Supporting Information

Liow *et al*. 10.1073/pnas.0709763105

SI Materials and Methods

Taxonomy. Our text makes use of familiar ordinal level groups of mammals as a shorthand (Artiodactyla, Carnivora, Creodonta, Hyracoidea, Perissodactyla, Primates, Proboscidea, Tubulidentata, Insectivora, Rodentia and Lagomorpha). However, note that cetaceans are not part of the dataset, so technically, either Cetartiodactyla or Artiodactyla could be used. Among small mammals, bats, being spottily represented in the NOW database, are excluded from the analysis. Insectivora is used as a higher taxon, although Lipotyphla would be more precise.

Survivorship and Preservation Probabilities. To grasp the dynamics of a population, ecologists often trap, mark, release and attempt to recapture those animals over a few subsequent sampling intervals (1, 2). Marked individuals, if not recaptured, can either be dead, have left the study site, or just did not enter the trap. The survival rates of the population can be estimated while formally taking into account sampling probabilities, estimated from the distribution of absences flanked by presences for an individual. Extending the capture-mark-recapture (CMR) approach to the fossil record (3, 4) implies that we assume that each species is equivalent to an individual and that we treat whole faunas as if they were a population (see ref. 5 for more details and references).

The models we present in our results are time-varying estimates of survival (Φ) and preservation (pr) probabilities with either no body size effect, or additive or multiplicative effects of body size. $\Phi\{gr*t\}$ pr$\{gr*t\}$ is the global model where both survival and preservation included interaction between group (gr, large and small) and time (t, where, if the observation windows are 1, 1.5, and 2 M.y. in length, then we have 22, 15, 11 occasions, respectively). We first ran all possible models with different combinations of Φ and pr and the two classes of body size as covariates. This strategy resulted in 25 models. However, the non-time-varying models performed very poorly and are therefore ignored for the rest of the analyses. Thus only 9 different combinations of survival and preservation with time effect are presented, namely:

$$\Phi\{gr*t\}pr\{gr*t\} \text{ Global model}$$

$$\Phi\{gr+t\}pr\{gr*t\}$$

$$\Phi\{gr*t\}pr\{gr+t\}$$

$$\Phi\{gr+t\}pr\{gr+t\}$$

$$\Phi\{t\}pr\{gr*t\}$$

$$\Phi\{gr*t\}pr\{t\}$$

$$\Phi\{t\}pr\{gr+t\}$$

$$\Phi\{gr+t\}pr\{t\}$$

$$\Phi\{t\}pr\{t\} \text{ pure time varying model.}$$

We constrained our estimates of survival and preservation probabilities to lie between 0 and 1 using a logit link function (8). The CMR approach makes assumptions (1, 7) that are usually not strictly met by data. The lack of fit can be adjusted however, by estimating a variance inflation factor ($\hat{c}$) and adjusting the ranks of models accordingly (1). We estimated $\hat{c}$ using "Test.2" and "Test.3" as detailed in ref. 7 and used the mean of the estimated $\hat{c}$'s for each global model for each data subset. We compared the models in each data subset using a model selection approach, as advocated by Burnham and Anderson (8). Akaike Information Criteria (AICs) were converted to QAICs such that

$$QAIC = -\{2 \log(L(\hat{\theta})/\hat{c})\} + 2K,$$

where $L(\hat{\theta})$ is the likelihood of the parameters given the data and K is the number of parameters estimated.

The model weight for the kth model is calculated as

$$\frac{\exp(-0.5(QAIC_k - QAIC_{min}))}{\sum\limits_{i=1}^{k} \exp(-0.5(QAIC_i - QAIC_{min}))}$$

where $QAIC_{min}$ is the QAIC for the best model given a data subset. Therefore, model weights sum to one for all nine models compared.

Discussion

As mentioned in the text, there are certainly exceptions to the rule of large mammals having truly greater body masses and vice versa. Moreover, body size is a trait that can evolve within clades. For example, extant beavers (Rodentia: Castoridae) are quite large, despite their being considered small mammals. However, many fossil beaver species were smaller. Other extant large "small" mammals are dominantly subtropical and tropical or outside our longitudinal limits and therefore not considered in our dataset.

There are, however, a few SLOH genera in the NOW dataset that are assigned a large body size category, and these are namely some carnivores (see Table S4 in Dataset S1) such as *Crocuta* (hyenas) and *Meles* (European badgers) that both dig or use burrows or dens. These exceptions do not, however, change our general conclusions.

It is also worth looking at the genera falling in the tail of the histogram presented in Fig. 1 (reproduced in Fig. S1*a*). In the 16-M.y. class, we have *Miodyromys, Glirulus, Blackia*, and *Miopetaurista,* arboreal to bush-loving dormice and gliding squirrels, which exhibit varying degrees of dormancy and probably use thermo-regulated hiding places. *Desmanella* is an extinct mole and so can be inferred to be a burrower. Without exception, these extinct small-bodied genera can be thought of as SLOH taxa. One extant genus, *Glirulus*, was not removed by our procedure because it is represented by a species that lives only in Japan today.

In the 15-M.y. class, *Parapodemus*, is a burrowing field mouse, *Galerix*, a standard nonderived erinaceid that is likely to have burrowed and to have been capable of dormancy; *Democricetodon,* a hamster predecessor; *Spermophilinus*, a ground squirrel and burrower; and *Paenelimnoecus*, a shrew (all small mammals and very likely all to have some form of SLOH behavior).

The 14-M.y. class includes *Miosorex* and *Allosorex*, both shrews; *Pliospalax*, a spalacid burrower; *Scaptonyx*, a mole; *Keramidomys* and *Armantomys*, both extinct eomyid rodents with unknown lifestyles but that were probably forest dwellers; and *Hylopetes*, a gliding arboreal squirrel. All these small mammals are again probably capable of SLOH behavior as far as we know. Two extant genera were not removed by our procedures because they are extinct in Europe: *Scaptonyx*, found today in China; and *Hylopetes*, found in South East Asia. The only large mammal we

have in this class is *Tapirus,* an extant genus retained in our "extinct" dataset because its occurrence is very rare, especially in Pleistocene deposits. In fact, *Tapirus* contributes to Fig. 1 and Fig. S1*a* but not Fig. S1*d*, where only well sampled taxa and sites were retained. Moreover, *Tapirus* is not found in Europe today and therefore was not removed by our procedure for producing Fig. 1 and Fig. S1 and Table 1.

1. Lebreton JD, Burnham KP, Clobert J, Anderson DR (1992) Modeling survival and testing biological hypotheses using marked animals: A unified approach with case-studies. *Ecol Monogr* 62:67–118.
2. Nichols JD, Boulinier T, Hines JE, Pollock KH, Sauer JR (1998) Estimating rates of local species extinction, colonization, and turnover in animal communities. *Ecol Applic* 8:1213–1225.
3. Nichols JD, Morris RW, Brownie C, Pollock KH (1986) Sources of variation in extinction rates, turnover, and diversity of marine invertebrate families during the Paleozoic. *Paleobiology* 12:421–432.
4. Kroger B (2005) Adaptive evolution in Paleozoic coiled cephalopods. *Paleobiology* 31:253–268.
5. Connolly SR, Miller AI (2001) Joint estimation of sampling and turnover rates from fossil databases: Capture-mark-recapture methods revisited. *Paleobiology* 27:751–767.
6. Venables WN, Ripley BD (2002) *Modern Applied Statistics With S* (Springer, New York).
7. Cooch E, White G (2006) *Program Mark: A Gentle Introduction*, www.phidot.org/software/mark/docs/book.
8. Burnham KP, Anderson DK (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York).
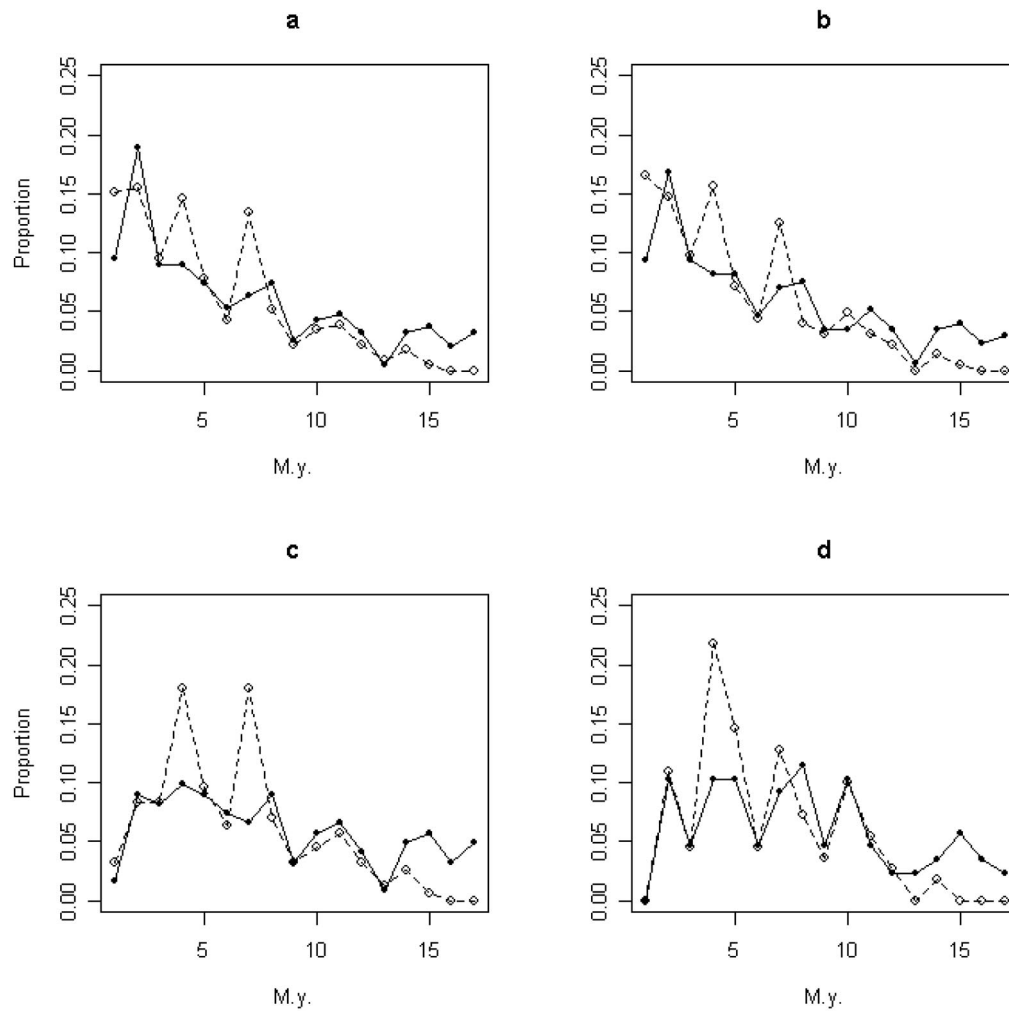
Fig. S1. Histograms of genus durations of extinct small and large mammal genera from the NOW database. Proportions are calculated for small and large mammals separately. The *x* axes are in millions of years (M.y.), and 1-M.y. bins are shown, with data points plotted at the higher limit. Solid circles and lines represent small mammals, and open circles indicate large mammals. (*a–d*) Data subsets for All, 5_occ, 5_taxa, and 10_occ_10_taxa, respectively.
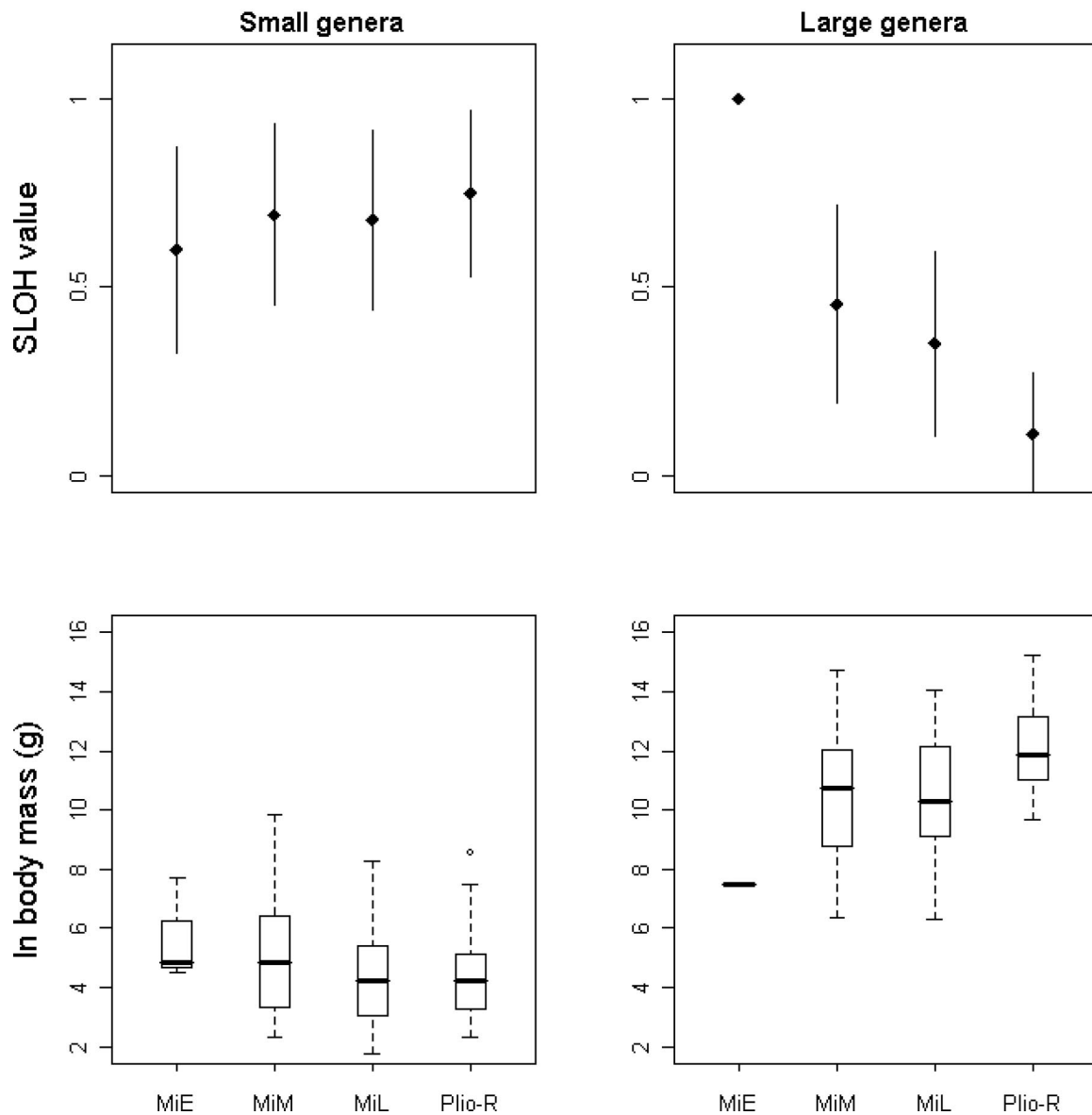
Fig. S2. The change in the proportion of SLOH genera and body mass over the Neogene. (*Upper*) The mean SLOH values of small and large genera over the four approximate time intervals with one standard deviation. MiE, early Miocene; MiM, middle Miocene; MiL, late Miocene; Plio-R, Pliocene to Recent. (*Lower*) Boxplots of the change in average ln body mass over the same period for small and large genera. In all cases, the data are for NOW genera for which there are still living species from which we could estimate average body masses and code SLOH values.
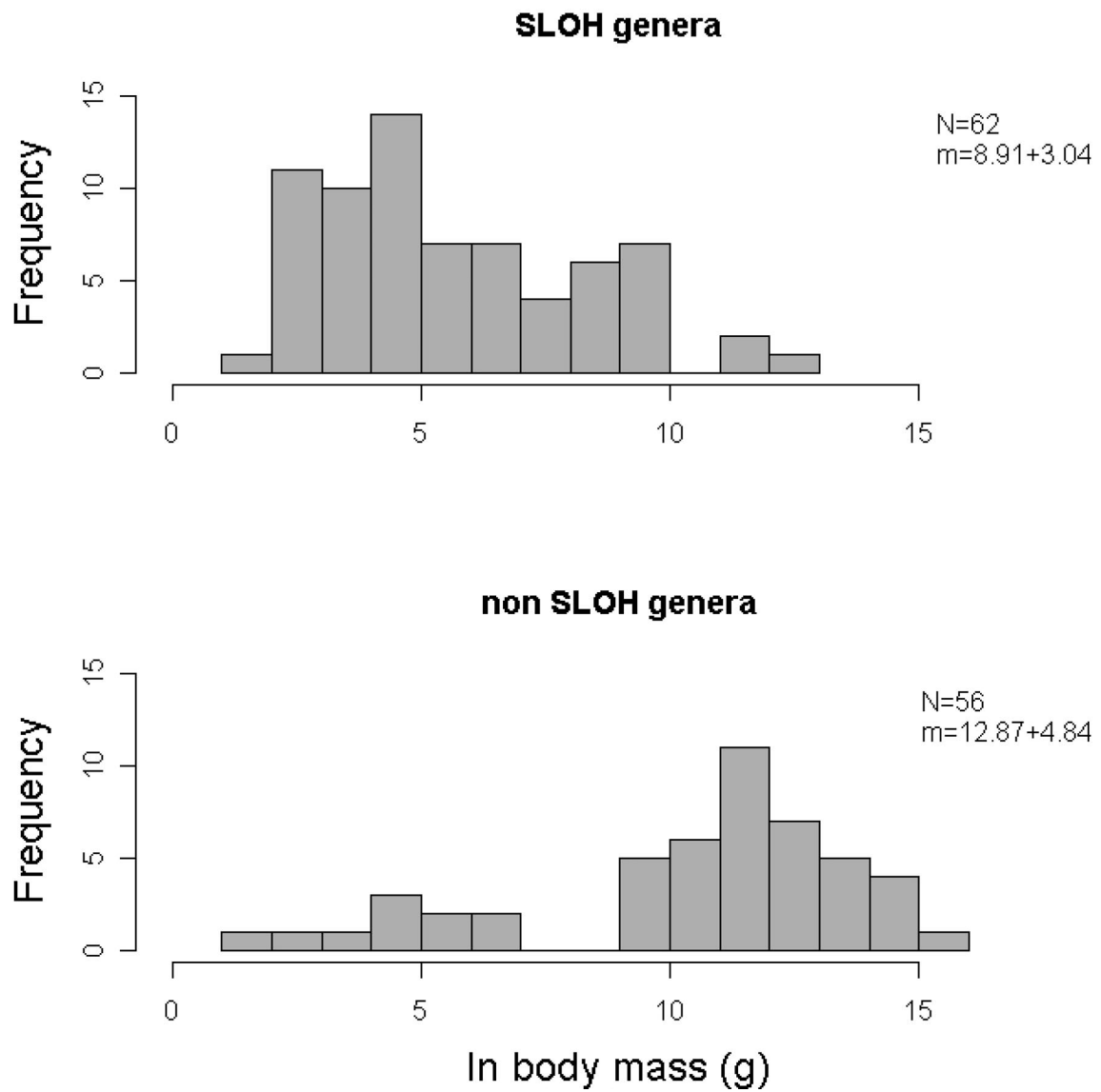
SLOH genera

N=62
m=8.91+3.04

non SLOH genera

N=56
m=12.87+4.84

ln body mass (g)

Fig. S3. Alternative plot of Fig. 3.

Other Supporting Information Files

Dataset S1(XLS)