primers to avoid the direct sequencing of the non-target oligonucleotide sequences (Affymetrix). The cleaned PCR products (3.5 uL) were prepared for direct sequencing with the addition of BigDye Terminator v3.1 Direct Cycle Sequencing Kit (Applied Biosystems, catalog no. 4337455) (1 uL), an oligonucleotide primer complementary to the target region (1 uL), and ddH\textsubscript{2}O for dilution (4.5 uL). The Seq2 thermocycling program used to incorporate
BigDye involved the following steps (95°C for 0:30, ramp to 55°C for 0:10, ramp to 60°C for 4 min., repeat cycle 49 more times, and held at 4°C for forever).

The Seq2 products were cleaned through two different processes including an Edge Bio column (Performa Spin Columns, catalogue no. 13622) cleaning procedure and the BigDye Xterminator Purification (Thermofisher, catalogue no. 4376486) protocol. The Edge Bio column procedure removed unwanted salts, amino acids, nucleotides, and solvents using a fully hydrated gel matrix filter. Each spin column was centrifuged (3000 rpm for 3-5 minutes) and then placed into a new 1.5 mL tube. Seq2 PCR product was added to the column and spun down (3000 rpm, 3-5 minutes) in a new tube. The flow-through was sent for processing. The BigDye Xterminator removes excess BigDye terminators and salts. SAM™ solution (45 uL) and the Xterminator solution (10 uL) were added to Seq2 product (10 uL). The DNA, BigDye and Xterminator mix was vortexed on a foam adapter for 30 minutes in 1.5 uL falcon tubes and centrifuged at 1000 rpm for 2 minutes. The direct Sanger sequencing products were processed at the Center for Molecular Biology using an automated machine (Applied Biosystems 3130xl Genetic Analyzer) that detects the fluorescently labeled dideoxynucleotides to yield DNA sequences.

**Molecular Markers for Species Identification**

The chloroplast intergenic region of rpl32-trnL<sub>UAG</sub> was used for species identification due to its documented use in *Rhipsalis* phylogeny studies (Bonatelli et al. 2013). Direct
sequencing results of the rpl32-trnL$^{UAG}$ region in the dataset were compared to the nucleotide database from National Center for Biotechnology Information (NCBI) with the basic local alignment search tools (Blast). This tool allows for comparisons across a database of biological sequence information for species identification purposes within our samples. Individuals with a high sequence identity (ranging from 94-100%) with *Rhipsalis baccifera* were confirmed to be *Rhipsalis baccifera*.

DNA sequencing data was used to determine the amount of genetic divergence between individuals from different populations. Increased genetic divergence is determined by accounting for single nucleotide polymorphisms, insertions, and deletions that differ from the consensus sequence of *Rhipsalis baccifera*. Phylogenetic studies use aligned DNA sequences to compare individual nucleotide and spacing differences to determine the relationships between individual samples. The quality of a polymorphism was determined using the individual peaks, derived from the signal from reporter fluorophores in the sequencing reaction, by comparing a sequence alignment with other sample sequences that have a homologous polymorphism, or by checking the consistency of the widths and heights of each peak in Seqman Pro v.6.1 (DNASTar, Madison, WI).

Seqman Pro was used to create assemblies of the sequencing data after species identification. All samples were named for their sampling location. Assemblies were made by trimming sequence ends, scanning for vectors, optimized for sequence
assembly order, and omitting single sequence contigs using the Classic Assembler. I exported the contig as a single file and transferred the assembled sequences into Seaview. An alignment of the *Rhipsalis baccifera* sequences was made using Seaview V.3.2, which is an open-sourced program used for sequence alignment and molecular phylogeny. G blocks were added, using Seaview, to eliminate poorly aligned sequences to provide a reproducible set of conditions before phylogenetic analysis was performed. Distance trees were made in Seaview using 100 bootstrap replicates, which acts as a confidence interval with percentages of getting the same phylogenetic groupings from the tree reconstructions (Figure 4). Branch lengths of the tree act as a metric for time of genetic divergence.

**Table 2.** PCR reaction mixture per 20 uL reactions for chloroplast DNA intergenic region amplification is described as volumes used of each reagent. These reagent stoichiometry values were based on the control PCR reaction included with the Phire Plant Direct PCR Kit.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x Phire Plant Buffer</td>
<td>10 uL</td>
<td>Includes salts and dNTPs</td>
</tr>
<tr>
<td>PCR water</td>
<td>8.7 uL</td>
<td>Nuclease free</td>
</tr>
<tr>
<td>Forward Primer (25 uM)</td>
<td>.2 uL</td>
<td>Diluted in Tris buffer</td>
</tr>
<tr>
<td>Reverse Primer (25 uM)</td>
<td>.2 uL</td>
<td>Diluted in Tris buffer</td>
</tr>
<tr>
<td>Phire Hotstart II Polymerase</td>
<td>.4 uL</td>
<td></td>
</tr>
<tr>
<td>Template plant material in buffer</td>
<td>.5 uL</td>
<td>Varied in concentration depending on how much plant material was diluted</td>
</tr>
</tbody>
</table>

I used chloroplast DNA for species identification because it is the portion of the plant genome with the fastest mutation rate, which is advantageous for an intraspecific
taxonomic study. The intergenic chloroplast region utilized for species identification was the rpl32-trnL^{UAG} region, which is around 1116 base pairs long for Rhipsalis baccifera and related species with about 70 informative characters in the form of insertions, deletions, and SNPs (Ornelas et al. 2015). It was identified as the chloroplast region with the largest number of informative characters in comparison to other intergenic regions (Korotkova et al. 2011). However, this region did not contain enough variation between individual sequences for distinctions at the intraspecific level. I first sequenced the outer regions of the rpl32-trnL^{UAG} region with the primers detailed in Shaw et al. (2007). However, these regions had homologous sequences across the individual samples. The sequences from Ornelas et al. (2015) suggest that the middle region of rpl32-trnL^{UAG} had the most informative characters. Joey Shaw’s chloroplast markers for low taxonomic phylogenetic studies included the sequences for the oligonucleotides used as primers in the polymerase chain reaction for amplifying the rpl32-trnL^{UAG} region, and are listed in Table 3 (Shaw et al. 2007). I designed the oligonucleotides, which are also listed in Table 3, in SeqBuilder to be specific to sequences of Rhipsalis baccifera for this study in order to prime the middle sequencing reactions that were not sufficiently sequenced using Shaw’s primers.

Table 3. Primer sequences for PCR and direct sequencing reactions

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequences 5’ to 3’</th>
<th>Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpl32F (Shaw et al. 2007)</td>
<td>CAGTTCCAAAAAAACGTACTTC</td>
<td>Chloroplast</td>
</tr>
<tr>
<td>trnL^{UAG} (Shaw et al. 2007)</td>
<td>CTGCTTCCTAAGAGCAGCGT</td>
<td>Chloroplast</td>
</tr>
<tr>
<td>seq1 rpl32-trnL (this study)</td>
<td>CGGAATTGCTATATACATACCTACC</td>
<td>Chloroplast</td>
</tr>
<tr>
<td>seq2 rpl32-trnL (this study)</td>
<td>GGAATTAATGCTTTGCATTCCC</td>
<td>Chloroplast</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Phosphoglucone isomerase forward (Terauchi et al. 1997)</td>
<td>TTYGCNTTTYGGAYTGGGT</td>
<td>Nuclear</td>
</tr>
<tr>
<td>Phosphoglucone isomerase reverse (Terauchi et al. 1997)</td>
<td>TCIACICCCCAITGRTCTAAIGARTTIAT</td>
<td>Nuclear</td>
</tr>
</tbody>
</table>

**Table 4.** PCR reaction mixture per 20 μL reactions for nuclear DNA phosphoglucone isomerase coding region amplification is described as volumes used of each reagent. These reagent stoichiometry values were based on the control PCR reaction included with the Phire Plant Direct PCR Kit.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x Phire Plant Buffer</td>
<td>10 μL</td>
<td>Includes salts and dNTPs</td>
</tr>
<tr>
<td>PCR water</td>
<td>8.2 μL</td>
<td>Nuclease free</td>
</tr>
<tr>
<td>Forward Primer (25 μM)</td>
<td>.2 μL</td>
<td>Diluted in Tris buffer</td>
</tr>
<tr>
<td>Reverse Primer (25 μM)</td>
<td>.2 μL</td>
<td>Diluted in Tris buffer</td>
</tr>
<tr>
<td>Phire Hotstart II Polymerase</td>
<td>.4 μL</td>
<td></td>
</tr>
<tr>
<td>Template plant DNA from Qiagen DNeasy Extraction Kit</td>
<td>1.0 μL</td>
<td>Varied in concentration depending on how much plant material was diluted</td>
</tr>
</tbody>
</table>

The highly similar sequences in the chloroplast data suggested that population divergence could not be determined using the rpl32-trnL\text{UAG} molecular marker.

Therefore, I looked towards amplifying the coding region of phosphoglucone isomerase in *Rhipsalis baccifera*, which may have a faster evolution rate than the intergenic regions of chloroplast DNA (Terauchi *et al.* 1997, Toulson Wimmer and Merritt, 2008).
I set out to amplify the coding region of phosphoglucone isomerase in *Rhipsalis baccifera* with a PCR using degenerate primers designed for any eukaryote. The PGI PCR product for *Rhipsalis baccifera* was expected to be around 750 base pairs. PCR amplification of the PGI coding region in *Rhipsalis baccifera* is different from PCR of chloroplast intergenic regions because nuclear gene copy number is significantly lower than the gene copy number in chloroplast genomes. The reagents and stoichiometry of reagents used in the phosphoglucone isomerase PCR reactions are described in Table 4. An increase in *Rhipsalis baccifera* template DNA was used to account for the difference in gene copy number. The DNA templates used in all PGI reactions and the cpDNA reactions on the samples received from the Kew Royal Botanic Gardens were extracted using the Qiagen DNeasy Extraction Kit since the direct PCR using macerated plant material likely did not provide sufficiently high quality DNA for PCR reactions for PGI. PCR reactions amplifying PGI DNA (20 uL) were run under the following thermocycling conditions (98°C for 5:00, 98°C for 1:00, 59°C for 0:30, 72°C for 1:00, repeat from step two 59 times, 72°C or 1:00, and 4°C or forever). An unusually high number of PCR cycles (60 cycles) was needed to amplify template DNA using the degenerate PGI primers from Terauchi *et al* (1997).
Histology and Microscopy

*Rhipsalis baccifera* root tissue was stained with the Feulgen-technique and observed with brightfield microscopy to estimate nuclear size and ploidy levels. Obtaining a consistently healthy and mitotically dividing source of root tissue that could be squashed on a microscope slide, yet avoid mold growth was difficult. Consequently, A grow-light chamber with 12 hour day: night cycling was utilized to regulate the mitotic rhythm of the *Rhipsalis baccifera* cutting samples that were grown in test tubes of water.

Root tissue was harvested in the late morning (9 am to 11 am) and fixed in 3:1 ethanol: acetic acid for at least 15 minutes at room temperature. The metaphase figures gleaned from the initial brightfield microscopy work appeared most often in the early morning samples. Studies using mitotic division and karyotyping spreads to understand replication in Cactaceae also concluded that the two hours after sunrise provided the mitotic figures as the most active time period for cellular division (Majure et al. 2012; Cota and Philbrick, 1994). Consistently finding samples with active mitotic division proved difficult. Instead, interphase nuclear area values were used to estimate ploidy levels with larger areas corresponding with high levels of polyploidy.

Growing mitotically dividing root tips in sand and perlite media proved difficult during the initial attempts to produce root squashes on histological glass stains for brightfield microscopy. The grains of sand prevented clean root tips from being secured without
damaging the zone of high division upon deracination. In addition, sand particles that could not be removed with sonication broke the coverslips, which prevented a thin root squash preparation needed for visualization with the light microscope. Despite this, some root tips grown in sand were imaged and used in data collection. Cuttings placed in water grew roots, but were prone to rotting. Indole-3-butyric acid acted as a rooting hormone and was added to the water for hydroponic growth of the roots. These water-grown roots acted as the best samples for producing slides for mitotic division, but also took longer to grow.

A standard Feulgen procedure was used to stain the root tip nuclei in *Rhipsalis baccifera* and two positive controls. *Lathyrus odoratus* (sweet pea) and *Allium sativum* (garlic) were germinated hydroponically, and their root tips were used as positive controls for the staining and slide preparation. These species were chosen as controls because *Allium sativum* has a relatively large genome size of 2C=32.46 pg according to Ohri et al. (1996), and *Lathyrus odoratus* has a relatively small genome size of 2C = 15.56, according to Murray et al. (1992) (Plant DNA c-values database, Kew Royal Botanic Gardens).

Roots were harvested from stem cuttings and placed in fixative (3:1 Ethanol:Acetic Acid) for at least 15 minutes at room temperature. After two sets of five-minute rinses in distilled water, the root tips were incubated in hydrochloric acid (1 M, 60 °C) in a circulated water bath for 10 minutes. The root tips were rinsed in distilled water twice
before being placed in Schiff’s reagent for aldehydes (Sigma-Aldrich, catalogue no. 84655) for at least 30 min. Schiff’s reagent binds fuchsin to the free-aldehydes created after a strong-acid hydrolysis for a DNA-specific staining technique. The stained meristematic tissue of the roots contained the highest density of nuclei, and this area was prepared for imaging. All root tips were squashed with 45% acetic acid and imaged on an Olympus BX-51 brightfield microscope with the 40x objective.

Feulgen staining binds stoichiometrically to the amount of DNA present, and can be used to identify nuclear genome size. In addition, Feulgen stain does not photobleach as much as other nuclear specific stains like DAPI and propidium iodide (Gallardo-Escarate et al. 2007). Slides prepared with the Feulgen technique can still be imaged and analyzed using ImageJ software. The free-hand cropping tool in ImageJ (an open-source image-analysis software developed by W. Rasband at the NIH, USA) was utilized to trace the perimeter of each stained nucleus. ImageJ software was used to calculate the area of each nucleus using the elliptical cropping tool. Each image was analyzed by setting the image scale equal to the length of the scale added using the Olympus microscope software. I averaged the area of at least 10 nuclei per individual root to yield data for nuclear size estimates in μm².
Results

Species Identification

The rpl32-trnL^{UAG} intergenic chloroplast region is sufficiently variable to distinguish for outgroup species in the *Rhipsalis* genus, but was not sufficient for detecting divergence at the intraspecies level for *Rhipsalis baccifera*. The samples identified as *Rhipsalis baccifera* using this marker are shown with all of the sampling locations (Table 1).

Morphological differences between samples from the New World and the Old World were observed despite their highly similar sequences (Table 1). I observed and recorded that the Old World samples have morphological features specific to the subspecies, such as increased spine density and darker green stems for the subspecies *mauritiana*.

The majority of the samples in my data set were BLASTed and showed sequence identities most similar to *Rhipsalis baccifera*. Four samples were initially characterized as *Rhipsalis baccifera* due to morphological characteristics, but they had sequence identities that were more similar to *Rhipsalis neves-armondii*, *Rhipsalis teres*, *Rhipsalis pilocarpa*, or *Rhipsalis puniceodiscus*. These samples lacked spines and had green, cylindrical, thin stems that looked similar to *Rhipsalis baccifera*. However, morphological characters such as flower positioning and fruit color can only be determined seasonally, and may not have been accounted for during the initial species identification process in various botanical gardens. Sample #2 from Brazil had BLAST results with sequence identities most similar to *Rhipsalis teres*, which is sister to *Rhipsalis baccifera*. The BLAST results for Brazil sample 2 collected in the wild by Rob Nicholson showed a sequence
identity of 98% and a query cover of 80% to *Rhipsalis teres*. The plants acquired from the Kew Gardens from Brazil included a set of four from the Bahia region. Samples Bahia_8, Bahia_6 and Bahia_14 were morphologically characterized as subspecies *hileiabaiana* by the Kew Royal Botanical Gardens. The samples morphologically appear like hybrids between the two sister species. The other two were identified as subspecies *baccifera* using morphological characteristics.

I sampled from the individuals from Brazil from the Kew collection, Brazil #1 (*Rhipsalis baccifera* ssp. *baccifera*), and Brazil #2 (*Rhipsalis teres*) to compare genetic divergence associated with subspecies morphologies that were not geographically separated. The BLAST results for the *hileiabaiana* subspecies samples are Bahia_14 (sequence identity of 94%, query cover 91%), Bahia_8 (sequence identity of 98%, query cover 95%), and Bahia_6 (sequence identity of 92%, query cover 88%) for sequence similarity to *Rhipsalis baccifera*. In comparison, Brazil #1 subspecies *baccifera* sample collected from the same biogeographic region has a sequence identity of 99% and a query cover of 90%. The sequence identity is not consistent across the subspecies *hileiabaiana* (ranging between 98% and 92%). However, the chloroplast sequence from Bahia_6 and Bahia_14 did have low sequence identity, defined as a similarity below 95%, which suggests speciation.

Subspecies *hileiabaiana* is morphologically distinct from other subspecies and only occurs in the Bahia region of Brazil, which suggests that it is geographically isolated from *Rhipsalis baccifera* individuals. Other outgroup taxa, such as *Rhipsalis teres* had sequence identity to *Rhipsalis baccifera* of around 90%, which suggests that subspecies
While *hileiabaiana* might also be merit species ranking. Sample Bahia_8 may be more accurately categorized molecularly as subspecies *baccifera*. The relationship between the Brazilian samples and the categorized subspecies are shown (Figure 5). The phylogeny presented in Figure 5 also suggest that *Rhipsalis teres* may be more similar to *Rhipsalis baccifera* subspecies *baccifera* in comparison to subspecies *hileiabaiana*, which was unexpected. If I rerooted the tree to position *Rhipsalis teres* as an outgroup, I would expect to find *R. baccifera* and *R. baccifera* ssp. *hileiabaiana* to be sister taxa, with *R. baccifera* subspecies *baccifera* furthest away from *Rhipsalis teres*. Subspecies *hileiabaiana* is putatively a hybrid between the two sister taxa, but more similarly related to *R. baccifera* subspecies *baccifera*, the divergent species to all *Rhipsalis baccifera* subspecies.

The subspecies *shaferi* has indeterminate stem lengths, and was previously considered to be its own species separate from *Rhipsalis baccifera* due to this morphological difference and was investigated separately from the other samples (Korotkova et al. 2011). Morphologically, these subspecies *shaferi* individuals have thinner, lighter colored stems that tapered at the tips. The individuals identified as subspecies *shaferi* from the Kew Botanical Gardens from Sao Paolo, Brazil (92% identical with 90% query cover) and the Kyoto Botanical Gardens (93% identical with a 90% query cover) both had dissimilar identity sequence percentages in comparison to other New World individuals that were BLASTed. These sequence identity differences still grouped the subspecies *hileiabaiana* and *shaferi* individuals with other *Rhipsalis baccifera* individuals (Figure 4).
The *R.neves-armondii* and *Rhipsalis puniceodiscus* samples formed a clade with a bootstrap support of 99 when compared to the *Rhipsalis baccifera* samples (Figure 4).

**Figure 4.** Distance tree for all samples used in this study for intergenic region rpl32-trnL.

*Figure 4.* A Jukes-Cantor distance tree (239 sites) using Bio-neighbor joining and 100 bootstrap replicates to support branch groupings. G-blocks were applied. The outgroup species are labelled in red boxes. The long branches showing the divergence of the outgroup species from *Rhipsalis baccifera* has been purposefully cut off so that the
sample labels are legible. The outgroup samples Colombia_97166 and Peru_Huntington are both *Rhipsalis neves-armondii* and form a clade that has a bootstrap support of 99. The low bootstrap support of other branches shows that the rpl32-trnL\textsubscript{(UAG)} sequences for *Rhipsalis baccifera* are highly similar and cannot be genetically differentiated on this tree.

**Figure 5.** Distance tree for Brazilian samples for the intergenic region rpl32-trnL\textsubscript{(UAG)}

**Figure 5.** A Jukes-Cantor distance tree (775 sites) using Bio-neighbor joining and 100 bootstrap replicates to support branch groupings. The subspecies group together for *hileiabaiana* and *baccifera*. The outgroup Brazil 2 is *Rhipsalis teres*, which is grouped with other *Rhipsalis baccifera* individuals in the larger dataset.

**Phosphoglucone isomerase**

I did not succeed in amplifying phosphoglucone isomerase in *Rhipsalis baccifera*.

Instead, the phosphoglucone isomerase nucleotide BLAST result suggests that the DNA amplified and sequenced was from a bacterial contaminant rather than a plant. The results included sequence identity ranging from 81-83% associated with bacterial strains including *microbacterium spXT11*, *cellvibrio gilvus*, and *propionibacterium*.

**Microscopy-**
Morphology of nuclei

I observed that root tip nuclei morphology varied across the different taxa sampled in this study. Regardless of the histology technique applied to the root squash, the nuclei of *Rhipsalis baccifera* stained lightly in a speckled pattern in the center and stained darkly on the perimeter of the nucleus (Figure 6). In comparison, *Allium sativum* (garlic) root tip nuclei stained darkly and densely throughout the entire area with clear regions indicating the presence of nucleoli. The *Lathyrus odoratus* nuclei also stained darkly throughout the entire area of the nucleus, but less prophase nuclei were observed in the prepared root squash in comparison to the garlic root tips.

*Rhipsalis baccifera* roots divided and grew much slower than *Allium cepa*. A time period of 24 hours would yield roots tips of 3-5 mm in length from garlic, which was stained, squashed, and imaged as a positive control in Figure 6. Contrastingly, the *Rhipsalis baccifera* roots would take up to a few weeks to divide, grow, and elongate in order for root tissue cells to be used in histology and image collection. Also, mitosis occurred more frequently in *Allium sativum* in comparison to *Rhipsalis baccifera*. Karyotyping was the first strategy used to determine ploidy levels in *Rhipsalis baccifera*, but the lack of mitotic figures led to a methodological switch to nuclear genome size estimates to determine ploidy level. The *Rhipsalis baccifera* root tips grown in sand were prone to desiccation, and the cuttings grown in water likely did not have enough gas exchange. These cuttings yielded roots with protoxylem and root hairs that prevented the preparation of a flat squash needed to visualize stained interphase nuclei. Ideally, a
microdensitometer, an optical device, could have been used to estimate the nuclear density of multiple focal planes (Greilhuber, 2008). Integrated optical densities could have been measured, but these data were ruled out due to variable factors for image acquisition such as exposure length, condenser lense height, and focal plane adjustments (Hardie et al. 2002).

**Nuclear Area as a Proxy for Ploidy Levels**

The *Rhipsalis baccifera* individuals showed ploidy level variations that correspond with the putative ploidy levels reported by Cota-Sanchez et al (2010). Octoploids populations putatively grow in Madagascar in the subspecies *horrida*, which is represented in the nuclear area averages in my data set where a Malagasy *Rhipsalis baccifera* sample shows the largest average nuclear area (at 105 um²) in comparison to New World individuals (Figure 8). The Vera Cruz individual of the subspecies *baccifera* and the sample Puerto Rico also showed increased nuclear area compared to the individual labeled as subspecies *shaferi* and the Brazilian sample (Figure 8). The Madagascar sample featured in this study is likely an octoploid, the Vera Cruz individual and the Puerto Rican individual are likely both tetraploids, and the subspecies *shaferi* and Brazilian samples are likely diploids. Nuclear area averages present a viable methodology for estimating ploidy counts.

This methodology was confirmed as a rough proxy by using plants with referenced genome sizes. The positive control *Allium sativum* (garlic) had the largest nuclear area
of 128 \( \mu m^2 \), which is larger than the 89 \( \mu m^2 \) nuclear area of *Lathyrus odoratus* (sweet pea). Proportionately, garlic has a larger nuclear genome (2C=32.46 pg) in comparison to sweet pea (2C = 15.56 pg) (Kew genome size database). Both measurements for picogram content of nuclear genomes for the control plants were taken using flow cytometry (Murray *et al.* 1992; Ohri *et al.* 1996). I found that the dramatic differences in nuclear area could also be observed qualitatively on the microscope.

**Figure 6.** Feulgen nuclear staining of *Rhipsalis baccifera* in comparison to garlic

![Figure 6](image)

**Figure 6.** Brightfield images were digitally taken with the 40x objective using brightfield on an Olympus BX-51 upright microscope. Scale bars were added by the Olympus camera controller program attached to the microscope system. (A) *Allium sativum* (garlic) positive control stained dark and dense nuclei. More mitotic figures were observed in the garlic root tips, which were grown hydroponically from a commonplace bulb. Many cells can be observed in prophase, and individual chromosomes can be made out (B) Madagascar *Rhipsalis baccifera* individual identified as an octoploid with patchily stained nuclei. Almost all of the cells were in interphase.

**Figure 7.** Feulgen stained *Rhipsalis baccifera* from Puerto Rico
Figure 7. A Feulgen stained root tip from Puerto Rico #2 with the hydrolysed nuclei appearing as bright fuschia. The image was digitally taken with the 40x objective using brightfield on an Olympus BX-51 upright microscope. The nuclei for Rhipsalis baccifera were much smaller than the nuclei in the root tip positive controls Allium sativum (garlic) and pisum sativum (pea). The structures surrounding the actively dividing mitotic zone includes proto-xylem and root hairs used for water transport and nutrient absorption, respectively.
**Figure 8.** Nuclear area measurements for *Rhipsalis baccifera*, garlic, and sweet pea

- Madagascar sample 85646 had the largest average nuclear area (105 $\mu m^2$, SE= 2.89) as a putative tetraploid, in comparison to New World individuals from Vera Cruz (80 $\mu m^2$, SE= 4.82), Puerto Rico (89 $\mu m^2$, SE= 3.88), ssp. *shaferi* (53 $\mu m^2$, SE= 3.33), and Brazil 1 (55 $\mu m^2$, SE= 3.27). The average nuclear area for garlic was relatively large (128 $\mu m^2$, SE= 5.53) and the other control sweet pea (89 $\mu m^2$, SE= 5.82) was around the same nuclear area as a tetraploid *Rhipsalis baccifera*. 
Discussion:
Molecular Phylogeny of *Rhipsalis baccifera*

*Rhipsalis baccifera* subspecies have distinct morphological characteristics, but most of the samples in this data set are genetically highly similar. The data suggests that the majority of the samples showed sequence identity similarity with *Rhipsalis baccifera*, but the GenBank dataset used for sequence comparison contains mostly the New World subspecies *baccifera* samples. Most of the sequences for the rpl32-trnL(UAG) intergenic region for *Rhipsalis baccifera* were submitted by Ornelas *et al.* (2015). Although no explicit notes about subspecies were included, these sequences were likely *Rhipsalis baccifera ssp. baccifera* because of the other subspecies endemic to the New World, ssp. *shaferi* individuals are morphologically distinct and ssp. *hileiabaiana* individuals are endemic to Brazil (Ornelas *et al.* 2015). No morphological distinctions were noted across the individual plants by Ornelas *et al.* (2015). Contrastingly, I observed dramatic morphological differences associated with biogeography based on stem and spine characteristics. There is sampling bias in the BLAST results from the *Rhipsalis baccifera* samples that I sequenced because my samples were largely only compared to the Ornelas *et al.* (2015) dataset. This sampling bias using BLAST occurred because Ornelas *et al.* (2015) sequenced 154 samples of *Rhipsalis baccifera*, which made up to the majority of the sequences in GenBank for the rpl32-trnL(UAG) marker. Other *Rhipsalis* species have been sequenced for the rpl32-trnL(UAG) marker and added to the GenBank database from the *Rhipsalideae* clade studies by Korotkova *et al.* (2011) and Calvente *et al.* (2011).
It is challenging to phylogenetically analyze intraspecific samples because few molecular markers have been identified that have a high genetic variation to look at closely related individuals. Intergenic regions of the chloroplast DNA in plants provide an opportunity for neutral mutations to occur including SNPs and indels (Shaw et al. 2007). Plastid DNA, present in both mitochondrial and chloroplast genomes, is inherited maternally through the cytoplasm (Shaw et al. 2007). The sequencing data of *Rhipsalis baccifera* featured in this study can only account for maternal genetic inheritance, and could not be used to determine genetic variation across individuals in the nuclear genome. Genome duplication refers only to the nuclear genome, which undergoes meiosis and mitosis with genes from both parentals. The chloroplast genome is the fastest evolving genome in plants, but the intergenic region used in this study was insufficient for detecting population differences.

The outgroup *Rhipsalis neves-armondii* samples were identified by BLAST results and in the constructed phylogeny using this study’s sequencing results, as expected, which confirms that the chloroplast molecular marker *rpl32-trnL* UAG can accurately distinguish between interclade species at a low taxonomic level. These two samples had observably thicker stems with a ribbed structure, in comparison to the smooth, thinner, cylindrical stems of the other individuals in my dataset. *Rhipsalis neves-armondii* is endemic to South America and usually found in southeastern Brazil as a species within the subgenera *Calamorhipsalisi* of the *Rhipsalis* genus (Calvente et al. 2011). *Rhipsalis*
*puniceodiscus* is a sister species to *Rhipsalis neves-armondii*, and appeared in some of the BLAST results for these outgroup samples (Calvente et al. 2011).

Species identification is important in this study because these samples presented morphological characteristics similar to *Rhipsalis baccifera*, but also showed sequence divergence as separate species. The morphology of thin, cylindrical, pendant stems occurs convergently within the monophyletic core *Rhipsalis* clade, and includes *Rhipsalis baccifera* (Calvente et al. 2012). *Calamorhipsalis* and *Rhipsalis* are distinguished by their flower morphology with the pericarpel immersed in the areole and lateral flowers, respectively (Calvente et al. 2012). These morphological differences were not immediately noticeable because the plant material was collected at time when the flowers were neither budding nor in bloom. The distance tree (Figure 4) with a bootstrap support of 99 showed the out-grouped *Rhipsalis* taxa that I sequenced, such as *Rhipsalis neves-armondii* and a *Rhipsalis juengeri* sample. The confidence interval for the phylogenies constructed in the tree is high, which strongly supports that despite their morphological similarities *Rhipsalis baccifera* is genetically different from the species within the subgenera *Calamorphalis*. The sequencing data from the samples collected from Bahia, Brazil show that the subspecies *hileiabaiana* individuals have lower sequence similarity to other *Rhipsalis baccifera* samples, in comparison to other subspecies. My data suggest that the subspecies *baccifera* individual from Bahia, Brazil is highly genetically similar to the
*Rhipsalis baccifera* sequences submitted by Ornelas et al. (2015). Sequence analysis of the other subspecies (*erythrocarpa, horrida, mauritiana,* and *baccifera*) revealed genetic similarity to the sequences in GenBank. Genetic similarity is observable in the sequencing data for subspecies *baccifera* samples collected from a large biogeographic range. For example, the Mexican sample from Oaxaca has a 98% sequence identity and the Brazil sample 1 has a sequence identity of 99% with other *Rhipsalis baccifera* ssp. *baccifera.*

The subspecies *shaferi* individuals were still grouped with other *Rhipsalis baccifera* New World individuals for the distance tree using the rpl32-trnL(UAG) marker in this study, despite these individuals having lower sequence identity with other *Rhipsalis baccifera* samples. This suggests that more molecular markers with an increased number of informative characters may be needed to distinguish the subspecies and subspecies *shaferi* may merit its own species ranking, in concordance with the suggestions put forward by Korotkova et al. (2011). Although the taxonomic study of the entire Rhipsalideae clade by Korotkova et al. (2011) used the rpl32-trnL(UAG) marker in conjunction with other chloroplast markers to determine that subspecies *shaferi* and *hileiabaiana* may justify species ranking, their data lacked a large dataset that included many *Rhipsalis baccifera* individuals. In conclusion, Figure 4 was proposed as a distance tree illustrating the chloroplast intergenic region sequencing results of all samples to show that the *Rhipsalis baccifera* samples were genetically highly similar. Upon closer analysis in the BLAST results, the individuals morphologically identified as subspecies
*hileiabaiana* and *shaferi* both have low sequence identity in comparison to the subspecies *baccifera* samples in NCBI.

Direct sequencing of phosphoglucone failed, but was pursued to understand genetic divergence in the nuclear genome in a variable region. The intronic regions of PGI are not functionally conserved and have fast evolution rates in plants with population specific variation (Toulson Wimmer and Merritt, 2008). The degenerate primers, designed from plant, yeast, and animal sequences of PGI to amplify PGI in *Dioscorea tokoro*, were not specific enough to the sequence of the coding region in *Rhipsalis baccifera*. A gradient PCR was performed to increase the annealing temperature to increase the sequence specificity of the primers binding to the template DNA. However, annealing temperatures above 59°C did not yield any observable PCR products. Degenerate primers designed from plant PGI sequences might serve as a better reference for the initial amplification and sequencing of PGI in *Rhipsalis baccifera*.

**Microscopy - Cytotyping Individual Plants**

The average nuclear areas confirmed that polyploidy levels correlate with biogeographic populations. The data suggest that octoploid populations of *Rhipsalis baccifera* can be found in Madagascar. Similarly, tetraploid populations can be found in *Rhipsalis baccifera* in the Caribbean. The subspecies *shaferi* individual similarly had a smaller nuclear area, which suggests it is a diploid, in congruence with the data in Cota-Sanchez *et al.* (2010). Some error in the average nuclear area measurements can be accounted
for in the data collection procedure in ImageJ. The nuclear area could appear larger than
expected from diffuse staining or variability in the orientation of the nuclei from the
squash. Another issue presented in nuclear areas include polyploids undergoing genome
downsizing after a polyploidization event. The mean 1C DNA value does not increase in
direct proportion with ploidy, as one might expect (Leitch and Bennett, 2004). Diploid
progenitors therefore have the highest mean DNA in comparison to polyploids, although
on rare occasions mean DNA can increase slightly in comparison to diploids (Leitch and
Bennett, 2004). Therefore, I expected that an octoploid individual would have slightly
less than four times the amount of DNA as a diploid individual in *Rhipsalis baccifera*
using high throughput methods like microdensitometry or flow cytometry for evaluating
c-values. However, observing these slight downsizing differences in polyploid genome
downsizing was not possible using nuclear area as a proxy for ploidy level.

Similarly, I hypothesized that autopolyploidy in *Rhipsalis baccifera* may be
advantageous to introduce more allelic variation within a homogenized population. I
hypothesize that polyploidy in *Rhipsalis baccifera* results in a larger fruit that is
advantageous and selected for with the different bird dispersal vectors in the Old World.
Also, the increased stem size might have been selected for to accommodate larger fruit
in polyploid plants and increased drought tolerance. The selection of polyploids in the
Old World is likely due to these morphological adaptations.
Table 5. Summary of results for both cytology and molecular data

<table>
<thead>
<tr>
<th>Polyploidy Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ploidy level relates to biogeography and morphology</td>
</tr>
<tr>
<td>Ploidy level variation does not result in genetic divergence in <em>Rhipsalis baccifera</em></td>
</tr>
<tr>
<td>Rhipsalis baccifera individuals are likely autopolyploids (not determined)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Molecular Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing data suggests that certain subspecies may warrant species ranking due to genetic divergence.</td>
</tr>
<tr>
<td>Old World populations genetically similar to New World ssp. <em>baccifera</em></td>
</tr>
<tr>
<td>New World subspecies are genetically divergent</td>
</tr>
<tr>
<td>Species ranking may be warranted in New World subspecies <em>shaferi</em> and <em>hileiabaiana</em> (only 92% sequence identity with subspecies <em>baccifera</em>)</td>
</tr>
</tbody>
</table>

Conclusion

The phylogenetic relationships of cacti elucidate the amazing biodiversity of a family of succulents that diversified and adapted to grow in a wide range of habitats and ecologies. The genetic mechanisms and physiology of the selfing and clonally propagating *Rhipsalis baccifera* can be applied to understanding characteristics that allow other tropical species to have a wide range of endemic growth. Although *Rhipsalis baccifera* is the focus of population genetics studies due to its unusual distribution, other cacti species exhibit the same genetic mechanisms such as polyploidization, bird-vector seed dispersal, and genetic drift. Furthermore, understanding both the
biochemical and physiological development of *Rhipsalis baccifera* may reveal how plants with an evolutionary history of rapid geographic radiation and diversification fared under conditions of climate and atmospheric changes.

Both the biological species concept and the ecological species concept apply to *Rhipsalis baccifera* as an epiphytic species that successfully adapts to a number of environments ranging in humidity and temperature. However, molecular techniques can affirm these differences with genomic analysis to understand evolutionary rates in a widely dispersed species. The chloroplast sequencing results suggest that *Rhipsalis baccifera* genetic divergence has only occurred in the subspecies shaferi and hileiabaiana, which are endemic to the New World as diploid populations. Both subspecies shaferi and hileiabaiana were grouped phylogenetically into the same clades as the other *Rhipsalis baccifera* subspecies, despite having lower sequence identity percentages. The similarity across individuals from both the Old World and *Rhipsalis bacciferassp. baccifera* suggest that the morphospecies are genetically highly similar. Polyploidy in *Rhipsalis baccifera* correlates with biogeography. The cytology results support the putative biogeographic ranges of polyploids presented by Cota-Sanchez et al (2010). The New World populations can be diploid or tetraploid. The *Rhipsalis baccifera ssp. horrida* individual displayed the largest nuclear area, which suggests that it is an octoploid.

In future studies, the genetic divergence of *Rhipsalis baccifera* from polyploidization may be observed in nuclear sequencing results. Sequencing phosphoglucose isomerase,
a nuclear gene with allelic variation specific to populations in a species, may be an effective molecular marker for finding nuclear genetic variation in *Rhipsalis baccifera*. Alternatively, looking at the intronic regions of phosphoglucose isomerase may reveal highly variable area for low taxonomic categorization. Future intraspecific projects that look at population divergence in *Rhipsalis baccifera* using molecular markers with higher variability may be able to provide genetic support for a molecular clock that accounts for dispersal events.

**Broader significance**

Insight into cacti autoployploidization events is necessary to determine species within a clade that are largely defined and organized based on morphologies. The use of variable genetic markers for species delimitations in chloroplast and nuclear genomes can also be used for other taxa that have radiated quickly. Additionally, the transcontinental *Rhipsalis baccifera* provides a particularly interesting example for the geographic dispersal of plants from the neotropics into Africa, South America, the Caribbean, and South Asia. *Rhipsalis baccifera* is noted as the only species of cactus endemic to the Old World, but it remains to be known when that dispersal initially occurred (Calvente et al. 2011). The use of nuclear genes and their introns may be helpful in providing evidence for genetic variation that even intergenic regions of chloroplast DNA data fail to detect at a level of sufficient significance as molecular markers for plant evolution.
Acknowledgements

I would like to offer my deepest gratitude to Robert B. Merritt and Laura A. Katz, my advisers and mentors during this research project and my education at Smith College. I would also like to thank my fellow Katzlab members, Richard T. Briggs, Judith Wopereis, and Nathan Derr for their loving support as I engaged with sentiments of both scientific frustration and wonderment. Thank you to Rob Nicholson for introducing me to this project and the wonderful field of plant systematics. I would like to thank the Lyman Plant House staff for their acquisition and maintenance of the botanical samples for this study. Lastly, thank you to all of my friends and family for supporting my interests in cactus evolution and jazz saxophone. There’s an old Texan tale of tenor saxophone players creating walls of sound so powerful that they are capable of blowing down cacti in the desert. I like to think that an enthusiasm for both jazz and science can accomplish such a feat. Funding for this research came from the Nancy Kershaw Tomlinson Memorial Fund for undergraduate research and from the Summer Undergraduate Research Fellowship (SURF) at Smith College.
References:


Love, Askell. (1976), IPOB Chromosome Number Reports LIV. Taxon 25, 644


Appendix

Protocol 1. DNA Extraction with the Qiagen DNEasy Plant Kit. Modifications to the protocol were made according to cactus DNA extractions prepared with the same kit to examine taxonomy in Cactaceae (Nyffeler, 2002).

1. Disrupted 0-20 mg of fresh stem cortex tissue with a pipette tip in a 1.5 mL falcon tube.
2. Added 400 µL of Buffer AP1 and 4 uL of RNase A. Vortexed and incubated for 10 minutes at 65 °C. The tube was inverted 2-3 times during the incubation step in a circulating water bath.
3. 130 µL of Buffer P3 was added. The tube was mixed and incubated for 5 minutes on ice.
4. The tubes were spun down at 6000 rpm for 1 minute.
5. The supernatant was centrifuged at 14,000 rpm for 5 minutes.
6. The tubes were spun down at 12,000 rpm for 5 min.
7. The supernatant was pipetted into a QIAshredder spin column placed in a 2 mL collection tube. The tubes were centrifuged at 14,000 rpm for 2 minutes.
8. The flow-through was transferred into a new tube without disturbing the pellet. 1.5 volumes of Buffer AW1 was added and mixed with the pellet by pipetting.
9. 650 µL was transferred into a DNeasy Mini Spin column placed in a 2 µL collection tube. The tubes were centrifuged at 8000 rpm for 1 minute. The flow through was discarded. This step was repeated with the remainder of the sample.
10. The spin column was transferred to a new collection tube. 500 µL of Buffer AW2 was added and the tubes were centrifuged at 8,000 rpm for 1 minute. The flow through was discarded.
11. Another 500 uL of Buffer AW2 was added and the tubes were centrifuged at 14,000 rpm for 2 minutes.
12. The spin column was transferred to a new 1.5 µL tube.
13. 100µL of Buffer AE was added for elution. The tubes were incubated at room temperature for 5 minutes. The tubes were centrifuged at 8,000 rpm for 1 minute.
14. Step 13 was repeated for a total volume of around 200 µL.