

# Delimiting Species without Nuclear Monophyly in Madagascar's Mouse Lemurs

David W. Weisrock<sup>1,\*</sup>, Rodin M. Rasoloarison<sup>2,3</sup>, Isabella Fiorentino<sup>1</sup>, José M. Ralison<sup>3</sup>, Steven M. Goodman<sup>4,5</sup>, Peter M. Kappeler<sup>2</sup>, Anne D. Yoder<sup>1,6</sup>

**1** Department of Biology, Duke University, Durham, North Carolina, United States of America, **2** Department of Behavioural Ecology and Sociobiology, German Primate Centre, Göttingen, Germany, **3** Département de Biologie Animale, Université d'Antananarivo, Antananarivo, Madagascar, **4** Field Museum of Natural History, Chicago, Illinois, United States of America, **5** Vahatra, Antananarivo, Madagascar, **6** Duke Lemur Center, Duke University, Durham, North Carolina, United States of America

## Abstract

**Background:** Speciation begins when populations become genetically separated through a substantial reduction in gene flow, and it is at this point that a genetically cohesive set of populations attain the sole property of species: the independent evolution of a population-level lineage. The comprehensive delimitation of species within biodiversity hotspots, regardless of their level of divergence, is important for understanding the factors that drive the diversification of biota and for identifying them as targets for conservation. However, delimiting recently diverged species is challenging due to insufficient time for the differential evolution of characters—including morphological differences, reproductive isolation, and gene tree monophyly—that are typically used as evidence for separately evolving lineages.

**Methodology:** In this study, we assembled multiple lines of evidence from the analysis of mtDNA and nDNA sequence data for the delimitation of a high diversity of cryptically diverged population-level mouse lemur lineages across the island of Madagascar. Our study uses a multi-faceted approach that applies phylogenetic, population genetic, and genealogical analysis for recognizing lineage diversity and presents the most thoroughly sampled species delimitation of mouse lemur ever performed.

**Conclusions:** The resolution of a large number of geographically defined clades in the mtDNA gene tree provides strong initial evidence for recognizing a high diversity of population-level lineages in mouse lemurs. We find additional support for lineage recognition in the striking concordance between mtDNA clades and patterns of nuclear population structure. Lineages identified using these two sources of evidence also exhibit patterns of population divergence according to genealogical exclusivity estimates. Mouse lemur lineage diversity is reflected in both a geographically fine-scaled pattern of population divergence within established and geographically widespread taxa, as well as newly resolved patterns of microendemism revealed through expanded field sampling into previously poorly and well-sampled regions.

**Citation:** Weisrock DW, Rasoloarison RM, Fiorentino I, Ralison JM, Goodman SM, et al. (2010) Delimiting Species without Nuclear Monophyly in Madagascar's Mouse Lemurs. PLoS ONE 5(3): e9883. doi:10.1371/journal.pone.0009883

**Editor:** Manfred Kayser, Erasmus University Medical Center, Netherlands

**Received:** August 25, 2009; **Accepted:** February 19, 2010; **Published:** March 31, 2010

**Copyright:** © 2010 Weisrock et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This project was supported by grant DEB 0516276 from the National Science Foundation, and by funding from the John D. and Catherine T. MacArthur Foundation, the German Primate Center (DPZ), the Deutsche Forschungsgemeinschaft (grant Ka 1082-8), and the Volkswagen Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dweis2@uky.edu

‡ Current address: Department of Biology, University of Kentucky, Lexington, Kentucky, United States of America

## Introduction

Through decades of diverging opinions, at least one component of the species problem – the disagreement over what exactly species are – has found resolution in the consensus view that species are solely defined as separately evolving metapopulation lineages [1,2,3]. As such, species exist from the very beginning of their separation and divergence from other lineages and the many available criteria for delimiting these lineages thus mark different points in this process [4,5]. Despite this reconciliation, evolutionary biologists still have great difficulty in recognizing species in the early stages of divergence due to the limited time for differences to evolve that satisfy most delimitation criteria [5]. Identifying lineages in the early stages of species divergence is, nonetheless, extremely important because the study of these lineages is

expected to be the most informative about the speciation process [6].

Until recently, species delimitation methods that provide valid biological evidence for the early stages of lineage divergence have been difficult to properly enumerate. Recent developments in the population genetics of speciation may offer new power to resolve recently diverged lineages, however [7]. Genetic patterns generated by population-level processes operating within diverging lineages are expected to contain the signal of speciation [8,9,10] even though divergence is not long enough to generate overt phylogenetic patterns of independent evolution, such as exclusive monophyly at multiple loci [11,12]. This is, in part, the basis for favoring the use of monophyly in mitochondrial DNA gene trees as evidence for speciation due to a reduced effective population size relative to nuclear loci [13]. This preference relies on the

expectation that an increased rate of lineage sorting leads to mtDNA monophyly prior to nuclear monophyly [14]. Even so, accuracy in species delimitation requires the use of more than a single locus [15] and information from multiple loci, ideally in combination with other types of data, are expected to provide a more robust estimate of independently evolving lineages. Methods that summarize population genetic and genealogical patterns across multiple loci are essential for diagnosing these young evolutionary lineages [16,17,18].

Lemurs (Lemuriformes: Primates) are a flagship group in the study of the evolutionary and biogeographic mechanisms that have lead to Madagascar's megadiverse biota [19,20,21], but face enormous pressure from human-related activity associated with the destruction of their natural forest habitat [22]. A realistic understanding of the species diversity and boundaries of lemurs is therefore fundamental to understanding the evolution of Malagasy biodiversity and in conserving these threatened primates.

In the past decade, the species diversity of mouse lemurs (genus *Microcebus*) has increased more than seven-fold, largely through the analysis of mtDNA sequence data [23,24,25,26,27,28,29]. This rapid increase in species numbers has led to a questioning of the true level of lemur species diversity [30] and debate over whether or not this represents overdiagnosis or a true representation of species diversity. The sole utilization of mtDNA, typical of these recent studies, is problematic in that such data do not address the lack of independence of substitutions among nucleotide positions and ignores the potential discord between gene trees and species trees [31,32,33]. For example, the exhibition of female philopatry in mouse lemurs [34,35] suggests that the maternally inherited mtDNA locus may be strongly biased in such cases.

In this paper, we present the most thoroughly sampled species delimitation study of Malagasy mouse lemurs to date. We present our analyses as a working example of species delimitation where many lineages are morphologically cryptic and recently diverged. These small nocturnal primates also typify the complexities that evolutionary biologists face in assessing diversity in an organismal group that is difficult to study. In the case of mouse lemurs, which are phenotypically very similar, morphological data in most cases cannot be used to separate putative cryptic species, available sample sizes are typically small, both within and among populations, and ecological and behavioral data are often lacking. Instead of viewing these groups as systematically intractable, we argue that the accumulation of substantial genetic data allows us to progress towards a general assessment of species delimitation. Given that lemurs are for the most part forest-dependent, the stakes are often very high, with habitat destruction perhaps erasing species as fast as we can identify and study them [36].

In the case of the present study, we geographically sampled mouse lemurs from the remaining forested areas across Madagascar, covering virtually all of the island's unique biomes and micro-endemic regions (Fig. 1) [21]. While the number of individuals from any given locality is sometimes reduced, the total sample of 216 individuals (286, with the inclusion of GenBank mtDNA data) represents the effort of numerous field biologists from a time span of more than a decade. As such, it is the most complete synthesis to date, though future sampling efforts will continue to augment our understanding of mouse lemur evolutionary diversity. Although previous studies of mouse lemur evolutionary diversity have included morphological data [28,37], here we focus primarily on genetic data sampled from mtDNA and four independent nuclear loci. We assess patterns of monophyly across independently reconstructed gene trees to identify lineages that exhibit genealogical exclusivity, an expected pattern for lineages with long durations of divergence [11,12,15].

However, the larger focus of this study is on the delimitation of population-level lineages that have a relatively recent history of divergence. We use a combined data approach to identify distinct nuclear genetic clusters using analyses of population structure [18,38]. We test hypotheses of diverging lineages identified through the mtDNA gene tree and nuclear clustering with recently developed statistics that quantify the degree of exclusive ancestry in an assemblage of gene trees [17]. As such, this study offers an opportunity to fully explore the strengths, as well as the limitations, of genetic data for the interpretation of species boundaries. In conclusion, we aim to show that a well-developed multi-locus genetic data set that is analyzed appropriately can yield penetrating insights into the history and reproductive boundaries both within and among the evolutionary lineages that we infer to be species.

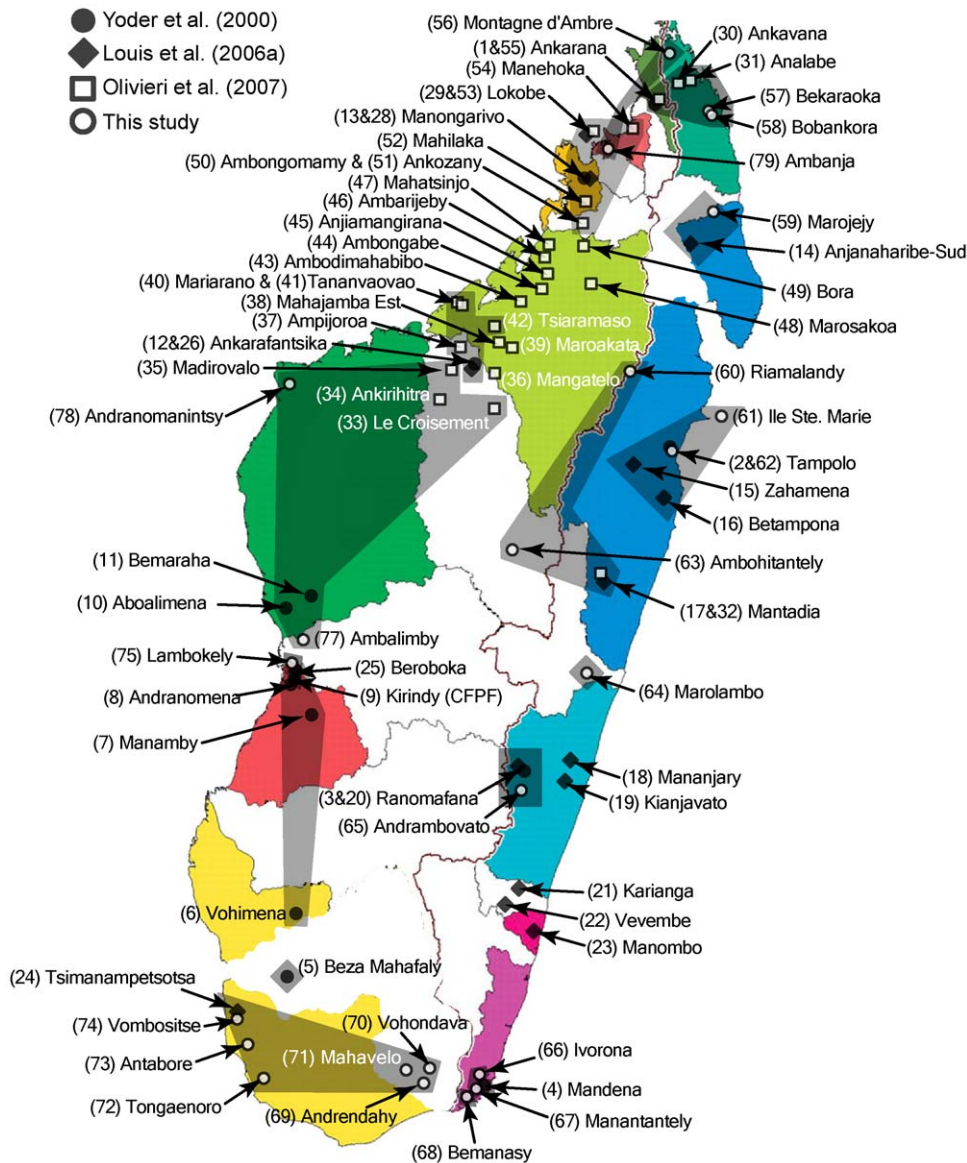
## Materials and Methods

### Organismal sampling

MtDNA sequence data were analyzed from 286 individual mouse lemurs distributed across Madagascar (Fig. 1). 102 of these represent new material from 23 previously unsampled localities distributed throughout Madagascar. MtDNA sequence data from the remaining individuals comes from previous work [25,26,37]. Nuclear sequence data were analyzed from a total of 216 mouse lemurs. As with the mtDNA data, 102 of these are new material from the 23 previously unsampled localities. Nuclear sequence data from the remaining samples are those of Heckman et al. [39], which are from the same localities presented in the mtDNA work of Yoder et al. (2000). Full details regarding the number of individuals per locality, locality names, and geographic coordinate data can be found in Table S1. The nuclear sequence data of Heckman et al. [39] required editing for the identification of previously unidentified heterozygous nucleotide positions, and for the refinement of allele phases, and are deposited in GenBank with new accession numbers. Outgroup sequences for phylogenetic analysis were obtained from GenBank for single individuals of each of the following species: *Cheirogaleus crossleyi*, *C. major*, *C. medius*, and *Mirza coquereli*. GenBank accession numbers for all outgroup individuals and for the mtDNA data of Louis et al. [25] and Olivieri et al. [26] can be found in Table S2. All new sequence data and sequence data from the studies of Yoder et al. [37] and Heckman et al. [39] are available in GenBank with the following accession numbers: mtDNA *cox2* (GU326974-GU327160), mtDNA *cob* (GU327161-GU327362), *adora3* (GU230899-GU231330), *eno* (GU231331-GU231716), *fga* (GU231717-GU232130), and *vref* (GU232131-GU232490).

### Genetic Sampling and Data Collection

MtDNA sequence data from two separate regions, *cox2* and *cob*, were generated using PCR and direct sequencing with the PCR primers L7553/H8320 [40] and L14724/H15915 [41], respectively. Nuclear DNA sequence data were generated from four independent loci (*adora3*, *fga*, *eno*, and *vref*). These are the same loci used in Heckman et al. [39] and PCR primer information are found within that reference except for those used in the amplification of *adora3*, which are from Horvath et al. [42]. PCR was performed in volumes of 20  $\mu$ l using 2  $\mu$ l template DNA (approximately 50-150 ng DNA), 25  $\mu$ M each dNTP, 1  $\mu$ M each primer, and 0.625U Taq polymerase in a standard 1x reaction buffer. Typical mtDNA amplification conditions were carried out with an initial 94°C denaturation for 2 min, followed by 35 cycles of 30 s at 94°C, 30 s at 50°C and 45 s at 68°C. A final 7 minutes extension was performed at 72°C. See Horvath et al. (2008) for specific annealing conditions for each nuclear locus. PCR products



including all individuals and their respective gene copies using MrBayes v3.1.2 [47]. Evolutionary models for each locus were assessed for the haplotype data sets using Akaike Information Criteria in MrModeltest v2.3 [48]. MtDNA data for individual mouse lemurs were concatenated and analyzed in a two-partition framework with model parameters estimated separately for the *cox2* and *cytb* genes. Nuclear gene data sets were analyzed as a single partition. Four Markov chains were used with the default temperature parameter of 0.2. Default priors were used in all analyses and random trees were used to start each Markov chain. Chains were run for 10 million generations with topology and model parameter estimates sampled every 1000 generations. The first five million generations were discarded as burn-in yielding a posterior distribution of 5000 sampled trees. Mean log likelihood (lnL), branch lengths and topologies were compared across four replicate analyses to insure that a stable posterior distribution was reached. Sampled trees from the posterior distributions of replicate analyses were pooled and parsed with MrBayes to construct a majority rule consensus tree and to calculate posterior probabilities (PPs) of all resulting branches. We assessed nuclear monophyly of terminal clades resolved in the mtDNA haplotype tree and for previously described taxa by filtering the nuclear Bayesian posterior distributions for trees that meet that constraint of monophyly for specified groups. Tree filtering was performed in PAUP\* version 4.0 [49].

### Bayesian Structure Analysis

Analysis of population structure was assessed using a genotype matrix of the nuclear loci in STRUCTURE v2.2 [50,51]. We ran a series of analyses under models assuming a specific number of populations ( $K$ ), with a range of  $K$  from 2 to 26. In each iteration, individuals were assigned probabilistically to a cluster based on their multilocus genotype. All analyses used one million MCMC generations to estimate the posterior distribution following a burnin period of one million generations. Our model incorporated the possibility that some individuals may have mixed population ancestry and the possibility that allele frequencies are correlated among populations due to migration or shared ancestry [51]. For each  $K$ , the log (ln) probability of the data ( $X$ ) was estimated [ $\ln \Pr(X|K)$ ] and used to calculate the posterior probability (PP) of  $K$  under the assumption of a uniform prior. We also calculated  $\Delta K$  [52], which is based on the rate of change in  $\ln \Pr(X|K)$  between successive  $K$  values. These two measures often identify different optimal measures of  $K$ , and  $\Delta K$  may favor smaller values of  $K$  that represent basal levels of hierarchical structure in systems that substantially deviate from an island model [52]. We visualized the

optimal  $K$  STRUCTURE plot using Microsoft Excel. Membership coefficients for individuals with posterior probabilities less than 0.05 were disregarded and proportionally assigned to the other cluster assignment coefficients.

### Genealogical Tests of Population Divergence

We assessed the level of genealogical divergence in our nuclear gene trees for hypothesized lineages identified in the mtDNA gene tree and nuclear STRUCTURE analysis using the genealogical sorting index ( $gsi$ ) [17]. For pre-defined groups in a gene tree, the  $gsi$  is a standardized measure of the degree to which they exhibit exclusive ancestry. The  $gsi$  statistic ranges from 1 (monophyly) to 0 (a complete lack of genealogical divergence with other groups). A major benefit of the  $gsi$  statistic for species delimitation is the ability to assess its statistical significance through the randomization of group labels across the tips in a gene tree. Consequently, hypothesized lineages can be tested against a null hypothesis of no divergence, as measured by coalescent patterns in their gene trees. For each individual nuclear locus, we calculated the  $gsi$  for 100 trees randomly sampled from the combined Bayesian posterior distribution of trees. These 100 individual  $gsi$  measurements were then used together with equal weight to calculate an ensemble  $gsi$  statistic ( $gsi_T$ ) for each locus. Thus,  $gsi_T$  measurements serve as a summary of the genealogical exclusivity across the Bayesian posterior distribution of trees for a given locus. In addition to the single-locus assessments of genealogical exclusivity, we also calculated  $gsi_T$  for the set of majority-rule consensus trees of all four nuclear loci. The significance of all  $gsi$  and  $gsi_T$  statistics was assessed using 1000 randomization permutations. All analyses were performed using the Genealogical Sorting Index web server ([www.genealogicalsorting.org](http://www.genealogicalsorting.org)).

### Results

Information regarding size, level of variability, molecular evolutionary models, and likelihood values from phylogenetic analyses for all mtDNA and nDNA genes can be found in Table 1. No evidence was found for recombination within any of the four nuclear loci. Furthermore, no locus yielded a significant signature of a departure from neutrality.

### MtDNA Gene Tree

The mtDNA haplotype tree resolved the majority of previously described species as clades (Fig. 2, Table 2; full mtDNA gene tree details are presented in Figure S1) with just two exceptions: *M. rufus* mtDNA haplotypes are paraphyletic and *M. jollyae* and

**Table 1.** Information for loci used in this study.

Locus	Size (bp)	Variable Sites	No. of Haplotypes	Favored Model <sup>a</sup>	Mean Bayesian lnL (haplotype data)	Mean Bayesian lnL (full data)	ML lnL (full data)
mtDNA ( <i>cox2</i> )	684	176	169 <sup>b</sup>	HKY+H+G	-13360.0 <sup>c</sup> (-13400.0, -13320.0)	—	—
mtDNA ( <i>cytb</i> )	1184	396	169 <sup>b</sup>	HKY+H+G	—	—	—
<i>adora3</i>	384	34	38	HKY+G	-1104.6 (-1118.3, -1093.08)	-1223.7 (-1244.3, -1204.9)	-852.9
<i>eno</i>	913	150	156	HKY+H+G	-4557.6 (-4582.1, -4531.5)	-4717.33 (-4754.1, -4684.4)	-3915.7
<i>fga</i>	632	74	82	HKY+H+G	-2232.9 (-2251.07, -2215.7)	-2433.1 (-2460.1, -2406.9)	-1953.0
<i>vwf</i>	824	117	140	HKY+H+G	-3980.1 (-4007.5, -3953.3)	-4164.8 (-4230.8, -4099.0)	-3379.7

<sup>a</sup>Model selected according to the Akaike Information Criterion in the program MrModelTest.

<sup>b</sup>Haplotype number includes sequence data from GenBank and is calculated for the concatenated set of mtDNA data.

<sup>c</sup>Scores are calculated for the concatenated mtDNA data.

doi:10.1371/journal.pone.0009883.t001





**Figure 2. Correspondence between clades in the mtDNA gene tree and nuclear STRUCTURE clusters.** The mtDNA gene tree results from Bayesian analysis of concatenated *cox2* and *cytb* sequence. It is presented as a maximum credibility topology with branch lengths averaged across the posterior distribution. Numbers on branches represent posterior probabilities. Relationships within terminal clades are collapsed for ease of presentation and clades are labeled according to present species designations. Locality numbers are given in parentheses and correspond to Fig. 1 and Table S1. Clades are mapped to corresponding clusters in the nuclear STRUCTURE plot. Each cluster is designated by a different color with horizontal bars representing individuals and the proportion of a bar assigned to a single color representing the posterior probability that an individual is assigned to that cluster. This can also be interpreted as the percentage of an individual's genome that is derived from that particular genetic cluster. Localities from which individuals are sampled from are given along the right side of the plot. MtDNA clades not mapped to the assignment plot represent individuals for which corresponding nDNA data is not available. The colors in the Bayesian assignment plot do not correspond to colored areas of micro-endemism in Fig. 1.  
doi:10.1371/journal.pone.0009883.g002

*Microcebus* sp. nova 3 [25] share a mtDNA haplotype. MtDNA variation is highly structured across the island with numerous sampling localities within currently recognized species exhibiting monophyly (dashed-line clades in Fig. 2). In addition, some sets of localities within described species are reconstructed as geographically paraphyletic groups (e.g. within *M. griseorufus* and *M. myoxinus*). Three novel mtDNA clades are resolved based on new field sampling: (1) haplotypes sampled from Marolambo (locality 64), (2) haplotypes sampled from the localities of Ivorona and Manantantely (localities 66 and 67), and (3) haplotypes sampled from Ambanja and Montagne d'Ambre (Localities 56 and 79). This latter clade may represent *M. amholdi*, a recently described species from Montagne d'Ambre that was diagnosed using patterns of mtDNA divergence [29]. However, that study used different mtDNA gene regions than those used here, preventing a direct link with the lineage resolved in our mtDNA gene tree. Therefore, we treat the Ambanja+Montagne d'Ambre lineage resolved here as an unnamed lineage with the understanding that it may represent *M. amholdi*.

## Nuclear Gene Trees

Bayesian phylogenetic analysis of haplotype data for the individual nuclear genes results in majority rule consensus trees that resolve only a subset of the clades resolved in the mtDNA gene tree (Figures S2, S3, S4, S5). Filtering of trees from the full-data Bayesian posterior distributions reveals that few of the mtDNA-based clades exhibit monophyly with posterior probabilities greater than 0.05 (Table 2). The exceptions to this are *M. griseorufus*, *M. sambiranensis*, *Microcebus* sp. (Ambanja+Montagne d'Ambre), and *Microcebus* sp. (Ivorona+Manantantely), all of which are monophyletic in one nuclear gene tree, *M. mamiatra*, which is monophyletic in two nuclear gene trees, *M. tavaratra* and *Microcebus* sp. (Marolambo), both of which are monophyletic in three nuclear gene trees, and *M. ravelobensis* and *M. simmonsii*, which exhibit monophyly in all four nuclear gene trees (Table 2). Finally, construction of a strict consensus tree from the ML nuclear gene trees does not yield reciprocal monophyly for additional groups that are not present in the mtDNA gene tree (results not shown), but instead yields reciprocal monophyly for just two species, *M. ravelobensis* and *M. simmonsii*.

## Nuclear Population Structure

Plots of the estimated log probability of the data [ $\log \Pr(X|K)$ ] for replicated STRUCTURE analyses reveal a general pattern of a plateau or decrease in values above a  $K=19$  (Figure S6). Four replicate analyses yield a posterior probability of 1.0 for a  $K=19$  (Figure S6). The remaining replicate analyses had posterior probabilities of 1.0 for  $K$ s of 20, 23, 24, and 26. In contrast, calculations of  $\Delta K$  produce a peak at  $K=2$  (Figure S6). This result likely stems from the identification of a basal level of hierarchical structure in the data [52]. The plateau in patterns of the  $\log \Pr(X|K)$  around a  $K=19$ , and the high posterior probabilities for  $K \geq 19$ , suggest an overall higher population cluster number. A plot

of individual membership coefficients for  $K=19$  reveals a high number of population clusters with average individual membership coefficients (i.e. posterior probabilities) greater than 0.9 (Figure 2, Table 2). STRUCTURE plots for  $K > 19$  do not yield additional clusters with high membership coefficients for more exclusive sets of populations or individuals, but instead further divide already admixed sets of individuals (e.g. populations within *M. murinus*) into additional admixed cluster assignments. Therefore, we place our focus on  $K=19$  as an estimate of the upper level of population clustering.

There is strong concordance between identified nuclear clusters and terminal mtDNA clades (Fig. 2). Of the 12 mtDNA clades representing described species, and for which we have corresponding nuclear data, 11 map to one or more nuclear genotypic clusters characteristic to that clade. Many populations within described species that exhibit mtDNA monophyly also map to distinct nuclear clusters with high individual membership coefficients (Table 2). For example, all individuals sampled from the *M. murinus* population of Mandena (locality 4) have a monophyletic assemblage of mtDNA haplotypes and an assignment to a single nuclear STRUCTURE cluster with average membership coefficients of 1.0. The three novel mtDNA clades from the populations of Marolambo, Ivorona+Manantantely, and Ambanja+Montagne d'Ambre also each map to their own respective nuclear clusters with average individual membership coefficients of 1.0, 0.93, and 1.0, respectively. In total, using corresponding patterns of nuclear STRUCTURE clustering and mtDNA monophyly, we diagnose a total of 16 hypothesized lineages of *Microcebus* (Table 2). These criteria focus on the assignment of mtDNA clade-specific localities or sets of localities to one or more characteristic nuclear clusters with limited signs of admixture (i.e. average cluster membership coefficients  $\leq 5\%$  for other localities).

In contrast to these patterns, individuals within the mtDNA haplotype clade that correspond to the recently described species, *M. mamiatra*, have nuclear STRUCTURE assignments identical to individuals of *M. sambiranensis* (Fig. 2, Table 2). Similar patterns are also seen for some localities within described species that exhibit mtDNA monophyly, but do not have high membership coefficients to distinct nuclear clusters (e.g. the locality of Vohimena within *M. murinus* and the locality of Ambalimby within *M. myoxinus*; Fig. 2, Table 2).

## Nuclear Genealogical Sorting

Overall, there is considerable variation in measures of genealogical divergence across hypothesized lineages, across gene trees within lineages, and across the Bayesian posterior distribution of individual gene trees (Table 3). Nonetheless, significant measures of exclusive ancestry are estimated at all of these levels of organization with just two exceptions (Table 3): (1) the *adora3* Bayesian posterior distribution contains  $>5$  individual trees in which the *gsi* for *M. lehilahytsara* is not significant; however, the *adora3 gsi-T* for *M. lehilahytsara* significantly rejects the null

**Table 2.** Probability of monophyly in the single-gene Bayesian posterior distributions and nuclear cluster assignments for *Microcebus* species and populations.

Lineage	Posterior probability for monophyly			Bayesian assignment cluster(s) w/average individual membership coefficients			mtDNA group + genotypic cluster
	mtDNA	<i>adora3</i>	<i>eno</i>	<i>fga</i>	<i>wvf</i>		
<i>M. griseorufus</i>	1.0	0	0.99	0	0.0004	1 (0.28), 2 (0.71), <b>1+2 (1.0)</b>	Yes
<i>M. murinus</i> – sensu lato	1.0	0	0.04	0	0	3 (0.17), 4 (0.29), 5 (0.28), 6 (0.25), <b>3+4+5+6 (0.99)</b>	-
– Bemanasy	1.0	0	0	0	0	<b>3 (0.99)</b> , 4 (0.01)	Yes
– And/Kir/Manam/Vohi	0.0 <sup>a</sup>	0	0	0	0	3 (0.05), 4 (0.47), 5 (0.45), <b>4+5 (0.92)</b> , 6 (0.03)	Yes <sup>a</sup>
– Mandena	1.0	0	0	0	0	<b>6 (1.0)</b>	Yes
<i>M. damfossi</i>	1.0	No Data	No Data	No Data	No Data	No Data	-
<i>M. bongolavensis</i>	<b>0.99</b>	No Data	No Data	No Data	No Data	No Data	-
<i>M. ravelobensis</i>	1.0	0.9781	1.0	1.0	0.9914	<b>7 (1.0)</b>	Yes
<i>Microcebus</i> sp. – Marolambo	1.0	0	1.0	0.576	0.205	<b>8 (1.0)</b>	Yes
<i>Microcebus</i> sp. – Ivorona/Manantantely	1.0	0	0	0.9754	0	1 (0.01), 8 (0.01), <b>9 (0.93)</b> , 14 (0.04), 15 (0.01)	Yes
<i>M. jollyae</i>	0.0	No Data	No Data	No Data	No Data	No Data	-
<i>M. simmonsii</i>	1.0	0.9957	1.0	1.0	0.9914	<b>10 (1.0)</b>	Yes
<i>M. tavaratra</i>	1.0	0.9318	1.0	0	0.16	<b>11 (0.99)</b> , 19 (0.01)	Yes
<i>M. mamiatra</i>	1.0	0.001	0	0	0	12 (1.0)	No
<i>Microcebus</i> sp. Ambanja/Montagne d'Ambre	1.0	0	0.301	0	0	<b>13 (1.0)</b>	Yes
<i>M. sambiranensis</i>	1.0	0.149	0	0	0	<b>12 (.98)</b> , 15 (0.02)	Yes
<i>M. mittermeieri</i>	1.0	0	0	0	0	<b>14 (0.97)</b> , 15 (0.03)	Yes
<i>M. lehilahysara</i>	1.0	0	0	0	0	14 (0.04), <b>15 (0.95)</b> , 18 (0.01)	Yes
<i>M. berthae</i>	1.0	0	0	0	0	<b>16 (0.91)</b> , 17 (0.04), 18 (0.02), 19 (0.03)	Yes <sup>c</sup>
<i>M. rufus</i>	0 <sup>b</sup>	0	0	0	0	8 (0.01), 16 (0.12), 17 (0.78), 18 (0.02), 19 (0.07)	Yes <sup>c</sup>
<i>M. myoxinus</i>	<b>0.77</b>	0	0	0	0	8 (0.01), 14 (-0), 15 (-0), 17 (0.01), 18 (0.36), 19 (0.62), <b>18+19 (0.98)</b>	Yes <sup>c</sup>
Total number of diagnosed lineages =							<b>16</b>

For definitions of the locality acronyms within *M. murinus* see Table S1.

<sup>a</sup>The populations of Andranomena+Kirindy(CFPF)+ Mamamby and the population of Vohimena each form separate mtDNA clades.

<sup>b</sup>*Microcebus rufus* is paraphyletic with respect to *M. myoxinus*.

<sup>c</sup>*Microcebus berthae* and *M. myoxinus* both are placed into distinct nuclear clusters that are also present in greater than 5% of the average cluster assignments of *M. rufus*. We diagnose these as three distinct clusters based on their geographic distributions (see discussion).  
doi:10.1371/journal.pone.0009883.t002

**Table 3.** Ensemble genealogical sorting indices ( $gsi_T$ ) for the combined set of majority-rule consensus trees from all four nuclear loci and for the Bayesian posterior distributions of each individual nuclear locus.

Lineage	All loci $gsi_T$	<i>adora3</i> $gsi_T$ (min-max)	<i>eno</i> $gsi_T$ (min-max)	<i>fib</i> $gsi_T$ (min-max)	<i>vwf</i> $gsi_T$ (min-max)
<i>M. berthae</i>	0.158	0.185 (0.141–0.269)	0.573 (0.377–0.843)	0.259 (0.207–0.343)	0.303 (0.222–0.412)
<i>M. griseorufus</i>	0.487	0.414 (0.257–0.519)	0.729 (0.683–0.814)	0.673 (0.664–0.717)	0.48 (0.359–0.576)
<i>M. lehilahytsara</i>	0.069	0.119 (0.074–0.187) <sup>a</sup>	0.243 (0.155–0.529)	0.407 (0.27–0.538)	0.226 (0.149–0.344)
<i>M. mittermeieri</i>	0.142	0.141 (0.103–0.211)	0.384 (0.255–0.643)	0.361 (0.179–0.639)	0.512 (0.304–0.739)
<i>M. murinus</i> (Bemanasy)	0.157	0.224 (0.153–0.327)	0.204 (0.168–0.247)	0.2 (0.141–0.28)	0.323 (0.246–0.487)
<i>M. murinus</i> (Mandena)	0.357	0.314 (0.255–0.377)	0.768 (0.765–0.862)	0.323 (0.246–0.45)	0.474 (0.4–0.571)
<i>M. murinus</i> (remaining pops)	0.406	0.469 (0.395–0.571)	0.892 (0.842–0.926)	0.36 (0.279–0.506)	0.45 (0.317–0.57)
<i>M. myoxinus</i>	0.416	0.341 (0.311–0.392)	0.522 (0.441–0.631)	0.429 (0.368–0.489)	0.736 (0.672–0.793)
<i>M. ravelobensis</i>	1.0	1.0 (0.956–1.0)	1.0 (1.0–1.0)	0.995 (0.727–1.0)	1.0 (1.0–1.0)
<i>M. rufus</i>	0.326	0.302 (0.25–0.357)	0.533 (0.45–0.67)	0.431 (0.369–0.5)	0.7 (0.62–0.829)
<i>M. sambiranensis</i>	0.491	0.913 (0.334–1.0)	0.407 (0.293–0.605)	0.344 (0.203–0.465)	0.756 (0.559–0.935)
<i>M. simmonsii</i>	1.0	1.0 (0.981–1.0)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	0.999 (0.944–1.0)
<i>M. tavaratra</i>	0.724	0.981 (0.495–1.0)	1.0 (1.0–1.0)	0.359 (0.194–0.511)	0.895 (0.761–1.0)
<i>Microcebus</i> sp. (Ambanja/Montagne d'Ambre)	0.270	0.259 (0.126–0.538)	0.892 (0.725–1.0)	0.516 (0.296–0.78)	0.383 (0.183–0.536)
<i>Microcebus</i> sp. (Ivorona/Manantantely)	0.478	0.141 (0.125–0.23)	0.669 (0.589–0.8)	1.0 (1.0–1.0)	0.383 (0.187–0.566)
<i>Microcebus</i> sp. (Marolambo)	0.713	0.097 (0.062–0.183) <sup>a</sup>	1.0 (1.0–1.0)	0.966 (0.831–1.0)	0.861 (0.71–1.0)

<sup>a</sup>>5 of the 100 trees sampled from the Bayesian posterior distribution had  $gsi$  values with  $p > 0.05$ .

<sup>b</sup>*Microcebus rufus* is not diagnosed as a diverging lineage according to the criteria used in this study, but is included here based on patterns of population differentiation from *M. berthae* and *M. myoxinus* and due to its delimitation based on morphological and ecological traits [37].

The minimum and maximum  $gsi$  for individual trees within the Bayesian posterior distributions are given in parentheses.

doi:10.1371/journal.pone.0009883.t003

hypothesis of no divergence, and (2) the *adora3* Bayesian posterior distribution contains >5 individual trees in which the  $gsi$  for the undescribed lineage (*Microcebus* sp.) from Marolambo is not significant. However, in this latter case the overall *adora3*  $gsi_T$  measurement is not significant.  $Gsi_T$  values for most lineages are lower for the *adora3* locus than the other three nuclear loci, a result consistent with the overall lower level of genetic variation in the *adora3* data set (Table 1).

## Discussion

### Species delimitation in mouse lemurs

Speciation begins when populations become genetically separated through a substantial reduction in gene flow, either through vicariance, or through selection and adaptation, and it is at this point that a genetically cohesive set of populations attain the sole property of species: the independent evolution of a population-level lineage [1,3]. The delimitation of lineages in these early stages of speciation is important because their study is most likely to yield insights into the mechanisms that drove their formation [6]. The comprehensive delimitation of all diverging lineages within hotspots of biodiversity, regardless of their level of divergence, is important for understanding the factors driving the diversification of biota [e.g. 21] and guiding biologically realistic conservation action [53]. However, delimiting recently diverged lineages is challenging due to insufficient time for the differential evolution of characters – including morphological differences, reproductive isolation, and monophyly in gene trees – that are typically used as evidence for separately evolving lineages [5]. The most likely patterns to evolve over short time scales of divergence are the population genetic patterns of differentiation. The reduction of gene flow among populations allows genetic drift to operate independently within cohesive sets of populations

yielding distinctive patterns in allele frequencies [54] and in gene trees [9,55]. The application of species delimitation criteria that identify lineages exhibiting the population genetic patterns of cohesion through gene flow [18,38] and genealogical patterns of divergence [16,17] can thus have significant benefits over criteria that are more often limited to identifying lineages that are well into the process of divergence.

In this study, we assembled multiple lines of evidence from the analysis of mtDNA and nDNA sequence data for the delimitation of numerous cryptically diverged lineages of mouse lemurs across the island of Madagascar. This evidence included one of the more standard phylogeographic components of lineage diagnosis: the resolution of clades in a mtDNA gene tree. However, there is an overall lack of corresponding nuclear monophyly for most mtDNA clades. Only two population-level lineages, the previously described *M. ravelobensis* and *M. simmonsii*, are delimited via monophyly in all mtDNA and nDNA gene trees (Table 2), and just nine lineages are delimited when this criterion was relaxed to monophyly in mtDNA and at least one nuclear gene tree. Interestingly, these monophyly criteria fail to delimit many lineages previously described as species on the basis of clear and concordant patterns of mtDNA, morphological, and ecological differentiation [28,37,56]. In contrast, STRUCTURE analysis of the nuclear data clustered individuals into groups that strongly align with mtDNA clades (Fig. 2). Furthermore, despite a lack of nuclear monophyly for most of these population groupings, estimates of their  $gsi$  significantly favor independent histories of lineage divergence (Table 3). Within the framework of the general lineage concept of species [1,3], which clarifies that new species arise at the very beginning of lineage divergence from an ancestor, corroborating lines of evidence that indicate the independent divergence of a lineage provide substantial support for its delimitation as a species [5]. We, therefore, present our study as



a working example of a comprehensive approach for delimitating recently diverged species and elaborate on the individual components of this approach below.

The resolution of a large number of geographically defined clades in the mtDNA gene tree (Fig. 2) provided a strong suggestion for a high diversity of mouse lemur lineages. Many mtDNA clades correspond to groups resolved in previous mtDNA-based studies [24,25,26,37]. This is in large part due to our inclusion of GenBank mtDNA data from these previous studies, which facilitated the placement of much of our nuclear results within the context of previous species delimitation work. Novel to this study, however, was the resolution of additional mtDNA clades resolved on the basis of two different sampling regimes. First, expanded field sampling led to the resolution of mtDNA clades within described species that are specific to one or a few sampling localities. For example, within the *M. murinus* mtDNA clade the localities of Bemanasy and Mandena were each resolved as monophyletic sets of haplotypes (Fig. 2). This fine scale pattern of mtDNA differentiation was also seen in paraphyletic groupings within described species. For example, both *M. griseorufus* and *M. myoxinus* contained paraphyletic groups with respect to other single-locality monophyletic groupings (Table 2). Both of these phylogenetic patterns represent a potentially fine-scale level of geographic genetic divergence not previously considered within mouse lemurs. Second, we also resolved three novel and relatively divergent mtDNA clades (each labeled *Microcebus* sp. in Fig. 2), each of which corresponds to new localities that had not previously been sampled. These patterns indicate that the upper limit of our understanding of mouse lemur lineage diversity may also be constrained by our ability to comprehensively sample across Madagascar.

Much of the recent flurry of mouse lemur species descriptions has been based primarily on patterns in mtDNA gene trees. While the sorting of mtDNA variation may, on average, track lineage divergence more rapidly than nDNA [14], there are important reasons for not relying on mtDNA as a sole source of evidence in species delimitation, including the potential to over diagnose mtDNA lineages influenced by selection [15] or sex-biased dispersal [57]. Accordingly, the strongest genetic evidence for speciation comes from evaluating patterns of variation across multiple loci [15,58]. We find additional evidence for mouse lemur lineage divergence in the striking degree of concordance between mtDNA clades and patterns of nuclear population genetic structure. Localities or sets of localities that are resolved as a clade in the mtDNA gene tree are predominantly assigned to a characteristic nuclear genotypic cluster with high average assignment probabilities (0.9–1.0) in the STRUCTURE assignment plot (Fig. 2, Table 2). There are only limited signs of admixture between some population clusters, a pattern consistent with mtDNA evidence for a lack of gene flow between groups. In contrast to the evolution of nuclear monophyly, which can require long durations of time [ $\sim 4\text{--}7N$  generations for just 50% of the nuclear genome's gene trees [15]], the evolution of allele frequency differences and genome-wide patterns of Hardy-Weinberg and linkage equilibrium within lineages is expected to occur more rapidly. Population assignment analyses based on these parameters may serve as useful methods for delimiting the early stages of lineage divergence [18,38].

One potential criticism of our interpretation of the STRUCTURE results is that these patterns simply represent the geographic structuring of intraspecific variation. For example, STRUCTURE analysis places human population genetic variation into geographically-defined clusters [59]. A major benefit to the use of STRUCTURE analysis is its ability to identify

individuals with admixed genomic profiles that arise through gene flow between previously isolated population clusters [51]. Human populations are undoubtedly cohesive through gene flow, as evidenced, in part, by signatures of admixture in their STRUCTURE profiles. However, this does not negate STRUCTURE's ability to identify cluster patterns that reflect *histories* of isolation and divergence. In fact, the models implemented in STRUCTURE specifically allow for the reconstruction of ancestral clusters even if most individuals are largely composed of admixed genotypes [51]. Consequently, we contend that, within natural systems, populations characterized by distinct genotypic clusters with limited signs of admixture exhibit at least one layer of evidence for lineage divergence.

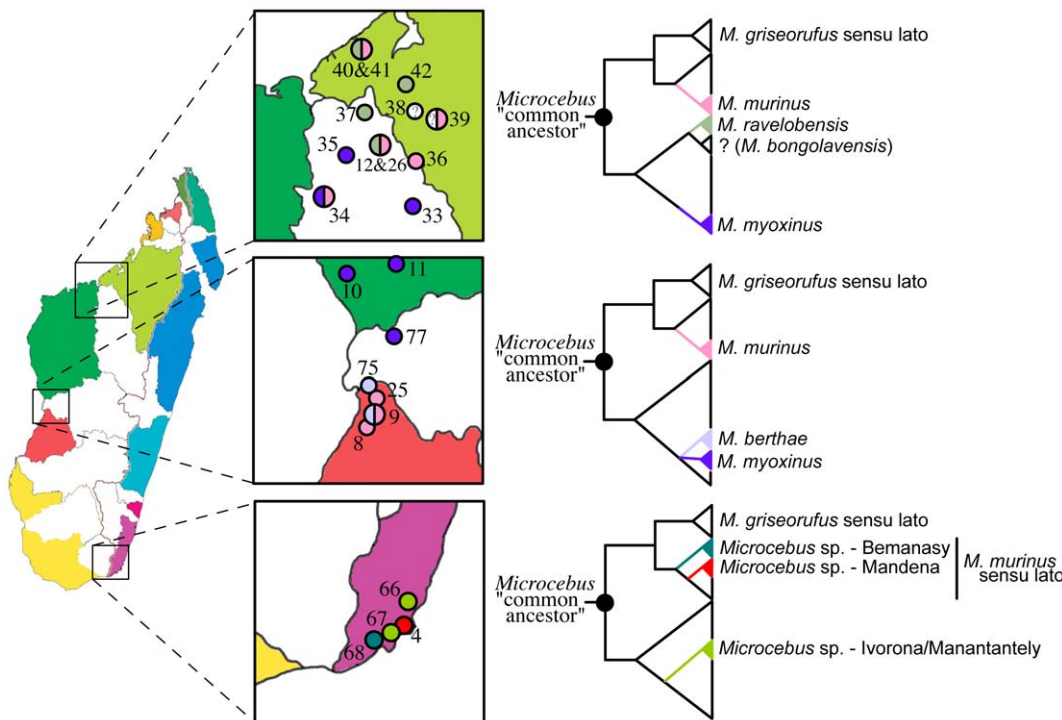
To further test the hypothesis that these sets of mouse lemur populations have an underlying history of lineage divergence, we quantified the magnitude of their genealogical divergence and tested whether or not genealogical patterns within lineages were significantly different from those expected under a history of no divergence. Accompanying the speciation process is an expected transition, driven by genetic drift, in the gene genealogies of a diverging lineage from polyphyletic sets of ancestral gene copies to monophyletic sets of unique gene copies [12,33]. Consequently, characteristic topological patterns are expected to evolve in the gene trees of a diverging lineage long before the evolution of monophyly, and these patterns can be used to distinguish independent and recently diverged sets of populations [16,17]. All mouse lemur lineages delimited via mtDNA monophyly and nuclear clustering have significant patterns of genealogical exclusivity in their nuclear gene trees, as measured by *gsi* values for the Bayesian posterior distributions of individual gene trees, and by the ensemble *gsi<sub>T</sub>*, which integrates patterns across loci (Table 3). These results clarify our interpretations of the nuclear STRUCTURE clusters by providing a genealogical perspective of their underlying history. Despite the fact that the majority of nuclear clusters do not represent monophyletic groups in their component nuclear gene trees, they nonetheless have genealogical patterns consistent with a history of lineage divergence. *Gsi<sub>T</sub>* values across loci indicate a broad range in the degree of genealogical exclusivity across loci (0.069–1.0). However, because the progression to monophyly is governed by both time and effective population size [33,55], it is difficult to place lineages on a relative temporal scale of divergence based on their overall *Gsi<sub>T</sub>*. Instead, the resolution of lineages with significant gene tree-wide *Gsi<sub>T</sub>* values provides a simple, but quantitative indication of divergence for many lineages that have not been diverging long enough to have evolved concordant patterns of monophyly across the majority of their gene trees.

Based on this total set of evidence, we diagnose 16 population-level lineages of mouse lemurs (Tables 2, 3). Within the framework of the metapopulation lineage concept of species [1,3], all of these lineages are recognized as species, regardless of their relative level of divergence. In contrast, species criteria that place greater emphases on more explicit properties of species (e.g. reproductive isolation or gene tree monophyly) would yield a substantially lower number of delimited and recognized species. Without placing emphasis on the number of species formally recognized by our results, we highlight the strong evidence for considerable lineage formation within *Microcebus*, including lineages with genetic patterns indicative of long durations of independent divergence, as well as those with patterns suggestive of a shallower depth of divergence. In total, these results are counter to a recently proposed argument that the escalation in lemur species diversity is largely due to oversplitting of geographically localized variants that are merely components of larger cohesive lineages [30]. Instead,

our estimates of a high lineage diversity of mouse lemurs is reflected in both a geographically fine-scale pattern of population divergence within established and geographically widespread species, as well as newly resolved patterns of micro-endemism revealed through expanded field sampling into previously poorly and well-sampled regions. Fine-scale patterns of lineage divergence are particularly pronounced within *M. murinus* in the southeastern populations of Bemanasy and Mandena, which are genetically diverged and distinct from each other and from western *M. murinus* populations. The geographic scale of divergence between these two southeastern lineages within *M. murinus* across a pronounced bioclimatic ecotone [60] is striking, with a physical separation of just 27 km (Fig. 3), but genetic patterns that seem to exclude any sign of genetic exchange. Moreover, new regions of micro-endemic lineages are revealed. From the newly-sampled Marolambo population (locality 64) on the central eastern side of the island, a novel lineage was discovered that displays no evidence of gene flow with *M. lehilahytsara* or *M. simmonsii* to the north, or *M. rufus* to the south (Fig. 2). In the well-sampled regions of Ivorona and Manantantely (localities 66 and 67) in the extreme southeast, we observe sympatric distributions of three distinct lineages (Fig. 3) that show no evidence of gene flow, despite the fact that these populations are separated by no more than 10 km. Yet, another novel lineage is detected in the northern populations of Ambanja (locality 79) and Montagne d'Ambre (locality 56). This latter lineage may equate with a recently described species from Montagne d'Ambre,

*M. amholdi*, which was described on the basis of mtDNA divergence from other northern populations [29]. However, the lack of overlapping mtDNA regions across studies limits a direct connection between the lineage diagnosed here and *M. amholdi*. Overall, these patterns of lineage divergence suggest that mouse lemur speciation can occur at a broad range of geographic levels, and indeed, indicates the probability of even more as-yet-undescribed lineage diversity. A number of additional species have been described in recent years that we cannot address with the nuclear data in this study [25,26,27,29] and future work will be needed to assess the validity of these described taxa as independent lineages.

In contrast, the analytical approach and criteria used here also identified patterns reflecting the signal of population divergence, but without straightforward evidence for independent evolution. This is best exemplified with a set of three previously described species: *M. berthae* and *M. myoxinus* in the western part of the island, and *M. rufus* in the east. Each is distinct in the mtDNA gene tree (*M. rufus* is paraphyletic), is largely comprised of a characteristic nuclear cluster in STRUCTURE analyses (Fig. 2), and has significant *gsi* estimates. However, while individuals within *M. berthae* and *M. myoxinus* are predominantly assigned to distinct nuclear clusters, the *M. rufus* population of Ranomafana has an average of 12% of its overall membership coefficients attributed to the *M. berthae* cluster and 9% attributed to *M. myoxinus* clusters (Table 2). As a comparison, no other species diagnosed in this study had an average membership coefficient of >5% from any



**Figure 3. Phylogenetic descriptions of areas of sympatry and fine-scale allopatry among *Microcebus* lineages.** Colored sections in the maps of Madagascar signify areas of micro-endemism and white sections signify retreat dispersion regions [21]. All known regions of sympatry in mouse lemurs involve populations of *M. murinus* sensu lato. The top two outlined boxes highlight regions of sympatry involving *M. murinus* and at least three other species (*M. bongolavensis* could not be assessed in this study), each of which shares a most recent common ancestor with *M. murinus* at the root of most gene trees. However, at least one case of sympatry is known that involves *M. murinus* and its sister lineage, *M. griseorufus*, in the Berenty Reserve in southern Madagascar [not sampled in this study and not depicted here [69]. As further support for the speciation model of Wilmé et al. (2006), areas of sympatry appear to be clustered in or near retreat dispersion regions, which are proposed to have influenced the expansion of species ranges during the Quaternary. In contrast, the bottom outlined box highlights lineage distributions within a single region of micro-endemism in the southeast.

doi:10.1371/journal.pone.0009883.g003

other species' nuclear cluster. These patterns are suggestive of gene flow into the range of *M. rufus*. In such cases, determining the evolutionary independence of these potential lineages will require additional sources of information. In this particular case, geography is particularly informative: the western lineages (*M. berthae* and *M. myoxinus*) are highly genetically distinct from each other (Fig. 2, Table 2), but are extremely close geographically (Figs. 1, 3). In contrast, both are separated from *M. rufus* by a high elevation north-south mountain system (Fig. 1). These geographic patterns indicate that if gene flow is the cause of the slightly mixed cluster assignment for *M. rufus*, then it is historical in nature, and currently all three lineages (*M. berthae*, *M. myoxinus*, and *M. rufus*) are genetically isolated. As such, these patterns suggest that all three are diverging population-level lineages.

Finally, we not only identify sets of independently evolving lineages with this approach, but also identify cohesive sets of populations, particularly in situations where mtDNA distinctiveness would suggest divergence. This is best exemplified with the recently described *M. mamiatra* from northwest Madagascar [24,26]. Our results find some individuals from Ambanja (locality 79) to contain mtDNA haplotypes placed within the *M. mamiatra* clade, which nonetheless have a nuclear structure profile indistinguishable from individuals of *M. sambiranensis* (Fig. 2). We therefore propose that the diagnosis of *M. mamiatra* is based on insufficient structure in mtDNA variation and should not be recognized. Similar patterns exist within the currently recognized taxa *M. murinus*, *M. myoxinus*, and *M. simonsi*. Within each of these taxa, subsets of populations exhibit monophyly in the mtDNA gene tree (Fig. 2), suggestive of lineage divergence; yet, these subpopulations also exhibit nuclear STRUCTURE profiles that indicate an admixed background of nuclear gene flow and cohesion with other populations (Fig. 2). These contrasting patterns between the mtDNA and nuclear genetic data may be related to the propensity for female philopatry in lemurs [34,35] and highlight the potential pitfalls of relying solely on mtDNA gene tree patterns in the delimitation of mouse lemur species.

### Mouse lemur evolution

What do these results have to say about the mechanism of speciation in *Microcebus*, and about lemur diversification in general? The isolated geographic distributions of most mouse lemur species (Fig. 2) indicate a strong role for allopatric speciation. Furthermore, while we are currently limited in our ability to infer a robust species tree and divergence time estimates for mouse lemur lineages, the predominant lack of reciprocal monophyly across nuclear gene trees (Figures S2, S3, S4, S5) suggests relatively recent divergence times for most. These two inferences align with a recent model for the origin of micro-endemic regions in Madagascar, which places emphasis on reduced precipitation during the Quaternary and formation of forest refugia in lower elevation river catchments around the periphery of the island [21]. A number of mouse lemur lineages lie within, or mostly within, these areas of micro-endemism (Fig. 1), indicating that this model may account for diversification at some level. Additional diversification models developed for Madagascar, including the roles of climate [61], rivers [19], and mountains [62] also remain to be tested with a robust estimate of the mouse lemur species tree [63].

Also of interest is the paralleled high level of recently recognized species diversity in other nocturnal lemur genera (e.g. *Avahi*, *Lepilemur*), suggesting a link between nocturnal activity and unrealized population genetic structure. An important behavioral mechanism in nocturnal lemurs is acoustic signaling [64], a necessary form of communication in the dark. Mouse lemur

species in particular have evolved distinctly different male advertisement calls [65], providing a potential mechanism for premating isolation that can limit admixture among populations that have previously experienced allopatric divergence. Contrastingly, however, a study of the closely related and nocturnal dwarf lemurs (*Chetrogaleus*) using the same markers and some of the methods applied here finds a disproportionately low level of taxonomic diversity [3 diagnosable species as compared to 7 proposed in recent years [66]], despite having a similar island-wide distribution [67].

We provide here the most comprehensive species delimitation study of lemurs ever performed and are the first to integrate gene tree analysis, Bayesian STRUCTURE analysis, and quantifications of genealogical divergence in the identification of population-level lineages. While our results confirm the high lineage diversity results of previous mtDNA-based studies, this should not serve as justification for limiting future studies to a single marker in a DNA-barcoding protocol. Instead, our results regarding the lack of corresponding nuclear evidence of divergence for some resolved mtDNA lineages and the under appreciation of fine-scale geographic patterns of lineage divergence not previously highlighted in mtDNA studies emphasizes the need for a more thorough approach utilizing appropriate genetic data and analytical methods. The approach outlined here may also serve as a general model for lineage diagnosis of other groups with a predominant pattern of cryptic diversity. For example, recent surveys of Malagasy frogs found an additional 129 mtDNA lineages that either lack clear morphological or acoustic differences, or that have not yet been studied with independent data that can be used to test their distinctiveness [68]. Testing these lineages with the population genetic and genealogical framework used here may provide an ideal alternative to further field and museum-based assessments of species delimitation.

### Supporting Information

**Table S1** Locality and sampling information for all *Microcebus* localities used in this study.

Found at: doi:10.1371/journal.pone.0009883.s001 (0.19 MB DOC)

**Table S2** GenBank accession numbers for all outgroup sequence data and the *Microcebus* mtDNA sequence data of Louis et al. [24] and Olivieri et al. [25].

Found at: doi:10.1371/journal.pone.0009883.s002 (0.12 MB DOC)

**Figure S1** MtDNA gene tree. The tree results from Bayesian phylogenetic analysis of the mtDNA haplotype data set and is presented as the maximum credible topology with branch lengths averaged across the posterior distribution (Mean  $-\ln L = 1.33 \times 10^4$ , Std. Dev. = 7.76). Tip labels include a species name if the haplotype was sampled from an individual identified to a species in a previous study. Haplotypes recovered from newly sampled individuals are indicated with the locality name.

Found at: doi:10.1371/journal.pone.0009883.s003 (1.84 MB TIF)

**Figure S2** *adora3* gene tree. The tree results from Bayesian phylogenetic analysis of the *adora3* haplotype data set and is presented as the maximum credible topology with branch lengths averaged across the posterior distribution (Mean  $-\ln L = 1104.67$ , 95% HPD = 1118.35-1093.08). Tip labels include a species name if the haplotype was sampled from an individual identified to a species in a previous study. Haplotypes recovered from newly sampled individuals are indicated with the locality name.

Found at: doi:10.1371/journal.pone.0009883.s004 (1.54 MB TIF)

**Figure S3** *emo* gene tree. The tree results from Bayesian phylogenetic analysis of the *emo* haplotype data set and is presented as the maximum credible topology with branch lengths averaged across the posterior distribution (Mean  $-\ln L = 4557.6$ , 95%HPD = 4582.15–4531.58). Tip labels include a species name if the haplotype was sampled from an individual identified to a species in a previous study. Haplotypes recovered from newly sampled individuals are indicated with the locality name.

Found at: doi:10.1371/journal.pone.0009883.s005 (1.74 MB TIF)

**Figure S4** *fga* gene tree. The tree results from Bayesian phylogenetic analysis of the *fga* haplotype data set and is presented as the maximum credible topology with branch lengths averaged across the posterior distribution (Mean  $-\ln L = 2232.92$ , 95%HPD = 2251.07–2215.76). Tip labels include a species name if the haplotype was sampled from an individual identified to a species in a previous study. Haplotypes recovered from newly sampled individuals are indicated with the locality name.

Found at: doi:10.1371/journal.pone.0009883.s006 (1.84 MB TIF)

**Figure S5** *wvf* gene tree. The tree results from Bayesian phylogenetic analysis of the *wvf* haplotype data set and is presented as the maximum credible topology with branch lengths averaged across the posterior distribution (Mean  $-\ln L = 3980.14$ , 95%HPD = 4007.55–3953.37). Tip labels include a species name if the haplotype was sampled from an individual identified to a species in a previous study. Haplotypes recovered from newly sampled individuals are indicated with the locality name.

## References

- de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH, eds. *Endless forms: species and speciation*. Oxford: Oxford University Press. pp 57–75.
- Mayden RL (1997) The denouement in the saga of the species problem. In: Claridge MF, Dawah HA, Wilson MR, eds. *Species: The units of biodiversity*. London: Chapman and Hall. pp 381–424.
- de Queiroz K (1999) The general lineage concept of species and the defining properties of the species category. In: Wilson RA, ed. *Species: New Interdisciplinary Essays*. Cambridge, MA: MIT Press. pp 49–89.
- Mayden RL (1999) Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. *J Nematol* 31: 95–116.
- de Queiroz K (2007) Species concepts and species delimitation. *Syst Biol* 56: 879–886.
- Wiens JJ (2004) What is speciation and how should we study it? *Am Nat* 163: 914–923.
- Wiens JJ (2007) Species delimitation: new approaches for discovering diversity. *Syst Biol* 56: 875–878.
- Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat Rev Genet* 3: 380–390.
- Degnan JH, Salter LA (2005) Gene tree distributions under the coalescent process. *Evolution* 59: 24–37.
- Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, et al. (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol* 55: 595–609.
- Baum DA, Shaw KL (1995) Genealogical perspectives on the species problem. In: Hoch PC, Stephenson AG, eds. *Experimental and Molecular Approaches to Plant Biosystematics*. St. Louis: Missouri Botanical Garden. pp 289–303.
- Avice JC, Ball RM (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxf Surv Evol Biol* 7: 45–67.
- Hebert PDN, Cywinka A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *P Roy Soc Lond B Bio* 270: 313–321.
- Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear gene trees. *Evolution* 49: 718–726.
- Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution* 56: 1557–1565.
- Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Syst Biol* 56: 887–895.
- Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution* 62: 2411–2422.
- Shaffer HB, Thomson RC (2007) Delimiting species in recent radiations. *Syst Biol* 56: 896–906.
- Martin RD (1972) Adaptive radiation and behavior of Malagasy lemurs. *Philos Trans R Soc Lond, B, Biol Sci* 264: 295–352.
- Thalmann U (2007) Biodiversity, phylogeography, biogeography and conservation: lemurs as an example. *Folia Primatol* 78: 420–443.
- Wilmé L, Goodman SM, Ganzhorn JU (2006) Biogeographic evolution of Madagascar's microendemic biota. *Science* 312: 1063–1065.
- Harper GJ, Steininger MK, Tucker CJ, Juhn D, Hawkins F (2007) Fifty years of deforestation and forest fragmentation in Madagascar. *Environ Conserv* 34: 325–333.
- Kappeler PM, Rasoloarison RM, Razafimanantsoa L, Walter L, Roos C (2005) Morphology, behaviour and molecular evolution of giant mouse lemurs (*Mirza* spp.) Gray, 1870, with description of a new species. *Primate Rep* 71: 3–26.
- Andriantompohavana R, Zaonarivelo JR, Engberg SE, Randriamampionona R, McGuire SM, et al. (2006) Mouse lemurs of northwestern Madagascar with a description of a new species at Lokobe Special Reserve. *Occas Pap Tex Tech Univ Mus* 259: 1–23.
- Louis EE, Coles MS, Andriantompohavana R, Sommer JA, Engberg SE, et al. (2006) Revision of the mouse lemurs (*Microcebus*) of eastern Madagascar. *Int J Primatol* 27: 347–389.
- Olivieri G, Zimmermann E, Randrianambinina B, Rasoloharijaona S, Rakotondravony D, et al. (2007) The ever-increasing diversity in mouse lemurs: three new species in north and northwestern Madagascar. *Mol Phylogenet Evol* 43: 309–327.
- Radespiel U, Olivieri G, Rasolofoson DW, Rakotondratsimba G, Rakotonirainy O, et al. (2008) Exceptional diversity of mouse lemurs (*Microcebus* spp.) in the Makira region with the description of one new species. *Am J Primatol* 70: 1–14.
- Rasoloarison RM, Goodman SM, Ganzhorn JU (2000) Taxonomic revision of mouse lemurs (*Microcebus*) in the western portions of Madagascar. *Int J Primatol* 21: 963–1019.
- Louis EE, Engberg SE, McGuire S, McCormick MJ, Randriamampionona R, et al. (2008) Revision of the mouse lemurs, *Microcebus* (Primates, Lemuriformes), of northern and northwestern Madagascar with descriptions of two new species at Montagne d'Ambre National Park and Antafondro Classified Forest. *Primate Conservation* 23: 19–38.
- Tattersall I (2007) Madagascar's lemurs: cryptic diversity or taxonomic inflation? *Evol Anthropol* 16: 12–23.
- Weisrock DW, Shaffer HB, Storz BL, Storz SR, Voss SR (2006) Multiple nuclear gene sequences identify phylogenetic species boundaries in the rapidly radiating clade of Mexican ambystomatid salamanders. *Mol Ecol* 15: 2489–2503.
- Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu Rev Ecol Syst* 34: 397–423.
- Maddison WP (1997) Gene trees in species trees. *Syst Biol* 46: 523–536.

34. Fredsted T, Pertoldi C, Schierup MH, Kappeler PM (2005) Microsatellite analyses reveal fine-scale genetic structure in grey mouse lemurs (*Microcebus murinus*). *Mol Ecol* 14: 2363–2372.
35. Wimmer B, Tautz D, Kappeler PM (2002) The genetic population structure of the gray mouse lemur (*Microcebus murinus*), a basal primate from Madagascar. *Behav Ecol Sociobiol* 52: 166–175.
36. Allnutt T, Ferrier S, Manion G, Powell G, Ricketts T, et al. (2008) A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conservation Letters* 1: 173–181.
37. Yoder AD, Rasoloarison RM, Goodman SM, Irwin JA, Atsalis S, et al. (2000) Remarkable species diversity in Malagasy mouse lemurs (primates, *Microcebus*). *Proc Natl Acad Sci USA* 97: 11325–11330.
38. Dumnil J, Caron H, Scotti I, Cazal SO, Petit RJ (2006) Blind population genetics survey of tropical rainforest trees. *Mol Ecol* 15: 3505–3513.
39. Heckman KL, Mariani CL, Rasoloarison R, Yoder AD (2007) Multiple nuclear loci reveal patterns of incomplete lineage sorting and complex species history within western mouse lemurs (*Microcebus*). *Mol Phylogenet Evol* 43: 353–367.
40. Adkins RM, Honeycutt RL (1994) Evolution of the primate Cytochrome *c* Oxidase Subunit II gene. *J Mol Evol* 38: 215–231.
41. Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the Cytochrome *b* gene of mammals. *J Mol Evol* 32: 128–144.
42. Horvath JE, Weisrock DW, Embry SL, Fiorentino I, Balhoff JP, et al. (2008) Development and application of a phylogenomic toolkit: resolving the evolutionary history of Madagascar's lemurs. *Genome Res* 18: 489–499.
43. Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68: 978–989.
44. Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111: 147–164.
45. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
46. Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
47. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
48. Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
49. Swofford DL (2002) PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). 4 ed. SunderlandMassachusetts: Sinauer Associates.
50. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
51. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
52. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611–2620.
53. Andreone F, Carpenter AI, Cox N, du Preez L, Freeman K, et al. (2008) The challenge of conserving amphibian megadiversity in Madagascar. *PLoS Biol* 6: 943–946.
54. Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* 236: 787–792.
55. Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Mol Biol Evol* 5: 568–583.
56. Yoder AD, Olson LE, Hanley C, Heckman KL, Rasoloarison R, et al. (2005) A multidimensional approach for detecting species patterns in Malagasy vertebrates. *Proc Natl Acad Sci USA* 102: 6587–6594.
57. Caparroz R, Miyaki CY, Baker AJ (2009) Contrasting phylogeographic patterns in mitochondrial DNA and microsatellites: evidence of female philopatry and male-biased gene flow among regional populations of the Blue-and-yellow Macaw (*Psittaciformes: Ara ararauna*) in Brazil. *Auk* 126: 359–370.
58. Hudson RR, Turelli M (2003) Stochasticity overrules the “three-times rule”: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* 57: 182–190.
59. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, et al. (2002) Genetic structure of human populations. *Science* 298: 2381–2385.
60. Goodman SM, Pidgeon M, Hawkins AFA, Schulenberg TS (1997) The birds of southeastern Madagascar. *Fieldiana* 87: 1–132.
61. Pearson RG, Raxworthy CJ (2009) The evolution of local endemism in Madagascar: watershed versus climatic gradient hypotheses evaluated by null biogeographic models. *Evolution* 63: 959–967.
62. Wollenberg KC, Vieites DR, van der Meijden A, Glaw F, Cannatella DC, et al. (2008) Patterns of endemism and species richness in Malagasy cophyline frogs support a key role of mountainous areas for speciation. *Evolution* 62: 1890–1907.
63. Vences M, Wollenberg KC, Vieites D, Lees DC (2009) Madagascar as a model region of species diversification. *Trends Ecol Evol* 24: 456–465.
64. Fichtel C (2007) Avoiding predators at night: antipredator strategies in red-tailed sportive lemurs (*Lepilemur ruficaudatus*). *Am J Primatol* 69: 611–624.
65. Braune P, Schmidt S, Zimmermann E (2008) Acoustic divergence in the communication of cryptic species of nocturnal primates (*Microcebus* spp.). *BMC Biol* 6: 19.
66. Groves CP (2000) The genus *Cheirogaleus*: unrecognized biodiversity in dwarf lemurs. *Int J Primatol* 21: 943–962.
67. Groeneveld LF, Weisrock DW, Rasoloarison RM, Yoder AD, Kappeler PM (2009) Species delimitation in lemurs: multiple genetic loci reveal low levels of species diversity in the genus *Cheirogaleus*. *BMC Evol Biol* 9: 30.
68. Vieites D, Wollenberg KC, Andreone F, Koehler J, Glaw F, et al. (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proc Natl Acad Sci USA* 106: 8267–8272.
69. Yoder AD, Burns MM, Genin F (2002) Molecular evidence of reproductive isolation in sympatric sibling species of mouse lemurs. *Int J Primatol* 23: 1335–1343.