



HAL
open science

Morphological variability or inter-observer bias? A methodological toolkit to improve data quality of multi-researcher datasets for the analysis of morphological variation

Dominik Schüßler, Marina B Blanco, Nicola K Guthrie, Gabriele M Sgarlata, Melanie Dammhahn, Refaly Ernest, Mamy Rina Evasoa, Alida Hasiniaina, Daniel Hending, Fabien Jan, et al.

► To cite this version:

Dominik Schüßler, Marina B Blanco, Nicola K Guthrie, Gabriele M Sgarlata, Melanie Dammhahn, et al.. Morphological variability or inter-observer bias? A methodological toolkit to improve data quality of multi-researcher datasets for the analysis of morphological variation. *American Journal of Biological Anthropology*, inPress, 10.1002/ajpa.24836 . hal-04300286

HAL Id: hal-04300286

<https://hal.science/hal-04300286>

Submitted on 23 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Morphological variability or inter-observer bias? A methodological toolkit to improve data quality of multi-researcher datasets for the analysis of morphological variation

Journal:	<i>American Journal of Biological Anthropology</i>
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Schübler, Dominik; University of Hildesheim, Institute of Biology and Chemistry Blanco, Marina B.; Duke University, Duke Lemur Center Guthrie, Nicola; University of Calgary, Anthropology and Archaeology Sgarlata, Gabriele; Instituto Gulbenkian de Ciencia Dammhahn, Melanie; University of Münster, Behavioural Biology Ernest, Refaly; QIT Madagascar Minerals Evasoa, Mamy; University of Veterinary Medicine Hannover, Institute of Zoology Hasaniaina, Alida; University of Veterinary Medicine Hannover, Institute of Zoology; School for International Training Hending, Daniel; University of Oxford, Biology Jan, Fabian; Gulbenkian Institute of Science Le Pors, Barbara; Gulbenkian Institute of Science Miller, Alex; Perth Zoo Olivieri, Gillian; University of Veterinary Medicine Hannover, Institute of Zoology Rakotonanahary, Ando; Madagascar Biodiversity Partnership Rakotondranary, Solofomalala; University of Antananarivo Faculty of Science, Mention Anthropobiologie et Développement Durable Rakotondravony, Romule; University of Mahajanga, Faculté des Sciences, de Technologies et de l'Environnement Ralantoharijaona, Tantely; University of Mahajanga, Faculté des Sciences, de Technologies et de l'Environnement Ramananjato, Veronarindra; University of California Berkeley, Integrative Biology; University of Antananarivo Faculty of Science, Mention Zoologie et Biodiversite Animale Randrianambinina, Blanchard; University of Mahajanga Raeolinjanakolona, Nancia; University of Antananarivo Faculty of Science, Mention Zoologie et Biodiversite Animale Rasoazanabary, Emilienne; Stony Brook University, Anthropology Rasoloarison, Rodin; German Primate Centre Leibniz Institute for Primate Research Rasolofoson, David; GERP Madagascar Rasoloharijaona, Solofo; Université de Mahajanga Rasolondraibe, Emmanuel; Université de Mahajanga</p>

	<p>Roberts, Sam; Duke University Teixeira, Helena; University of Veterinary Medicine Hannover, Institute of Zoology; University of Reunion Island; University of New Caledonia van Elst, Tobias; University of Veterinary Medicine Hannover, Institute of Zoology Johnson, Steig; University of Calgary, Department of Anthropology & Archaeology Ganzhorn, Jörg; Universität Hamburg Chikhi, Lounes; Gulbenkian Institute of Science; University of Toulouse III Paul Sabatier Libraries Kappeler, Peter; Arbts. Grp. Verhaltenforschung, Okologie Deutsches Primatenzentrum; University of Göttingen Louis, Edward; Henry Doorly Zoo, Center for Conservation and Research; Madagascar Biodiversity Partnership Salmona, Jordi; Université Toulouse III Paul Sabatier Radespiel, Ute; TiHo Hannover, Institute of Zoology</p>
Key Words:	Allen's rule, Bergmann's rule, Microcebus, cryptic species, sexual size dimorphism
Subfield: Please select 2 subfields. Select the main subject first.:	Primate biology [behavior, ecology, physiology, anatomy], Theory

SCHOLARONE™
 Manuscripts

Tables

Table 1: Quality filtering pipeline for the reduction of inter-observer bias in multi-observer datasets.

Step	Aim	Description	Example
1	Define geographic and taxonomic units	Hierarchical levels in the dataset (e.g., species or population level) must be defined. Individuals should be grouped accordingly for the next steps.	The taxonomy of mouse lemurs is still a topic of ongoing debate (e.g., Poelstra et al., 2021; Tiley et al., 2022), and geographic variation in morphological traits for taxa with wide distributions are possible (e.g., Schüßler et al., 2020). We defined populations per taxon as our units of investigation.
2	Exclude certain individuals from the population	It must be clarified whether there is strong sexual size dimorphism or intra-annual variability of certain variables (e.g., body mass, girth, fat deposits) in any of the study taxa. Can adults/juveniles or pregnant/non-pregnant females be distinguished from each other? Is there strong seasonal variation in the composition of the population? In these cases, the dataset may be split into male/female sub-categories, some individuals may be excluded completely (e.g., juveniles), or certain variables set to NA (e.g., body mass of pregnant females).	Mouse lemurs are fast-growing seasonal breeders with high population turnover rates and a relatively high mortality of young individuals (Radespiel et al., 2001; Lutermann et al., 2006; Kraus et al., 2008; Blanco et al., 2015; Evasoa et al., 2018). To remove the confounding effect of age, juveniles were excluded from the dataset. Furthermore, we excluded measurements of individuals from the nonbreeding season (as indicated by the lack of males with enlarged testes per species; Evasoa et al., 2018) to control for the high proportion of not yet fully-grown individuals in the population before the reproductive season starts. For pregnant females (identified during capture or as numerical outliers, i.e., if heavier than heaviest male), body mass was set to NA.
3	Exclude outliers	Outliers are data points that are inconsistent with the majority of data and their statistical distribution (Pearson, 2002; Liu et al., 2004). These may stem from methodological errors or may be actual biological variation. We argue that for the analysis of interspecific differences both kinds of outliers should be excluded. Biological outliers are of low value, because they do not contribute much to the statistical distribution of a certain measurement. Instead, their inclusion may impair	We set this breakpoint to the median $\pm 3 \times \text{MAD}$ (Davies & Gather, 1993; Pearson, 2005), and identified outliers were replaced with NA.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

		the power of subsequent statistical procedures. Therefore, we propose to use the generic but powerful Hampel identifier (Davies & Gather, 1993; Pearson, 2002) to exclude outliers of any kind. It uses the median and the median absolute deviation (MAD; Hampel, 1985) to define breakpoints below/above which measurements are identified as outliers (Pearson, 2002).	
4	Exclude researchers that contribute too high variation	It should be assessed whether some researchers were less consistent than others in measuring. Such an observer effect introduces additional noise that may dilute subtle differences among the units of investigation, particularly in small-sized and cryptic species. We propose the use of the coefficient of variation (CV) as a measure of variability in a sample in relation to its mean (e.g., Pélabon et al., 2020). It is often applied to quantify the plasticity in a phenotypic trait and to compare it across different traits (Houle, 1992; Pélabon et al., 2020). To distinguish between actual phenotypic plasticity of species and high measurement variation of certain researchers, we assumed that (1) CV values per species-researcher combination follow a normal distribution, and (2) researchers with high CV values generate a long right tail in the distribution. In practice, the CVs for all researchers per unit of investigation were calculated separately for each variable.	We determined the upper limit of the Hampel identifier (see above) and used it as a truncation value, above which variables measured by the respective researchers were replaced with NA (Figure 2B).
5	Exclude systematic deviations among researchers	If there are several researchers who contributed to measurements of the same unit of investigation (i.e., species or population), variable-wise comparisons may help to identify systematic measurement errors. For this, we recommend visualizing the distributions of specific variables across all units of investigation via boxplots, with researchers represented as color-coded boxes.	For example, certain researchers measured mouse lemur hind foot length consistently different from others across different populations, which may indicate a different application of calipers (Figure 2A). We applied the following rules to exclude such inter-observer biases from our dataset: (1) If < 3 researchers contributed measurements, it is not possible to assess which measurements were more accurate; (2) if ≥ 3 researchers contributed measurements, boxes that do not overlap with each other and represent the smaller number of researchers (e.g., one deviating from the others) indicate a researcher whose values for this variable should be set to NA. Examples are illustrated in Figure 2A.
6	Define the dataset for	By replacing filtered values with NA (see previous steps), we aimed at retaining as much information as possible. Depending on the analysis of interest,	

downstream analyses	it is possible to use the whole dataset (e.g., for univariate analyses) or one that omits all individuals containing NA values (i.e., for multivariate analyses). In the latter case, it should be assessed whether remaining variables are still sufficient for testing the question under study.
------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

4

Morphological variability or inter-observer bias?

A methodological toolkit to improve data quality of multi-researcher datasets for the analysis of morphological variation

Running title: Quality filtering of morphological data

Authors:

Dominik Schüßler^{1#},

Marina B. Blanco², Nicola K. Guthrie³, Gabriele M. Sgarlata⁴,

Melanie Dammhahn⁵, Refaly Ernest⁶, Mamy Rina Evasoa⁷, Alida Hasiniaina^{7,8}, Daniel Hending⁹, Fabien Jan⁴, Barbara le Pors⁴, Alex Miller¹⁰, Gillian Olivieri⁷, Ando N. Rakotonanahary¹¹, Solofomalala Jacques Rakotondranary¹², Romule Rakotondravony¹³, Tantely Ralantoharijaona¹³, Veronarindra Ramananjato^{14,15}, Blanchard Randrianambinina¹³, Nancia N. Raoelinjanakolona¹⁵, Emilienne Rasoazanabary¹⁶, Rodin M. Rasoloarison¹⁷, David W. Rasolofoson¹⁸, Solofonirina Rasoloharijaona¹³, Emmanuel Rasolondraibe¹³, Sam Hyde Roberts¹⁹, Helena Teixeira^{7,20,21}, Tobias van Elst⁷,

Steig E. Johnson³, Jörg U. Ganzhorn²², Lounès Chikhi^{4,23}, Peter M. Kappeler^{17,24}, Edward E. Louis Jr.¹¹, Jordi Salmons²³, Ute Radespiel^{7#}

#: Corresponding authors:

Dominik Schüßler, dominik.schuessler@posteo.de

Ute Radespiel, ute.radespiel@tiho-hannover.de

Affiliation:

1: Institute of Biology and Chemistry, University of Hildesheim, Germany

2: Duke Lemur Center, Durham, NC, USA

3: Department of Anthropology and Archaeology, University of Calgary, Canada

4: Instituto Gulbenkian de Ciência, Oeiras, Portugal

5: Behavioural Biology, University of Münster, Germany

6: QIT Madagascar Minerals, Fort-Dauphin, Madagascar

7: Institute of Zoology, University of Veterinary Medicine Hannover, Germany

8: School for International Training, Antananarivo, Madagascar

- 1
2
3 32 9: Department of Biology, University of Oxford, John Krebs Field Station, Wytham, Oxford, UK
4
5 33 10: Perth Zoo, 20 Labouchere Road, South Perth, WA, Australia.
6
7 34 11: Madagascar Biodiversity Partnership (MBP), Manakambahiny, Antananarivo, Madagascar
8
9 35 12: Mention Anthropobiologie et Développement Durable, Faculté des Sciences, Université
10 36 d'Antananarivo, Antananarivo, Madagascar
11
12 37 13: Faculté des Sciences, de Technologies et de l'Environnement, Ecole Doctorale Ecosystèmes Naturels,
13 38 Université de Mahajanga, Madagascar
14
15 39 14: Department of Integrative Biology, University of Berkeley, CA, USA
16
17 40 15: Mention Zoologie et Biodiversité Animale, Faculté des Sciences, Université d'Antananarivo,
18 41 Madagascar
19
20 42 16: Department of Anthropology, University of Massachusetts, Amherst, MA, USA
21
22 43 17: Behavioral Ecology and Sociobiology Unit, German Primate Center, Leibniz Institute for Primate
23 44 Research, Göttingen, Germany
24
25 45 18: Groupe d'Étude et de Recherche sur les Primates de Madagascar (GERP), Antananarivo, Madagascar
26
27 46 19: Duke University, Durham, NC, USA
28
29 47 20: UMR ENTROPIE, Université de La Réunion, La Réunion, France
30
31 48 21: Université de Nouvelle-Calédonie, Noumea, New Caledonia, France
32
33 49 22: Animal Ecology and Conservation, University of Hamburg, Hamburg, Germany
34
35 50 23: Laboratoire Évolution et Diversité Biologique, UMR5174, IRD, CNRS, Université Toulouse III Paul
36 51 Sabatier, Toulouse, France
37
38 52 24: Department of Anthropology/Sociobiology, University of Göttingen, Göttingen, Germany
39
40
41 53
42
43 54

55 Research highlights

- 56
57
58
59
60
61
62
63
64
- Morphological variation is commonly used to infer environmental adaptations and phylogenetic relationships in animal taxa, but multi-researcher datasets can hamper analyses due to inter-observer biases.
 - Our filtering pipeline markedly improved data quality for downstream analyses and highlighted key steps in data quality assessment for generating reliable results.
 - Across species, mouse lemurs are sexually dimorphic (with larger females), they follow neither Rensch's rule for sexual size dimorphism nor Bergmann's rule (larger mouse lemur taxa live in warmer and wetter climates), but are concordant with Allen's rule with regard to having shorter tails in colder environments.

65

Abstract

Objectives

The investigation of morphological variation in animals is widely used in taxonomy, ecology, and evolution. Using large datasets for meta-analyses has dramatically increased, raising concerns about dataset compatibilities and biases introduced by contributions of multiple researchers.

71

Materials and Methods

We compiled morphological data on 13 variables for 3,073 individual mouse lemurs (*Cheirogaleidae*, *Microcebus* spp.) from 25 taxa and 153 different sampling locations, measured by 48 different researchers. We introduced and applied a filtering pipeline and quantified improvements in data quality (Shapiro-Francia statistic, skewness, and excess kurtosis). The filtered dataset was then used to test for genus-wide presence of sexual dimorphism and the applicability of Rensch's, Allen's and Bergmann's rules.

79

Results

Our filtering pipeline helped to reduce inter-observer bias (i.e., increased normality of data distributions). Inter-observer reliability of measurements was notably variable, highlighting the need to reduce data collection biases. Although subtle, we found a consistent pattern of sexual dimorphism across the genus *Microcebus*, with females being the larger sex in 6 of 13 variables. Sexual size dimorphism was isometric, providing no support for Rensch's rule. Variations in tail length but not in ear size were consistent with the predictions of Allen's rule. Body mass and length followed a pattern contrary to predictions of Bergmann's rule.

88

Discussion

We highlighted the usefulness of large multi-researcher datasets for testing ecological hypotheses, after correcting for inter-observer biases. Using genus-wide tests, we outlined generalizable patterns of morphological variability across all mouse lemurs and generated a set of new hypotheses for further investigation.

94

95

Keywords: Allen's rule, Bergmann's rule, *Microcebus*, cryptic species, sexual size dimorphism

97

98 Introduction

99 Morphological variation is central to the quantification of within- and between-species diversity and thus
100 at the heart of evolutionary biology and ecology (Darwin, 1859; Mayr, 1963; Wright, 1984). In times of
101 increasing computational power that allows the analyses of large genomic datasets, combining
102 morphological data from multiple sources will become an increasingly important task, for example, when
103 aiming to combine multiple lines of evidence to derive conclusions in taxonomy (i.e., integrative
104 taxonomy; de Queiroz, 2007; Padial et al., 2010; Leaché et al., 2019). In this and other contexts,
105 distinguishing between biologically meaningful signals and methodological biases arising from the
106 compilation of multi-observer datasets is of utmost importance (Burghardt et al., 2012; Howard, 2002;
107 Morrison 2015).

108 In general, processing of large datasets should involve the following steps: 1) contextualization of the
109 dataset, 2) assessment, and 3) improvement of data quality (Van den Berghe 2017). In morphological
110 studies, in particular, data quality can be severely constrained and impaired by intra- and inter-observer
111 biases. These biases can principally arise from a differential use of measurement definitions, variable
112 care or stringency while handling an animal, and the use of different tools by different researchers
113 (Howard, 2002; Roitberg et al., 2011). A similar problem may arise if one researcher collects data over
114 longer time periods and slowly and unconsciously shifts the measurement technique by developing
115 routines (Roitberg et al., 2011). Particularly for small-bodied species, small measurement errors can
116 introduce large relative biases. A problematic case in this respect is the study of cryptic species that
117 display only subtle phenotypic variation by definition (Bickford et al., 2007), which may become easily
118 masked when such biases are strong (Howard, 2002). Therefore, the detection and reduction of intra-
119 and inter-observer biases should be a priority for studies aiming to investigate species- or population-
120 specific morphological variability.

121 Different approaches can be taken to ensure dataset reliability and to counteract confounding effects of
122 inter- or intra-observer biases. Although transparency about measurement routines for morphological
123 traits (e.g., information on how many researchers were involved, how training was conducted, how
124 precise measurements were conducted) would be useful, it is remarkably rare in the literature. Yet,
125 detailing which measurement definitions were used is rather common, either by citing other sources or
126 by directly defining morphological variables (Chornelia et al., 2022; Michaloudi et al., 2018; Radespiel et
127 al., 2008, 2012). Using these definitions and measuring individuals by only one researcher is occasionally
128 stated to dispel doubts about the reliability of a dataset (e.g., Macdonald et al., 2018; Benites et al.,
129 2020), even though intra-observer bias can still exist. If more than one researcher is taking
130 measurements, the researcher's experience is sometimes reported (e.g., Pilloud et al., 2019), while joint
131 researcher trainings are considered particularly useful by others (e.g., Morrison, 2015). Further
132 approaches include taking measurements of each individual by different researchers and calculating an
133 agreement or mean value for later analyses (e.g., Shirley et al., 2014). Such measures are typically
134 applied in studies that require a high sampling effort and are designed prior to any data collection. In
135 reality, however, huge datasets are often compiled after data collection has taken place. In these cases,
136 *a posteriori* methods for reducing observer biases are needed but largely absent in the literature
137 (Burghardt et al., 2012). A standardized framework of how to assess the quality of compiled
138 morphological data and how to improve data quality has not been defined so far.

139 After data compatibility and quality of a dataset have been secured, patterns of morphological variation
140 among taxa can be investigated. Variation itself can then be attributed to a variety of potentially

1
2
3 141 interacting factors (Hallgrímsson & Hall, 2005). These factors could be phylogenetic signals with
4 142 relevance to species delimitation studies (e.g., Cadena et al., 2018; Chornelia et al., 2022; Schüßler et al.,
5 143 2020), large scale ecological adaptation to different environments or small scale niche differentiation
6 144 (e.g., Agostini et al., 2017; Meachen et al., 2015; Santini et al., 2018). Variation can further be attributed
7 145 to mechanisms of sexual selection acting differently on males and females (e.g., Jennions et al., 2001) or
8 146 natural selection for sex-dependent environmental niches (Kamilar & Pokempner, 2008; Williamson et
9 147 al., 2021). An established principle in the context of sexual size dimorphism is Rensch's rule (Rensch,
10 148 1950), which predicts that at the genus level "the ratio of female to male size is larger the smaller the
11 149 body size of a species" (Meiri & Liang, 2021). Regardless of sex, ecogeographic relationships such as
12 150 those formulated in Allen's and Bergmann's rules (Mayr, 1942), could also be the underlying forces of
13 151 morphological variation. According to both, thermoregulatory constraints in endothermic animals are a
14 152 dominant driver in size variation, either in body size or in the proportions of appendages. In colder
15 153 climates, body appendages like ears and a tail tend to be smaller (Allan's rule) and body size larger
16 154 (Bergmann's rule; Mayr, 1942) than in warmer climates. Deviation from these "rules" can in turn provide
17 155 insights into the evolutionary history of a lineage or relevant physiological processes (Maestri et al.,
18 156 2016; Rodríguez et al., 2008; Webb & Freckleton, 2007).

19 157 This study aims to (1) develop and apply an objective multi-stage quality filtering pipeline to eliminate
20 158 problematic data points and reduce the confounding effects of sampling, inter-observer bias and intra-
21 159 observer variation, (2) evaluate the improvement of data quality across different filtering steps, and (3)
22 160 use the filtered dataset to exemplarily test hypotheses concerning morphological variation in a genus of
23 161 small, cryptic primates in Madagascar (i.e., *Microcebus*). Specifically, we test for (3a) the presence of
24 162 sexual size dimorphism and predictions of (3b) Rensch's rule, (3c) Allen's rule and (3d) Bergmann's rule.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

163 Methods

164

165 Study system

166 Mouse lemurs (genus *Microcebus*) are among the most speciose genera of lemurs, containing >20 cryptic
167 species (Zimmermann & Radespiel, 2014; Hotaling et al., 2016; Sgarlata et al., 2019; Schüßler et al.,
168 2020). These small-bodied and omnivorous primates can be found in all major ecosystems across
169 Madagascar (Mittermeier et al., 2010). Although subtle, morphological variation has often been used to
170 support taxonomic decisions in mouse lemurs (e.g., Rasoloarison et al., 2000; Radespiel et al., 2012;
171 Schüßler et al., 2020), genus-wide tests to assess morphological variability among mouse lemurs have
172 not been conducted so far. Female-biased sexual dimorphism has been reported for *Microcebus murinus*
173 body mass and head length (Jenkins & Albrecht, 1991; Kappeler, 1991; Schmid & Kappeler, 1998). Body
174 mass variation, e.g., with heavier females, was also found in several other recent and extinct lemur
175 species (Godfrey et al., 1993; Kappeler et al., 2022), and sex differences in canine length were found in
176 *M. griseorufus* (Rasoazanabary, 2011). Mouse lemurs have been studied by various research groups over
177 the years. Due to small distributions (i.e., micro-endemism) scattered across all forest habitats of the
178 island, the high levels of endangerment of some species, the remoteness of their remaining forest
179 habitats, and the difficulties of capturing larger numbers of mouse lemurs in various habitat types, it is
180 not possible for one or a few researchers to generate a large morphological dataset encompassing all
181 taxa. Consequently, the assembly of multiple datasets generated by multiple research groups, as it was
182 done here, is the only way to study morphological variation in this genus. However, we believe that this
183 problem is a general one and applies to a large variety of cryptic species worldwide.

184

185 Dataset characteristics

186 We assembled a morphological dataset comprised of 3,073 individual mouse lemurs belonging to 25
187 different taxa (Hotaling et al., 2016; Schüßler et al., 2020; Sgarlata et al., 2019) that were captured and
188 measured by 48 researchers in 153 study sites across Madagascar (Figure 1). Representation of the
189 different taxa was highly variable with a maximum of 748 individuals for *M. ravelobensis* and only two or
190 four individuals for *M. marohita* and *M. manitatra*, respectively (Table S1). We collected measurements
191 for the 13 most widely used morphological variables (ear length, ear width, head length, head width,
192 inter- and intra-orbital distance, snout length, lower leg length, hind foot length, third toe length, body
193 length, tail length, body mass; Hafen et al., 1998; Zimmermann et al., 1998) as defined in Table S2.
194 However, these were not available for all individuals. We carefully checked which definitions were
195 applied by different researchers and excluded all measures that did not conform to the most common
196 definitions (Table S2). Most researchers provided additional data for each individual: sex was available
197 for 99.5%, stage of maturity for 100.0%, GPS location for 99.6%, capture date for 79.4%, and
198 reproductive state for 40.2% of all individuals.

199
200 <<< Figure 1 here >>>

201

202

203 Generalized quality filtering pipeline

204 We designed a six-step data quality filtering pipeline to reduce inter-observer bias and illustrate its use
205 with examples stemming from our dataset (Table 1). Steps 1, 2 and 6 were used to contextualize the
206 dataset (Van den Berghe et al., 2017). For this, geographic and taxonomic units were defined first,
207 characteristics of the study taxon were acknowledged second, and the data were defined for
208 downstream analyses. In steps 3-5, we aimed at improving overall data quality (i.e., to reduce inter-
209 observer biases; see Van den Berghe et al., 2017) by excluding outliers, researchers who contributed
210 unusually high variation in their measurements, and those who deviated systematically from the
211 majority of measurements per taxon (Table 1).

212

213 <<< Table 1 here >>>

214 <<< Figure 2 here >>>

215

216 Assessment of data quality

217 Morphological variation among individuals of the same species or taxon should theoretically follow a
218 normal distribution (Templeton, 2006; Cadena et al., 2018) if the sample is drawn at random from a
219 randomly mating population and if the inheritance of a polygenic phenotypic trait is close to Hardy-
220 Weinberg equilibrium (Templeton, 2006). Furthermore, cryptic species are per definition characterized
221 by a similar and non-diverging morphology (Bickford et al., 2007), consistent with the idea of a unimodal
222 normal distribution. This basic assumption for the distribution of measurements of a specific trait for a
223 certain taxon should hold irrespective of the researcher. Accordingly, biases originating from multiple
224 researchers should produce a deviation from a normal distribution for the trait under study.
225 Any assessment of data quality should be based on well-founded and comprehensible metrics (Heinrich
226 et al., 2018).

227 To test the effectiveness of the data quality filtering pipeline, we first quantified the deviation from a
228 normal distribution by subtracting the Shapiro-Francia test statistic W' from 1 (i.e., deviation from
229 normality = $1 - W'$; “nortest” package in R; Gross & Ligges, 2015) for each filtering step and variable
230 separately. This modification of the widely used Shapiro-Wilk test for assessment of data normality
231 represents the squared correlation between the ordered data values and the expected normal order
232 statistics (i.e., the theoretical quantiles in a qqplot; Shapiro & Francia, 1972, Royston, 1991). In contrast
233 to the typical Shapiro-Wilk test statistic W (Royston 1991), W' has been shown to powerfully capture the
234 deviation from a normal distribution (Mbah & Paothong, 2014), while making the test statistic directly
235 interpretable without relying on a P-value. Values of $1 - W'$ range from 0 (perfectly normally distributed
236 data) to 1 (no normal distribution).

237 Second, we calculated the skewness, which measures asymmetry or horizontal deviation from a normal
238 distribution, and excess kurtosis of the distribution, which measures “tailedness” or vertical deviation
239 from normality (Hopkins & Weeks, 1990) for each filtering step and variable separately. We used
240 absolute values for both measures to establish comparability across metrics, with 0 indicating perfect fit
241 to a normal distribution. For both measures, deviations of 0 - 2 units are considered acceptable for a
242 normal distribution (Trochim & Donnelly, 2006; Gravetter & Wallnau, 2014).

243 Finally, we assessed the inter-observer reproducibility of measures for each variable after the last quality
244 filtering step: we used the coefficients of variation (CV) calculated for each population-researcher
245 combination and compared them across the 13 different morphological variables. CV-values < 0.1 were
246 defined as very good quality and CV = 0.1-0.2 as good quality (Aronhime et al. 2013).

247

248 Testing hypotheses of morphological variation

249 Using the quality-filtered dataset, we tested for (1) the presence of sexual size dimorphism, as well as
250 predictions of (2) Rensch's, (3) Allen's and (4) Bergmann's rule in the genus *Microcebus*:

251 (1) Genus-wide sexual size dimorphism was tested in all 13 morphometric variables using linear mixed
252 effects models, including each variable individually as the response variable, sex as a fixed effect and
253 taxonomic affiliation (i.e., population of a species) as a random effect. Modelling was done using the
254 "lme4" and "lmerTest" packages in R, while the "performance" package was used to check the
255 assumptions of the models (Bates et al., 2015; Kuznetsova et al., 2017; Lüdecke et al., 2021). Variables
256 were standardized prior to analysis to improve comparability.

257 (2) Support for Rensch's rule (i.e., female-biased sexual size dimorphism is larger the smaller the body
258 size) is usually inferred if the slope of an allometric line is significantly >1 (i.e., male size on the y-axis and
259 female size on the x-axis, using standardized major axis regression; Meiri & Liang, 2021; Webb &
260 Freckleton, 2007). A slope equal to 1 would instead indicate an isometric relationship in which sexual size
261 dimorphism does not change in magnitude when overall size of the taxon changes. This effect was tested
262 first by regressing, for each species, female and male mean body length and mean body mass separately.
263 Statistical deviation from regression slopes of 1 was then tested by a standardized major axis regression
264 method as implemented in the "smatr" R package (Warton et al., 2011). For comprehensive illustration
265 purposes, we further estimated a linear model for both variables (i.e., body length and body mass).

266 (3) Conformity with Allen's rule (i.e., body appendages are smaller in colder climates) was tested for ear
267 size (= ear width * ear length) and tail length by taking the mean of the respective variable for each
268 population and extracting six different bioclimatic variables (i.e., mean annual temperature, minimum
269 temperature of the coldest month and maximum temperature of the warmest month, mean annual
270 precipitation, precipitation of wettest and driest month; CHELSA 2.1, Karger et al., 2017). Means per
271 population for each bioclimatic variable were calculated to reduce spatial autocorrelation. We followed
272 the proposed methodology by Alhajeri et al. (2020) who conducted a global test for Allen's rule in
273 rodents. As temperature and precipitation are known to covary across latitudes, including both allows to
274 disentangle purely thermoregulatory aspects from other climatic factors. Mean annual temperature and
275 annual precipitation were not used due to high correlations with other variables ($r \geq 0.84$). Linear models
276 (LMs) were then fitted using tail length and ear size separately as response variables and the log-scaled
277 bioclimatic variables as explanatory variables. To control for body size effects on appendage length, we
278 included head and body length as further explanatory variables (Alhajeri et al., 2020). We also computed
279 a null model excluding the bioclimatic variables. All possible variable combinations were calculated using
280 the "MuMIn" package in R (Barton, 2022), which ranks the best models based on their AIC (Akaike
281 Information Criterion). Only models with $\Delta AIC < 2$ and the null model are reported here. Akaike weights
282 (Wagenmakers & Farrell, 2004) were further calculated using the "qpcR" R package (Spiess, 2018).

283 (4) Conformity with Bergmann's rule (i.e., body size is larger in colder climates) was tested by using body
284 mass and body length separately as response variables and the bioclimatic variables as explanatory
285 variables following the same modeling approach as described above for the test of Allen's rule (Alhajeri
286 et al., 2020; Alhajeri & Steppan, 2016).

288 Results

289 Assessment of data quality

290 After quality filtering, a total of 1,679 individuals (54.6%) remained in our dataset, while the number of
291 individuals represented in the filtered dataset further differed among morphological variables (Figure
292 3A). The sharpest decline in sample size was observed from step 1 to step 2, in which juvenile individuals
293 and those captured outside of the reproductive season were dropped from the dataset. The consecutive
294 steps then led to minor declines in sample size for certain variables, taxa and/or researchers. The highest
295 decline was found for ear and third toe length (reduction of 43.3% and 42.1%, respectively; Figure 3A).
296 Changes in the Shapiro 1-W' statistic, skewness and kurtosis across the five filtering steps were different
297 among variables and across filtering steps (Figure 3). However, after applying our filtering pipeline, the
298 distributions of most variables showed substantially improved quality, i.e., higher conformity to a normal
299 distribution (Figure 3).

300
301 <<< Figure 3 here >>>
302
303 After quality filtering, mean inter-observer reproducibility was $CV < 0.1$ for most variables indicating very
304 good quality (Figure 4), and $CV < 0.2$ for the others (i.e., good quality): particularly reliable variables were
305 head length and width, interorbital distance, lower leg length, hind foot length and tail length. Others
306 like intraorbital distance, snout length and body mass were still associated with higher levels of variation
307 across researchers (Figure 4).

308
309 <<< Figure 4 >>>

310
311
312 **Morphological variation in mouse lemurs**
313 *Sexual dimorphism*
314 Significant and consistent size differences between the two sexes were found for eight out of 13
315 variables, with females being the larger sex concerning head length and width, as well as snout length,
316 hind foot length, body length and tail length. Males exhibited larger ears (i.e., ear length and width).
317 Most pronounced was the difference in ear dimensions, head and body length (Figure 5; Table S3).
318 Significant differences were visible as a congruent direction in male vs. female means in 58.8-84.2% of
319 the studied taxa (Table S3).

320
321 <<< Figure 5 >>>

322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360

1
2
3 323 *Rensch's rule: Female-biased sexual size dimorphism increases with decreasing body size*

4
5 324 The test for Rensch's rule revealed that the regressions between female (x-axis) and male (y-axis) body
6 325 length and mass in log-scale had slopes of $\beta = 0.969$ (95% CI: 0.817-1.149) and $\beta = 1.002$ (95% CI: 0.865-
7 326 1.160), respectively (Figure 6). In both cases, the relationship was not significantly different from 1 (body
8 327 length: test statistic $r = -0.090$, $df = 18$, $P = 0.705$; body mass: $r = 0.006$, $df = 24$, $P = 0.978$) and therefore
9 328 our data does not support Rensch's rule. Female and male body length and body mass (investigated
10 329 separately) were significantly and positively correlated with each other (body length: $r^2 = 0.880$, $P <$
11 330 0.001 , $n = 20$; body mass: $r^2 = 0.878$, $P < 0.001$, $n = 26$, respectively), indicating that taxa with larger and
12 331 heavier males also had larger and heavier females (Figure 6).

13
14
15 332

16
17 333 <<< Figure 6 >>>

18
19 334

20
21 335 *Allen's rule: The size of body appendices decreases in colder climates*

22
23 336 Conformity with Allen's rule was found for tail length but not for ear size. The best model for tail length
24 337 explained 68.3% of the variation by including different bioclimatic variables (Figure 7, Table S4). Tail
25 338 length increased with increasing minimum temperature of the coldest month and decreased with
26 339 increasing precipitation of the wettest and driest month (Figure 7A-C). The best model for ear size
27 340 explained 45.0% of the variation but did not include temperature-related variables. Instead, ear size
28 341 decreased with increasing precipitation of the driest month (Figure 7E). In both cases, null models
29 342 received much larger $\Delta AICc$ values and smaller Akaike weights than the preferred models (Table S4).
30 343 Positive relationships were found for tail length with head length and for ear size with body length
31 344 (Figure 7D, F; Table S4).

32
33
34 345

35
36 346 <<< Figure 7 >>>

37
38 347

39
40 348 *Bergmann's rule: Body size increases in colder climates*

41
42 349 Body length and mass did not support Bergmann's rule. Variation in body mass was best explained by a
43 350 positive relationship with minimum temperature in the coldest month together with precipitation of the
44 351 driest month (Figure 8A-B; Table S4). Variation in body length was best explained by a positive
45 352 relationship with maximum temperature of the warmest month and precipitation in the driest month
46 353 (Figure 8C-D), contrasting predictions under Bergmann's rule. These models explained 23.9% and 20.6%
47 354 of the variation, respectively (Table S4). For body mass, the null model was far worse than the preferred
48 355 models, but for body length, the preferred model was only slightly better than the null model.

49
50
51 356

52
53 357 <<< Figure 8 here >>>

54
55 358

359 Discussion

360 A new methodological toolkit to assess data quality

361 The application of our data quality filtering pipeline resulted in a marked decrease in available individuals
362 and measurements per individual for downstream analyses. However, the process resulted in a strong
363 increase in data quality through improved metrics of data normality after filtering. Excluding non-adult
364 individuals and those from the non-breeding season resulted in the largest loss of available data points
365 (step 1), but notably improved the overall quality as indicated by changes in skewness, kurtosis and
366 Shapiro-Francia W' test statistic. In comparison, filtering steps 2-5 led only to minor changes in the
367 sample size per variable, with the exception of ear length and third toe length. These two variables
368 appeared to be difficult to measure accurately across researchers. Both measures measured particularly
369 fragile structures of mouse lemurs and may therefore be particularly error-prone. For all other variables,
370 steps 2-5 led only to the exclusion of single researcher-variable combinations due to unusually high
371 variation in the measurements or to systematic deviations from measurements made by other
372 researchers. It must be noted that biases could have remained undetected in cases where only one or
373 two researchers provided data for a certain taxon or where inter-observer bias followed a normal
374 distribution.

375 Inter-observer reproducibility differed notably between the 13 different morphological variables, but
376 was within an acceptable quality range for most of them. Body mass, snout length and intraorbital
377 distance had the smallest reproducibility among all variables. Body mass can be measured with high
378 accuracy but is also well known to vary across individuals due to different reasons other than seasonality
379 (e.g., body condition, resource availability; Randrianambinina et al., 2003; Schmid & Kappeler, 1998).
380 Whether this variable is included in subsequent morphological analyses in species like mouse lemurs
381 should be well considered and justified as it introduces a high amount of variation if treated like other
382 variables. In contrast, snout length and intraorbital distance seem to be particularly difficult to measure
383 in mouse lemurs, especially if taken on animals that are not anesthetized. Snout length is difficult to
384 measure accurately since the proximal end point of this measure, the inflection near the inner corner of
385 the eyes, is sometimes not easy to locate. The most reliable measures, i.e., those with low values of
386 inter-observer variation, were head length and width, interorbital distance, lower leg length, hind foot
387 length, and tail length (Figure 4). As morphological variation between cryptic species can be subtle (e.g.,
388 Radespiel et al., 2012; Schüßler et al., 2020; Sgarlata et al., 2019), methodological biases can easily dilute
389 biological variation, especially in multivariate analyses.

390

391 Guidelines for future morphological studies

392 We have presented a pipeline to reduce inter-observer biases after data collection has been completed.
393 Besides illustrating the usefulness of the pipeline for the improvement of data quality, we make the
394 following suggestions for future data collection and a streamlined quality assessment based on our
395 findings:

- 396 1) Multi-observer studies should be designed in such a way that larger sample sizes (morphometric
397 measurements for > 30 individuals; Yusoff & Wah 2012) are collected by several researchers ($n >$
398 3) for each taxon in the dataset, so that inter-observer biases can be reliably identified and
399 reduced.

- 2) Several measures to improve inter-observer reproducibility should be implemented before data collection (e.g., defining standard protocols, co-training of researchers)
- 3) We suggest calculating the inter-observer reproducibility before any further analyses, and to use only those measures that are most reliably measured by the set of researchers involved.
- 4) Measurements should be used that don't rely on ambiguous start or endpoints for caliper placement. Start and endpoints of measures must be clearly defined and identifiable across taxa.
- 5) Sampling protocols and measures undertaken for quality filtering must be transparently reported. Ideally, data quality and inter-observer reproducibility for every variable is briefly outlined to warrant reliability of the results in downstream analyses.

409

410

411 Morphological variation in mouse lemurs

412 *Sex-related differences*

413 Sexual size dimorphism in primates is most likely driven by sexual selection, by natural selection of niche
414 differentiation and by differential growth patterns (Cassini, 2020; Kappeler et al., 2019; Plavcan, 2001;
415 Thomas et al., 2015). Strepsirrhine primates and particularly the lemurs of Madagascar have generally
416 been considered as sexually monomorphic compared to most other primate families (Jenkins & Albrecht,
417 1991; Kappeler, 1991; Plavcan, 2001). Counterexamples were only rarely reported but included few
418 mouse lemur species like *Microcebus murinus* (Kappeler 1991; Kappeler et al., 2022), warranting further
419 investigation. We found a consistent pattern of sexual size dimorphism across the whole genus that was
420 visible in the majority of the studied mouse lemur species, with females being significantly larger in six of
421 the 13 morphological variables investigated. In contrast, males had significantly larger ears than females,
422 and no sexual dimorphism was detected in the remaining five variables.

423 Female dominance is a trait found in a variety of species from all five extant lemur families (Evasoa et al.,
424 2019; Kappeler et al., 2022). However, variation among species exists, suggesting some level of species
425 divergence in this behavioral trait (Evasoa et al., 2019). Although a relationship between body mass
426 dimorphism and female dominance was not supported in a recent review study on 37 lemur species
427 (including nine mouse lemur species; Kappeler et al., 2022), subtle size differences between sexes may
428 have gone undetected due to the non-exclusive confounding effects of small sample size, large error-to-
429 measurement ratio in small animals, inter-observer bias and seasonal variations in body condition
430 (making body mass a rather unreliable variable to assess sexual size dimorphism).

431 Sex differences in skull size may also be explained by sex differences in ecological niches. A larger head
432 size has been empirically associated with increased bite force in *M. murinus* (Chazeau et al., 2012;
433 Thomas et al. 2015). Together with differences in canine length, as found in *M. griseorufus*
434 (Rasoazanabary 2011), this could indicate subtle niche differentiation between male and female mouse
435 lemurs concerning feeding habits or preferred food items. Sex-dependent feeding and foraging
436 strategies have so far not been reported for mouse lemurs. However, female *M. murinus* in Kirindy (W
437 Madagascar) were observed in higher densities around gum trees than males, which was interpreted as a
438 consequence of female dominance (Génin 2003). The solitary foraging mode of the omnivorous mouse
439 lemurs (e.g., Dammhahn & Kappeler 2008; Radespiel et al. 2006) certainly conveys considerable

1
2
3 440 flexibility with respect to individualized or sex-dependent feeding habits. Detailed studies comparing the
4 441 feeding ecology of both sexes across mouse lemur species are needed to further test these hypotheses.

6 442 Interestingly, males had significantly larger ears than females, which may also have biological relevance,
7 443 since mouse lemurs rely extensively on acoustic communication in addition to olfaction (Glatston et al.
8 444 1993; Zimmermann 2018). Sensory experiments with mouse lemurs in in Mandena, SE Madagascar, and
9 445 in captivity showed that auditory cues can initiate and guide insect hunting besides its relevance for
10 446 intraspecific communication such as during mate localization and advertisement (Buesching et al., 1998;
11 447 Goerlitz & Siemers, 2007; Piep et al., 2008; Siemers et al., 2007; Zimmermann & Lerch, 1993). Whether
12 448 the detected differences in ear size are responsible for different acoustic perceptions of males and
13 449 females needs to be investigated in future studies.

16 450 The slight female-biased sexual size dimorphism was uniform across the whole genus as indicated by its
17 451 isometric relationship (slope ≈ 1). Rensch's rule, according to which the ratio of female to male size is
18 452 expected to be larger the smaller the body size of the species (slope > 1), was not supported for mouse
19 453 lemurs. This finding is in line with the absence of Rensch's rule across a larger set of strepsirrhine
20 454 primates (Kappeler, 1990) and with the results of a large-scale study in birds (Webb & Freckleton 2007).
21 455 This deviation from Rensch's rule in mouse lemurs rather suggests an equal selection pressure on male
22 456 and female body size within the genus, as already proposed by Schmid & Kappeler (1998).

25 457

27 458 *Climate-related morphological patterns*

29 459 Climatic constraints are dominant factors shaping the morphology of organisms (Mayr 1942).
30 460 Bergmann's and Allen's ecogeographic rules state that body size and appendages are shaped by
31 461 thermoregulatory principles. Small-bodied endothermic animals like rodents (Alhajeri et al. 2020) and
32 462 birds (Symonds & Tattersall 2010) largely conform to Allen's rule, while showing a more complex
33 463 association with but general conformity to Bergmann's rule (Alhajeri & Stepan 2016; Meiri & Dayan
34 464 2003; Salewski & Watt 2016).

37 465 In mouse lemurs, climatic niche separation has been shown for a few taxa (Kamilar et al., 2016). We
38 466 found statistical support for an influence of climate on morphology: tail length in mouse lemurs followed
39 467 the predictions of Allen's rule. Animals living in warmer and drier environments (i.e., higher
40 468 temperatures during the coldest months, lower precipitation in the driest and wettest months) had
41 469 longer tails than those living in habitats of opposite characteristics. Prevention of heat loss via shorter
42 470 tails is a likely explanation for this observation (Serrat, 2014; Sgarlata et al., 2019). As more than 68% of
43 471 the variation in tail length was explained by climate, this adaptation appears to have a close evolutionary
44 472 link to the surrounding environment.

47 473 Ear size could also be explained by climatic variables but to a much lesser degree than tail length.
48 474 Temperature-related variables, however, were not included in the best performing models. Instead, we
49 475 found that large-eared mouse lemurs were more likely to occur in areas of low amounts of rainfall during
50 476 the dry season. This relationship most likely echoes the divide between the western dry deciduous and
51 477 the eastern rain forests of Madagascar. These two macro-habitats differ in many aspects: dry forests are
52 478 subject to stronger seasonal changes in climatic conditions as well as resource availability and differ in
53 479 vegetation structure and composition from rain forests (Blanco et al., 2015; Génin, 2008). Acoustic
54 480 communication may be particularly constrained in rain forest habitats where high precipitation and high
55 481 vegetation density may lead to sound attenuation and diffusion (Luther & Gentry, 2013; Rasoloharijaona

1
2
3 482 et al., 2010; Taylor et al., 2022). Whether mouse lemur ear sizes are causally linked to the relevance of
4 483 intraspecific acoustic communication in different habitat types or show other adaptations, e.g., to
5 484 different foraging habits, cannot be tested with our dataset but warrants further investigation. The
6 485 difference between small- and large-eared mouse lemurs may further indicate an ancestral phylogenetic
7 486 constraint associated with an ancient split between dry and rain forest taxa (Yoder & Heckman, 2006;
8 487 Vences et al., 2009) and may not necessarily be driven by or highly interlinked with climatic adaptations.

9
10
11 488 In our test of Bergmann's rule, we found a relationship between body mass and length and climatic
12 489 variables, with larger/heavier individuals being more likely to live in warmer and wetter environments.
13 490 This finding contrasts the prediction of Bergmann's rule that individuals from colder environments
14 491 should be larger (Salewski & Watt, 2016). Although the association between the variables was relatively
15 492 low ($R^2 \leq 0.213$), its direction agrees with results from other studies on small mammals in tropical
16 493 environments (Maestri et al., 2016; Rodríguez et al., 2008; Sgarlata et al., 2022). In all cases, body size
17 494 appeared to increase in warmer environments (e.g., at lower altitudes) probably due to an increase in
18 495 resource availability, suggesting that Bergmann's rule may only apply to animals in colder regions but not
19 496 in the tropics (Rodríguez et al. 2008). This also conforms to general patterns of ecogeographic size
20 497 variation across lemurs of all sizes, both living and recently extinct (Albrecht et al., 1990).

21
22
23
24 498 Several mouse lemur species (e.g., *M. berthae* and *M. murinus* at Kirindy, Schmid, 2000; *M. griseorufus*,
25 499 Génin, 2008; *M. lehilahytsara*, Randrianambinina et al., 2003, Andriambelason et al., 2020; *M.*
26 500 *ravelobensis*, Thorén et al., 2011; *M. rufus*, Atsalis, 1999, Ramananjato et al., 2022) can cope with
27 501 thermoregulatory challenges under low temperatures using torpor or even hibernation (Blanco et al.
28 502 2018; Salewski & Watt 2016). However, more than 78% of the variation in body mass and body length
29 503 could not be explained by climatic factors. It is known that resource availability, seasonality and most
30 504 likely also phylogeny are additional drivers of mouse lemur body mass variation (Atsalis 1999; Génin
31 505 2008; Randrianambinina et al. 2003). Phylogenetic constraints on body size of mouse lemurs are
32 506 indicated by the observation that several smaller taxa (*M. berthae*, *M. myoxinus*, *M. lehilahytsara*, *M.*
33 507 *rufus*) form a distinctive clade of small-bodied mouse lemur species (Louis & Lei 2016) that can occur not
34 508 only in sympatry with other larger species but also across a variety of environments in eastern and
35 509 western Madagascar. These include the highly seasonal western dry deciduous forests with their clear
36 510 separation into a 4–6 month hot rainy season and a 6–8-month cooler dry season lacking precipitations
37 511 almost entirely, the central highlands that undergo substantial seasonal temperature fluctuations, and
38 512 the eastern lowland rainforests that exhibit the least climatic variations across the year (Andriambelason
39 513 et al. 2020; Blanco et al. 2015, Evasoa et al. 2018). A comprehensive evaluation of the relative impact of
40 514 environmental variables and phylogenetic constraints on the morphology of mouse lemurs is currently
41 515 hampered by a lack of a fully resolved phylogeny of all currently described mouse lemur lineages.

42
43
44
45
46 516

47 48 517 **Conclusions**

49
50 518 Our new filtering pipeline was developed in order to enable analyzing a large multi-researcher dataset.
51 519 We detailed on the six different filtering steps and illustrated their usefulness for reducing potential
52 520 inter-observer biases. With our generalized suggestions for future morphological studies, we aim to
53 521 improve data collection procedures and overall quality in these kind of studies. Using this pipeline, we
54 522 filtered the largest currently available dataset of morphometric measurements for the species-rich genus
55 523 of mouse lemurs and generated a consolidated dataset with which four biologically meaningful
56 524 hypotheses were tested.

1
2
3 525 Morphometric variation was partly explained by sexual size dimorphism and by various bioclimatic
4 526 variables which were partly in agreement with the predictions of Allen's rule (tail length) but not with
5 527 those of Rensch's or Bergmann's rule. The remaining unexplained variation could be due to yearly and
6 528 monthly variations in resource availability, possible differences in resource use across lineages and
7 529 between the sexes, variation in intraspecific acoustic communication patterns and sociality, as well as
8 530 phylogenetic constraints. Our genus-wide analysis could therefore outline general patterns of
9 531 morphological variation in mouse lemurs. We further generated hypotheses for later testing in order to
10 532 fully understand the drivers of morphological variation within a genus of nocturnal, cryptic and small-
11 533 bodied primates.
12
13
14
15 534
16
17
18 535
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

We sincerely thank the national and regional Malagasy authorities who granted permits to conduct years of research on mouse lemurs across the island. Immense thanks go to all collaborators, students and field assistants who facilitated and contributed to data collection over the years, in particular Augusta, Yves Rostant Andriamalala, Jean-Basile Andriambelason, Bertrand Andriatsitohaina, Henri Andriamampionona, Angelika Beck, Amidou Bensouleimany, Melanie Barnavon, Editho Bellarmain Beandakana, Marius Gabriellah Fanomboavonjy, Vivien Gabbillaud, Laurie R. Godfrey, Thomas Hafen, Mohamed Thani Ibouroi, Irène, Eric Tiana Jacques, Sharon Kessler, Frederik Kiene, Annette Klein, Célia Kun-Rodriguez, Ange Nandrianina, Dhurham Said Ali Oussen, Naina Ratsimba Rabemananjara, Deborah Rabesamihanta, Zaranaina Radavison, Tahiriniaina Radriarimanga, Stephan Michel Rafamantanantsoa, Fils Donna Rakotoarizaka, Sylvestre Rakotomalala, Adolphe Rakotomanantena, Augustin Rakotonandrianina, José M. Ralison, Roger Ramarokoto, Elodi Rambelason, Malcolm S. Ramsay, Roger Doménico Randimbiharirina, Herman Randriamampionona, Francois Jacques Randriamaroson, Herve Donnace Rasamivelona, Hoby Rasoanaivo, Norbert Randrianasolo Padou Cressant Razafindravelo, Louis Joelson Stanislas Razafindraletsy, Valisoa Sinia Torille Rovanirina, Brunot Tsviverimana, Patrice Velomitady, Vera Zietemann, and the late Elke Zimmermann.

552

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

555

556

DATA AVAILABILITY STATEMENT

The quality filtering pipeline, the assessment of data quality (R scripts), and all supplementary information are available at: <https://doi.org/10.25625/6EINJN>

560

561

References

- 562 Agostini, G., Rasoazanabary, E., & Godfrey, L.R. (2017). The befuddling nature of mouse lemur hands
563 and feet at Bezà Mahafaly, SW Madagascar. *American Journal of Primatology*, 79(9), e22680.
- 564 Albrecht, G.H., Jenkins, P.D., & Godfrey, L.R. (1990). Ecogeographic size variation among the living and
565 subfossil prosimians of Madagascar. *American Journal of Primatology*, 22(1), 1-50.
- 566 Alhajeri, B.H., Fourcade, Y., Upham, N.S. & Alhaddad, H. (2020). A global test of Allen's rule in rodents.
567 *Global Ecology and Biogeography*, 29(12), 2248-2260. doi:10.1111/geb.13198
- 568 Alhajeri, B.H. & Stepan, S.J. (2016). Association between climate and body size in rodents: A
569 phylogenetic test of Bergmann's rule. *Mammalian Biology*, 81(2), 219-225.
570 doi:10.1016/j.mambio.2015.12.001
- 571 Andriambeloson, J.B., Greene, L.K. & Blanco, M.B. (2020). Prolonged torpor in Goodman's Mouse Lemur
572 (*Microcebus lehilahytsara*) from the high-altitude forest of Tsinjoarivo, Central-Eastern Madagascar. *Folia*
573 *Primatologica*, 91(6), 697-710. doi:10.1159/000509102
- 574 Aronhime, S., Calcagno, C., Jajamovich, G.H., Dyvorne, H.A., Robson, P., Dieterich, D., Fiel, M.I., Martel-
575 Laferrriere, V., Chatterji, M., Rusinek, H. & Taouli, B. (2013). DCE-MRI of the liver: Effect of linear and
576 nonlinear conversions on hepatic perfusion quantification and reproducibility. *Journal of Magnetic*
577 *Resonance Imaging*, 40(1), 90-98. doi:10.1002/jmri.24341
- 578 Atsalis, S. (1999). Seasonal fluctuations in body fat and activity levels in a rain-forest species of mouse
579 lemur, *Microcebus rufus*. *International Journal of Primatology*, 20(6), 883-910.
580 doi:10.1023/A:1020826502103
- 581 Barton, K. (2022). MuMIn: Multi-Model Inference. R package version 1.47.1 Bates, D., Mächler, M.,
582 Bolker, B. & Walker, S. (2015). Fitting linear mixed effects models using lme4. *Journal of Statistical*
583 *Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01
- 584 Benites, P., Eaton, M.D., Garcia-Trejo, E.A. & Navarro-Sigüenza, A.G. (2020). Environment influences the
585 geographic phenotypic variation in Velazquez's woodpecker (*Centurus santacruzii*). *Journal of*
586 *Ornithology*, 161(3), 621-634. doi:10.1007/s10336-020-01748-x
- 587 Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007).
588 Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22(3), 148-155
- 589 Blanco, M.B., Dausmann, K.H., Faherty, S.L., & Yoder, A.D. (2018). Tropical heterothermy is "cool": The
590 expression of daily torpor and hibernation in primates. *Evolutionary Anthropology: Issues, News, and*
591 *Reviews*, 27(4), 147-161.
- 592 Blanco, M.B., Rasoazanabary, E. & Godfrey, L.R. (2015). Unpredictable environments, opportunistic
593 responses: Reproduction and population turnover in two wild mouse lemur species (*Microcebus rufus*
594 and *M. griseorufus*) from eastern and western Madagascar. *American Journal of Primatology*, 77(9), 936-
595 947.
- 596 Buesching, C.D., Heistermann, M., Hodges, J.K. & Zimmermann, E. (1998). Multimodal oestrus
597 advertisement in a small nocturnal prosimian, *Microcebus murinus*. *Folia Primatologica*, 69(7), 295-308.

- 1
2
3 598 Burghardt, G.M., Bartmess-LeVasseur, J.N., Browning, S.A., Morrison, K.E., Stec, C.L., Zachau, C.E. &
4 599 Freeberg, T.M. (2012). Perspectives - Minimizing observer bias in behavioral studies: A review and
5 600 recommendations. *Ethology*, 118(6), 511-517. doi:10.1111/j.1439-0310.2012.02040.x
- 7 601 Cadena, C.D., Zapata, F. & Jiménez, I. (2018). Issues and perspectives in species delimitation using
8 602 phenotypic data: Atlantean evolution in Darwin's finches. *Systematic Biology*, 67(2), 181-194.
9 603 doi:10.1093/sysbio/syx071
- 11 604 Cassini, M.H. (2020). Sexual size dimorphism and sexual selection in primates. *Mammal Review*, 50(3),
13 605 231-239. doi:10.1111/mam.12191
- 15 606 Chazeau, C., Marchal, J., Hackert, R., Perret, M. & Herrel, A. (2012). Proximate determinants of bite force
16 607 capacity in the mouse lemur. *Journal of Zoology*, 290(1), 42-48. doi:10.1111/jzo.12011
- 18 608 Chornelia, A., Lu, J. & Hughes, A.C. (2022). How to accurately delineate morphologically conserved taxa
19 609 and diagnose their phenotypic disparities: Species delimitation in cryptic *Rhinolophidae* (*Chiroptera*).
20 610 *Frontiers in Ecology and Evolution*, 10. doi:10.3389/fevo.2022.854509
- 22 611 Coimbra, R.T., Winter, S., Kumar, V., Koepfli, K.P., Gooley, R.M., Dobrynin, P., Fennessy, J. & Janke, A.
23 612 (2021). Whole-genome analysis of giraffe supports four distinct species. *Current Biology*.
24 613 doi:10.1016/j.cub.2021.04.033
- 26 614 Dammhahn, M. & Kappeler, P.M. (2008). Small-scale coexistence of two mouse lemur species
27 615 (*Microcebus berthae* and *M. murinus*) within a homogeneous competitive environment. *Oecologia*,
28 616 157(3), 473-483. doi:10.1007/s00442-008-1079-x
- 30 617 Darwin, C. (1859). *On the Origin of Species*. Routledge 2004. doi:10.4324/9780203509104
- 32 618 Davies, L. & Gather, U. (1993). The identification of multiple outliers. *Journal of the American Statistical*
33 619 *Association*, 88(423), 782-792. doi:10.1080/01621459.1993.10476339
- 35 620 De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, 56(6), 879-886.
36 621 doi:10.1080/10635150701701083
- 38 622 den Berghe, S.V. & Gaeveren, K.V. (2017). Data quality assessment and improvement: A vrije universiteit
39 623 brussel case study. *Procedia Computer Science*, 106, 32-38. doi:10.1016/j.procs.2017.03.006
- 41 624 Evasoa, M.R., Zimmermann, E., Hasiniaina, A.F., Rasoloharijaona, S., Randrianambinina, B. & Radespiel,
42 625 U. (2019). Sources of variation in social tolerance in mouse lemurs (*Microcebus* spp.). *BMC Ecology*,
43 626 19(1). doi:10.1186/s12898-019-0236-x
- 45 627 Evasoa, R.M., Radespiel, U., Hasiniaina, A.F., Rasoloharijaona, S., Randrianambinina, B., Rakotondravony,
46 628 R. & Zimmermann, E. (2018). Variation in reproduction of the smallest-bodied primate radiation, the
47 629 mouse lemurs (*Microcebus* spp.): A synopsis. *American Journal of Primatology*, e22874.
48 630 doi:10.1002/ajp.22874
- 50 631 Génin, F. (2003). Female dominance in competition for gum trees in the grey mouse lemur *Microcebus*
51 632 *murinus*. *Revue d'Ecologie, Terre et Vie*, 58(4), 397-410.
- 53 633 Génin, F. (2008). Life in unpredictable environments: First investigation of the natural history of
54 634 *Microcebus griseorufus*. *International Journal of Primatology*, (29), 303-321. doi:10.1007/s10764-008-
55 635 9243-z

- 1
2
3 636 Glatston, A.R. (1979). Reproduction and behaviour of the lesser mouse lemur (*Microcebus murinus*,
4 637 Miller 1777) in captivity. Ph.D. Thesis, University of London
5
6 638 Godfrey, L.R., Lyon, S.K., & Sutherland, M.R. (1993). Sexual dimorphism in large-bodied primates: the
7 639 case of the subfossil lemurs. *American Journal of Physical Anthropology*, 90(3), 315-334.
8
9 640 Goerlitz, H.R., & Siemers, B.M. (2007). Sensory ecology of prey rustling sounds: acoustical features and
10 641 their classification by wild grey mouse lemurs. *Functional Ecology*, 21(1), 143-153.
11
12 642 Gravetter, F. & Wallnau, L. (2014). Essentials of statistics for the behavioral sciences. Belmont, CA.
13 643 Wadsworth, 8 Edition
14
15 644 Gross, J. & Ligges, U. (2015). nortest: Tests for Normality. URL [https://CRAN.R-](https://CRAN.R-project.org/package=nortest)
16 645 [project.org/package=nortest](https://CRAN.R-project.org/package=nortest), R package version 1.0-4
17
18 646 Hafen, T., Neveu, H., Rumpler, Y., Wilden, I. & Zimmermann, E. (1998). Acoustically dimorphic
19 647 advertisement calls separate morphologically and genetically homogenous populations of the grey
20 648 mouse lemur (*Microcebus murinus*). *Folia Primatologica*, 69(1), 342-356. doi:10.1159/000052723
21
22
23 649 Hallgrímsson, B. & Hall, B.K. (2005). *Variation and variability: central concepts in biology*. In Variation.
24 650 Elsevier, (1-7)
25
26 651 Hampel, F.R. (1985). The breakdown points of the mean combined with some rejection rules.
27 652 *Technometrics*, 27(2), 95-107. doi:10.1080/00401706.1985.10488027
28
29 653 Heinrich, B., Hristova, D., Klier, M., Schiller, A. & Szubartowicz, M. (2018). Requirements for data quality
30 654 metrics. *Journal of Data and Information Quality*, 9(2), 1-32. doi:10.1145/3148238
31
32 655 Hopkins, K.D. & Weeks, D.L. (1990). Tests for normality and measures of skewness and kurtosis: Their
33 656 place in research reporting. *Educational and Psychological Measurement*, 50(4), 717-729.
34 657 doi:10.1177/0013164490504001
35
36 658 Hotaling, S., Foley, M.E., Lawrence, N.M., Bocanegra, J., Blanco, M.B., Rasoloarison, R., Kappeler, P.M.,
37 659 Barrett, M.A., Yoder, A.D. & Weisrock, D.W. (2016). Species discovery and validation in a cryptic radiation
38 660 of endangered primates: coalescent-based species delimitation in Madagascar's mouse lemurs.
39 661 *Molecular Ecology*, 25, 2029-2045. doi:10.1111/mec.13604
40
41
42 662 Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130(1), 195-204
43
44 663 Howard, B.R. (2002). Control of variability. *ILAR Journal*, 43(4), 194-201. doi:10.1093/ilar.43.4.194
45
46 664 J, L. (2006). Plotrix: a package in the red light district of r. *R-News*, 6(4), 8-12.
47
48 665 Jenkins, P.D. & Albrecht, G.H. (1991). Sexual dimorphism and sex ratios in Madagascan prosimians.
49 666 *American Journal of Primatology*, 24(1), 1-14. doi:10.1002/ajp.1350240102
50
51 667 Jennions, M.D., Moller, A.P. & Petrie, M. (2001). Sexually selected traits and adult survival: A meta-
52 668 analysis. *The Quarterly Review of Biology*, 76(1), 3-36. doi:10.1086/393743
53
54 669 Kamilar, J. M., Blanco, M. B., & Muldoon, K. M. (2016). Ecological niche modeling of mouse lemurs
55 670 (*Microcebus* spp.) and its implications for their species diversity and biogeography. In S.M. Lehman, U.
56 671 Radespiel & E. Zimmermann (Editors), *The Dwarf and Mouse Lemurs of Madagascar: Biology, Behavior*
57 672 *and Conservation Biogeography of the Cheirogaleidae*, Chapter 3. Cambridge University Press, (451-463).
58
59
60

- 1
2
3 673 Kamilar, J. & Pokempner, A. (2008). Does body mass dimorphism increase male-female dietary niche
4 674 separation? a comparative study of primates. *Behaviour*, 145(9), 1211-1234.
5 675 doi:10.1163/156853908785387601
- 6
7 676 Kappeler, P.M. (1990). The evolution of sexual size dimorphism in prosimian primates. *American Journal*
8 677 *of Primatology*, 21(3), 201-214. doi:10.1002/ajp.1350210304
- 9
10 678 Kappeler, P.M. (1991). Patterns of sexual dimorphism in body weight among prosimian primates. *Folia*
11 679 *Primatologica*, 57(3), 132-146. doi:10.1159/000156575
- 12
13 680 Kappeler, P.M., Fichtel, C. & Radespiel, U. (2022). The island of female power? Intersexual dominance
14 681 relationships in the lemurs of Madagascar. *Frontiers in Ecology and Evolution*, 10.
15 682 doi:10.3389/fevo.2022.858859
- 16
17 683 Kappeler, P.M., Nunn, C.L., Vining, A.Q., & Goodman, S.M. (2019). Evolutionary dynamics of sexual size
18 684 dimorphism in non-volant mammals following their independent colonization of Madagascar. *Scientific*
19 685 *Reports*, 9(1), 1-14. doi: 10.1038/s41598-018-36246-x
- 20
21 686 Karger, D.N., Conrad, O., Böhrer, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder,
22 687 H.P. & Kessler, M. (2017). Data from: Climatologies at high resolution for the earth's land surface areas.
23 688 doi:10.5061/DRYAD.KD1D4
- 24
25 689 Kraus, C., Eberle, M. & Kappeler, P.M. (2008). The costs of risky male behaviour: sex differences in
26 690 seasonal survival in a small sexually monomorphic primate. *Proceedings of the Royal Society B: Biological*
27 691 *Sciences*, 275(1643), 1635-1644. doi:10.1098/rspb.2008.0200
- 28
29 692 Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. (2017). lmerTest package: Tests in linear mixed
30 693 effects models. *Journal of Statistical Software*, 82(13), 1-26. doi:10.18637/jss.v082.i13
- 31
32 694 Lahann, P., Schmid, J. & Ganzhorn, J.U. (2006). Geographic variation in populations of *Microcebus*
33 695 *murinus* in Madagascar: Resource seasonality or Bergmann's rule? *International Journal of Primatology*,
34 696 27(4), 983-999. doi:10.1007/s10764-006-9055-y
- 35
36 697 Leache, A., Zhu, T., Rannala, B. & Yang, Z. (2019). The spectre of too many species. *Systematic Biology*,
37 698 68(1), 168-181. doi:10.1093/sysbio/syy051
- 38
39 699 Liu, H., Shah, S. & Jiang, W. (2004). On-line outlier detection and data cleaning. *Computers & Chemical*
40 700 *Engineering*, 28(9), 1635-1647. doi:10.1016/j.compchemeng.2004.01.009
- 41
42 701 Louis, E.E. & Lei, R. (2016). Mitogenomics of the family Cheirogaleidae and relationships to taxonomy
43 702 and biogeography in Madagascar. In S.M. Lehman, U. Radespiel & E. Zimmermann (Editors), *The Dwarf*
44 703 *and Mouse Lemurs of Madagascar: Biology, Behavior and Conservation Biogeography of the*
45 704 *Cheirogaleidae*, Chapter 3. Cambridge University Press, (54-93)
- 46
47 705 Lüdecke, D., Ben-Shachar, M.S., Patil, I., Waggoner, P. & Makowski, D. (2021). performance: An R
48 706 Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*,
49 707 6(60), 3139. doi:10.21105/joss.03139
- 50
51 708 Lutermann, H., Schmelting, B., Radespiel, U., Ehresmann, P. & Zimmermann, E. (2006). The role of
52 709 survival for the evolution of female philopatry in a solitary forager, the grey mouse lemur (*Microcebus*
53
54
55
56
57
58
59
60

- 1
2
3 710 *murinus*). *Proceedings of the Royal Society B: Biological Sciences*, 273(1600), 2527-2533.
4 711 doi:10.1098/rspb.2006.3603
5
6 712 Luther, D. & Gentry, K. (2013). Sources of background noise and their influence on vertebrate acoustic
7 713 communication. *Behaviour*, 150(9-10), 1045-1068. doi:10.1163/1568539x-00003054
8
9 714 Macdonald, S.L., Llewelyn, J. & Phillips, B.L. (2017). Using connectivity to identify climatic drivers of local
10 715 adaptation. *Ecology Letters*, 21(2), 207-216. doi:10.1111/ele.12883
11
12 716 Maestri, R., Luza, A.L., de Barros, L.D., Hartz, S.M., Ferrari, A., de Freitas, T.R.O. & Duarte, L.D.S. (2016).
13 717 Geographical variation of body size in sigmodontine rodents depends on both environment and
14 718 phylogenetic composition of communities. *Journal of Biogeography*, 43(6), 1192-1202.
15 719 doi:10.1111/jbi.12718
16
17
18 720 Mayr, E. (1942). *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard
19 721 University Press.
20
21 722 Mbah, A.K. & Paothong, A. (2014). Shapiro-Francia test compared to other normality test using expected
22 723 p-value. *Journal of Statistical Computation and Simulation*, 85(15), 3002-3016.
23 724 doi:10.1080/00949655.2014.947986
24
25 725 Meiri, S. & Dayan, T. (2003). On the validity of Bergmann's rule. *Journal of Biogeography*, 30(3), 331-351.
26 726 doi:10.1046/j.1365-2699.2003.00837.x
27
28 727 Meiri, S. & Liang, T. (2021). Rensch's rule - Definitions and statistics. *Global Ecology and Biogeography*,
29 728 30(3), 573-577. doi:10.1111/geb.13255
30
31 729 Michaloudi, E., Papakostas, S., Stamou, G., Neděla, V., Tihlaíková, E., Zhang, W. & Declerck, S.A.J. (2018).
32 730 Reverse taxonomy applied to the *Brachionus calyciflorus* cryptic species complex: Morphometric analysis
33 731 conforms species delimitations revealed by molecular phylogenetic analysis and allows the
34 732 (re)description of four species. *PLOS ONE*, 13(9), e0203168. doi:10.1371/journal.pone.0203168
35
36
37 733 Mittermeier, R.A., Louis, E.E.J., Richardson, M., Schwitzer, C., Langrand, O., Rylands, A.B., Hawkins, F.,
38 734 Rajaobelina, S., Ratsimbazafy, J., Rasoloarison, R., Roos, C., Kappeler, P. & Mackinnon, J. (2010). *Lemurs*
39 735 *of Madagascar*. Conservation International, 3. Edition
40
41 736 Morrison, L.W. (2015). Observer error in vegetation surveys: a review. *Journal of Plant Ecology*, 9(4), 367-
42 737 379. doi:10.1093/jpe/rtv077
43
44 738 Padiá, J.M., Miralles, A., De la Riva, I. & Vences, M. (2010). The integrative future of taxonomy. *Frontiers*
45 739 *in Zoology*, 7(1), 16. doi:10.1186/1742-9994-7-16
46
47 740 Pearson, R. (2002). Outliers in process modeling and identification. *IEEE Transactions on Control Systems*
48 741 *Technology*, 10(1), 55-63. doi:10.1109/87.974338
49
50 742 Pearson, R.K. (2005). *Mining imperfect data: Dealing with contamination and incomplete records*. SIAM
51
52 743 Pélabon, C., Hilde, C.H., Einum, S. & Gamelon, M. (2020). On the use of the coefficient of variation to
53 744 quantify and compare trait variation. *Evolution Letters*, 4(3), 180-188. doi:10.1002/evl3.171
54
55 745 Piep, M., Radespiel, U., Zimmermann, E., Schmidt, S. & Siemers, B.M. (2008). The sensory basis of prey
56 746 detection in captive-born grey mouse lemurs, *Microcebus murinus*. *Animal Behaviour*, 75(3), 871-878.
57
58
59
60

- 1
2
3 747 Pilloud, M.A., Adams, D.M. & Hefner, J.T. (2018). Observer error and its impact on ancestry estimation
4 748 using dental morphology. *International Journal of Legal Medicine*, 133(3), 949-962. doi:10.1007/s00414-
5 749 018-1985-3
- 7 750 Plavcan, J.M. (2001). Sexual dimorphism in primate evolution. *American Journal of Physical*
8 751 *Anthropology*, 116(S33), 25-53. doi:10.1002/ajpa.10011
- 10 752 Poelstra, J.W., Salmons, J., Tiley, G.P., Schüßler, D., Blanco, M.B., Andriambeloson, J.B., Bouchez, O.,
11 753 Campbell, C.R., Etter, P.D., Hohenlohe, P.A., Hunnicutt, K.E., Iribar, A., Johnson, E.A., Kappeler, P.M.,
12 754 Larsen, P.A., Manzi, S., Ralison, J.M., Randrianambinina, B., Rasoloarison, R.M., Rasolofoson, D.W.,
13 755 Stahlke, A.R., Weisrock, D.W., Williams, R.C., Chikhi, L., Louis, E.E., Radespiel, U. & Yoder, A.D. (2021).
14 756 Cryptic patterns of speciation in cryptic primates: Microendemic mouse lemurs and the multispecies
15 757 coalescent. *Systematic Biology*, 70(2), 203-218. doi:10.1093/sysbio/syaa053
- 18 758 Radespiel, U., Olivieri, G., Rasolofoson, D.W., Rakotondratsimba, G., Rakotonirainy, O., Rasoloharijaona,
19 759 S., Randrianambinina, B., Ratsimbazafy, J.H., Ratelolahy, F., Randriamboavonjy, T., Rasolofoharivelo, T.,
20 760 Craul, M., Rakotozafy, L. & Randrianarison, R.M. (2008). Exceptional diversity of mouse lemurs
21 761 (*Microcebus* spp.) in the Makira region with the description of one new species. *American Journal of*
22 762 *Primatology*, 70(11), 1033-1046. doi:10.1002/ajp.20592
- 25 763 Radespiel, U., Ratsimbazafy, J.H., Rasoloharijaona, S., Raveloson, H., Andriaholinirina, N.,
26 764 Rakotondravony, R., Randrianarison, R.M. & Randrianambinina, B. (2012). First indications of a highland
27 765 specialist among mouse lemurs (*Microcebus* spp.) and evidence for a new mouse lemur species from
28 766 eastern Madagascar. *Primates*, 53(2), 157-170. doi:10.1007/s10329-011-0290-2
- 30 767 Radespiel, U., Reimann, W., Rahelinirina, M. & Zimmermann, E. (2006). Feeding ecology of sympatric
31 768 mouse lemur species in northwestern Madagascar. *International Journal of Primatology*, 27(1), 311-321.
32 769 doi:10.1007/s10764-005-9005-0
- 35 770 Radespiel, U., Sarikaya, Z., Zimmermann, E. & Bruford, M.W. (2001). Sociogenetic structure in a free-
36 771 living nocturnal primate population: sex-specific differences in the grey mouse lemur (*Microcebus*
37 772 *murinus*). *Behavioral Ecology and Sociobiology*, 50(6), 493-502. doi:10.1007/s002650100402
- 39 773 Ramananjato, V., Randimbarison, F., Rabarijaonina, T.H.N.P. & Razafindratsima, O.H. (2022). Arboreal
40 774 mouse lemurs discovered sleeping in a burrow on the ground. *Ecology and Evolution*. doi:
41 775 10.1002/ece3.9543
- 43 776 Randrianambinina, B., Rakotondravony, D., Radespiel, U. & Zimmermann, E. (2003). Seasonal changes in
44 777 general activity, body mass and reproduction of two small nocturnal primates: a comparison of the
45 778 golden brown mouse lemur (*Microcebus ravelobensis*) in Northwestern Madagascar and the brown
46 779 mouse lemur (*Microcebus rufus*) in Eastern Madagascar. *Primates*, 44(4), 321-331. doi:10.1007/s10329-
47 780 003-0046-8
- 50 781 Rasoazanabary, E. (2011). The human factor in mouse lemur (*Microcebus griseorufus*) conservation:
51 782 Local resource utilization and habitat disturbance at Beza Mahafaly, SW Madagascar. Ph.D. Thesis,
52 783 University of Massachusetts Amherst
- 54 784 Rasoloarison, R.M., Goodman, S.M. & Ganzhorn, J.U. (2000). Taxonomic revision of mouse lemurs
55 785 (*Microcebus*) in the western portions of Madagascar. *International Journal of Primatology*, 21(6), 963-
56 786 1019. doi:10.1023/A:1005511129475

- 1
2
3 787 Rasoloharijaona, S., Randrianambinina, B. & Joly-Radko, M. (2010). Does nonnutritive tree gouging in a
4 788 rainforest-dwelling lemur convey resource ownership as does loud calling in a dry forest-dwelling lemur?
5 789 *American Journal of Primatology*, 72(12), 1062-1072. doi:10.1002/ajp.20865
6
7 790 Rensch, B. (1950). Die Abhängigkeit der relativen Sexualdifferenz von der Körpergrösse. *Bonner*
8 791 *Zoologische Beiträge* 1, 58-69
9
10 792 Rodríguez, M.Á., Olalla-Tárraga, M.Á. & Hawkins, B.A. (2008). Bergmann's rule and the geography of
11 793 mammal body size in the Western Hemisphere. *Global Ecology and Biogeography*, 17(2), 274-283.
12 794 doi:10.1111/j.1466-8238.2007.00363.x
13
14 795 Roitberg, E.S., Orlova, V.F., Kuranova, V.N., Bulakhova, N.A., Zinenko, O.I., Ljubisavljevic, K., Shamgunova,
15 796 R.R., Carretero, M.A., Clasen, A., Fokt, M. & Böhme, W. (2011). Inter-observer and intra-observer
16 797 differences in measuring body length: a test in the common lizard, *Zootoca vivipara*. *Amphibia-Reptilia*,
17 798 32(4), 477-484. doi:10.1163/156853811x601636
18
19 799 Royston, P. (1991). Estimating departure from normality. *Statistics in Medicine*, 10(8), 1283-1293.
20 800 doi:10.1002/sim.4780100811
21
22 801 Salewski, V. & Watt, C. (2016). Bergmann's rule: A biophysiological rule examined in birds. *Oikos*, 126(2).
23 802 doi:10.1111/oik.03698
24
25 803 Schmid, J. (2000). Daily torpor in the gray mouse lemur (*Microcebus murinus*) in Madagascar: Energetic
26 804 consequences and biological significance. *Oecologia*, 123(2), 175-183. doi:10.1007/s004420051003
27
28 805 Schmid, J. & Kappeler, P.M. (1998). Fluctuating sexual dimorphism and differential hibernation by sex in
29 806 a primate, the gray mouse lemur (*Microcebus murinus*). *Behavioral Ecology and Sociobiology*, 43(2), 125-
30 807 132
31
32 808 Schüßler, D., Blanco, M.B., Salmona, J., Poelstra, J., Andriambeloson, J.B., Miller, A., Randrianambinina,
33 809 B., Rasolofoson, D.W., Mantilla-Contreras, J., Chikhi, L., Louis, E.E., Yoder, A.D. & Radespiel, U. (2020).
34 810 Ecology and morphology of mouse lemurs (*Microcebus* spp.) in a hotspot of microendemism in
35 811 northeastern Madagascar, with the description of a new species. *American Journal of Primatology*, 82(9),
36 812 e23180. doi:10.1002/ajp.23180
37
38 813 Serrat, M.A. (2014). Environmental temperature impact on bone and cartilage growth. *Comprehensive*
39 814 *Physiology*, 4, 621-655. doi:10.1002/cphy.c130023
40
41 815 Sgarlata, G.M., Rasolondraibe, E., Salmona, J., Pors, B.L., Ralantoharijaona, T., Rakotonanahary, A., Jan,
42 816 F., Manzi, S., Iribar-Pelozuelo, A., Zaonarivelo, J.R., Andriaholinirina, N.V., Rasoloharijaona, S. & Chikhi, L.
43 817 (2022). The genomic diversity of the *Eliurus* genus in northern Madagascar with a putative new species.
44 818 *BioRxiv*. doi:10.1101/2022.10.21.513246
45
46 819 Sgarlata, G.M., Salmona, J., Pors, B.L., Rasolondraibe, E., Jan, F., Ralantoharijaona, T., Rakotonanahary,
47 820 A., Randriamaroson, J., Marques, A.J., Aleixo-Pais, I., Zoeten, T., Oussen, D.S.A., Knoop, S.B., Teixeira, H.,
48 821 Gabillaud, V., Miller, A., Ibouroi, M.T., Rasoloharijaona, S., Zaonarivelo, J.R., Andriaholinirina, N.V. &
49 822 Chikhi, L. (2019). Genetic and morphological diversity of mouse lemurs (*Microcebus* spp.) in northern
50 823 Madagascar: The discovery of a putative new species? *American Journal of Primatology*, 81(12).
51 824 doi:10.1002/ajp.23070
52
53
54
55
56
57
58
59
60

- 1
2
3 825 Shapiro, S.S. & Francia, R.S. (1972). An approximate analysis of variance test for normality. *Journal of the*
4 826 *American Statistical Association*, 67(337), 215-216. doi:10.1080/01621459.1972.10481232
- 6 827 Shirley, M.H., Vliet, K.A., Carr, A.N. & Austin, J.D. (2014). Rigorous approaches to species delimitation
7 828 have significant implications for African crocodylian systematics and conservation. *Proceedings of the*
8 829 *Royal Society B: Biological Sciences*, 281(1776), 20132483. doi:10.1098/rspb.2013.2483
- 10 830 Siemers, B.M., Goerlitz, H.R., Robsomanitrاندrasana, E., Piep, M., Ramanamanjato, J.B.,
11 831 Rakotondravony, D., Ramilijaona, O. & Ganzhorn, J.U. (2007). Sensory basis of food detection in wild
12 832 *Microcebus murinus*. *International Journal of Primatology*, 28(2), 291-304. doi:10.1007/s10764-007-
13 833 9135-7
- 16 834 Symonds, M.R.E. & Tattersall, G.J. (2010). Geographical variation in bill size across bird species provides
17 835 evidence for Allen's rule. *The American Naturalist*, 176(2), 188-197. doi:10.1086/653666Taylor, M.C.,
18 836 Travouillon, K.J., Andrew, M.E., Fleming, P.A. & Warburton, N.M. (2022). Keeping an ear out: size
19 837 relationship of the tympanic bullae and pinnae in bandicoots and bilbies (Marsupialia:
20 838 Peramelemorphia). *Current Zoology*, 68(3), 251-264.
- 22 839 Templeton, A.R. (2006). *Population Genetics and Microevolutionary Theory*. John Wiley & Sons, Inc.
23 840 doi:10.1002/0470047356
- 26 841 Thomas, P., Pouydebat, E., Hardy, I., Aujard, F., Ross, C.F. & Herrel, A. (2015). Sexual dimorphism in bite
27 842 force in the grey mouse lemur. *Journal of Zoology*, 296(2), 133-138. doi:10.1111/jzo.12225
- 29 843 Thompson, J.N. (2005). *The Geographic Mosaic of Coevolution*. University of Chicago Press
- 31 844 Thorén, S., Quietzsich, F., Schwochow, D., Sehen, L., Meusel, C., Meares, K., & Radespiel, U. (2011).
32 845 Seasonal changes in feeding ecology and activity patterns of two sympatric mouse lemur species, the
33 846 gray mouse lemur (*Microcebus murinus*) and the golden-brown mouse lemur (*M. ravelobensis*), in
34 847 northwestern Madagascar. *International Journal of Primatology*, 32(3), 566-586
- 36 848 Tiley, G.P., van Elst, T., Teixeira, H., Schüßler, D., Salmons, J., Blanco, M.B., Ralison, J.M.,
37 849 Randrianambinina, B., Rasoloarison, R.M., Stahlke, A.R., Hohenlohe, P.A., Chikhi, L., Louis, E.E., Radespiel,
38 850 U. & Yoder, A.D. (2022). Population genomic structure in Goodman's mouse lemur reveals long-standing
39 851 separation of Madagascar's Central Highlands and eastern rainforests. *Molecular Ecology*, 31, 4901-4918.
40 852 doi:10.1111/mec.16632
- 42 853 Trochim, W. & Donnelly, J. (2006). *The research methods knowledge base*. 3rd Edition.
- 44 854 Vences, M., Wollenberg, K.C., Vieites, D.R. & Lees, D.C. (2009). Madagascar as a model region of species
45 855 diversification. *Trends in Ecology & Evolution*, 24(8), 456-465
- 47 856 Wagenmakers, E.J. & Farrell, S. (2004). AIC model selection using akaike weights. *Psychonomic Bulletin &*
48 857 *Review*, 11(1), 192-196. doi:10.3758/bf03206482
- 50 858 Warton, D.I., Duursma, R.A., Falster, D.S. & Taskinen, S. (2011). smatr 3: R package for estimation and
51 859 inference about allometric lines. *Methods in Ecology and Evolution*, 3(2), 257-259. doi:10.1111/j.2041-
52 860 210x.2011.00153.x
- 55 861 Webb, T.J. & Freckleton, R.P. (2007). Only half right: Species with female-biased sexual size dimorphism
56 862 consistently break Rensch's rule. *PLOS ONE*, 2(9), e897. doi:10.1371/journal.pone.0000897

- 1
2
3 863 Williamson, R.E., Webb, S.E., Dubreuil, C., Lopez, R., Hernandez, S.C., Fedigan, L.M. & Melin, A.D. (2021).
4 864 Sharing spaces: niche differentiation in diet and substrate use among wild capuchin monkeys. *Animal*
5 865 *Behaviour*, 179, 317-338. doi:10.1016/j.anbehav.2021.06.002
6
7 866 Wright, S. (1984). *Evolution and the genetics of populations: Variability within and among natural*
8 867 *populations*, vol. 4. University of Chicago Press.
9
10 868 Yoder, A.D. & Heckman, K.L. (2006). Mouse lemur phylogeography revises a model of ecogeographic
11 869 constraint in Madagascar. In *Primate Biogeography*. Springer US, (255-268). doi:10.1007/0-387-31710-49
12
13 870 Yusoff, S.b. & Wah, Y.B. (2012). Comparison of conventional measures of skewness and kurtosis for small
14 871 sample size. In *International Conference on Statistics in Science, Business and Engineering (ICSSBE)*. IEEE.
15 872 doi:10.1109/icssbe.2012.6396619
16
17 873 Zimmermann, E. (2018). High frequency/ultrasonic communication in basal primates: The mouse and
18 874 dwarf lemurs of Madagascar. In *Handbook of Ultrasonic Vocalization - A Window into the Emotional*
19 875 *Brain*. Elsevier, (521-533). doi:10.1016/b978-0-12-809600-0.00048-2
20
21 876 Zimmermann, E., Cepok, S., Rakotoarison, N., Zietemann, V. & Radespiel, U. (1998). Sympatric mouse
22 877 lemurs in north-west Madagascar: A new rufous mouse lemur species (*Microcebus ravelobensis*). *Folia*
23 878 *Primatologica*, 69(2), 106-114.
24
25 879 Zimmermann, E. & Lerch, C. (1993). The complex acoustic design of an advertisement call in male mouse
26 880 lemurs (*Microcebus murinus*, Prosimii, Primates) and sources of its variation. *Ethology*, 93(3), 211-224.
27
28 881 Zimmermann, E. & Radespiel, U. (2014). Species concepts, diversity, and evolution in primates: Lessons
29 882 to be learned from mouse lemurs. *Evolutionary Anthropology: Issues, News, and Reviews*, 23(1), 11-14.
30 883 doi:10.1002/evan.21388
31
32
33
34 884
35
36 885
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

886 Figure captions

887
888 **Figure 1:** Distribution of *Microcebus* spp. populations sampled across Madagascar. Overlapping symbols
889 were rarified to improve display. Minimum temperature of the coldest month according to Karger et al.
890 (2017) is shown in shadings from white to black.

891
892 **Figure 2:** Illustration of steps 4 and 5 of the generalized data filtering pipeline. A: Boxplots per lineage
893 and researcher for one variable (hind foot length), boxes encircled in red indicate examples outlined in
894 Table 1. B and C: Histograms for ear length and interorbital distance, respectively, with truncation
895 threshold (median + 3 * median absolute deviation) introduced as vertical dashed red line. Truncation
896 (i.e., the exclusion of measurements for this variable for the respective researcher) takes place in B but
897 not in C.

898
899 **Figure 3:** Quality assessment following the five data filtering steps indicated by different color code in B-
900 D. A: Completeness of the dataset refers to the number of available values per variable compared to the
901 number of individuals in the raw data ($n = 3,073$). B: $1-W'$ measure of normality, C: absolute skewness, D:
902 absolute kurtosis. Values closer to zero in the ridge plots B-D indicate increased conformity to a normal
903 distribution. For skewness (C) and kurtosis (D), the range of acceptable values is indicated with the
904 vertical line at 2 units from 0.

905
906 **Figure 4:** Inter-observer reproducibility quantified as coefficient of variation (CV) across the 13 different
907 morphometric parameters. Only data after quality filtering was included. The horizontal line indicates
908 the upper limit of the CV values interpreted as very good reproducibility (CV=0.0-0.1, Aronhime et al.,
909 2013).

910
911 **Figure 5:** Sexual size dimorphism across all *Microcebus* taxa displayed via model estimates with 95%
912 confidence intervals. The differences in all variables except inter- and intraorbital distances, lower leg
913 length, third toe length, and body mass were statistically significant ($P \leq 0.029$; Table S3).

914
915 **Figure 6:** Relationship between female and male body length (A) and body mass (B). Rensch's rule is not
916 supported in both cases ($P \geq 0.705$). Red line: linear model for observed data (including 95% confidence
917 interval), dashed line: slope = 1. Note: model estimates slightly differ between the standardized major
918 axis and the linear model; the former was used for hypothesis testing and the latter for visualization.

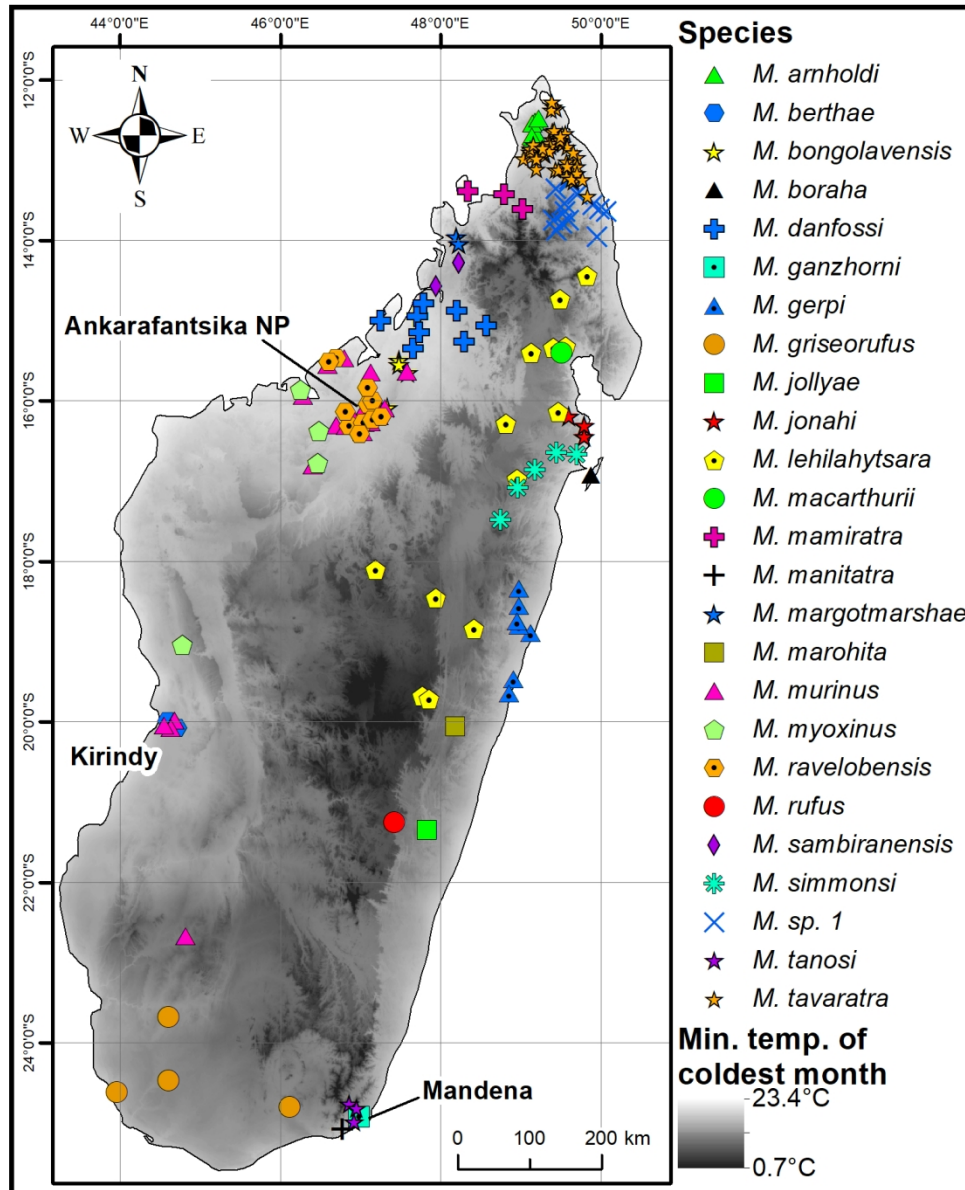
919
920 **Figure 7:** Partial regression plots for the linear model best explaining the variation in tail length (A-D, $R^2 =$
921 0.683) and ear size (E-F, $R^2 = 0.450$) across mouse lemur taxa. Note: the y-axis represents the marginal
922 impact of the response variable on the explanatory variable.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

923

924 **Figure 8:** Partial regression plots for the linear models best explaining the variation in body mass (A-B, R^2
925 = 0.239) and body length (C-D, $R^2 = 0.206$) across mouse lemur taxa. Note: the y-axis represents the
926 marginal impact of the response variable on the explanatory variable.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Distribution of *Microcebus* spp. populations sampled across Madagascar. Overlapping symbols were rarified to improve display. Minimum temperature of the coldest month according to Karger et al. (2017) is shown in shadings from white to black.

134x165mm (300 x 300 DPI)

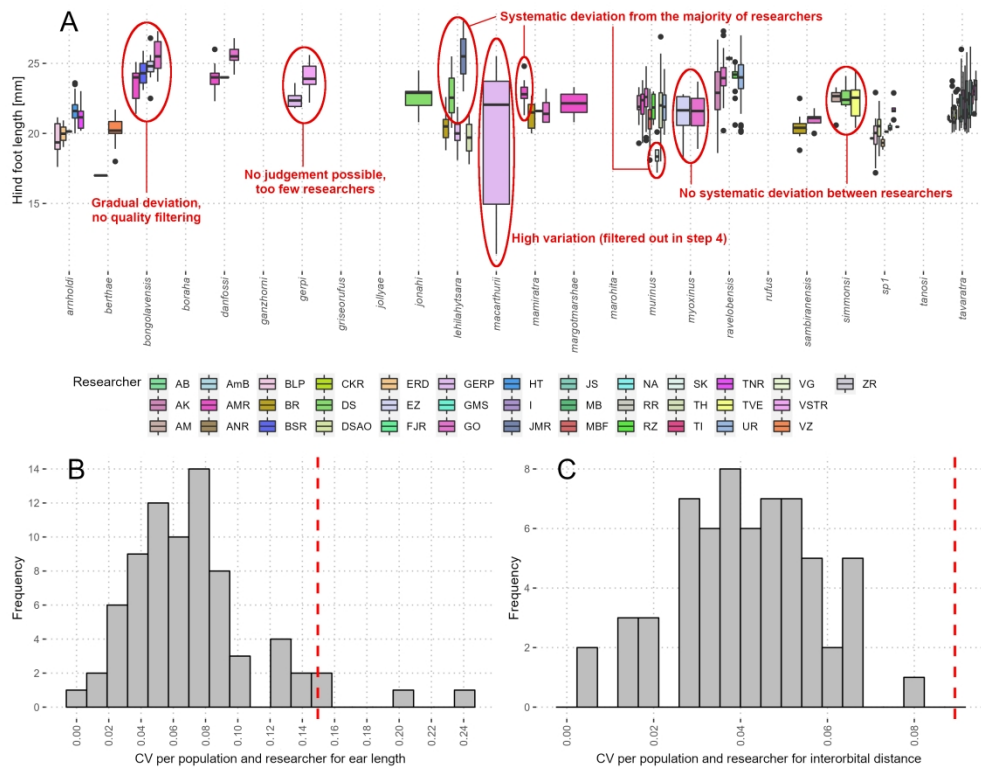
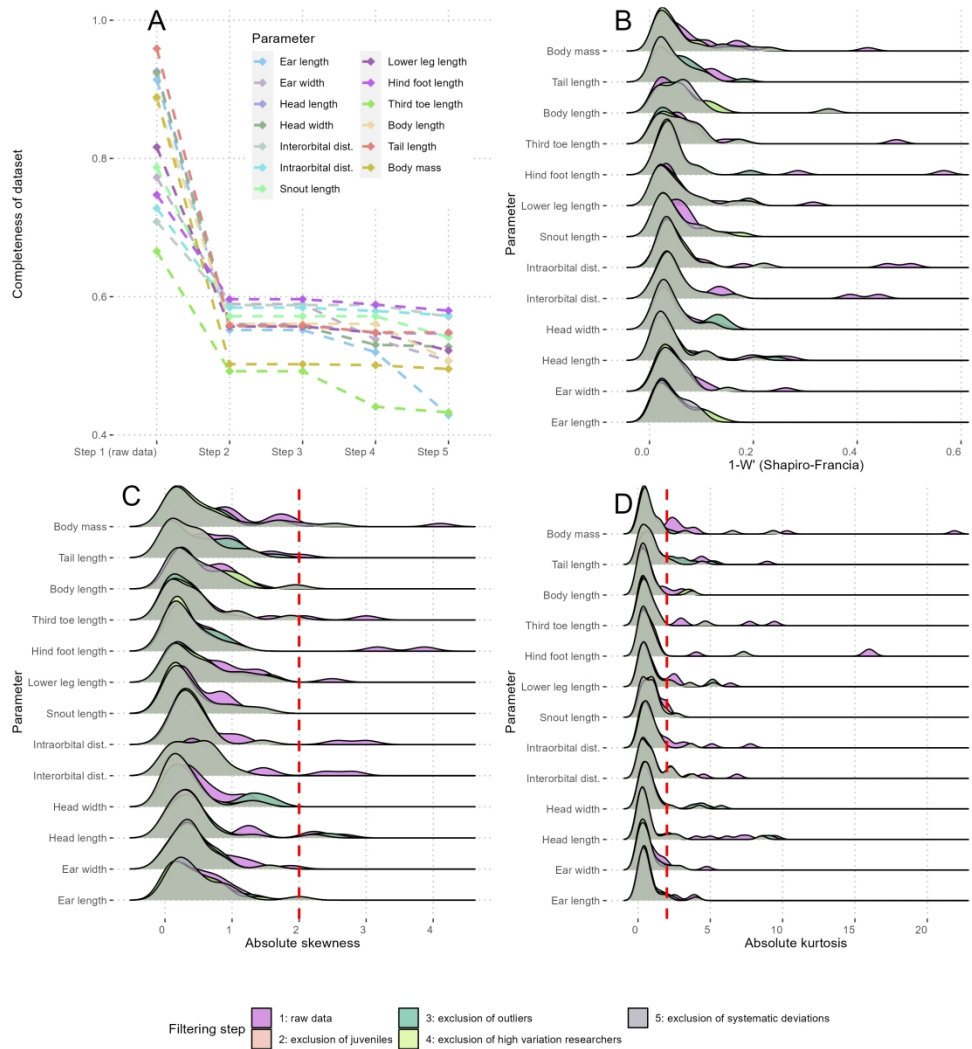


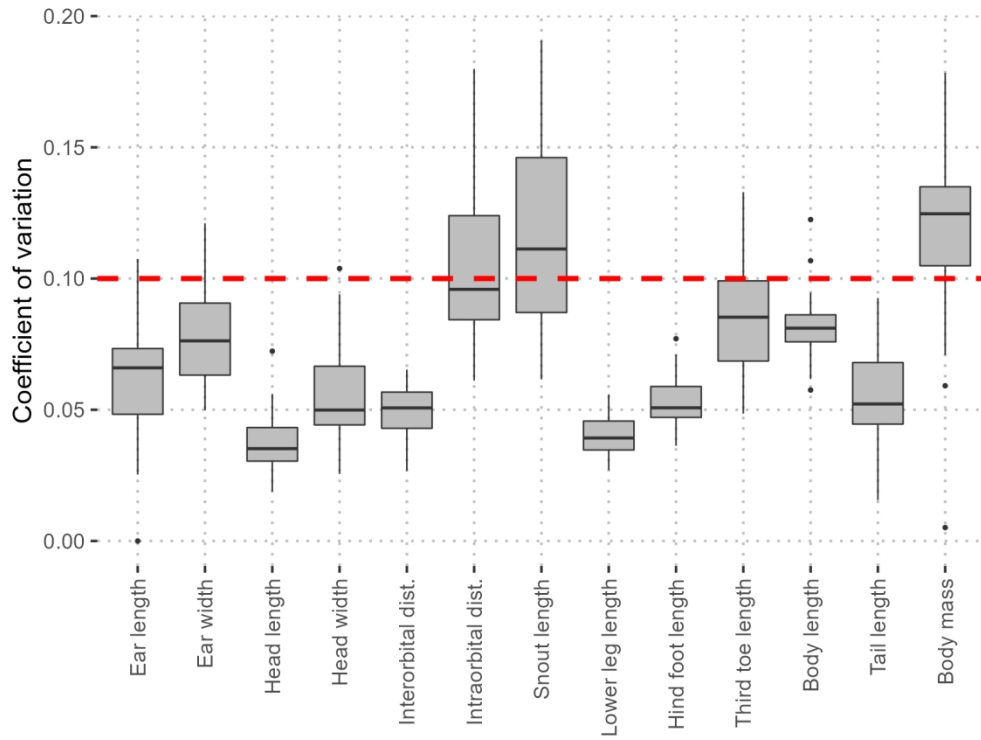
Illustration of steps 4 and 5 of the generalized data filtering pipeline. A: Boxplots per lineage and researcher for one variable (hind foot length), boxes encircled in red indicate examples outlined in Table 1. B and C: Histograms for ear length and interorbital distance, respectively, with truncation threshold (median + 3 * median absolute deviation) introduced as vertical dashed red line. Truncation (i.e., the exclusion of measurements for this variable for the respective researcher) takes place in B but not in C.

287x223mm (300 x 300 DPI)



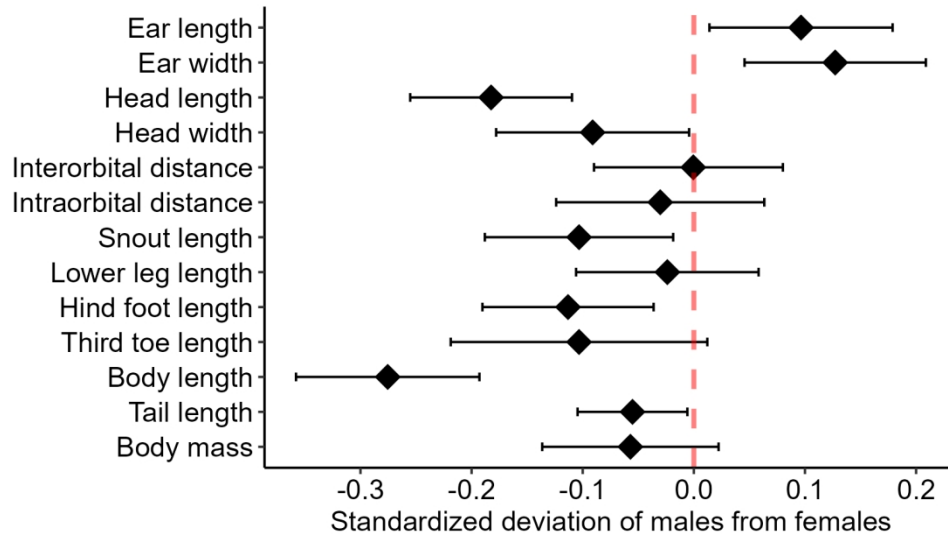
Quality assessment following the five data filtering steps indicated by different color code in B-D. A: Completeness of the dataset refers to the number of available values per variable compared to the number of individuals in the raw data (n = 3,073). B: 1-W' measure of normality, C: absolute skewness, D: absolute kurtosis. Values closer to zero in the ridge plots B-D indicate increased conformity to a normal distribution. For skewness (C) and kurtosis (D), the range of acceptable values is indicated with the vertical line at 2 units from 0.

287x319mm (300 x 300 DPI)



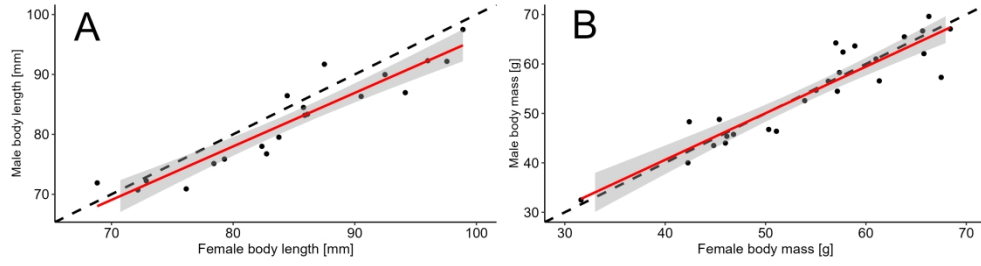
Inter-observer reproducibility quantified as coefficient of variation (CV) across the 13 different morphometric parameters. Only data after quality filtering was included. The horizontal line indicates the upper limit of the CV values interpreted as very good reproducibility (CV=0.0-0.1, Aronhime et al., 2013).

143x111mm (300 x 300 DPI)



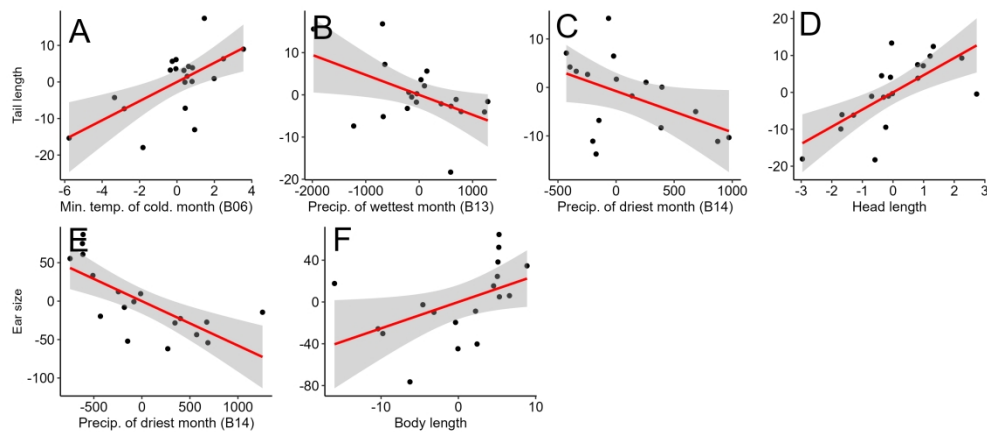
Sexual size dimorphism across all *Microcebus* taxa displayed via model estimates with 95% confidence intervals. The differences in all variables except inter- and intraorbital distances, lower leg length, third toe length, and body mass were statistically significant ($P \leq 0.029$; Table S3).

152x84mm (300 x 300 DPI)



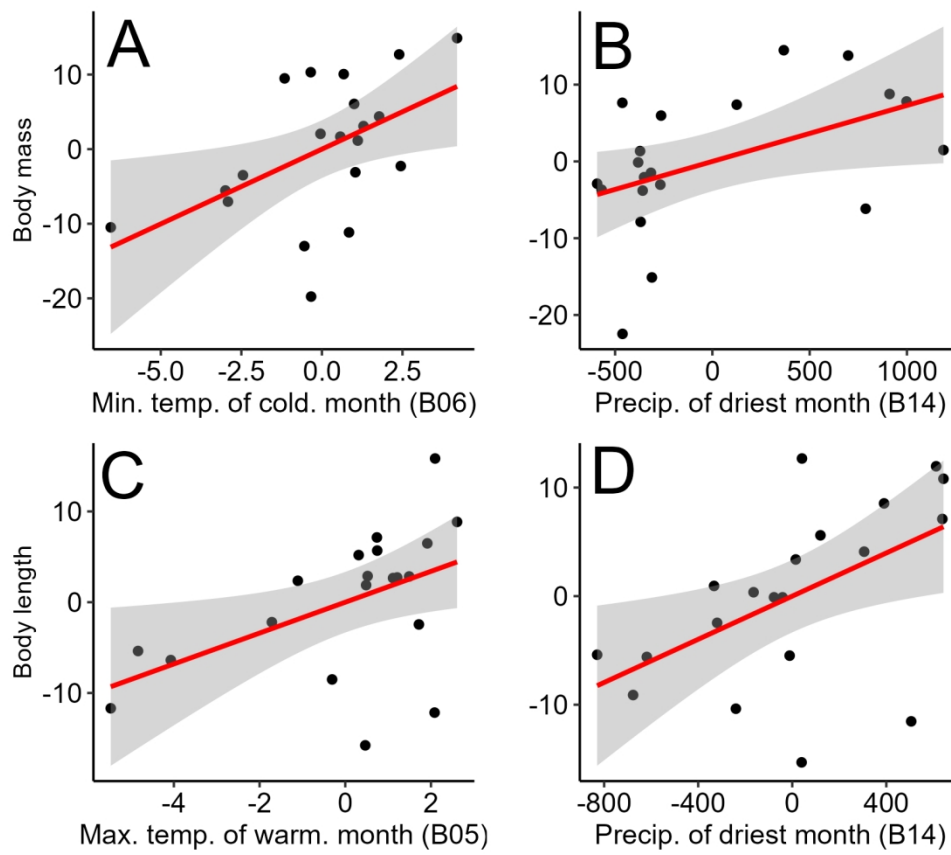
Relationship between female and male body length (A) and body mass (B). Rensch's rule is not supported in both cases ($P \geq 0.705$). Red line: linear model for observed data (including 95% confidence interval), dashed line: slope = 1. Note: model estimates slightly differ between the standardized major axis and the linear model; the former was used for hypothesis testing and the latter for visualization.

305x84mm (300 x 300 DPI)



Partial regression plots for the linear model best explaining the variation in tail length (A-D, $R^2 = 0.683$) and ear size (E-F, $R^2 = 0.450$) across mouse lemur taxa. Note: the y-axis represents the marginal impact of the response variable on the explanatory variable.

305x135mm (300 x 300 DPI)



Partial regression plots for the linear models best explaining the variation in body mass (A-B, $R^2 = 0.239$) and body length (C-D, $R^2 = 0.206$) across mouse lemur taxa. Note: the y-axis represents the marginal impact of the response variable on the explanatory variable.

152x135mm (300 x 300 DPI)