PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Check for updates

Cite this article: Poelstra J, Montero BK, Lüdemann J, Yang Z, Rakotondranary SJ, Hohenlohe P, Stetter N, Ganzhorn JU, Yoder AD. 2022 RADseq data reveal a lack of admixture in a mouse lemur contact zone contrary to previous microsatellite results. *Proc. R. Soc. B* 20220596. https://doi.org/10.1098/rspb.2022.0596

Received: 12 April 2022
Accepted: 11 July 2022

Subject Category:
Evolution

Subject Areas:
evolution

Keywords:
RADseq, admixture, cryptic species, mouse lemurs, Madagascar, microsatellites

Author for correspondence:
Anne D. Yoder
e-mail: anne.yoder@duke.edu

RADseq data reveal a lack of admixture in a mouse lemur contact zone contrary to previous microsatellite results

Jelmer Poelstra[1,2], B. Karina Montero[3], Jan Lüdemann[3], Ziheng Yang[4], S. Jacques Rakotondranary[3,5], Paul Hohenlohe[6], Nadine Stetter[3,7], Jörg U. Ganzhorn[3,†] and Anne D. Yoder[1,†]

[1]Department of Biology, Duke University, Durham, NC 27708, USA
[2]Molecular and Cellular Imaging Center, Ohio State University, Wooster, OH 44691, USA
[3]Institute of Zoology, Department Animal Ecology and Conservation, Universität Hamburg, Hamburg, 20146, Germany
[4]Department of Genetics, Evolution and Environment, University College London, London, UK
[5]Département Biologie Animale, Faculté des Sciences, Université d'Antananarivo, P.O. Box 906, Antananarivo 101, Madagascar
[6]Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA
[7]Bernhard Nocht Institute for Tropical Medicine, 20359 Hamburg, Germany

ZY, 0000-0003-3351-7981; PH, 0000-0002-7616-0161; ADY, 0000-0002-1781-9552

Microsatellites have been a workhorse of evolutionary genetic studies for decades and are still commonly in use for estimating signatures of genetic diversity at the population and species level across a multitude of taxa. Yet, the very high mutation rate of these loci is a double-edged sword, conferring great sensitivity at shallow levels of analysis (e.g. paternity analysis) but yielding considerable uncertainty for deeper evolutionary comparisons. For the present study, we used reduced representation genome-wide data (RADseq) to test for patterns of interspecific hybridization previously characterized using microsatellite data in a contact zone between two closely related mouse lemur species in Madagascar (*Microcebus murinus* and *M. griseorufus*). We revisit this system by examining populations in, near, and far from the contact zone, including many of the same individuals that had previously been identified as hybrids with microsatellite data. Surprisingly, we find no evidence for admixed nuclear ancestry. Instead, re-analyses of microsatellite data and simulations suggest that previously inferred hybrids were false positives and that the program `NewHybrids` can be particularly sensitive to erroneously inferring hybrid ancestry. Combined with results from coalescent-based analyses and evidence for local syntopic co-occurrence, we conclude that the two mouse lemur species are in fact completely reproductively isolated, thus providing a new understanding of the evolutionary rate whereby reproductive isolation can be achieved in a primate.

1. Introduction

Microsatellites are tandem repeats of repetitive DNA that typically range in length from one to six nucleotides and occur at thousands of locations within the genomes of most organisms [1,2]. Individual microsatellite loci contain from as few as five to as many as 40 or more repeats, with copy number changes caused by slip-strand mispairing during DNA replication. Mutation rates for microsatellites are orders of magnitude higher than for other types of variants, including SNPs, with the overall rate being a balance between the generation of replication errors and the correction of errors by proofreading and mismatch repair, all of which can vary by species [3]. Given their high rate of change,

[†]Co-senior authors.

Electronic supplementary material is available online at rs.figshare.com.

THE ROYAL SOCIETY PUBLISHING

microsatellite loci have high allelic richness, often in excess of 10 alleles within humans and other primates [4]. This rich allelic diversity, combined with relatively low genotyping costs, have made microsatellites a popular genetic marker for applications ranging from paternity analysis to historical demography. In particular, they have proven useful for identifying conservation units in endangered species (e.g. [5]) as well as for revealing the presence of homoploid hybrid speciation (e.g. [6]).

Yet, their extreme sensitivity can also be cause for concern. The high rate of recurrent mutations (i.e. homoplasy) makes them poor indicators of long-term population history [2,4]. For example, the combination of homoplasy and potentially inappropriate models of mutational dynamics can yield highly inflated estimates of gene flow between populations and species [7,8]. Thus, inferences above all but the shallowest evolutionary levels should be treated with caution.

In this study, we revisit hypotheses of hybridization between two named species of mouse lemur, *Microcebus murinus* (*sensu lato*) and *M. griseorufus*, reported from previous studies using microsatellite data [9–11]. These previous studies focused on two contact zones in the southeast of Madagascar wherein hybrids were reported to occur.

To date, seven different pairs of mouse lemur species have been shown to co-occur locally at various localities throughout Madagascar. One widespread species, *M. murinus*, is involved in five of these cases. In all but one of these seven cases of sympatry, no hybridization has been detected thus suggesting that co-occurring species are reproductively isolated. Sources of reproductive isolation among sympatric mouse lemurs are poorly known, but factors that may contribute to prezygotic isolation via differential mate choice may include divergence in acoustic [12,13] and olfactory signalling [14,15]. Additionally, opportunities for reproductive interaction may be reduced by ecological divergence manifesting, for example, in differential timing of the highly seasonal and temporally constrained reproductive season seen in mouse lemurs [16–18].

It is thus intriguing that hybridization has only been detected between *M. murinus* and *M. griseorufus*, using microsatellite loci [9,10], which is also unique among the seven cases of sympatry in consisting of a pair of sister lineages. Using the programs STRUCTURE [19] and GeneClass [20], Gligor *et al.* [9], p. 529) concluded that 'most individuals within the transition zone' had mixed ancestry (no individual-level assignments were made). Hapke *et al.* [10] studied a contact zone 40 km further north, and used the same set of microsatellite loci for a total of 159 mouse lemurs, with STRUCTURE and NewHybrids [21] identifying a total of 18 admixed individuals. Of these, 15 individuals showed signs of nuclear admixture (i.e. among microsatellites) whereas 3 had a mismatch between microsatellite and mitochondrial ancestry.

Here, we use RADseq data to revisit the contact zone area studied by Hapke *et al.* [10] and follow-up work in Lüdemann [11] that used the same microsatellites and methods. We have included a total of 130 individuals, including 18 of the individuals that were inferred to be hybrids by these studies in addition to samples from nearby and distant allopatric populations. To ensure that non-admixed individuals from parental species were present, as is critical for accurately identifying either the presence or absence of hybrids [22], we also include samples from nearby and distant allopatric populations. We examine individual-level admixture in the northern contact zone and used coalescent modelling to ask whether there is evidence for ongoing and/or ancestral gene flow between the species. To our surprise, we found no evidence for admixed individuals in the contact zone—including among the individuals previously identified as hybrids—and also infer a lack of ongoing gene flow between the two species more generally.

2. Methods

(a) Sampling

Hapke *et al.* [10] and follow-up work in Lüdemann [11] detected hybridization between *M. murinus* (hereafter referred to as *murinus*) and *M. griseorufus* (hereafter referred to as *griseorufus*) using 9 microsatellites and a fragment of the HV1 mitochondrial locus from individuals in the Andohahela area in southeastern Madagascar. We made use of a selection of 94 of their samples and augmented this dataset with 33 samples from distant, allopatric sites, and 3 *M. rufus* samples that were used as an outgroup (electronic supplementary material, table S1, table S2).

At two of the sites examined by Hapke *et al.* [10], they detected unadmixed individuals of both parental species as well as individuals with admixed ancestry (individuals inferred to be admixed by Hapke *et al.* [10] and Lüdemann [11] are hereafter referred to as 'putative hybrids'). From these two contact zone sites, Mangatsiaka and Tsimelahy, which we refer to as 'sympatric' sites, we selected 78 samples (electronic supplementary material, table S1), including 15 individuals for which Hapke *et al.* [10] or Lüdemann [11] had detected nuclear admixture, and an additional 3 with a mitonuclear ancestry mismatch. We additionally selected samples from nearby sites at which Hapke *et al.* [10] had exclusively (or nearly so) detected unadmixed individuals of only one of the two species: 8 *griseorufus* from Hazofotsy and 8 *murinus* from Ambatoabo (electronic supplementary material, table S1). We refer to these contact zone sites as 'parapatric' sites. 'Allopatric' samples, taken well away from the contact zone, were represented by 14 *griseorufus*, 8 *murinus* and 11 *M. ganzhorni*, a species that was recently split from *murinus* [23], from Mandena in far southeastern Madagascar (electronic supplementary material, table S2, figure 1). Below, we show that *M. ganzhorni* diverged very recently from the Andohahela area *murinus* populations, while a much deeper split occurs between western and other southeastern Madagascar populations, all of which continue to be classified as *murinus.* Therefore, we here include *M. ganzhorni* under the nomer '*M. murinus s.l.*'.

We used the following geographically defined population groupings for analyses where individuals are assigned to predefined groups (figure 1): western *griseorufus* (abbreviated '*gri*-W'), central/contact zone area *griseorufus* (abbreviated '*gri*-C'), western *murinus* (abbreviated '*mur*-W'), central/contact zone area *murinus* (abbreviated '*mur*-C') and eastern *murinus s.l.* (abbreviated '*mur*-E'; this population corresponds to *M. ganzhorni sensu* Hotaling *et al.* [23]).

(b) Sequencing and genotyping

We prepared Restriction-site Associated DNA (RAD) sequencing libraries following the protocol of Ali *et al.* [24]. Libraries were sequenced using paired-end 150 bp sequencing on an Illumina HiSeq 4000 at Duke University's Center for Genomic and Computational Biology sequencing facility.

After read flipping, demultiplexing, trimming and mapping to the *M. murinus* reference genome ('Mmurinus 3.0', [25]), we

royalsocietypublishing.org/journal/rspb Proc. R. Soc. B 20220596

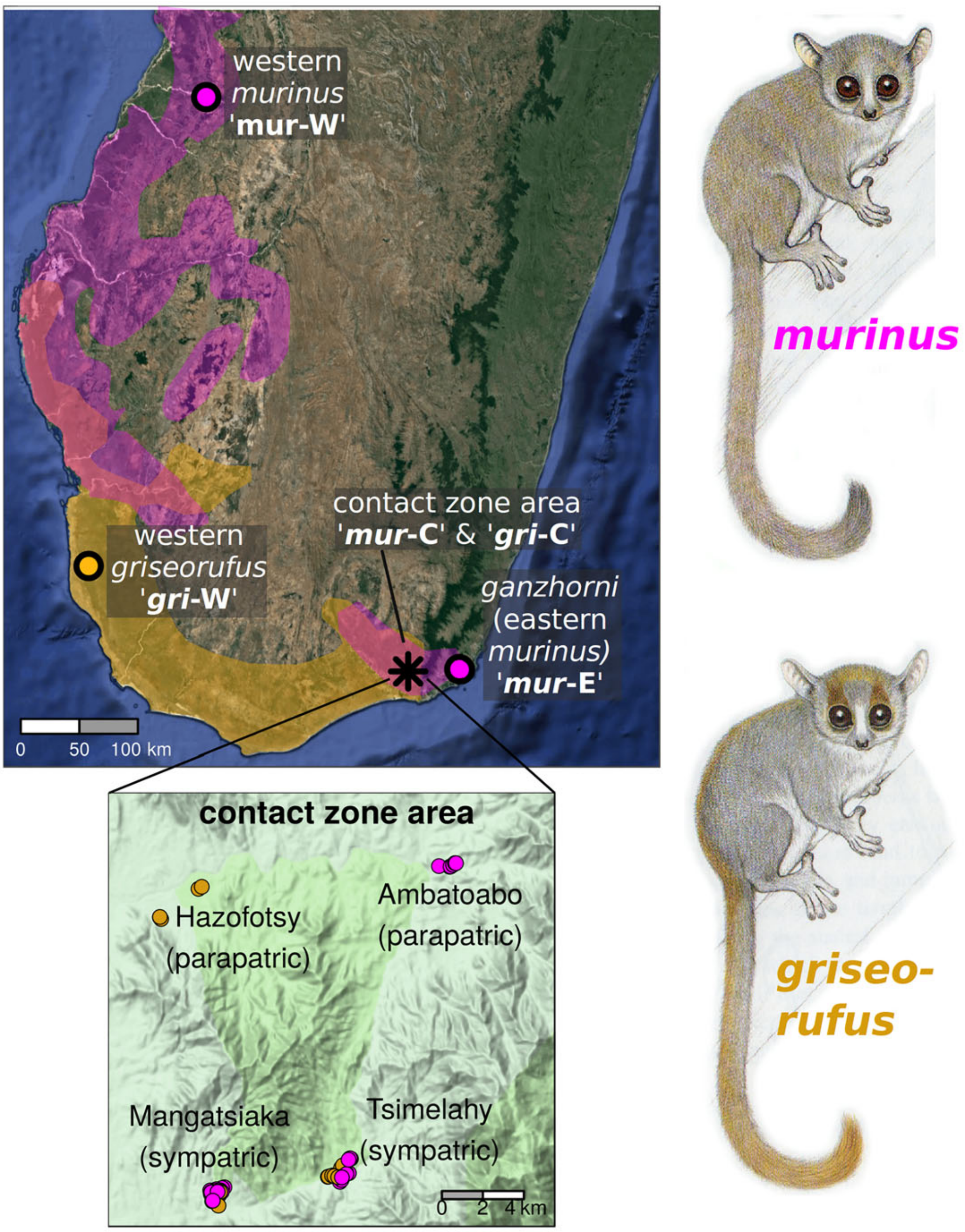

Figure 1. Distributions and sampling sites of *murinus* and *griseorufus* in southern Madagascar. The distribution of murinus is shown in purple and that of griseorufus in gold. A population in southeastern Madagascar was recently split from murinus as *M. ganzhorni*, but is here included within *murinus s.l.* The range of *M. murinus* extends to the north of the area shown in the map, whereas the entire distribution of *M. griseorufus* is shown. Inset: Overview of sampling in the contact zone area (corresponding to the study site of [10]), showing two parapatric (Hazofotsy with *griseorufus* and Ambatoaba with *murinus*) and two sympatric (Mangatsiaka and Tsimelahy) sites. *Microcebus* illustrations courtesy of Stephen Nash. (Online version in colour.)

performed genotype calling with GATK v. 4.0.7.0 [26], and we filtered SNPs and individuals largely according to the 'FS6' filter of O'Leary *et al*. [27] (see electronic supplementary materials for details).

For the set of individuals from the contact zone area, we additionally produced two datasets using more lenient filtering procedures to be able to examine admixture using more individuals and SNPs: (1) a dataset produced by omitting the last round of removal of SNPs and individuals based on missing data; (2) a dataset produced using the FS6 filter without the individual-filtering steps that retained two additional putative hybrids and two individuals with mitonuclear discordance.

Based on GATK-called genotypes, we also produced full-sequence FASTA files for each RAD locus (see electronic supplementary materials for details).

(c) Detection of hybrids using clustering approaches

For the detection of admixed individuals, we used complementary model-free and model-based approaches. First, we used Principal Component Analysis (PCA) as implemented in the SNPRelate R package v. 1.17.2 [28], using the snpgdsPCA() function. Second, we used the program ADMIXTURE v. 1.3.0 [29] to detect clusters and assign individual-level ancestry proportions from each cluster. Third, we used the program NewHybrids v. 1.1 [21], which identified the majority of admixed individuals in Hapke *et al.* [10] and Lüdemann [11]. NewHybrids was used to estimate, for each sample, the posterior probability of it belonging to each of six predefined categories: *griseorufus*, *murinus*, F1 hybrid (*griseorufus* × *murinus*), F2 hybrid (F1 x F1), *griseorufus* backcross (F1 x *griseorufus*) and *murinus* backcross (F1 x *murinus*). 500 000 iterations were used as burn-in, with another 1 500 000 iterations after that, using Jaffereys-like priors. A run was considered successful if it passed a test for convergence implemented in the hybriddetective R package [30].

(d) Reanalysis of microsatellite data

We reanalyzed the Hapke *et al*. [10] and Lüdemann [11] microsatellite data using only the samples included in this study. Like in Hapke *et al*. [10], we used the Bayesian classification methods STRUCTURE v. 2.3.4 ([19]; see the electronic supplementary materials for details) and NewHybrids v. 1.1 to detect hybrids. For STRUCTURE, 20 runs using $K = 2$ were used to calculate the average membership coefficients by creating an optimal alignment using the full-search algorithm implemented in CLUMPP v. 1.1.2 [31]. To keep the results directly comparable with Hapke *et al*. [10], we used the same threshold for the detection of hybrids: a sample was considered a hybrid when the posterior probability for assignment to the species of their mitochondrial haplotype was ≤ 0.9 for Structure or ≤ 0.5 in NewHybrids, and part of a specific hybrid category when the corresponding probability was greater than 0.5.

(e) Comparison of microsatellites and SNPs using simulations

Using simulations, we compared the performance of microsatellites and SNPs for detecting hybrids. The hybriddetective R package [30] was used to generate multi-generational hybrids from both the microsatellite and SNP data. First, unadmixed *murinus* and *griseorufus* individuals were created by randomly drawing two alleles per locus from the allopatric reference populations, without replacement. For subsequent F1 samples, one allele per locus was drawn from an unadmixed individual of each species. This procedure, drawing from the appropriate population, was continued for F2 and backcross individuals. In total, 60 simulated individuals were created: 20 each of unadmixed *griseorufus* and *murinus*, and 5 each of F1, F2, F1 x unadmixed *griseorufus*, and F1 x unadmixed *griseorufus*. Ancestry assignment was compared between microsatellites and SNPs by running STRUCTURE and NewHybrids, as described above, on the simulated genotypes.

(f) Phylogenetic inference

To enable subsequent tests of gene flow and demographic modelling, we determined relationships among all *murinus s.l.* and *griseorufus* individuals sampled by our study, using three *M. rufus* individuals as an outgroup. First, we used the NeighborNet method implemented in Splitstree v. 4.14.4 [32]. This method visually displays phylogenetic conflict in an unrooted tree and thus shows phylogenetic relationships while also allowing for the detection of potentially admixed populations and individuals. Second, we used Treemix v. 1.13 [33] to estimate relationships among predefined populations (*gri*-W, *gri*-C, *mur*-W, *mur*-C and *mur*-E) both with and without admixture events among populations.

(g) Formal admixture statistics

The D-statistic and related formal statistics for admixture use phylogenetic invariants to infer post-divergence gene flow between non-sister populations. We used the qpDstat and F4RatioTest programs of admixtools v. 4.1 [34] to compute four-taxon D-statistics and f_4-ratio tests, respectively, to test for gene flow among the predefined mouse lemur populations. For all tests, *M. rufus* was used as the outgroup. Significance of D-values was determined using the default Z-value reported by qpDstat, which uses weighted block jackknifing.

(h) Demographic modelling

We ran the coalescent-based approaches implemented in G-PhoCS v. 1.3 [35] and BPP v. 4.2 [36], using Markov Chain Monte Carlo (MCMC) to jointly infer population sizes, divergence times and migration rates for the three *murinus* populations (*mur*-W, *mur*-C and *mur*-SE) and the two *griseorufus* populations (*gri*-W and *gri*-SE). While G-PhoCS implements an isolation-with-migration model with continuous gene flow during potentially long periods, the multispecies-coalescent-with-introgression (MSCi) model in BPP models discrete introgression events.

As input for G-PhoCS and BPP, we created full-sequence FASTA files with loci for three individuals per population based on the GATK genotypes (See electronic supplementary materials for details).

We converted the migration rate parameter m to the population migration rate (2 Nm), which is the number of haploid genomes (i.e. twice the number of migrants) in the source population that arrive each generation by migration from the target population. Divergence times, population sizes and the proportion of migrants per generation ($m \times \mu$) were converted using empirical estimates of the mutation rate (1.52×10^{-8}, [37]) and generation time. For the generation time, we used a lognormal distribution with a mean of ln(3.5) and a standard deviation of ln(1.16) based on two available estimates for *Microcebus* (4.5 years from [38] and 2.5 years from [39]).

3. Results

(a) Genotyping

GATK genotyping followed by the standard (FS6) filtering procedure for all individuals resulted in a VCF file with 83 individuals and 60 460 SNPs. The equivalent VCF file with only samples from sympatric and parapatric sites in the contact zone area (Andahohela area, figure 1) contained 69 individuals, 12 of which were putative hybrids, and 7,180 SNPs. The two less stringent filtering procedures (see Methods) for the contact zone set resulted in the retention of 78 individuals (13 putative hybrids) and 48 556 SNPs and 79 individuals (18 putative hybrids) and 1360 SNPs, respectively. 16 individuals, among which 2 putative hybrids, did not survive the filtering steps for any of the final VCF files. The full-sequence FASTA file produced for G-PhoCS analyses contained 12 952 loci with an average length of 475 bp. For a comparison of QC and filtering statistics among populations, see the electronic supplementary materials.

(b) No evidence for ongoing hybridization in the contact zone

ADMIXTURE identified $K = 2$ as the optimal number of clusters among individuals from the contact zone area (figure 2*a* - top). All individuals, including the 12 putative hybrids that passed filtering, were entirely assigned to one of the two clusters (figure 2*a* - bottom), with no signs of admixture. Results were also plotted for $K = 3$, for which a third cluster corresponded to differentiation between sympatric (Mangatsiaka, Tsimelahy) and parapatric (Hazofotsy) sites in *griseorufus* (electronic supplementary material, figure S11).

Principal component analysis (PCA) with individuals from the contact zone revealed a wide separation between two groups along the first principal component axis (PC1), which explained around tenfold more of the variation

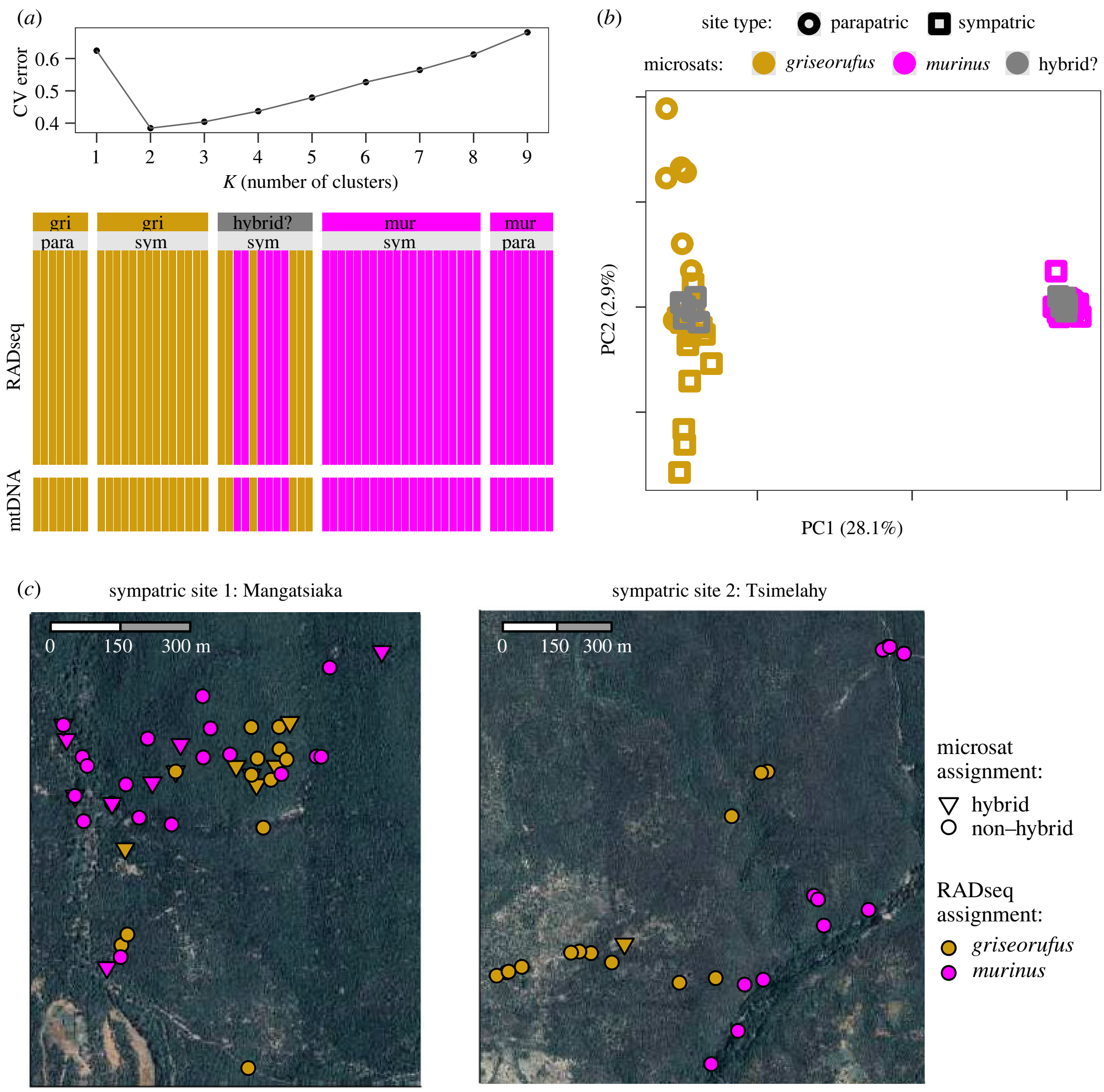

Figure 2. No evidence for hybridization in the contact zone. Nuclear RADseq data from the contact zone area was used for all analyses, including 12 individuals that had been identified as admixed in a previous microsatellite study (dark gray in panels *a* and *b*). (*a*) ADMIXTURE results. Top: a cross-validation error plot identifies $K = 2$ as the optimal number of clusters. Bottom: Ancestry components for each individual for $K = 2$ reveal a lack of admixture: all individuals were inferred to have 100% ancestry from only a single species. Individuals were previously characterized using mtDNA (bottom bars) and microsatellites (labels at top). (*b*) A PCA analysis reveals two clusters that are well-separated along PC1, corresponding to *griseorufus* and *murinus*, with no individuals that are intermediate along this axis. (*c*) Map showing spatial distribution of *murinus* and *griseorufus* individuals at the two contact sites. (Online version in colour.)

compared to PC2. The separation along PC1 corresponded to differentiation between *griseorufus* and *murinus*, and importantly, all putative hybrids fell within one of those two groups, with none occupying an intermediate position (figure 2*b*). Similar to the ADMIXTURE results at $K = 3$, PC2 mostly corresponded to differentiation between sympatric and parapatric sites in *griseorufus* (see also electronic supplementary material, figure S12 for a within-species PCA).

NewHybrids was run with and without assigning individuals from the parapatric populations to reference parental species, and in both cases, all individuals were assigned to one of the two parental species and none were assigned to one of the hybrid categories. Assignment to species matched perfectly with ADMIXTURE assignments and PCA results.

Datasets produced by less stringent filtering procedures included an additional 4 putative hybrids that did not pass all filtering steps but could still be assessed using a more limited number of SNPs (electronic supplementary material, figure S13). ADMIXTURE and NewHybrids analyses of these datasets similarly showed no evidence for admixed individuals with the exception of mitonuclear discordance: for two of the individuals for which Lüdemann [11] had detected *griseorufus* ancestry in nuclear DNA but *murinus* mtDNA haplotypes mitonuclear discordance, we could confirm that the nuclear DNA has pure *griseorufus* ancestry (electronic supplementary material, figure S13). The third sample for which Lüdemann [11] detected mitonuclear discordance did not pass filtering at all. No other cases of mitonuclear discordance were found (figure 2*a*, electronic supplementary material, table S1.)

(c) False positives in hybrid detection using microsatellites with newHybrids

In a reanalysis of the Hapke *et al.* [10] microsatellite data for only the individuals that were included in this study, 11 individuals identified as hybrids in Hapke *et al.* [10] were no

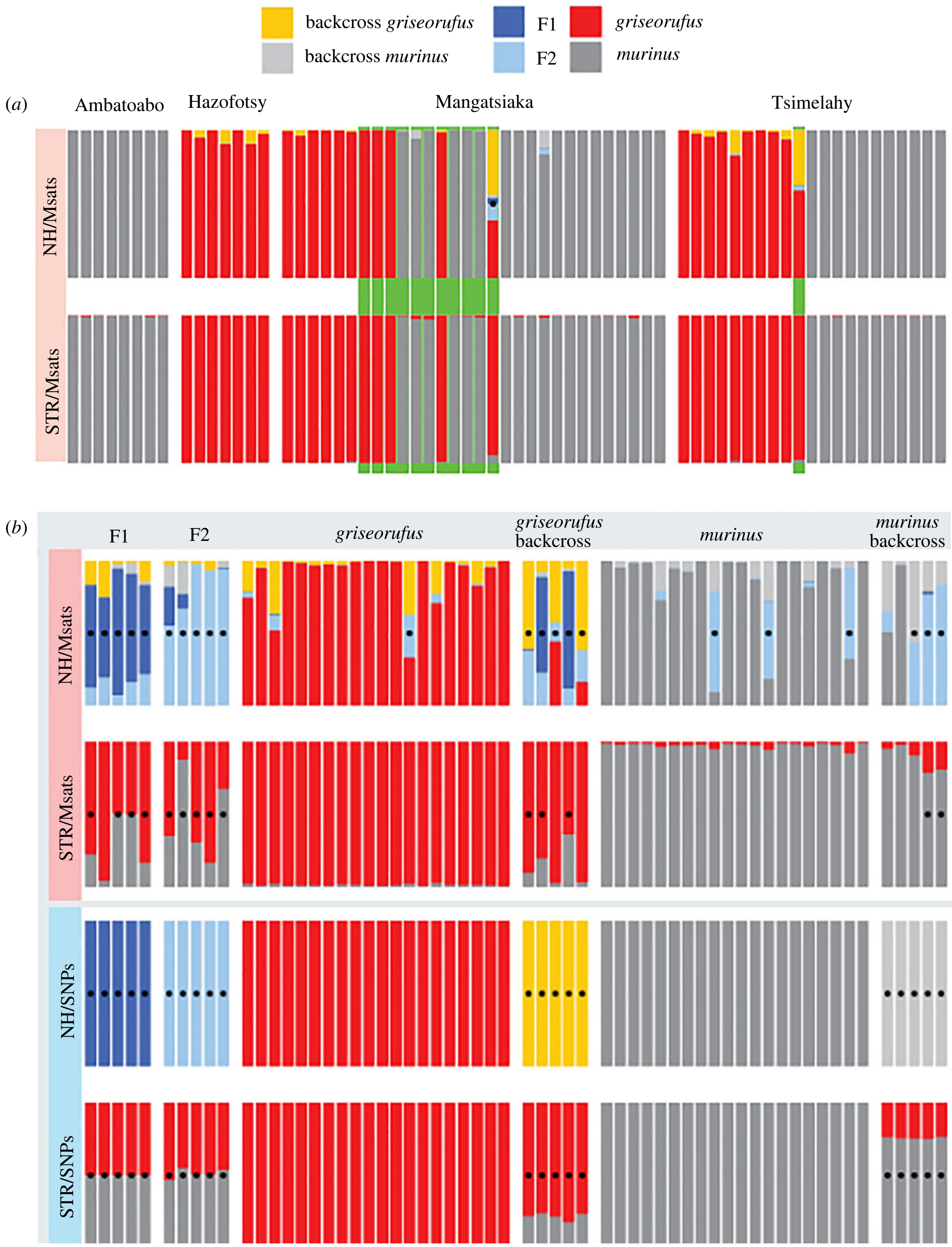

Figure 3. Re-analysis of microsatellite data and analysis of simulated individuals. (*a*) Re-analysis of microsatellite data with NewHybrids (NH; top row) and STRUCTURE (STR; bottom row). Among the 12 individuals previously identified as hybrids (green background bars), NewHybrids now identifies only a single individual as a hybrid (black dot), with several further *griseorufus* individuals showing non-significant signs of admixed ancestry (yellow ancestry). (*b*) Analysis of simulated individuals. Dots indicate detected hybrids. Using SNPs (bottom two rows), both NewHybrids and STRUCTURE correctly inferred ancestry for all individuals. Using microsatellites (top two rows), NewHybrids was prone to falsely inferring hybrids (4 out of 40 unadmixed individuals), and false negatives occurred both with NewHybrids (2 out of 20) and STRUCTURE (6 out of 20). (Online version in colour.)

longer identified as such by either NewHybrids or STRUCTURE. Only a single sample was now identified as a hybrid by NewHybrids, but STRUCTURE did not support this inference (figure 3*a*, electronic supplementary material, figure S14). As noted above, admixture was not detected for any individuals in the RADseq data, including those that had been identified as hybrids in the original microsatellite analyses figure 4.

In analyses of simulated microsatellite data, NewHybrids inferred that 4 out of 40 unadmixed individuals were hybrids, whereas STRUCTURE found no false positives. False negatives occurred with both NewHybrids (2 out of

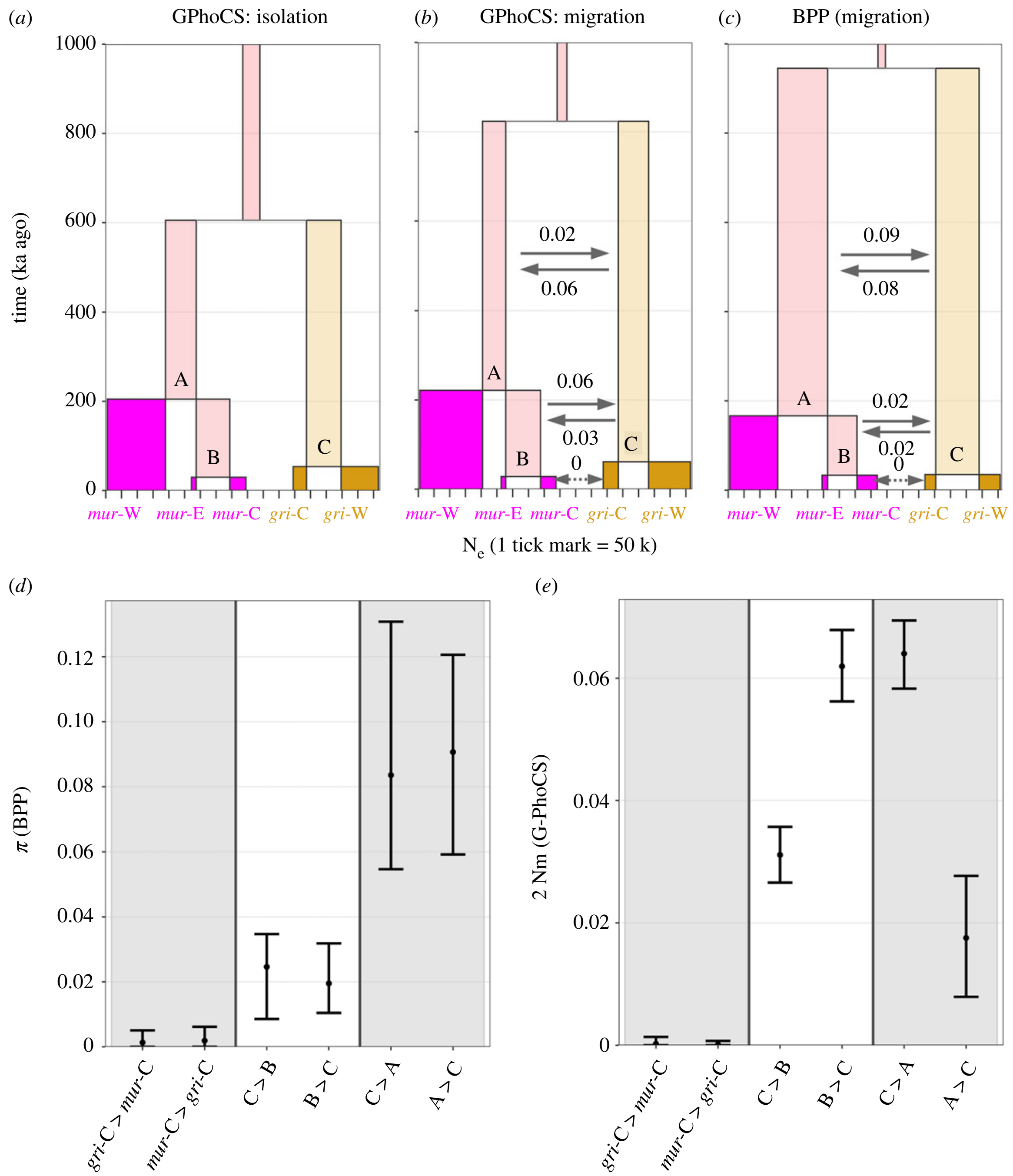

Figure 4. Demographic inferences using G-PhoCS and BPP. (*a–c*) Summary of results for G-PhoCS models without (*a*) and with (*b*) gene flow and for BPP (*c*; with gene flow). Each box represents an extant (bright colours: gold for griseorufus, purple for murinus) or ancestral (faded colours) lineage, with box width indicating Ne and box height indicating time. Gene flow was estimated reciprocally between three pairs of lineages, as depicted by the arrows (using the same units as panels *d* and *e*). (*d*) Point estimates and 95% HPDs of BPP introgression probabilities (phi). (*e*) Point estimates and 95% HPDs of G-PhoCS population migration rates (2 Nm). (Online version in colour.) **Q2**

20) and STRUCTURE (6 out of 20) for microsatellite data. On the other hand, NewHybrids and STRUCTURE analyses of simulated RADseq data were 100% accurate in inferring ancestry (figure 3*b*, electronic supplementary material, figure S15).

(d) Phylogenetic approaches clarify relationships within *murinus*

A SplitsTree NeighborNet phylogenetic network (electronic supplementary material, figure S16A) of the SNP data showed a very clear separation between *griseorufus* and *murinus* with little phylogenetic conflict, and strong intraspecific structure in *murinus*. All putative hybrids fell squarely within one of the two clades, with individual assignments in perfect agreement with clustering approaches. Similarly, a NeighborNet network using only contact zone individuals showed little to no phylogenetic conflict (electronic supplementary material, figure S17).

Treemix (electronic supplementary material, figure S16B) was run with *murinus* and *griseorufus* individuals assigned to the five populations and *M. rufus* as the outgroup, and confirmed the relationships within *murinus* suggested by Splitstree: *mur*-W was the most divergent and *mur*-C and *mur*-E were sister. No significant migration edges were found between *murinus* and *griseorufus*, with instead several significant edges between *M. rufus* and *griseorufus* and *M. rufus* and *murinus* (electronic supplementary material, figure S18). When *M. rufus* was excluded, significant migration edges between *griseorufus* and *murinus* did emerge, but did not include any between contact zone area populations (*gri*-C and *mur*-C) (electronic supplementary material, figure S19).

(e) No current – but some ancestral – interspecific gene flow

D-statistics showed an over-representation of shared derived sites between both *griseorufus* populations (*gri*-W and *gri*-C) and the two southeastern *murinus* populations (*mur*-C and *mur*-E; relative to their sister *mur*-W, western *murinus*) (electronic supplementary material, figure S20A). Values of D were highly similar regardless of which of the *griseorufus* or southeastern *murinus* populations were used, which suggests historical admixture between the ancestral *griseorufus* and southeastern *murinus* lineages, as well as a lack of ongoing gene flow in the contact zone. A lack of ongoing gene flow was further supported by values of D very close to (and not significantly different from) zero for comparisons testing for excess derived allele sharing between contact zone populations of both species relative to their sister populations (electronic supplementary material, figure S20A).

F_4-ratio tests similarly indicated ancestral admixture between *griseorufus* and the ancestor of contact zone (*mur*-C) and eastern *murinus (mur*-E) populations, specifically estimating that after divergence from western *murinus*, this ancestral southeastern *murinus* population experienced about 4.0–4.4% admixture with *griseorufus* (electronic supplementary material, figure S20B).

Demographic modelling using G-PhoCS and BPP supported the presence of non-zero but low levels of historical gene flow between ancestral *murinus* and *griseorufus* populations, but a lack of gene flow between extant contact zone area populations of *griseorufus* and *murinus* (Figure 6A-B).

4. Discussion

We re-examined a contact zone between two species of mouse lemur in southeastern Madagascar, where significant hybridization had previously been reported based primarily on evidence from microsatellite data [10]. With RADseq data, we found no evidence for the presence of admixed individuals, and using simulations and re-analyses of microsatellite data, we showed that previously detected hybrids were likely false positives. By including allopatric populations and performing multispecies coalescent analyses, we furthermore found a general lack of ongoing gene flow, and very low levels of ancestral gene flow, between these two species.

(a) Reconciling the lack of evidence for hybrids with microsatellite results

We found no admixed nuclear ancestry in any of the individuals from the contact zone. Our RADseq data are expected to have high power in species assignment and hybrid detection, given the combination of the relatively high number of genetic markers used [40,41] and the pronounced genetic differentiation between these two species (estimated divergence time in a no-migration scenario: approximately 600 ka ago, Figure 6; average F_{ST} in the contact zone area: 0.40, electronic supplementary material, table S5). Furthermore, in a re-analysis of microsatellite data using the same methods as the original studies [10,11], though restricted to the individuals used in this study, all but one of the previously detected hybrids were no longer classified as such (figure 3*a*).

Considering the clear and robust RADseq results, it is highly unlikely that true hybrids were missed in our analyses. Instead, our results suggest that the hybrids inferred in Hapke *et al.* [10] were false positives, and more generally, that the inference of hybridization using microsatellites can be sensitive to such false positives, particularly when using the program `NewHybrids`.

In our simulations with microsatellites, `STRUCTURE` suffered from false negatives only, whereas `NewHybrids` produced 4 false positives among 40 simulated unadmixed individuals (figure 3*b*). Additionally, in our reanalysis of the microsatellite data, the single individual that `NewHybrids` continued to assign hybrid ancestry to did not show signs of admixture using `STRUCTURE` (figure 3*a*). In Hapke *et al.* [10], their Figure 5), `STRUCTURE` did not consistently infer admixed ancestry for several of the putative hybrids. This was especially apparent when parapatric populations were included, in which case only 4 out of the 12 `NewHybrids` positives showed admixed ancestry using `STRUCTURE` (and 3 out of those 4 were still assigned less than 10% admixed ancestry by `STRUCTURE`, [10], their Figure 5). Even though `NewHybrids` appears considerably more prone to false positives than `STRUCTURE`, the latter did show admixed ancestry for 7 individuals in an analysis using only individuals from the contact zone site Mangatsiaka (versus 9 with `NewHybrids`). At the same time, both programs had 100% accurate assignments with simulated SNP data, suggesting that the false positives found in the microsatellite analysis stem mostly from challenges with this type of molecular marker, to which `NewHybrids` appears to be more sensitive than `STRUCTURE`.

(b) Evolutionary resolution of microsatellite versus SNP data

The results of our simulation analysis suggest that microsatellite data are vulnerable to both false positive and false negative detection of admixture between species. This effect will be especially significant when parental lineages are sufficiently phylogenetically diverged such that the rate of recurrent or backward mutation will obscure the true evolutionary signal [2,4]. To our knowledge, this study is the first to directly compare microsatellite and SNP data in a population genetic analysis within mammals. As reviewed by Sunde *et al.* [42], such 'head-to-head' studies are extremely rare and are presently limited to plants and fish. Nonetheless, relative strengths and weaknesses of the two data types are emerging. Whereas earlier assessments of microsatellite data posited that their extremely high evolutionary rate would make them ideal for revealing subtle population genetic parameters [4,7], direct comparison with SNP data is showing the opposite to be true. Indeed, these studies indicate that SNP data are more sensitive across a broad range of evolutionary parameters, including phylogenetic structure, admixture, population subdivision and measures of heterozygosity [42–44]. Recent work is also clarifying the degree to which SNP data are robust to small organismal datasets, even those with as few as $N = 2$ [44]. These observations and assessments are further supported by both the simulation and empirical results reported in this study.

(c) Lack of ongoing gene flow and implications for speciation

The presence of at least two individuals with mitonuclear discordance (a *griseorufus*-type mitochondrial haplotype, and *murinus* nuclear DNA) may suggest some ongoing or recent gene flow between the two species. However, consistent with the lack of evidence for nuclear admixture in contact zone sites, we found no evidence for ongoing gene flow using multiple methods, including a phylogenetic network (electronic supplementary material, figure S16A), `Treemix` (electronic supplementary material, figure S16B), formal admixture statistics (electronic supplementary material, figure S20) and two multispecies coalescent methods (`G-PhoCS` and `BPP`, Figure 6). Combined with syntopic occurrence at least one of the contact zone sites (figure 2), these findings strongly suggest that *murinus* and *griseorufus* are currently reproductively isolated, which is striking giving the estimated divergence time of less than 1 million years (see also [45]).

Little is known about the relative importance of different types of reproductive isolation in mouse lemurs. Across their ranges, *murinus* and *griseorufus* occur in distinct habitat types, with *griseorufus* mostly limited to spiny forests that appear to be too arid for *murinus* [46,47]. Separation by habitat (e.g. [48]) at larger scales could therefore minimize or even prevent syntopic co-occurrence despite nominal sympatry in the contact zone area, thus limiting interactions between the species. At one of the two sympatric sites included in this study, Tsimelahy, species-specific sampling locations are indeed consistent with separation by habitat, but at the other, Mangatsiaka, the two species co-occur even at a very fine spatial scale ([46]; figure 2*c*). Therefore, the observed lack of gene flow is unlikely to simply be a by-product of separation by habitat, and additional sources of pre- and/or postzygotic reproductive isolation need to be invoked.

5. Conclusion

Using RADseq data, we found no evidence for admixture between two species of mouse lemurs in a contact zone in southern Madagascar. This is in sharp contrast to a previous study that found widespread hybridization among the same samples using microsatellites. Our results suggest that the hybrids inferred by the previous study were likely false positives, and we urge caution when using microsatellites to infer hybridization. Thus, our results support concerns around the usage of microsatellites—most importantly, that rates of evolution in microsatellites are simply too high for use at interspecific levels given their propensity for homoplasy beyond the intrapopulation level [7,49]. Finally, we estimate a divergence time of less than 1 million years and a lack of historical gene flow, which in combination with local syntopic occurrence and no evidence for admixture, suggests the rapid development of reproductive isolation between these species.

Data accessibility. Sample metadata can be found in electronic supplementary material, table S1. Additional metadata and processed data, such as VCF files and analysis input and output files can be found at the Dryad Digital Repository at https://doi.org/10.5061/dryad.1jwstqjx3 (**current link for reviewers**: https://datadryad.org/stash/share/zQUUxWkecTSZPdbgm_5OMp_LhvGxKwqpd4ZAgbuWgtk). All code used to run the analyses and produce the figures in this manuscript can be found on GitHub at https://github.com/jelmerp/lemurs_contactzone_grimur. Raw sequence available through the NCBI (Accession numbers pending).

The data are provided in electronic supplementary material [50].

Authors' contributions. J.P.: conceptualization, formal analysis, methodology, writing—original draft, writing—review and editing; K.M.: data curation, formal analysis, resources, writing—review and editing; J.L.: data curation, formal analysis, writing—review and editing; Z.Y.: formal analysis, writing—review and editing; S.J.R.: resources, writing—review and editing; P.A.H.: data curation, formal analysis, resources, writing—review and editing; N.S.: data curation, resources, writing—review and editing; J.U.G.: conceptualization, funding acquisition, investigation, project administration, resources, supervision, writing—original draft, writing—review and editing; A.D.Y.: conceptualization, funding acquisition, investigation, project administration, resources, supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests. **Q1**

Funding. Financial support has been provided by the Deutsche Forschungsgemeinschaft (DFG Ga 342/19) and the Landesforschungsförderung Hamburg to J.G. Manuscript preparation was supported by NSF DEB-2148914 to A.D.Y. The project would not have been possible without an Alexander von Humboldt Foundation Award to A.D.Y.

Acknowledgements. We would like to thank Ryan Campbell, Peter Larsen and Kelsie Hunnicutt for their help with the planning of the RADseq, Anne Veillet for help with RADseq library preparation, and Tobias L. Lenz for comments on an earlier version of the manuscript. We are grateful to Stephen Nash for providing illustrations of the parental species, and we thank Maurine Neiman, the Proceedings B editorial team, and two anonymous reviewers for critical review that significantly improved our manuscript. Field work has been carried out under the Accord de Collaboration between Madagascar National Parks, the University of Antananarivo and the University of Hamburg. All samples were exported and imported with appropriate CITES permits (permit numbers available upon request).

References

1. Field D, Wills C. 1996 Long, polymorphic microsatellites in simple organisms. *Proc. R. Soc. Lond. B* **263**, 209–215. (doi:10.1098/rspb.1996.0033)
2. Ellegren H. 2004 Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* **5**, 435–445. (doi:10.1038/nrg1348)
3. Eisen JA. 1999 Mechanistic basis for microsatellite instability. In *Microsatellites: evolution and applications* (eds DB Goldstein, C Schlötterer), pp. 34–48. Oxford, UK: Oxford University Press.
4. Guichoux E *et al.* 2011 Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* **11**, 591–611. (doi:10.1111/j.1755-0998.2011.03014.x)
5. Pacheco-Sierra G, Vázquez-Domínguez E, Pérez-Alquicira J, Suárez-Atilano M, Domínguez-Laso J. 2018 Ancestral hybridization yields evolutionary distinct hybrids lineages and species boundaries in crocodiles, posing unique conservation conundrums. *Front. Ecol. Evol.* **6**, 138. (doi:10.3389/fevo.2018.00138)

6. Gompert Z, Fordyce JA, Forister ML, Shapiro AM, Nice CC. 2006 Homoploid hybrid speciation in an extreme habitat. *Science* **314**, 1923–1925. (doi:10.1126/science.1135875)
7. Selkoe KA, Toonen RJ. 2006 Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9**, 615–629. (doi:10.1111/j.1461-0248.2006.00889.x)
8. Wan QH, Wu H, Fujihara T, Fang SG. 2004 Which genetic marker for which conservation genetics issue? *Electrophoresis* **25**, 2165–2176. (doi:10.1002/elps.200305922)
9. Gligor M, Ganzhorn JU, Rakotondravony D, Ramilijaona OR, Razafimahatratra E, Zischler H, Hapke A. 2009 Hybridization between mouse lemurs in an ecological transition zone in southern Madagascar. *Mol. Ecol.* **18**, 520–533. (doi:10.1111/j.1365-294X.2008.04040.x)
10. Hapke A, Gligor M, Rakotondranary SJ, Rosenkranz D, Zupke O. 2011 Hybridization of mouse lemurs: different patterns under different ecological conditions. *BMC Evol. Biol.* **11**, 297. (doi:10.1186/1471-2148-11-297)
11. Lüdemann J. 2018 Re-assessment of a lemur hybrid zone using mitochondrial and nuclear data. BSc thesis, Universität Hamburg.
12. Braune P, Schmidt S, Zimmermann E. 2008 Acoustic divergence in the communication of cryptic species of nocturnal primates (*Microcebus* ssp.). *BMC Biol.* **6**, 19. (doi:10.1186/1741-7007-6-19)
13. Hasiniaina AF, Radespiel U, Kessler SE, Evasoa MR, Rasoloharijaona S, Randrianambinina B, Zimmermann E, Schmidt S, Scheumann M. 2020 Evolutionary significance of the variation in acoustic communication of a cryptic nocturnal primate radiation (*Microcebus* spp.). *Ecol. Evol.* **10**, 3784–3797. (doi:10.1002/ece3.6177)
14. Kollikowski A, Zimmermann E, Radespiel U. 2019 First experimental evidence for olfactory species discrimination in two nocturnal primate species (*Microcebus lehilahytsara* and *M. murinus*). *Sci. Rep.* **9**, 20386. (doi:10.1038/s41598-019-56893-y)
15. Hunnicutt KE *et al.* 2020 Comparative genomic analysis of the pheromone receptor class 1 family (V1R) reveals extreme complexity in mouse lemurs (Genus, *Microcebus*) and a chromosomal hotspot across mammals. *Genome Biol. Evol.* **12**, 3562–3579. (doi:10.1093/gbe/evz200)
16. Schmelting B. 2000 Reproduction of two sympatric mouse lemur species (*Microcebus murinus* and *M. ravelobensis*) in northwest Madagascar: first results of a long term study. In *Diversité et endémisme à Madagascar - mémoires de la société de biogéographie*. Paris, France: Société de Biogéographie.
17. Evasoa MR, Radespiel U, Hasiniaina AF, Rasoloharijaona S, Randrianambinina B, Rakotondravony R, Zimmermann E. 2018 Variation in reproduction of the smallest-bodied primate radiation, the mouse lemurs (*Microcebus* spp.): A synopsis. *Am. J. Primatol.* **80**, e22874. (doi:10.1002/ajp.22874)
18. Schüßler D *et al.* 2020 Ecology and morphology of mouse lemurs (*Microcebus* spp.) in a hotspot of microendemism in northeastern Madagascar, with the description of a new species. *Am. J. Primatol.* **82**, e23180. (doi:10.1002/ajp.23180)
19. Pritchard JK, Stephens M, Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959. (doi:10.1093/genetics/155.2.945)
20. Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A. 2004 GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J. Hered.* **95**, 536–539. (doi:10.1093/jhered/esh074)
21. Anderson EC, Thompson EA. 2002 A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217–1229. (doi:10.1093/genetics/160.3.1217)
22. Pacheco-Sierra G, Gompert Z, Dominguez-Laso J, Vazquez-Dominguez E. 2016 Genetic and morphological evidence of a geographically widespread hybrid zone between two crocodile species, *Crocodylus acutus* and *Crocodylus moreletii*. *Mol. Ecol.* **25**, 3484–3498. (doi:10.1111/mec.13694)
23. Hotaling S *et al.* 2016 Species discovery and validation in a cryptic radiation of endangered primates: coalescent-based species delimitation in Madagascar's mouse lemurs. *Mol. Ecol.* **25**, 2029–2045. (doi:10.1111/mec.13604)
24. Ali OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, Miller MR. 2016 RAD capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics* **202**, 389–400. (doi:10.1534/genetics.115.183665)
25. Larsen PA *et al.* 2017 Hybrid de novo genome assembly and centromere characterization of the gray mouse lemur (*Microcebus murinus*). *BMC Biol.* **15**, 110. (doi:10.1186/s12915-017-0439-6)
26. DePristo MA *et al.* 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498. (doi:10.1038/ng.806)
27. O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. 2018 These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Mol. Ecol.* **27**, 3193–3206. (doi:10.1111/mec.14792)
28. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012 A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328. (doi:10.1093/bioinformatics/bts606)
29. Alexander DH, Novembre J, Lange K. 2009 Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664. (doi:10.1101/gr.094052.109)
30. Wringe BF, Stanley RRE, Jeffery NW, Anderson EC, Bradbury IR. 2017 . hybriddetective: A workflow and package to facilitate the detection of hybridization using genomic data in r. *Mol. Ecol. Resour.* **17**, e275–e284. (doi:10.1111/1755-0998.12704)
31. Jakobsson M, Rosenberg NA. 2007 CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806. (doi:10.1093/bioinformatics/btm233)
32. Huson DH, Bryant D. 2006 Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267. (doi:10.1093/molbev/msj030)
33. Pickrell JK, Pritchard JK. 2012 Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967. (doi:10.1371/journal.pgen.1002967)
34. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012 Ancient admixture in human history. *Genetics* **192**, 1065–1093. (doi:10.1534/genetics.112.145037)
35. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. 2011 Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034. (doi:10.1038/ng.937)
36. Flouri T, Jiao X, Rannala B, Yang Z. 2020 A Bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. *Mol. Biol. Evol.* **37**, 1211–1223. (doi:10.1093/molbev/msz296)
37. Campbell CR, Tiley GP, Poelstra JW, Hunnicutt KE, Larsen PA, Lee H-J, Thorne JL, dos Reis M, Yoder AD. 2021 Pedigree-based and phylogenetic methods support surprising patterns of mutation rate and spectrum in the gray mouse lemur. *Heredity* **127**, 233–244. (doi:10.1038/s41437-021-00446-5)
38. Zohdy S, Gerber BD, Tecot S, Blanco MB, Winchester JM, Wright PC, Jernvall J. 2014 Teeth, sex, and testosterone: aging in the world's smallest primate. *PLoS ONE* **9**, e109528. (doi:10.1371/journal.pone.0109528)
39. Radespiel U, Lutermann H, Schmelting B, Zimmermann E. 2019 An empirical estimate of the generation time of mouse lemurs. *Am. J. Primatol.* **81**, e23062. (doi:10.1002/ajp.23062)
40. Vähä J-P, Primmer CR. 2006 Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* **15**, 63–72. (doi:10.1111/j.1365-294X.2005.02773.x)
41. McFarlane SE, Pemberton JM. 2019 Detecting the true extent of introgression during anthropogenic hybridization. *Trends Ecol. Evol.* **34**, 315–326. (doi:10.1016/j.tree.2018.12.013)
42. Sunde J, Yildirim Y, Tibblin P, Forsman A. 2020 Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Front. Genet.* **11**, 218. (doi:10.3389/fgene.2020.00218)
43. Bohling J, Small M, Von Bargen J, Louden A, DeHaan P. 2019 Comparing inferences derived from microsatellite and RADseq datasets: a case study involving threatened bull trout. *Conserv. Genetics* **20**, 329–342. (doi:10.1007/s10592-018-1134-z)
44. Sturm AB, Eckert RJ, Mendez JG, Gonzalez-Diaz P, Voss JD. 2020 Population genetic structure of the great star coral, *Montastraea cavernosa*, across the Cuban archipelago with comparisons between microsatellite and SNP markers. *Sci. Rep.* **10**, 15432. (doi:10.1038/s41598-020-72112-5)

45. Poelstra JW *et al.* 2021 Cryptic patterns of speciation in cryptic primates: microendemic mouse lemurs and the multispecies coalescent. *Syst. Biol.* **70**, 203–218. (doi:10.1093/sysbio/syaa053)
46. Rakotondranary SJ, Ganzhorn JU. 2011 Habitat separation of Sympatric *Microcebus* spp. in the dry spiny forest of south-eastern Madagascar.'. *Folia Primatol.* **82**, 212–223. (doi:10.1159/000334816)
47. Rakotondranary SJ, Hapke A, Ganzhorn JU. 2011 Distribution and morphological variation of *Microcebus* spp. along an environmental gradient in southeastern Madagascar.'. *Int. J. Primatol.* **32**, 1037–1057. (doi:10.1007/s10764-011-9521-z)
48. Wuesthoff EF, Fuller TK, Sutherland C, Kamilar JM, Ramanankirahina R, Rakotondravony R, Rouse S, Radespiel U. 2021 Differential habitat use by sympatric species of mouse lemurs across a mangrove–dry forest habitat gradient. *J. Mammal.* **102**, 1266–1278. (doi:10.1093/jmammal/gyab082)
49. Estoup A, Tailliez C, Cornuet JM, Solignac M. 1995 Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* **12**, 1074–1084. (doi:10.1093/oxfordjournals.molbev.a040282)
50. Poelstra J, Montero BK, Lüdemann J, Yang Z, Rakotondranary SJ, Hohenlohe P, Stetter N, Ganzhorn JU, Yoder AD. 2022 Data from RADseq data reveal a lack of admixture in a mouse lemur contact zone contrary to previous microsatellite results. Figshare.