

RESEARCH
PAPER



Biological correlates of description date in carnivores and primates

Ben Collen^{1,2,*}, Andy Purvis¹ and John L. Gittleman³

¹Department of Biological Sciences, Imperial College London, Silwood Park campus, Ascot SL5 7PY, UK, ²Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK, ³Department of Biology, Gilmer Hall, University of Virginia, Charlottesville, VA 22904, USA. E-mail: ben.collen@imperial.ac.uk

ABSTRACT

Aim To examine which aspects of primates and carnivore biology can be used to predict attributes of species yet to be discovered.

Location Global.

Methods Multiple regressions of phylogenetically independent contrasts and non-phylogenetic species date of description, on multiple biological predictor variables, formed from previous hypotheses tested in the literature.

Results Orders differ, but both carnivore and primate species with a large geographical range tend to have been discovered earlier. When geographical range is controlled for, body mass is also significantly correlated with description date in carnivores, but remains a poor predictor in primates. No multiple-predictor model is apparent in the primates, but diurnal species are on average more likely to be described first. Carnivores not endemic to the tropics are more likely to be discovered earlier, reflecting a northern bias in description patterns.

Main conclusions Geographical range is by far the most important predictor variable. The study may have ramifications for conservation hotspot selection: species possessing a small geographical range are least likely to have been described, yet are most heavily weighted in some hotspot selection algorithms.

Keywords

Body size, conservation, geographical range, hotspots, independent contrasts, mammal.

*Correspondence: Ben Collen, Department of Biological Sciences, Imperial College London, Silwood Park campus, Ascot SL5 7PY, UK. E-mail: ben.collen@imperial.ac.uk

INTRODUCTION

It is estimated that we have described just 1.5–1.8 million of the approximately 10 million extant species (Wilson, 2003), and there is still considerable uncertainty over how many species exist (Godfray, 2002). If our description of species is inherently non-random, with species in some taxa more likely to be described than those in others, then our view of diversity is correspondingly distorted. This matters if, for instance, conservation policies are based on skewed reflections of true diversity patterns. Across higher taxa, studies consistently show that probability of description is not equal for all species within a taxon (Gaston, 1991; Allsop, 1997; Cabrero-Sanudo & Lobo, 2003). Broad comparisons among lower taxa have suggested that certain groups may receive a greater degree of taxonomic scrutiny (May, 1988), perhaps because they appeal to us more (Purvis *et al.*, 2003), and that some taxa are more likely to be seen due to larger

size (Gaston, 1991). Even within taxa, accumulating evidence suggests that some species are more likely to be described than others (Fig. 1), though explanations are more subtle and vary among groups (Gaston, 1991; Gaston & Blackburn, 1994; Gaston *et al.*, 1995b; Allsop, 1997; Reed & Boback, 2002).

If we assume that species remaining to be discovered are more similar to recently described species than to species named in the past, we can gain some insight into which biological attributes undescribed species might possess. The published literature contains several tests of hypotheses relating species attributes to dates of description in a variety of taxa (including butterflies, beetles, reptiles, amphibians and birds). These hypotheses include:

1 Larger animals are more apparent to taxonomists and collectors, are perhaps easier to collect, and are therefore discovered earlier (Gaston & Blackburn, 1994; Gaston *et al.*, 1995a; Reed & Boback, 2002; Cabrero-Sañudo & Lobo, 2003).

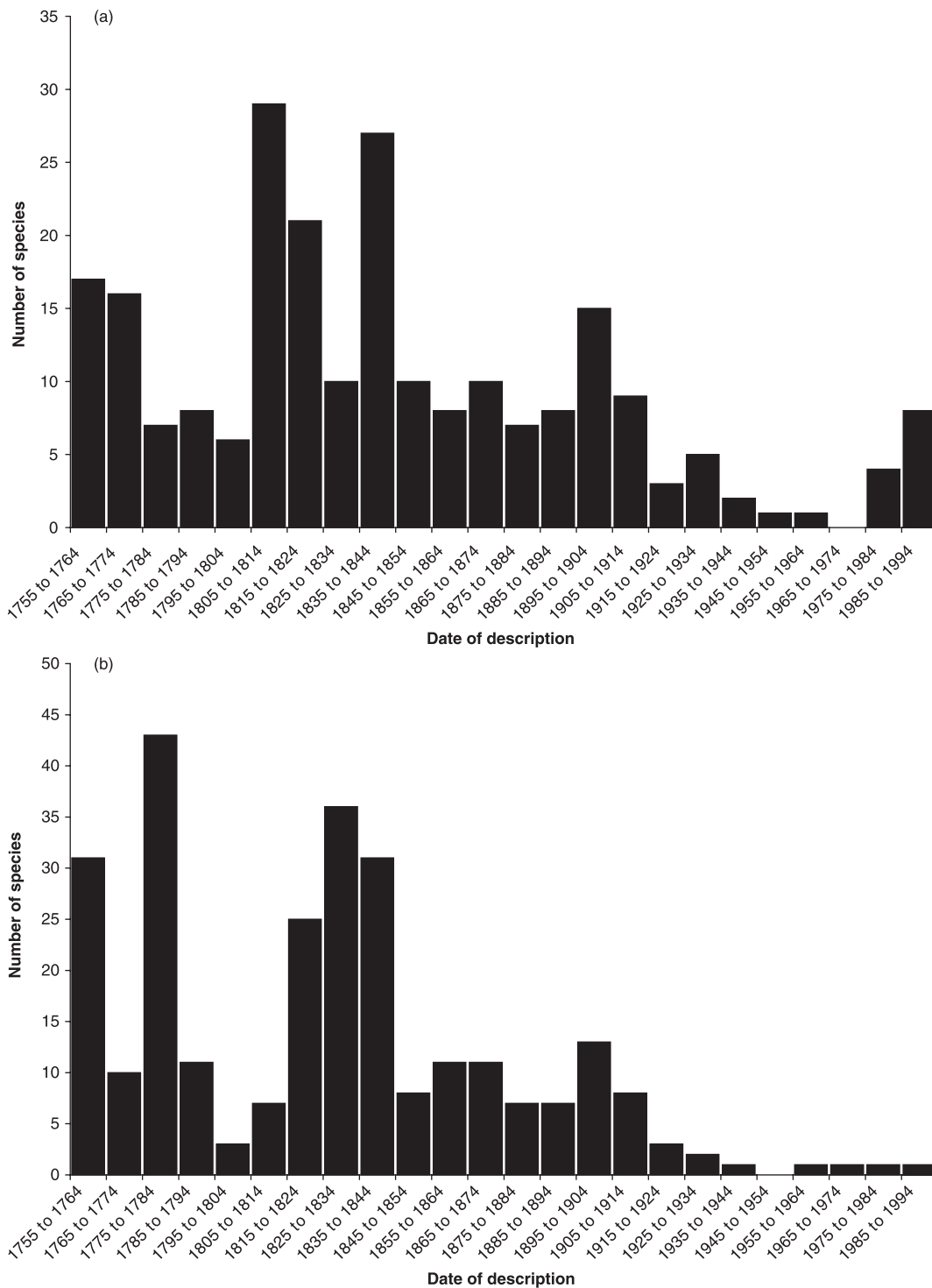


Figure 1 Frequency distribution for discovery date in (a) primates and (b) carnivores.

2 Species with large geographical ranges are more likely to be encountered and thus described (Patterson, 1994; Blackburn & Gaston, 1995; Gaston *et al.*, 1995a; Allsop, 1997; Cabrero-Sañudo & Lobo, 2003). In the same way, restricted range species are less likely to be discovered, e.g. South American endemic birds were found after those that are not restricted to the continent (Blackburn & Gaston, 1995).

3 Animals that attain higher densities are more overt or apparent and thus more likely to be described early (Gaston *et al.*, 1995a).

4 Conspicuous animals (for example those that have large home and day ranges, maintain large group sizes, social animals or species with diurnal activity timings that coincide with human activity) are more likely to be discovered. Many aspects of conspicuousness

have been highlighted, dependent on the taxa under study (e.g. plumage in birds: Blackburn & Gaston, 1995), but few tested.

5 Geographical location determines encounter rate, with early taxonomists centred in the northern hemisphere, and description patterns reflecting settlement by Europeans (e.g. Australian scarab beetles: Allsop, 1997; Palaearctic scarab beetles: Cabrero-Sañudo & Lobo, 2003).

6 Species with smaller numbers of related species within the same genus may take a shorter time to distinguish taxonomically (Gaston *et al.*, 1995a).

Even within a relatively intensely studied, charismatic taxon such as the mammals, species description still continues at a high rate. For example, between 1982 and 1993, the total number of world primate species recognized was raised from 179 (Honacki *et al.*, 1982) to 232 (Wilson & Reeder, 1993), an increase of 29.6%. This has since been elevated to 358 recognized species (Groves, 2001a), a further 34.9% change. Though much of the increase in species number is attributable to taxonomic revision and the application of new species concepts (Groves, 2001b), nevertheless, mammal species new to science continue to be discovered.

Statistical analyses on multispecies datasets can suffer from incomplete information about species' biology and the potential interrelationship between hypothesized predictor variables. Additionally, in a comparative study, treating species as statistically independent data points may be invalid as it can lead to pseudoreplication and hence elevated Type I error rates (Harvey & Pagel, 1991). However, continued inquiry of the validity of comparative analyses (e.g. Bjorklund, 1997; Harvey & Rambaut, 2000) means that it is sensible to present both phylogenetic and non-phylogenetic types of analysis (Freckleton *et al.*, 2002).

In this study we examine the relationships between description date and its possible biological correlates in carnivores and primates using phylogenetic comparative methods. Primates and carnivores make an excellent test case for analysing what determines date of description (i.e. when a species new to science is first formally described) because they are both widely distributed and biologically better understood than many other groups, although some species do remain enigmatic. Like some previous studies, we use single predictor regressions of independent contrasts of biological attributes on date of description to ensure statistical and phylogenetic independence across taxa. However, many biological predictor variables intercorrelate (for example body mass with geographical range size, latitude, density and home range, e.g. Gaston & Blackburn (2000)). For this reason we control for inter-related and confounding independent variables, and examine clade differences, using multiple regression. We find that around one third of the variation in date of description can be explained by our models, and discuss whether the results for primates and carnivores are similar to those identified previously for other taxa.

METHODS

Data

Species lists and description dates for primates (232 species) and carnivores (270 species) were taken from Wilson & Reeder

(1993), with dates of description spanning the period 1758–1991 for primates and 1758–1986 for carnivores. The following continuous predictor variables were tested: adult body mass (kg), current geographical range (km²), home range area (ha), day range length (km), population density (individuals per km²), average population group size and number of congeners. Four binary traits were also tested: sociality (1 = species commonly found in groups larger than one or both parents plus their litter, 0 = otherwise), activity timing (1 = diurnal species, 0 = other activity schedules), island endemicity (1 = island endemic, 0 = not wholly restricted to islands), and for carnivores whether the species is endemic to the tropics (1 = tropical endemic, 0 = range not wholly restricted to the tropics). Information on these species traits came from Purvis *et al.* (2000a). Continuous predictor variables were logarithmically transformed to equalize variances, except for number of congeners which was $\log(n + 1)$ transformed (some species have no congeners). Complete data were not available for all species (see Table 2 for sample sizes).

Phylogenetic analysis

Phylogenetically independent contrasts (Felsenstein, 1985; Pagel, 1992) were generated using the CAIC computer program (Purvis & Rambaut, 1995). The phylogenies used (Purvis, 1995 — modified as in Bininda-Emonds *et al.*, 1999; Purvis *et al.*, 2000b) were composite estimates of all extant carnivore and primate species recognized by Wilson & Reeder (1993). Neither phylogeny was completely resolved, so contrasts at multiple nodes were computed by calculating a difference between two subnodes after the daughter taxa of each multiple node is split into two monophyletic groups (Pagel, 1992; see Purvis & Rambaut, 1995 for full explanation). Branch lengths were set to be equal when generating the contrasts because preliminary analysis (not reported), showed that homogeneity of variances was more closely approached when branches were set to equal length, rather than when set to estimates of divergence time. All statistical tests were two tailed (except for sign tests where a *priori* predictions of the relationship between the independent variable and date of description were used to design the analyses, and to rationalize the use of one-tailed tests) and were performed using R version 1.6.1 (Ihaka & Gentleman, 1996).

Using the BRUNCH algorithm of CAIC to select non-overlapping and non-nested sets of taxa across the phylogeny, categorical variables were tested in turn against the response variable (date of description). Using this algorithm, all the contrasts in the predictor are set to positive so that under H₀, roughly half the contrasts should be positive and half negative. Sign tests and Wilcoxon signed rank tests were carried out on these contrasts to test for significance, for both orders separately and then for the orders combined.

Continuous variables were tested against description date using least-squares regression through the origin (Garland *et al.*, 1992). For primates and carnivores separately and then for the two orders combined, each variable was first regressed as a single predictor against date of description, as has been the common approach in previous studies examining description date

correlates (Blackburn & Gaston, 1995; Reed & Boback, 2002). Data points with Studentized *t*-residuals of greater than ± 3 were removed and the regressions repeated to assess if data with large leverage had unduly affected the results (following Jones & Purvis, 1997); however, removing these data never made a difference to the results.

For body mass and geographical range, we tested the linearity of the relationship by testing whether slope of description date was constant along the *X*-axis. At each node in the phylogeny, the average body mass of the taxon comparison and the relationship at the node between body mass and date of description, i.e. the quotient of the body mass contrast and the date contrast — termed the ‘contrast slope’ (Jones & Purvis, 1997) — were computed. Spearman’s rank correlation was used to test if these two quantities covaried. Further, runs tests were applied to the signs of the contrast slopes once ranked by average body mass, to test the null hypothesis that the signs of the contrast slopes are distributed randomly with respect to body mass. We then did the same for geographical range.

For each order we calculated the residuals of the regression of description date on each predictor variable in turn, and examined differences between orders using *t*-tests on these residuals. Multiple regression through the origin was then used in order to factor out geographical range, which proved the most important single predictor. Multiple models were constructed using forwards step-wise regression, sequentially adding each predictor variable to geographical range, and evaluating if additional variance was explained. Finally, to test for significant difference between the multiple models for the two orders, we fitted a model to the combined data using the predictors implicated in either model and the cross-product order \times trait interactions (Garland *et al.*, 1992).

Species analysis

For comparison, non-phylogenetic analyses were performed, treating each species as an independent data point. Variables were first regressed as single predictors. Multiple models were constructed using model simplification by first removing the high order interaction terms (order \times trait interaction), and the main effects (Crawley, 2002). Linearity was tested by including a quadratic function and examining the model to see if there was any improvement in explanatory power.

RESULTS

Contrasts analysis

Wilcoxon signed rank tests on categorical variables (Table 1) showed activity timing significantly predicts description date in primates, with diurnal species on average described earlier ($P = 0.008$). Tropical endemic carnivores are on average described later ($P < 0.0001$), and there is evidence that non-social ($P = 0.03$) and island endemic ($P = 0.05$) species are on average described later. When both orders were combined, no significant result was apparent for island status, but sociality and activity

Table 1 Wilcoxon signed rank tests of categorical variables predicting description date (all tests one tailed). For island status, 0 = not endemic to islands, 1 = endemic to islands. For activity timing, 0 = not diurnal, 1 = diurnal. For sociality, 0 = not social, 1 = social. For range in tropics, 0 = not tropical endemic, 1 = tropical endemic. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

Order	Predictor	Positive/negative contrasts	Wilcoxon
Carnivores	sociality	8/22	330*
	activity timing	8/17	162
	island status	11/16	80*
	tropical endemic	35/40	200***
Primates	sociality	6/15	150
	activity timing	1/8	56***
	island status	6/12	72
Combined	sociality	14/37	888**
	activity timing	9/25	413**
	island status	17/28	308

timing remained significant. Such conservative tests however, contain fewer contrasts than regression analysis.

The regression analysis of single predictors (Table 2) shows that, within both the primates and carnivores, species with larger geographical ranges tend to be described earlier (primates: $b = -20.74$, $t_{158} = -10.12$, $P < 0.001$; carnivores: $b = -12.63$, $t_{181} = -8.91$, $P < 0.001$). Slopes in the two orders are significantly different (geographical range: $b = -28.86$, $t_{384} = -6.91$, $P < 0.0001$; order \times geographical range: $b_{338} = 8.11$, $t = 3.27$, $P < 0.01$, see Fig. 2a). Within the carnivores, large body mass ($b = -11.63$, $t_{196} = -3.12$, $P < 0.001$) and large home range ($b = -5.07$, $t_{83} = -2.03$, $P = 0.04$) are also significantly correlated with early description date as single predictors.

Within the primate single predictor analysis, in addition to those with large geographical range, species with fewer congeners are described significantly earlier ($b = 25.66$, $t_{159} = 2.69$, $P = 0.01$). A more conservative analysis using the BRUNCH algorithm of CAIC (for a full description of the algorithm see Purvis & Rambaut, 1995) was performed as prior inspection of the variable revealed homogeneity of variance not closely approached. This proved marginally significant (sign test: $n + 22$, $n - 10$, $P = 0.05$).

Spearman’s rank correlations showed that in both orders, neither geographical range (primates $r_s = 0.09$, carnivore $r_s = -0.25$) nor body mass (primates $r_s = -0.03$, carnivore $r_s = -0.04$) significantly covaried with the contrast slope. Further, no significant results were observed in the runs tests on the signs of the contrast slopes.

Most of the significant correlates of description date did not explain significant additional variance when geographical range was controlled for in a multiple regression. Body size was significant in carnivores, with species more likely to be described early if they occupy a large geographical range and are large bodied (Table 3). No multiple model was apparent for the primates however. Differences between orders therefore concerned body mass (significant only in carnivores). Table 4a shows that in the

Table 2 Results of single predictor regressions of contrast data and species data for each order. N_c number of contrasts, N_s number of species, b_{contrast} coefficient of contrast analysis, t_{contrast} t -value of contrast analysis, b_{species} coefficient of species analysis, t_{species} t -value of species analysis, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. All tests are two tailed

Predictor	Primates						Carnivores					
	N_s	N_c	b_{contrast}	t_{contrast}	b_{species}	t_{species}	N_s	N_c	b_{contrast}	t_{contrast}	b_{species}	t_{species}
Population density	96	86	3.03	0.54	0.76	0.21	88	75	2.64	1.00	2.85	1.58
Body mass	179	138	-11.49	-1.35	-3.56	-1.30	240	197	-11.63	-3.12**	-3.72	-2.61**
Geographical range	211	159	-20.74	-10.00***	-14.21	-8.32***	236	182	-12.63	-8.71***	-12.89	-10.54***
Home range	131	110	-1.72	-0.43	-1.89	-0.87	99	84	-5.07	-2.15*	-3.96	-2.67**
Day range	98	84	-16.29	-1.32	-16.37	-2.50*	47	43	-3.60	-0.93	-2.01	-0.67
No. of congeners	217	160	25.66	2.69*	9.53	2.15*	270	211	4.15	0.71	2.60	0.81
Group size	170	136	-4.41	-0.60	-5.09	-1.51	109	89	-7.65	-1.15	4.61	1.01

combined model, this difference was itself significant. Order significantly interacted with geographical range, but not with body mass.

Species analysis

When species were treated as independent data points in the single predictor non-phylogenetic analysis (Table 2) an additional correlate, day range, was found to be significant in primates ($t_{97} = -2.50, P = 0.01$). In carnivores the same correlates

were significant as in the contrasts analysis. There was no evidence of nonlinearity. However, when the orders were combined, a multiple regression model including geographical range and body mass was found to significantly predict date of description (Table 4b), explaining around 25% of the variance.

DISCUSSION

Our results show that taxonomic description of carnivores and primates has been a biased process, with several biological correlates of description date apparently predisposing species to a likelihood of earlier discovery and description. Our findings strongly suggest that species with large geographical ranges tend to be described earliest, both in primates and carnivores. In carnivores for example, the median geographical range size for the first 50% of described species is $5.22 \times 10^6 \text{ km}^2$ compared to just $3.37 \times 10^5 \text{ km}^2$ for the most recently described 25%. An obvious mechanism is apparent whereby species with larger geographical ranges are more likely to be encountered by collectors and taxonomists, and the result is supported by several previous studies on groups as varied as North American butterflies (Gaston *et al.*, 1995a), South American oscine passerine birds (Blackburn & Gaston, 1995) and Australian scarab beetles (Allsop, 1997). One possibility confounding the result is that more recently described species have smaller ranges simply because they are more recently described, thus enabling little time for their ranges to be delineated (Blackburn & Gaston, 1995). However, this is unlikely to explain the magnitude of the effect found in carnivores and primates.

Our results provide limited support for the other hypotheses listed in the introduction. Large home range size correlates with early description in carnivores, but not independently of geographical range size. In primates, species with fewer congeners were discovered earlier. This regression analysis was repeated using the more conservative BRUNCH algorithm of CAIC and a sign test, and still remained marginally significant. Gaston *et al.* (1995a) highlight two potential explanations for a similar result in North American butterflies. Firstly, the taxonomy of smaller genera may take less time to distinguish, being less complex and

Table 3 Multiple regression model for carnivores predicting description date. Sample size: 206 species, 168 contrasts, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

Predictor	coefficient	se	t
Geographical range	-11.89	1.55	-7.64***
Body mass	-8.69	4.03	-2.16*

Table 4 Multiple regression model across carnivores and primates. (a) Contrast model, sample size: 385 species and 308 contrasts, model explains 34% of the total variance (b) Species model, d.f. = 383, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

(a)			
Predictor	coefficient	se	t
Geographical range	-10.77	1.59	-6.76***
Body mass	-7.67	3.66	-2.10*
Order \times geographical range	-9.15	2.55	-3.59***
(b)			
Predictor	coefficient	se	t
Geographical range	-10.90	1.02	-10.72***
Body mass	-3.62	1.50	-2.41*

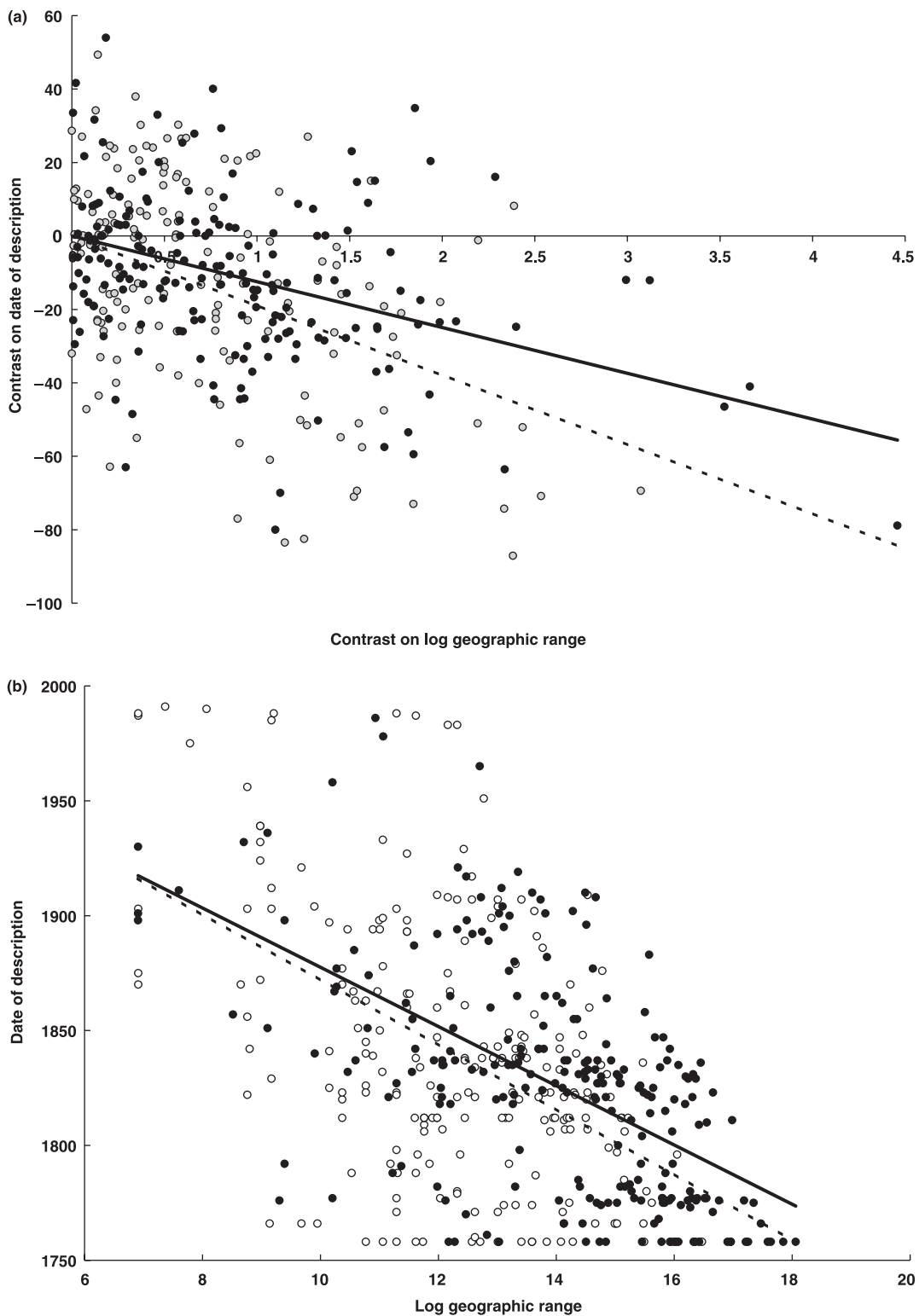


Figure 2 (a) Contrast plot showing the relationship between date of description and geographical range. Solid circles denote carnivores (solid line is regression line); open circles denote primates (dotted line is regression). An ANCOVA test showed a significant effect of order. (b) Species data plot.

involved. The result could also arise though as an artefact of taxonomists choosing to work on species-rich groups. Secondly, there could be a relationship between number of congeners and another trait that confounds the result. Number of congeners

does correlate with geographical range size, but drops out in the multiple regression model when other predictor variables are controlled for. There is further limited support for primate species with larger day ranges being described earlier. Day range

however, is not significant in the multiple model, nor as a single predictor when phylogeny is controlled for.

We find in the single predictor analysis that large bodied carnivores are described earlier, though not larger bodied primates. In the combined analysis, species with large geographical ranges and large body mass were described earliest, combining to explain around 34% of the variation in description date. When analysed separately by order, large bodied carnivores of a given geographical range were described earlier, however, body mass remains a poor predictor of description date in primates. The median body mass of the first 50% of carnivores described was 7.36 kg (mean value 9.75 kg) compared to a median value of 4.12 kg (mean value 6.56 kg) for total carnivores described to date. The pattern was less pronounced in primates, with a median body mass of 3.51 kg for the first 50% described, but 3.31 kg (mean values 5.28 kg vs. 5.60 kg, respectively) for the total described by 1993 (not significantly different: $t_{238} = -0.09$, $P > 0.05$).

The general lack of consistency between the orders, and carnivores' more size-dependent date of description, can perhaps be explained by habitat. With primates being primarily forest based, size may be of relatively little importance, and other factors (for example behavioural traits) may be more likely to determine human encounter rate and thus description. In the more habitat-diverse carnivores, this may not be the case, though body mass is obviously of only secondary importance to geographical range. Reed & Boback (2002) found similar inconsistencies of body size effect across reptiles and amphibians, and the trait has been found to predict description date in British beetles (Gaston, 1991), birds of the world (described 1966–90: Gaston & Blackburn, 1994), North American butterflies (wingspan: Gaston *et al.*, 1995a) and extinct mammals (Alroy, 2003), but not in South American birds (Blackburn & Gaston, 1995), Australian scarab beetles (Allsop, 1997), nor Western Palaearctic beetles (Cabrero-Sañudo & Lobo, 2003).

A study by Alroy (2003) on fossilized mammal species bears on the carnivore result that larger bodied species were discovered earlier. He explains a similar pattern in fossilized taxa, both in terms of taxonomists wanting to make the *big* catch, but also that certain mammal taxonomists (e.g. Cope, Matthew, Hibbard) were much more influential than others. In our data set there is no obvious effect of author, though a few authors dominate the description of most species. The lack of consistency of body size results between carnivores and primates is striking, especially given that it seems to pertain in different contexts (e.g. species richness: Gittleman & Purvis, 1998; extinction threat: Purvis *et al.*, 2000b). There are no doubt multiple reasons — carnivores have greater intrinsic variability in body mass; with increasing size, carnivores move more quickly and over larger areas; with vast travel distances, habitat diversity is much greater than in primates; and last, in the 19th century there was considerable reward and fame for bagging a large carnivore, and many type specimens were from hunting. All would contribute to large carnivores being discovered first.

Bergmann's rule states that, within endothermic vertebrate genera, larger species are found at higher latitudes (Bergmann,

1847; Mayr, 1956, 1963; Freckleton *et al.*, 2003), because larger bodies are beneficial in colder climates. We find that carnivore species with ranges solely in the tropics were on average described later, reflecting a northern bias in carnivore description with most early taxonomists coming from Europe and North America. So there exists a potential bias, whereby large carnivores were discovered early, but they were looked for in areas that contain larger carnivores. However, when the relationship between body mass and whether the species are tropical endemics, was examined it was not significant (sign test: $n + 13$, $n - 21$, $P > 0.05$) and neither was the relationship between body mass and percentage of total range in tropics ($b = -0.08$, $t_{158} = -0.59$, $P > 0.05$). To some extent we can therefore rule this out as a confounding factor.

Of the binary variables tested in primates, only activity timing proved significantly predictive of date of description, with diurnal primates on average being described earlier. The one opposing comparison is between *Eulemur mongoz* (not diurnal) and *E. coronatus* (diurnal). It is predicted that diurnal organisms have smaller geographical ranges than those with other activity timings, which could confound the diurnal result. *E. coronatus* has a geographical range of 6.63×10^3 km², *E. mongoz* a range of 2.09×10^4 km² however, there is no general tendency for range size to covary with activity timing (using the BRUNCH algorithm, Sign test: $n + 5$, $n - 3$, $P > 0.05$). In carnivores, sociality and island status were significantly predictive of description date, with social species described earlier on average, and description of species endemic to islands expected later than those that are less restricted (Allsop, 1997). The remaining variable that aimed to capture conspicuousness (i.e. big groups — continuous variable) did not significantly predict description date in either order.

Other potentially important factors that were not examined may include other aspects of conspicuousness, both behavioural and morphological. However, habitat may confound such traits (Blackburn & Gaston, 1995). An obvious candidate predictor for primates and carnivores is pelage colour, with the hypothesis that brighter coats are more perceptible. However, species with a bright pelage may just live in difficult-to-survey habitat — particularly likely in forest dwelling primates where many arboreal species within families have brighter coat colours (e.g. *Pongo pygmaeus* vs. *Gorilla gorilla*; *Presbytis rubicunda* vs. *Semnopithecus entellus*: Treves, 1997).

There has been debate about whether to control for phylogeny in comparative analyses, with some authors arguing that doing so makes no difference (see Freckleton *et al.*, 2002). In this study, phylogeny appears to matter, and phylogenetic methods have been shown to be important in some other taxa, e.g. South American birds (Blackburn & Gaston, 1995), North American butterflies (Gaston *et al.*, 1995a), but the methods have not been used in the majority of studies of species description to date.

No discussion of mammalian taxonomy is complete without reference to the phylogenetic species concept and its application. Taxonomic shuffling in the primates in particular has meant an increase over the past 10 years from 232 (Wilson & Reeder, 1993)

to 358 species (Groves, 2001a). Doubtless this increase in species number is primarily due to taxonomic reassessment and application of the phylogenetic, rather than biological species concept (Groves, 2001a). Right or wrong, its application may affect analyses like ours (see Agapow *et al.*, 2004). Although species numbers are increasing, many are being raised from synonymy or subspecies (specifically, three quarters of newly recognized species of Neotropical mammals described in the period 1982–94 had already been discovered, and were raised from synonymy: Patterson, 1994). So although their ranges will most likely be small, they will not be particularly recently described.

We can look to recent discoveries in primate taxonomy to see if the correlative traits predicted are accurate. Since Wilson & Reeder (1993), night monkeys (*Aotus*), Galagos (*Galagoideus udzungwensis* and *G. rondoensis*; Honess & Bearder, 1998) and mouse lemurs (*Microcebus ravelobensis*; Zimmerman *et al.*, 1997) have spawned additional species descriptions — all are nocturnal clades. As a cautionary note however, diurnal groups such as the Titi monkeys (*Callicebus*) have undergone extensive revision with the addition of 2 species new to science (van Roosmalen *et al.*, 2002) as well as the elevation of 6 taxa listed by Groves (2001a) as sub species to species level — an issue of species concept is therefore apparent.

The findings of this study may be of particular relevance for hotspot selection algorithms, and reflect a potential avenue for future research. Conservation priority areas are often defined using numbers of endemic species — or at the very least, area selection algorithms are most sensitive to such restricted groups. Small-range specialists or endemics are major contributors to areas of tropical species diversity (Stevens, 1989; Pagel *et al.*, 1991; Hughes *et al.*, 2002). Myers *et al.* (2000), for example, define hotspots using endemic plant diversity. So the very species to which some priority-setting algorithms attach most conservation value are least likely to have been described.

ACKNOWLEDGEMENTS

This work was funded by NERC studentship NER/A/S/2002/10377 (BC), NERC grant NER/A/S/2001/00581 (AP) and the National Science Foundation grant DEB-0129009 (JLG).

REFERENCES

- Agapow, P.-M., Bininda-Emonds, O.R.P., Crandall, K.A., Gittleman, J.L., Mace, G.M., Marshall, J.C. & Purvis, A. (2004) The impact of species concept on biodiversity studies. *Quarterly Review of Biology*, **79**, 161–179.
- Allsop, P.G. (1997) Probability of describing an Australian scarab beetle: influence of body size and distribution. *Journal of Biogeography*, **24**, 717–724.
- Alroy, J. (2003) Taxonomic inflation and body mass distributions in north American fossil mammals. *Journal of Mammalogy*, **84**, 431–443.
- Bergmann, C. (1847) Über die Verhältnisse der Wärmeökonomie der Thiere zu Ihrer Grösse. *Göttinger Studien*, **3**, 595–708.
- Bininda-Emonds, O.R.P., Gittleman, J.L. & Purvis, A. (1999) Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews*, **74**, 143–175.
- Bjorklund, M. (1997) Are ‘comparative methods’ always necessary? *Oikos*, **80**, 607–612.
- Blackburn, T.M. & Gaston, K.J. (1995) What determines the probability of discovering a species? A study of South American oscine passerine birds. *Journal of Biogeography*, **22**, 7–14.
- Cabrero-Sanudo, F.J. & Lobo, J.M. (2003) Estimating the number of species not yet described and their characteristics: the case of the Western Palaearctic dung beetle species (Coleoptera, Scarabaeoidea). *Biodiversity and Conservation*, **12**, 147–166.
- Crawley, M.J. (2002) *Statistical computing: an introduction to data analysis using S-Plus*. Wiley, Chichester.
- Felsenstein, J. (1985) Phylogenies and the comparative method. *American Naturalist*, **125**, 1–15.
- Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *American Naturalist*, **160**, 712–726.
- Freckleton, R.P., Harvey, P.H. & Pagel, M. (2003) Bergmann’s rule and body size in mammals. *American Naturalist*, **161**, 821–825.
- Garland, T., Harvey, P.H. & Ives, A.R. (1992) Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Systematic Biology*, **41**, 18–32.
- Gaston, K.J. (1991) Body size and probability of description: the beetle fauna of Britain. *Ecological Entomology*, **16**, 505–508.
- Gaston, K.J. & Blackburn, T.M. (1994) Are newly described bird species small bodied? *Biodiversity Letters*, **2**, 16–20.
- Gaston, K.J. & Blackburn, T.M. (2000) *Pattern and process in macroecology*. Blackwell Science, Oxford.
- Gaston, K.J., Blackburn, T.M. & Loder, N. (1995a) Which species are described first? The case of North American butterflies. *Biodiversity and Conservation*, **4**, 119–127.
- Gaston, K.J., Scoble, M.J. & Crook, A. (1995b) Patterns in species description: a case study using the Geometridae (Lepidoptera). *Biological Journal of the Linnean Society*, **55**, 225–237.
- Gittleman, J.L. & Purvis, A. (1998) Body size and species richness in carnivores and primates. *Proceedings of the Royal Society of London*, **265**, 113–119.
- Godfray, H.C.J. (2002) Challenges for taxonomy. *Nature*, **417**, 17–19.
- Groves, C.P. (2001a) *Primate taxonomy*. Smithsonian Institution Press, Washington DC.
- Groves, C.P. (2001b) Why taxonomic stability is a bad idea, or why are there so few species of primates (or are there?). *Evolutionary Anthropology*, **10**, 192–198.
- Harvey, P.H. & Pagel, M.D. (1991) *The comparative method in evolutionary biology*. Oxford University Press, Oxford.
- Harvey, P.H. & Rambaut, A. (2000) Comparative analyses for adaptive radiations. *Philosophical Transactions of the Royal Society of London, B*, **355**, 1599–1605.

- Honacki, J.H., Kinnman, K.E. & Koepl, J.W. (1982) *Mammal species of the world: a taxonomic and geographic reference*. Allen Press, Inc. and the Association of Systematics Collections, Kansas.
- Honess, P.E. & Bearder, S.K. (1998) Two new primate species from Tanzania: *Galagoides udzungwensis* and *Galagoides rondoensis*. *African Primates*, **2**, 11–15.
- Hughes, T.P., Bellwood, D.R. & Connolly, S.R. (2002) Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. *Ecology Letters*, **5**, 775–784.
- Ihaka, R. & Gentleman, R. (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299–314.
- Jones, K.E. & Purvis, A. (1997) An optimum body size for mammals? Comparative evidence from bats. *Functional Ecology*, **11**, 751–756.
- May, R.M. (1988) How many species are there on earth? *Science*, **241**, 1441–1449.
- Mayr, E. (1956) Geographical character gradients and climatic adaptation. *Evolution*, **10**, 105–108.
- Mayr, E. (1963) *Animal species and evolution*. Harvard University Press, Cambridge, Massachusetts.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., de Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
- Pagel, M.D. (1992) A method for the analysis of comparative data. *Journal of Theoretical Biology*, **156**, 431–442.
- Pagel, M.D., May, R.M. & Collie, A.R. (1991) Ecological aspects of the geographic distribution and diversity of mammalian species. *American Naturalist*, **137**, 791–815.
- Patterson, B.D. (1994) Accumulating knowledge on the dimensions of biodiversity: systematic perspectives on neotropical mammals. *Biodiversity Letters*, **2**, 79–86.
- Purvis, A. (1995) A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London, B*, **348**, 405–421.
- Purvis, A., Agapow, P.-M., Gittleman, J.L. & Mace, G.M. (2000a) Non-random extinction and the loss of evolutionary history. *Science*, **288**, 328–330.
- Purvis, A., Gittleman, J.L., Cowlshaw, G. & Mace, G.M. (2000b) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London, B*, **267**, 1947–1952.
- Purvis, A., Orme, C.D.L. & Dolphin, K. (2003) Why are most species small bodied? A phylogenetic view. *Macroecology: concepts and consequences* (ed. by T.M. Blackburn and K.J. Gaston), pp. 155–173. Blackwell Science, Oxford.
- Purvis, A. & Rambaut, A. (1995) Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Computational Applications in Bioscience*, **11**, 247–251.
- Reed, R.N. & Boback, S.M. (2002) Does body size predict dates of species description among North American and Australian reptiles and amphibians? *Global Ecology and Biogeography*, **11**, 41–47.
- van Roosmalen, M.G.M., van Roosmalen, T. & Mittermeier, R.A. (2002) A taxonomic review of the Titi monkeys, genus *Callicebus* Thomas, 1903, with the description of two new species, *Callicebus bernhardi* and *Callicebus stephennashi*, from Brazilian Amazonia. *Neotropical Primates*, **10** (Suppl.) 1–52.
- Stevens, G.C. (1989) The latitudinal gradient in geographic range: how so many species coexist in the tropics. *American Naturalist*, **133**, 240–256.
- Treves, A. (1997) Primate natal coats: a preliminary analysis of distribution and function. *American Journal of Physical Anthropology*, **104**, 47–70.
- Wilson, E.O. (2003) The encyclopedia of life. *Trends in Ecology and Evolution*, **18**, 77–80.
- Wilson, D.E. & Reeder, D.M. (1993) *Mammal species of the world: a taxonomic and geographic reference*. Smithsonian Institution Press, Washington.
- Zimmermann, E., Ehresmann, P., Zietemann, V. & Radespiel, U. (1997) A new primate species in north-western Madagascar: the golden-brown mouse lemur (*Microcebus ravelobensis*). *Primate Eye*, **63**, 26.

BIOSKETCHES

Ben Collen has interests in mammalian conservation and evolution. Currently, Ben is a PhD student in the Department of Biological Sciences at Imperial College London and the Institute of Zoology at the Zoological Society of London. Recent projects include investigations into correlates of population density in Asian primates and an analysis of consistency of Red List criteria application.

Andy Purvis is Reader in Biodiversity in the Department of Biological Sciences at Imperial College London. He takes a phylogenetic approach to a range of questions in macroevolution and conservation biology. Among current projects, he is coediting a symposium volume entitled 'Phylogeny & Conservation' and allegedly writing a book on macroevolution.

John Gittleman is Professor of Biology at the University of Virginia, Charlottesville. He is the editor of *Carnivore Behaviour, Ecology and Evolution* and coeditor of *Carnivore Conservation* and the journal *Animal Conservation*. His research focuses on macroevolutionary problems related to speciation, extinction and the evolution of biodiversity.