Genetic Diversity and Structure in the Philippine *Rafflesia lagascae* Complex
inform its Taxonomic Delimitation and Conservation

**Pieter B. Pelser,**<sup>1,4</sup> **Daniel L. Nickrent,**<sup>2</sup> **Chrissen E. C. Gemmill,**<sup>3</sup> and **Julie F. Barcelona**<sup>1</sup>

<sup>1</sup>School of Biological Sciences, University of Canterbury, Private Bag 4800,
Christchurch 8140, New Zealand.

<sup>2</sup>Department of Plant Biology, Southern Illinois University Carbondale, 1125 Lincoln
Drive, Carbondale, Illinois 62901-6509, U. S. A.

<sup>3</sup>School of Science, University of Waikato, Private Bag 3105, Hamilton 3240, New
Zealand.

<sup>4</sup>Author for correspondence (pieter.pelser@canterbury.ac.nz)
**Abstract**—*Rafflesia lagascae* (Rafflesiaceae) is a rare endo-holoparasitic plant species with a disjunct distribution on Luzon Island in the Philippines. It is morphologically very similar to *R. manillana*, which is only known from a single population on nearby Samar Island. This study aims to contribute to the taxonomy and conservation of *R. lagascae* and *R. manillana* (i.e., the *R. lagascae* complex) by resolving their patterns of genetic diversity and genetic structure. The results of population genetic analyses of a microsatellite data set indicate that despite their frequently extremely small sizes and geographic isolation, *Rafflesia* populations display moderate genetic diversity and do not show evidence of pronounced inbreeding. Most populations appear to be strongly genetically differentiated from each other, suggesting limited gene flow between them. Patterns of genetic diversity of staminate and pistillate *Rafflesia* flowers growing on the same *Tetrastigma* host plants indicate that the *R. lagascae* complex is monoecious and that host plants are regularly infected by more than one *Rafflesia* plant. Both might enable *Rafflesia* to maintain viable populations in areas with low host densities. PCoA and Bayesian cluster analyses show that the *R. lagascae* complex is composed of three genetically isolated taxa. One of these constitutes *R. manillana*, supporting the morphology-based hypothesis that it is taxonomically distinct from *R. lagascae*. The second taxon in this complex is composed of the single population of *R. lagascae* from Mt. Labo. This morphologically cryptic population is genetically distinct from the 11 remaining *R. lagascae* populations that were included in this study and that form the third taxon. We recommend that *R. lagascae* s. s., the Mt. Labo *R. lagascae*, and *R. manillana* should be managed as different conservation entities.

**Keywords**—Microsatellites, parasitic plants, rare species, species complex, taxonomy.
The host species of parasitic plants form essential components of their environment (Sandner and Matthies 2017). Their distribution patterns constitute the maximum potential distribution area of the parasitic plant species that depend on them for their existence (Costea and Stefanović 2009) in the same way that specialists of a rare habitat are confined to areas where this habitat is found (e.g., Medrano and Herrera 2008; Williams et al. 2016) and the distribution of elevation specialists is constrained by topography (e.g., Laurance et al. 2011). When the prerequisite conditions on which specialist plant species depend are rare, so are these species. Consequently, some specialist plants have small population sizes and are restricted to a single small area (e.g., de Lange 2003) or have a disjunct distribution pattern (e.g., Medrano and Herrera 2008). This makes these species vulnerable to inbreeding and genetic drift as a result of a loss of genetic connectivity between populations due to habitat fragmentation and degradation (e.g., Ellstrand and Elam 1993; Schmidt and Jensen 2000; Stanton et al. 2009; Williams et al. 2016). This might ultimately result in a loss of genetic diversity and thereby increase the probability that small and isolated populations go extinct (Schmidt and Jensen 2000). Indeed, many parasitic plant species are rare and listed as threatened or endangered (Marvier and Smith 1997; Sandner and Matthies 2017). Parasitic species that are very host specific are especially at risk, because they are particularly sensitive to changes in host-availability and typically have small distribution areas (Costea and Stefanović 2009).

*Rafflesia* R.Br. (*Rafflesia*aceae) is a Southeast Asian plant genus of ca. 30 species, all of which are obligate endo-holoparasites of *Tetrastigma* (Miq.) Planch. (*Vitaceae; Meijer 1997; Nais 2001; Pelser et al. 2016). The Philippines is one of the centers of *Rafflesia* diversity (Barcelona et al. 2009) and is home to 13 currently recognized species (Pelser et al. 2011 onwards; Barcelona et al. 2014; Galindon et al. 2016),
which are all endemic to the country. *Rafflesia lagascae* Blanco (1845) is arguably the most common *Rafflesia* species in the Philippines and has the largest distributional area (Barcelona et al. 2009). However, although it can be found throughout the island of Luzon (Fig. 1), it has a highly disjunct distribution and is only known from 15 small areas. In some of these areas, it is only known from one or a few individual host plants belonging to three species of *Tetrastigma* (Pelser et al. 2016). Because of the small number of *R. lagascae* populations and the extremely small size of many of them, as well as its narrow host range, this species might be particularly prone to genetic diversity loss following habitat fragmentation and degradation. This is potentially a significant conservation factor, because the Philippine tropical rainforest, to which this species is confined, has been reduced to an estimated 3–6% of its original cover (Mittermeier et al. 1998; Ong et al. 2002). Moreover, remaining tropical forest habitat has become highly degraded and fragmented as a result of human activities such as logging, mining, and land conversion for agriculture (Barcelona et al. 2009). To better understand the conservation needs of *R. lagascae* and to inform its management, information about its patterns of genetic diversity and genetic structure are needed. These patterns would, for instance, enable the identification of the populations that have the lowest genetic diversity and least genetic connectivity with other populations and that are therefore most at risk of inbreeding and genetic drift.

At present, nothing is known about patterns of genetic diversity and structure in any *Rafflesia* species. Previous molecular genetic studies of *Rafflesia* have focused on horizontal gene transfer from their hosts (Davis and Wurdack 2004; Nickrent et al. 2004; Xi et al. 2012, 2013), their evolutionary relationships and diversification (Barkman et al. 2004, 2008; Nickrent et al. 2004; Davis et al. 2007; Bendiksby et al.
2010), the molecular regulation of flower development (Nikolov et al. 2013; Ramamoorthy et al. 2013), and the possible loss of their chloroplast genome (Molina et al. 2014), but none have addressed research questions at the infraspecific level. Because of the absence of molecular genetic data for *Rafflesia* at the levels of individuals and populations, not only are their patterns of genetic diversity and genetic structure unknown, but other aspects of their biology also remain to be discovered. For example, most *Rafflesia* species, including *R. lagascae*, have unisexual flowers, but it is not known if they are dioecious or monoecious (Nais 2001). It is difficult to determine plant sexuality without genetic data, because of their endoparasitic nature (Balete et al. 2010). In other words, it is not possible to determine if staminate and pistillate flowers that emerge from the same host belong to the same genetically identical individual endophyte. Because self-incompatible plants might be more vulnerable to a loss of genetic diversity resulting from habitat fragmentation (Stanton et al. 2009), knowing if *R. lagascae* is monoecious or dioecious is relevant for its conservation management. Similarly, it is presently unclear if *Tetrastigma* vines can be host to more than one *Rafflesia* individual. This basic information about *Rafflesia* will enable more accurate estimates of population sizes and thereby contribute to their conservation management.

Population-level genetic data are also needed to inform the taxonomic delimitation of *R. lagascae*. This species is morphologically very similar to *R. manillana* Teschem. (Teschemacher 1844; Fig. 2) and both taxa were considered synonymous until relatively recently (Madulid et al. 2008; Pelser et al. 2013). Whereas *R. lagascae* is endemic to Luzon, *R. manillana* has only been reported from Basey municipality on Samar Island. The two allopatric species are of similar sizes (11–24 vs. 11–16 cm diam.; Madulid and Agoo 2008; Barcelona et al. 2009) and primarily differ in the
color of the floral diaphragm and the relative size of the diaphragm aperture (Pelser et al. 2013; Fig. 2), but it is presently unclear if these morphological differences are indicative of genetic differences between the two species. *Rafflesia manillana* is currently only known from a single population and is considered critically endangered by Madulid and Agoo (2008). Therefore, if despite their morphological differences both taxa are genetically indistinct, the conservation status of *R. manillana* would need to be revised.

In this study, we provide a first insight in the patterns of genetic diversity and genetic structure of *R. lagascae* and *R. manillana* (from here on collectively referred to as the *R. lagascae* complex). Specifically, we aim 1) to identify populations that have low genetic diversity, 2) to determine if taxa in the *R. lagascae* complex are monoecious or dioecious and if *Tetrastigma* plants can host more than one *Rafflesia* individual, and 3) to resolve the taxonomic delimitation of the *R. lagascae* complex.

**MATERIALS AND METHODS**

**Specimen Sampling**—Tissue samples on silica gel and voucher specimens preserved in 70% denatured alcohol were collected from 13 of the 16 areas from which members of the *R. lagascae* complex have been reported. For the three unsampled areas (Adams, Ilocos Norte Prov.; Barlig, Mountain Prov.; Calanasan, Apayao Prov.; Fig. 1), collecting permits or samples could not be secured in time for this study. Tissue was taken from at least one flower or flower bud from each infected host plant that was encountered, with a maximum of six samples taken per site (i.e., one or more infected host plants in close vicinity of each other and geographically isolated from other such sites). Nearby host plants were considered distinct individuals if they were
unambiguously spatially separated from each other. When possible, damage to *Rafflesia* plants was minimized by removing a small amount of tissue from the diaphragm or the perigone lobes of the flowers, which is not expected to affect their reproductive function. When flower buds were encountered, we sampled tissue from the perigone lobes or disk of young flower buds while carefully avoiding sampling from the cupula area where the tissues of the parasite and host intertwine. *Rafflesia* species have a high bud mortality rate (50–90%; Nais 2001 and references therein) and we therefore considered sampling from young buds as having a lower negative impact on the reproductive success of plants than sampling from mature buds that have escaped the causes of early mortality (e.g., nutrient limitation, predation by animals, infection by wasps; Nais 2001) and are likely to survive until the flowering stage. Voucher specimens (Appendix 1) are deposited at CAHUP, CANU, PNH, and SIU (acronyms follow Thiers, continuously updated).

**Development of Microsatellite Primers**—Contigs from a previously published Illumina sequence library of *R. lagascae* obtained from a sample of this species from Bolos Point (see Molina et al. 2014 for assembly methodology; NCBI Sequence Read Archive SRX434531, Bioproject PRJNA235228; Fig. 1) were screened using MSATCOMMANDER 1.0.8-beta (Faircloth 2008) for microsatellite loci of two or three nucleotides with a minimum of 15 repeats. This resulted in the identification of 1254 microsatellite loci. The embedded Primer3 software (Rozen and Skaletsky 1999) subsequently designed M13-tagged universal primers for these using the following settings: melting temperatures (Tm) of 58–60°C, GC content of 40–60%, and fragment sizes of 100–450 bp. This resulted in primers for 37 loci, which were screened for amplification success and specificity using DNA of four *R. lagascae* specimens. A total of 17 polymorphic loci (Table 1) were subsequently used for
genotyping the additional DNA samples of _R. lagascae_ and _R. manillana_ available to us. A blastn search using the contig DNA sequences that contain these 17 loci did not reveal resemblance with available Vitaceae genomic data in GenBank. We therefore assume that none of these loci are present in the genome of _R. lagascae_ as a result of horizontal gene transfer from their host plants.

**DNA Extraction and Microsatellite Genotyping**—Total genomic DNA for the microsatellite genotyping analyses was extracted using the Qiagen DNeasy plant mini kit (Qiagen, Germantown, Maryland) following the manufacturer’s protocol. Multiplex PCR analyses using up to four primer combinations per PCR sample were performed in a total volume of 4 µl using the Qiagen Type-it microsatellite PCR kit. These samples contained 1 µl of genomic DNA (ca. 2–60 ng per sample), 2 µl of 2× PCR master mix, 0.16 pmol of M13-tagged primer, 0.64 pmol of untagged primer, 2 pmol of M13 fluorescent-labeled primer (6FAM, NED, PET, or VIC; 5′-GGAAACAGCTATGACCAT-3’) and nuclease-free water to volume. The PCR conditions were as follows: an initial denaturation of 5 min at 95°C followed by 30 sec at 95°C, 90 sec at 60°C, 30 sec at 72°C, during 28 cycles, and a final extension of 30 min at 60°C. The PCR products were run on an ABI 3130xL Genetic Analyzer at the University of Canterbury. Geneious 6.1.7 (Biomatters Ltd, Auckland, New Zealand) was used to determine fragment lengths.

**Data Analyses**—We genotyped 143 DNA samples of the _R. lagascae_ complex obtained from 80 host plants. Microsatellite data of 98 of these samples were included in our analyses. A total of 45 DNA samples were excluded, because these had identical genotypes to those of other flowers or buds obtained from the same host plant and might therefore represent the same _Rafflesia_ individual.
Exact tests for deviations from Hardy-Weinberg equilibrium and linkage disequilibrium were performed with GENEPOP v.4.2 (Raymond and Rousset 1995) as implemented in GenePop on the web (http://genepop.curtin.edu.au/) and significance levels were adjusted for multiple tests using the B-Y method FDR at $p = 0.05$ (Narum 2006). Only loci found to display significant linkage disequilibrium across all populations were considered to be linked. Micro-checker v.2.2.3 (Van Oosterhout et al. 2004) was used to check for null alleles (95% confidence interval; 1,000 repetitions). Only loci indicated to display null alleles across all populations were considered to be truly producing null alleles.

GenAlEx v. 6.5 (Peakall and Smouse 2012) was used to assess the genetic diversity of each population by calculating the percentage of polymorphic loci, allelic richness (mean number of alleles and number of effective alleles), observed ($H_o$) and expected ($H_e$) heterozygosity, unbiased heterozygosity ($uH_e$), and the inbreeding coefficient. BOTTLENECK 1.2.0.2. (Cornuet and Luikart 1997) was used to test for recent genetic bottlenecks in the six Rafflesia populations from which at least nine genetically unique samples were available. Wilcoxon sign-rank tests were used to determine if excesses in heterozygosity are statistically significant. Mode-shift tests in BOTTLENECK were employed to identify shifts in allele frequencies.

Genetic structure was studied by determining the number of private alleles for each population in GenAlEx. In addition, an analysis of molecular variance (AMOVA; $F_{st}$, 999 permutations) was carried out in GenAlEx with a data set from which samples from Mt. Banahaw and Mt. Mingan were excluded because only one Rafflesia sample was collected from each of these two areas. Bayesian model-based clustering analyses were performed with STRUCTURE v.2.3.4 (Pritchard et al. 2000; Falush et al. 2003, 2007) using the admixture model and correlated allele frequencies. These analyses
were run with K values from 1–20 and with 20 iterations for each K value. Each analysis was run for 200,000 generations of which the first 100,000 were discarded as burn-in. The STRUCTURE output was summarized using STRUCTURE HARVESTER (Earl and Von Holdt 2012) to determine the value of K that best explains the genetic structure in the data. For this, both the Evanno et al. (2005) method (K with the highest value of ΔK) and the method of Pritchard et al. (2000; K with the highest Pr(X|K)) were used. CLUMPAK 1.1 (Kopelman et al. 2015) was used for the summation and graphical presentation of the STRUCTURE results. GenAlEx was subsequently used for a Principal Coordinate Analysis (PCoA) using a covariance matrix of co-dominant genotypic pairwise distances between individual samples with data standardization.

Mantel tests (999 replicates) were used in GenAlEx to test for a correlation between geographic distance and co-dominant genotypic distance between individuals. These tests were executed with non-transformed geographic distances as well as with a data set in which these distances were transformed using the natural logarithm.

RESULTS

Patterns of genetic diversity and genetic structure in the *R. lagascae* complex were studied using microsatellite data from 17 loci. These loci yielded a total of 249 different alleles. The number of alleles per locus across all populations ranged from 5 to 33 (mean 14.59). After B-Y correction (Narum 2006), two of the 13 populations that were included in our study (Bolos Point and Mt. Irid) showed a significant (*p = 0.05*) deviation from Hardy-Weinberg proportions because of a deficit of
heterozygotes (Table 2). Overall, however, the observed heterozygosity ($H_o = 0.60$) was only slightly higher or lower than estimates of the expected heterozygosity ($H_e = 0.53$, $uH_e = 0.64$). None of the pairs of loci displayed significant linkage disequilibrium across all populations and none of the 17 loci showed evidence of producing null alleles across all populations.

For 32 of the 80 Tetrastigma host plants of the *R. lagascae* complex that were examined, more than one *Rafflesia* flower or flower bud was genotyped. Two different *Rafflesia* genotypes were obtained from 17 of these hosts (53%), and three different parasite genotypes were recovered from one host plant. A total of 57 of the 93 (61%) *Rafflesia* flowers for which sex could be determined were staminate.

Staminate and pistillate flowers with identical genotypes were found on four host plants from the Basey, Bolos Point, Maria Aurora, and Mt. Malinao populations. Non-identical *Rafflesia* genotypes from the same host have a mean co-dominant genetic distance of 16.5 ($n = 18$; range 1–32). *Rafflesia* samples that were obtained from different hosts were all genetically different.

Populations at Bolos Point, Mt. Irid, and Salazar show the highest levels of genetic diversity as measured by the percentage of polymorphic loci, their allelic richness, and heterozygosity (Table 2). The inbreeding coefficient is low for the species complex as a whole ($F_{IS} = -0.14$). It is highest for the *R. manillana* population (Basey; $F_{IS} = 0.11$) and the Mt. Irid population of *R. lagascae* ($F_{IS} = 0.12$). Wilcoxon tests in BOTTLENECK did not reveal a significant excess of heterozygotes in the six populations that were included in these tests when a Stepwise Mutation Model was assumed. However, significant heterozygosity excess was found for the Aurora Memorial Natural Park (AMNP) and Basey populations under the Infinite Alleles Model and the Two Phase Model. The Bolos Point, Mt. Irid, and Mt Labo populations
only displayed significant heterozygosity excess under an Infinite Alleles Model. The AMNP population shows evidence of a shift in allelic frequencies.

All 13 populations of the *R. lagascae* complex have at least one private allele (Table 2). The Bolos Point and Mt. Irid populations have the most private alleles (21 and 11, respectively). Overall genetic differentiation within the *R. lagascae* complex is significant (*F’*<sub>ST</sub> = 0.651, *p* < 0.001). A total of 22% of the genetic variation was found among populations and 9% among individuals. All pairwise *F*<sub>ST</sub> values are statistically significant at *P* = 0.05 after B-Y correction (Narum 2006), except for the population pairs Burgos-Salazar and Burgos-Mt. Irid (Table 3). Using the Pritchard et al. (2000) and Evanno et al. (2005) methods, the results of the STRUCTURE analysis identified three primary genetic clusters (*K* = 3). One of these clusters is solely composed of specimens of the Basey population (*R. manillana*), the second cluster exclusively contains *R. lagascae* specimens from Mt. Labo, and all specimens of the remaining 11 *R. lagascae* populations form the third cluster (Fig. 3). The admixture proportions (*q* values) for individual specimens indicate very little admixture between the three clusters. The same genetic clusters were also recovered in a PCoA of pairwise genetic distances (Fig. 4).

To study patterns of genetic structure within *R. lagascae* s. s., STRUCTURE analyses were also performed with a data set that contains only populations of this group. For this data set, the Evanno et al. (2005) method favored *K* = 2, whereas the Pritchard et al. (2000) method preferred *K* = 8. The STRUCTURE plots of *K* = 2 (Fig. 5a) show that the Bolos Point population in the northern part of Luzon (Fig. 1) is placed in the same genetic cluster as populations from the southern part of mainland Luzon (Mt. Banahaw, Mt. Makiling, Mt. Natib), and the Bicol peninsula (Mt. Malinao). The remaining populations of *R. lagascae* s. s. are found in central Luzon.
Pelser, Nickrent, Gemmill, and Barcelona 14

(AMNP, Burgos, Maria Aurora, Mt. Irid, Mt. Mingan, Salazar) and show a pattern of admixture between both genetic clusters (Fig. 5a). The K=8 STRUCTURE plots (Fig. 5b; only the major clustering pattern presented) show that some nearby populations have similar genetic profiles (AMNP, Maria Aurora, and Mt. Mingan; Burgos and Salazar), that some genetic clusters are predominantly found in individual populations (e.g., Mt. Makiling, Mt. Natib, Mt. Irid), and that some populations are mostly composed of specimens with high $q$ values for a single genetic cluster (e.g., AMNP, Mt. Makiling, Mt. Malinao). Overall genetic differentiation within the $R. lagascae$ s. s. cluster is significant ($F'_{ST} = 0.420, P < 0.001$). A total of 12% of the genetic variation was found among populations and 9% among individuals. Within $R. lagascae$ s. s., Mantel tests did not reveal a statistically significant positive correlation at $p = 0.05$ between geographic and genetic distances.

**DISCUSSION**

Because of the size of their flowers, which are among the largest in the world (Nais 2001), and their intriguing parasitic life style, $Rafflesia$ species speak to the imagination of biologists and the general public alike. This fascination is further heightened by the fact that many $Rafflesia$ species are very rare. For example, $R. schadenbergiana$ Göpp. ex Hieron. (Hieronymus 1885) is currently known from only two living host plants (Barcelona et al. 2008; Pelser et al. 2016). As such, $Rafflesia$ are flagships of plant conservation: the giant pandas of the plant world (Josephson 2000). Yet, to the best of our knowledge, conservation genetic studies have not previously been attempted and, consequently, nothing was thus far known about the patterns of genetic diversity and structure of $Rafflesia$ species. This is not because of a
lack of interest by biologists, but rather because of difficulties typically associated
with studying rare species with small population sizes. Even *R. lagascae*, the most
common and widespread species of *Rafflesia* in the Philippines, is only known from
15 areas and few *Tetrastigma* host plants. For that reason, it is difficult to attain
sample sizes that provide enough statistical power to conclusively answer population
genetic questions. Although we sampled 13 of the 16 known populations of the *R.
lagascae* complex, and went through great efforts to locate as many host plants as
possible, many *Rafflesia* populations appear to consist of only a single or very few
host plants that produce only a few flowers and, consequently, yielded few samples
for our molecular genetic analyses. For example, from two populations (Mt. Banahaw
and Mt. Mingan) only a single sample could be obtained (Table 2). This means that,
despite extensive sampling, our results should be interpreted with care and that our
conclusions are tentative. However, although the absolute number of *Rafflesia*
specimens that we sampled was small, the percentage of individuals sampled in each
population is very high. Therefore, despite of low sample numbers in some
populations, our research resulted in very accurate estimates of population allele
frequencies.

**Breeding System and Co-infection of Hosts**—Genetically identical *Rafflesia*
samples were exclusively obtained from the same host plant and *Rafflesia* samples
from the same host plant that are not genetically identical have a mean co-dominant
genetic distance of 16.5. This indicates that our microsatellite data set of 17 loci
potentially provides enough resolution at the level of individuals to determine if plants
of the *R. lagascae* complex are dioecious or monoecious. Under this assumption, our
data indicate that this complex is monoecious, because genetically identical staminate
and pistillate flowers were found on the same host plants. Previously, however,
*Rafflesia* was hypothesized to be dioecious (e.g., Olah 1960; Beaman et al. 1988; Renner and Feil 1993; Bellot and Renner 2013; Wicaksono et al. 2016; but see Nais 2001 for more cautious statements). This reproductive strategy was considered more likely because *Rafflesia* species are rare and populations are far apart. Dioecy was therefore regarded as a likely adaptation to avoid inbreeding depression (Beaman et al. 1988). Two species, *R. baletei* Barcelona and Cajano (Barcelona et al. 2006) and *R. verrucosa* Balete, Pelser, Nickrent & Barcelona (2010), however, have both stamens and a developed ovary on the same flower, although it is unknown if these are also functionally bisexual. In addition, non-functional stamens have been reported for some female flowers of *R. consueloae* Galindon, Ong and Fernando (2016). If plants of the *Rafflesia lagascae* complex and other *Rafflesia* species are monoecious, it remains to be determined if they are self-compatible and geitonogamous or exclusively outcrossing. A geitonogamous breeding system might benefit small *Rafflesia* populations, because it increases the number of available mates, however, it carries with it the potentially negative consequences of inbreeding (Byers and Meagher 1992; Ellstrand and Elam 1993). Some studies have suggested that dioecy is overrepresented in parasitic plants (Renner and Ricklefs 1995; Bellot and Renner 2013), but this hypothesis rests in part on the assumption that the presence of pistillate and staminate flowers of endophytic parasites on the same host belong to different parasite individuals (Bellot and Renner 2013). If endophytic parasites with unisexual flowers such as those in Apodanthaceae and Rafflesiaceae are indeed monoecious, then the incidence of dioecy among all parasites has been overestimated.

For *Tetrastigma* host plants from which more than one *Rafflesia* sample was taken, genetically different *Rafflesia* flowers were recovered from about half. This suggests that *Tetrastigma* host plants are regularly home to multiple *Rafflesia* plants.
Consequently, even populations composed of a single host plant can maintain genetic diversity. Potentially, the monoecious breeding system and co-infection of host plants enable the *R. lagascae* complex to maintain viable populations in areas with low host densities.

**Genetic Diversity**—The numbers of samples taken from each *Rafflesia* population can be considered as rough estimates of relative population sizes, because we sampled *Rafflesia* from nearly all encountered infected host plants that produced flower buds or flowers during our field work. Although smaller populations of the *R. lagascae* complex typically display less genetic diversity than larger populations, as measured by the percentage of polymorphic loci, allelic richness, and heterozygosity estimates, they are not genetically depauperate and do not show evidence of marked inbreeding (Table 2), as might be expected for small and isolated populations (Ellstrand and Elam 1993; although see Binks et al. 2015 and references therein). Overall, heterozygosity estimates are similar to those of other perennial plant species (Nybom 2004). Six populations for which at least nine genetically unique samples were available were included in analyses to test for indications of recent genetic bottlenecks. Although we did not find evidence for a genetic bottleneck in the Salazar population, there are indications that the other populations (AMNP, Basey, Bolos Point, Mt. Irid, Mt. Labo) might have experienced a genetic bottleneck in the recent past. However, if they indeed went through a genetic bottleneck, it is at present not known what caused this, because all five populations are located in areas with relatively intact forests.

**Genetic Structure**—The results of the genetic structure analyses indicate low genetic connectivity between most populations of the *R. lagascae* complex. All populations have one or more private alleles (Table 2), most pairwise *F*<sub>ST</sub> values and
the overall $F'_{ST}$ value are statistically significant (Table 3), and a substantial proportion of the genetic variation in the complex is found among populations. The results of the STRUCTURE and PCoA analyses show that there are three distinct genetic groups that align with *R. manillana*, the Mt. Labo *R. lagascae* population, and the 11 remaining populations of *R. lagascae* (Figs. 3 and 4). These patterns suggest that these three groups are genetically isolated from each other.

The distinctiveness of *R. manillana* as indicated by the microsatellite data used in this study is supported by morphological differences between this taxon and *R. lagascae* (Madulid and Agoo 2008; Madulid et al. 2008; Pelser et al. 2013; Fig. 2). Except for *R. speciosa* Barcelona and Fernando (2002), which is found on the island of Panay and the adjacent northern part of Negros Island (Barcelona et al. 2009), all Philippine *Rafflesia* species are endemic to individual islands in the archipelago. *Rafflesia manillana* is endemic to Samar Island whereas *R. lagascae* is only known from Luzon Island and their genetic and morphological differences therefore support the hypothesis that the sea straits that separate the Philippine islands form significant barriers to gene flow between *Rafflesia* populations. The genetic isolation of *R. manillana* might be a contributing factor to its lower genetic diversity relative to most other populations from which a similar number of samples was obtained (Table 2).

Also the Mt. Labo *R. lagascae* population is genetically isolated from other populations of the *R. lagascae* complex and likewise shows relatively low genetic diversity (Table 2). In contrast with *R. manillana*, however, the Mt. Labo population does not appear to be morphologically different from the other *R. lagascae* populations (Fig. 2). Its floral characters (e.g., size, coloration, ornamentation of the diaphragm and perigone lobes) seem to fall within the range of variation observed in *R. lagascae* s. s., but detailed morphometric studies are needed to determine if the Mt.
Labo population is indeed morphologically indistinct from *R. lagascae* s. s. Pelser et al. (2016) identified *Tetrastigma loheri* Gagnepain (1910) as the host species of Mt. Labo *Rafflesia* plants. This species is the most common host of the *R. lagascae* complex (Pelser et al. 2016) and differences in host preference can therefore not explain the genetic distinctiveness of the Mt. Labo population.

Mt. Labo is located on Bicol Peninsula, which is currently connected to mainland Luzon by a relatively narrow land bridge, and at the southeastern end of the distribution area of *R. lagascae* (Fig. 1). According to Hall (1998, 2002), the Bicol Peninsula was part of a cluster of islands that included Samar and eastern parts of the current Visayas and Mindanao and that has progressively moved towards mainland Luzon over the last 45 my. The Bicol Peninsula only connected to southeastern Luzon c. 6 mya (Hall 2002). The genetic isolation of the Mt. Labo population can therefore potentially be explained by the geographical isolation of the Bicol Peninsula prior to 6 mya. Molecular dating studies by Bendiksby et al. (2010) suggest that *Rafflesia* was present in the Philippines as early as 5.2-9.3 mya. Thus, it is indeed possible that the dynamic geological history of the Philippines shaped the patterns of genetic structure of *R. lagascae*. The Mt. Malinao population, however, is also located on the Bicol Peninsula (Fig. 1), yet genetically more similar to populations from mainland Luzon than to the Mt. Labo population (Fig. 3). If the previously mentioned geographic hypothesis is correct, then long-distance dispersal of *Rafflesia lagascae* s. s. between mainland Luzon to Mt. Malinao could explain the observed genetic patterns.

Zoochory has been proposed as the dispersal method of *Rafflesia* seeds (Teijsmann 1856; Justesen 1922; Kuijt 1969; Emmons et al. 1991; Nais 2001; Bänziger 2004). Although various animals have been considered as dispersal agents, observations by Pelser et al. (2013) suggest seed dispersal by ants. Myrmecochory is potentially
plesiomorphic in Rafflesiaceae, because it is also common in Peraceae (Bond and Slingsby 1983; Lengyel et al. 2009, 2010), which form the sister group of the clade composed by Euphorbiaceae and Rafflesiaceae (Davis et al. 2007; Endress et al. 2013). If *Rafflesia* is indeed myrmecochorous, long-distance dispersal is presumably very rare, and this might explain the high island-endemicity of *Rafflesia* in the Philippines. However, other observations and experiments suggest that mammalian frugivores might play a role in endozoochorous seed dispersal (Emmons et al. 1991; Bänziger 2004) and this mode of dispersal might be able to transport *Rafflesia* seeds more frequently over greater distances. Alternatively, the Mt. Malinao area might be a fragment of mainland Luzon whose position has been scrambled by tectonic movement of the Bicol fragment. Although we are not aware of geological evidence that supports this hypothesis, this part of the Philippines clearly has a very complex and dynamic geological history that is presently only poorly known (Hall 2002).

Patterns of genetic structure within *R. lagascae* s. s. are less pronounced than those among *R. lagascae* s. s., Mt. Labo *R. lagascae*, and *R. manillana*, but still indicate low genetic connectivity among most populations (Table 2; Figs. 4 and 5). However, particularly some populations in central Luzon (i.e., Burgos, Mt. Irid, Salazar) show admixture between different genetic clusters (Fig. 5) and might therefore have experienced higher levels of recent gene flow than some of the other populations (e.g., Mt. Makiling and Mt. Malinao). Although evidence for isolation by distance in the *R. lagascae* s. s. cluster is not statistically significant, pairwise F\textsubscript{ST} values (Table 3) and the results of Bayesian cluster analyses in STRUCTURE (Fig. 5) suggest stronger genetic connectivity within two groups of nearby populations in central Luzon (i.e., among the AMNP, Maria Aurora, and Mt. Mingan populations and between Burgos and Salazar) than among other populations. This indicates that nearby populations are
perhaps less likely to become genetically isolated than those that are further away or that genetic connectivity between populations in central Luzon is higher than in other parts of the distribution area of *R. lagascae*.

**Taxonomic Delimitation of the *R. lagascae* Complex**—The results of our genetic structure analyses inform the taxonomic delimitation of the *R. lagascae* complex. In addition to confirming that *R. manillana* is not only morphologically different from *R. lagascae* (Fig. 2), but also genetically distinct, a genetically different, but morphologically cryptic population of *R. lagascae* was discovered on Mt. Labo. If a species concept is applied in which genetic distinctiveness or the absence of interbreeding is considered evidence for recognizing groups of individuals as different species (Coyne and Orr 2004), then the Mt. Labo population should be described as a distinct species. We refrain here from formally naming this species, awaiting the completion of detailed morphological studies in search of characters in which the Mt. Labo plants are different from those of *R. lagascae* s. s.

**Conservation Implications**—The *R. lagascae* complex is known from relatively few and disjunct populations, some of which are only composed of a few host plants. Despite this, our results do not provide evidence of pronounced inbreeding and populations show moderate levels of genetic diversity. This potentially means that the disjunct distribution of the *R. lagascae* complex and the small sizes of some of its populations are a result of natural differences in host presence and abundance among areas and that the *R. lagascae* complex has adapted to this. In other words, the *R. lagascae* complex is possibly naturally rare. It is extremely likely, however, that the number of populations of the *R. lagascae* complex has decreased over the past century and that the geographic distances between them have increased as a result of widespread habitat destruction and degradation in the Philippines (Mittermeier et al.
Especially in the southern part of its distribution in Luzon, the natural forests ecosystems in which *R. lagascae* is found are confined to small areas separated by vast areas in which forests have been converted into farmland. This has undoubtedly resulted in decreased genetic connectivity between the remaining populations.

Our results show substantial differences in allelic composition and frequencies between populations of the *R. lagascae* complex (Table 2). This means that considerable genetic diversity might be lost if any of its populations would go extinct (Schmidt and Jensen 2000). Small or isolated populations outside of protected areas are most at risk of extinction. These include the populations from Maria Aurora, Mt. Malinao, and Mt. Mingan (Table 2). However, also the larger populations outside of protected areas (Bolos Point, Mt. Irid, and Mt. Labo) are of conservation concern, because they harbor the highest levels of genetic diversity and need to be protected from habitat destruction and degradation to remain centers of genetic diversity.

Because of their pronounced genetic differences and the lack of gene flow between them, *R. lagascae* s. s., Mt. Labo *R. lagascae*, and *R. manillana* should be managed as distinct conservation entities. Mt. Labo *R. lagascae* and *R. manillana* are each only known from a single population, which shows lower genetic diversity than similar sized populations of *R. lagascae* s. s. (Table 2). Both taxa therefore have a higher risk of losing genetic diversity and of becoming extinct than *R. lagascae* s. s. The habitat of *R. manillana* is protected as part of the Samar Island Natural Park and is relatively intact, but the Mt. Labo area does not enjoy formal protection. The Mt. Labo *Rafflesia* is therefore particularly prone to experiencing a reduction in the size of its single population and should thus be managed as a critically endangered species (IUCN 2012).
ACKNOWLEDGEMENTS—This project was funded by the Marsden Fund Council from Government funding administered by the Royal Society of New Zealand and by the National Geographic Foundation. We would like to thank the three anonymous reviewers of this manuscript for their valuable comments. We are grateful to the National Museum of the Philippines, the Department of Environment and Natural Resources (DENR) and in particular M.S. Lim, Director of the Biodiversity Management Bureau (BMB) and staff of the DENR regional and local offices for facilitating the issuance of collecting, transport and export permits. Prior Informed Consents were issued by the Protected Areas and Management Boards (PAMB) of Aurora Memorial National Park, Bataan Natural Park (Mt. Natib), Pantabangan-Carranglan Watershed Reserve, Mt. Banahaw-Mt. San Cristobal Protected Area, Samar Island Natural Park, the Mayors of Buhi, Camarines Sur; Gattaran and Lal-o, Cagayan; Gabaldon, Nueva Ecija; Labo and San Lorenzo Ruiz, Camarines Norte; Rodriguez, Rizal, and the Makiling Center for Mountain Ecosystems. We would also like to thank our NGO partners: Cagayan Valley Partners in People Development and the Philippine Native Plants Conservation Society, Inc. For sharing information on distributional ranges and fieldwork assistance, we thank J. Aliposa, D.S. Balete, M.O. Cajano†, J.R. Callado, L.L. Co†, V. Dacumos, A. Diesmos, N.B. Gapas, J. Sarmiento, and D. Tandang. Numerous field guides/assistants and staff of DENR and LGU deserve our sincerest gratitude for without their help, fieldwork would have not been possible. Finally, M.L. Hale and T.E. Steeves provided helpful feedback on the results of the population genetic analyses and M. Walters helped with preparing figures for this publication.


*Environmental Science and Technology* 34: 130A–135A.


Global warming, elevational ranges and the vulnerability of tropical biota.

*Biological Conservation* 144: 548–557.


Pelser, Nickrent, Gemmill, and Barcelona 30


