The extended Price equation quantifies species selection on mammalian body size across the Palaeocene/Eocene Thermal Maximum

Brian D. Rankin1,2, Jeremy W. Fox1, Christian R. Barrón-Ortiz1, Amy E. Chew3, Patricia A. Holroyd2, Joshua A. Ludtke1, Xingkai Yang1 and Jessica M. Theodor1

1Department of Biological Sciences, University of Calgary, 2500 University Drive, Calgary, Canada AB T2N 1N4
2Museum of Paleontology, University of California, 1101 Valley Life Sciences Building, Berkeley, CA 94720, USA
3Department of Anatomy, Western University of Health Sciences, 309 E. Second Street, Pomona, CA 91766, USA

Species selection, covariation of species’ traits with their net diversification rates, is an important component of macroevolution. Most studies have relied on indirect evidence for its operation and have not quantified its strength relative to other macroevolutionary forces. We use an extension of the Price equation to quantify the mechanisms of body size macroevolution in mammals from the latest Palaeocene and earliest Eocene of the Bighorn and Clarks Fork Basins of Wyoming. Dwarfing of mammalian taxa across the Palaeocene/Eocene Thermal Maximum (PETM), an intense, brief warming event that occurred at approximately 56 Ma, has been suggested to reflect anagenetic change and the immigration of small bodied-mammals, but might also be attributable to species selection. Using previously reconstructed ancestor–descendant relationships, we partitioned change in mean mammalian body size into three distinct mechanisms: species selection operating on resident mammals, anagenetic change within resident mammalian lineages and change due to immigrants. The remarkable decrease in mean body size across the warming event occurred through anagenetic change and immigration. Species selection also was strong across the PETM but, intriguingly, favoured larger-bodied species, implying some unknown mechanism(s) by which warming events affect macroevolution.

1. Introduction

Species selection is analogous to evolution by natural selection [1–3]. If individual organisms vary in traits that affect reproduction and mortality, and if offspring tend to resemble their parents, the average trait value of the population will change over time. Similarly, if species vary in traits that affect their propensity to speciate and become extinct, and if descendant species tend to resemble their immediate ancestors, then the average trait value of a set of species will change over time. Indirect evidence suggests the operation of species selection [1,2]; however, interpretation of this evidence is challenging because other factors can create similar patterns. An ideal, direct test for species selection requires knowledge of ancestor–descendant relationships, extinction events and ancestral character states. The palaeontological record can provide this information, although determining ancestor–descendant relationships for extinct taxa is challenging unless the fossil record is dense and reasonably continuous [4,5].

Lack of appropriate data has prevented the application of analytical methods for quantifying species selection and distinguishing it from other mechanisms of macroevolutionary change. Price [6,7] derived a general equation that partitioned directional evolutionary change in mean phenotype into components attributable to selection and other causes. Originally, the Price equation was applied to...
evolving populations of individual organisms, but it also can be used to separate directional macroevolutionary change into components attributable to species selection and anagenetic change [8]. Recent reviews [1,3] have identified the Price equation as a promising tool for quantifying species selection, but attempts to apply it remain few [9,10].

One limitation of the original Price equation is that it assumes a population closed to immigration. Kerr & Godfrey-Smith [11] extended the Price equation to incorporate the effects of immigration, with the insight that immigration that generates directional phenotypic evolution is the ‘mirror image’ of selection. In the context of macroevolution, immigrants are descendant species without ancestors in the dataset, making them the mirror image of ancestral species with no descendants in the dataset. This greatly expands the range of cases to which the Price equation applies.

Here we use the extended Price equation to examine the macroevolution of mean body size of mammals from the Bighorn and Clarks Fork Basins of Wyoming across the Palaeocene/Eocene Thermal Maximum (PETM), when temperatures may have increased by as much as 5 to 9°C [11]. Consider a set of $n^a$ ancestral species present at time 1 and a set of $n_d$ descendant species present at some later time 2. Let $C_j$ be an indicator variable that equals 1 if descendant species $j$ descended from, or is the same species as, ancestral species $i$, and 0 otherwise. Ancestral species $i$ thus has $C_j = 1$ if ancestral species $i$ becomes extinct before time 2 without giving rise to any descendant species, or if it persists at time 2, along with all species that it left, normalized by the average number of descendant species, so that $C_j \geq 1$. Let $z_j$ be the body size of ancestral species $i$ and let $z_i$ be the body size of descendant species $j$.

We quantify macroevolutionary change from time 1 to time 2 as the change in average size of the species present: $\Delta z = \bar{z}_2 - \bar{z}_1 = \sum_{i=1}^{n^a} (z_i/n_a) - \sum_{i=1}^{n^d} (z_j/n_d)$. Total macroevolutionary change $\Delta z$ can then be partitioned into three additive components [11]

$$
\Delta z = \cov(C_j, z_j) / (C_j/n_d^a) - \cov(C_j, z_j) / (C_j/n_d^d) + \text{ave}((\Delta z)_j)
$$

where $\cov$ is the covariance operator, $\text{ave}$ is the average (arithmetic mean) operator, $(\Delta z)_j = C_j (z_j - \bar{z})$ is the difference in body size between descendant species $j$ and the ancestral species $i$ from which it descended (defined as 0 if species $j$ did not descend from species $i$), and $\text{ave}((\Delta z)) = \sum_{i=1}^{n^a} \sum_{j=1}^{n^d} C_j (z_j - \bar{z}) / C_j$ is the average difference in body size between descendant species and the ancestors from which they descended.

The first term in equation (2.1) is the covariance between ancestral species body size and the number of descendant species they left, normalized by the average number of descendant species per ancestral species. This term captures the effect of species selection. It equals zero when all ancestral species are the same size (leaving no variation for species selection to act on), or when the number of descendants an ancestral species leaves is independent of its body size (no species selection on body size) (figure 1a).

Our definition of species selection implicitly incorporates both the selection gradient on body size and the heritability of body size, thereby capturing the macroevolutionary effect of species selection, in contrast to Simpson [10]. This definition also places the mechanisms of macroevolution, including species selection, on a common scale. Separating species selection into subcomponents further clarifies its interpretation. The first term in equation (2.1) can be rewritten as

$$
\frac{\cov(C_j, z_j)}{(C_j/n_d^a)} = \frac{\beta_{z_j} (z_j)}{(C_j/n_d^a)}.
$$

Equation (2.2) follows from equation (2.1) by the definition of linear regression: the (population) covariance of any two variables $x$ and $y$ equals the slope $\beta_{x,y}$ of the regression of $y$ on $x$, multiplied by the (population) variance of $x$. Equation (2.2) states the change in mean body size attributable to species selection increases with the variance among ancestral species in
body size, \( \sigma^2_z \), and the selection gradient, the slope of the regression of the number of descendant species on ancestral body, \( \beta_{C,zt} \). Both components of equation (2.2) are normalized by the mean number of descendants per ancestral species, \( C_z / n^z \). If the average ancestral species leaves few descendants (e.g. \( C_z / n^z \) close to zero), then species selection can have a large effect on mean body size even if there is little variation among ancestral species in body size, and even if the regression between ancestral body size and number of descendants has a shallow slope. Conversely, if the average ancestral species leaves many descendants, then there must be substantial variation among ancestral species in body size, and a steep regression between ancestral body size and number of descendants, for species selection to have a large effect on mean body size.

The second term in equation (2.1) is the covariance between descendant body size and the number of ancestors each descendant species bears, scaled by the average number of ancestors per descendant species, \( C_z / n^z \). Both components of equation (2.2) are normalized by the mean number of descendants per ancestral species, \( C_z / n^z \). This term captures the effect of immigration. This term equals zero when there are no immigrant species, and when immigrants are the same size on average as non-immigrant descendant species (figure 1b). The second term can be rewritten as

\[
\text{cov}(C_z, z^d) = \left( \frac{\beta_{C,zt}}{C_z/n^z} \right) \left( \frac{\sigma^2_z}{C_z/n^z} \right).
\]

Equation (2.3) states that the change in mean body size owing to non-random immigration increases with the variance among descendant species in body size, \( \sigma^2_z \), and the slope of the regression of the number of ancestral species on descendant body size, \( \beta_{C,zt} \). Both components of equation (2.3) are normalized by the mean number of ancestors per descendant species, \( C_z / n^z \). This equals the proportion of descendant species that are not immigrants, and necessarily is greater than 0 and less than or equal to 1. Thus, the greater the proportion of descendant species that are immigrants (i.e. the smaller \( C_z / n^z \)), the greater the effect of non-random immigration on mean body size.

The third term in equation (2.1) captures anagenetic change in body size, and represents the average difference in body size between a descendant species and its ancestor. It includes differences in body size between descendant species and their ancestors that evolve at the time of speciation and differences in body size that evolve after speciation (termed ‘cladogenetic’ and ‘anagenetic’ change, respectively, by Simpson [10]). Immigrant species and ancestral species with no descendants do not contribute to this term. This term equals zero when descendant species are the same size on average as their ancestors (figure 1c).

All analyses were conducted using R v. 2.15.1 (R Foundation for Statistical Computing, Vienna, Austria; for R code, see electronic supplementary material, appendix S1).

(b) Regional setting and localities

Latest Palaeocene and earliest Eocene sediments from the Bighorn and Clarks Fork Basins, Wyoming, USA (from approx. 57 to 55 Ma) document the most complete record of mammalian evolution across the PETM (e.g. [14,17–20,22,23]). Approximately 2000 fossil localities, yielding more than 5000 mammalian specimens, have been discovered (figure 2), with most of these assemblages directly tied to local composite stratigraphic sections [17,22,25]. Fossil remains from this region have been principally collected by surface prospecting of erosional-surface lag accumulations, and slightly less than 25% of samples in the latest Palaeocene and earliest Eocene were quarried or screen-washed.

On the basis of fossil mammals, strata within the Bighorn and Clarks Fork Basins are temporally divided into North American Land Mammal ‘ages’ (NALMAs) [22,25,26]. The latest Palaeocene consists of the Clarkforkian NALMA, which is further divided into three biozones (i.e. from oldest to youngest: C1, C2 and C3), and the earliest Eocene comprises the Wasatchian NALMA, which contains seven biozones (i.e. Wa0, Wa1/2, Wa3, etc.); sediments representing Wa1 are stratigraphically thin in the southern Bighorn Basin [23], and it is not possible to distinguish Wa1 and Wa2 in several sources of data (e.g. [27]; therefore, we use the slightly more inclusive Wa1/2 biozone. Here, all Clarkforkian sub-divisions were examined but only the two earliest biozones of the Wasatchian were included.

Oxygen isotope and palaeobotanical records indicate a general long-term trend of warming through the Clarkforkian, with estimates of the regional mean annual temperature during the earliest interval reaching 13°C and commonly exceeding 16°C during the late Clarkforkian [28]. Mean annual temperatures within the Bighorn and Clarks Fork Basins may have reached as high as 23°C at the PETM, and range from 16°C to 20°C during the Wa1/2, although the temperature may have been cooling throughout this biozone [29].

(c) Evolutionary relationships

Twenty-two lineages, encompassing 10 orders, 24 genera and 50 species, spanning the five biozones, were included in the analyses (electronic supplementary material, datasets S1–S4). Evolutionary relationships and occurrence data from the published literature were used to identify resident and immigrant taxa (electronic supplementary material, datasets S1–S4). The vast majority of these studies used a stratophenetic approach to reconstruct genealogical relationships [4], and although these relationships are not the results of explicit cladistic analyses, they represent falsifiable phylogenetic hypotheses [30].

For instances in which two or more species of a lineage occur within a single time interval (e.g. Azygonx during the latest Clarkforkian and earliest Wasatchian), each species was treated as a discrete data point with a separate connection to the ancestor.
in the preceding biozone. Lineages with incomplete data, including lineages in which a species is known to have been present in the Bighorn and Clarks Fork Basins during a specific time interval but measurements of the appropriate molar are not known, or those with uncertain evolutionary relationships, were excluded from the analyses.

(d) Body size
We estimated body size as the natural logarithm of the lower first molar area (or second molar area in the case of rodents), with measurements gathered through personal observation or from the published literature (e.g. [14,20,22,23]). The mean of the natural logarithm was determined for each species and for each biozone from which the species is known (electronic supplementary material, table S2).

3. Results
The mean mammalian body size of the lineages included in this study changed considerably across the five time intervals (figure 3a–c). Overall mean mammalian body size decreased slightly from the early to middle Clarkforkian before increasing from the middle to late Clarkforkian. Further, and in agreement with other studies [14,17–20], mean body size decreased substantially across the Clarkforkian/Wasatchian boundary but subsequently increased slightly between Wa0 and Wa1/2.

The substantial decrease in mean mammalian body size in the Bighorn and Clarks Fork Basins across the PETM is principally attributable to the immigration of small species and, to a lesser extent, anagenetic evolution towards smaller body sizes (figure 3b,c). The effect of immigration across the PETM was equal in magnitude but opposite in direction to the effect of anagenetic change. The smaller changes in mean mammalian body size detected across the remaining time intervals predominantly reflect the result of species selection, with anagenetic change and immigration modestly adding or subtracting to the effect of species selection.

Variation over time in the direction and magnitude of non-random immigration reflects variation in both the proportion of descendant species that were immigrants and the association between descendant species body size and their number of ancestors (i.e. whether or not the taxa were immigrants; electronic supplementary material, figure S2). The effect of non-random immigration was much larger in magnitude across the PETM than in other time intervals as this is the only interval in which there were both a high proportion of immigrants among the descendant species and an appreciable difference in average size between immigrant and non-immigrant descendant species (electronic supplementary material, figure S2).

4. Discussion
(a) Change in mean mammalian body size
The decrease in mean mammalian body size across the PETM of the Bighorn and Clarks Fork Basins is predominantly
attributable to immigration during the earliest Wasatchian, with a smaller contribution from anagenetic change. The immigration of small-bodied taxa into the basins, such as the condylarth *Ectocion parvus* and the euprimate *Teilhardina bradlii*, as well as the reduction in body size of resident taxa, including the carnivoran *Vixeraurus politus*, contribute towards these terms. These findings corrobore previous studies [18–21], while separating and quantifying the relative importance of these two mechanisms for the first time. The results also reveal an important and unexpected influence of species selection on the change in mean mammalian body size across the PETM. Species selection across the PETM was approximately equal in magnitude to anagenetic change, but favoured larger-bodied resident taxa. Both the extinction of the small-bodied plesiadapiform primate *Carpolestes simpsoni* and the speciation of the large-bodied tillodont *Azygonyx* contribute towards this term (electronic supplementary material, figure S1). Interestingly, the decline and ultimate demise of plesiadapiform primates is often associated with the appearance of rodents and euprimates during the late Palaeocene and earliest Eocene, respectively [31]. Although discriminating competitive interactions in the fossil record remains difficult, the pattern revealed here provides some support for competition as a means of species selection [32–34].

That species selection and anagenetic change opposed one another is not necessarily surprising. From a conceptual perspective, species selection is an example of ‘multilevel selection’ [35], in which fitnesses and traits are attributes of groups rather than the individuals comprising the groups, and group fitness is not defined in terms of the fitnesses of the individuals comprising the group. Covariation between group traits and group fitness (i.e. species selection) is different than changes in group traits values (i.e. anagenetic change). There is no reason to expect that natural selection within species will favour the same traits as species selection among species [2]. Smaller-bodied individuals, for example, might be favoured by natural selection in warm environments with nutrient-poor plants [18,19], but smaller-bodied species would not necessarily be more prone to speciate or be less prone to extinction.

By quantifying the effects of anogenesis, immigration and species selection, the extended Price equation constrains the search for underlying biological mechanisms. Both anagenetic change towards smaller body sizes and immigration of smaller-bodied species may reflect Bergmann’s rule [20]; yet these two mechanisms alone are not a complete explanation for mammalian body size change as species selection across the PETM favoured larger body sizes. Body size is correlated with numerous other individual- and species-level properties that might be subject to both individual and species selection, and determining causality is difficult. The biological mechanisms by which increasing temperatures, higher pCO$_2$ levels and drier climates might cause anagenetic changes in body sizes were recently reviewed [36]. Intraspecific immigration also could generate anagenetic change, with the body size of a species in our study area decreasing because of the immigration of small-bodied conspecifics from more southerly populations. Our results highlight the need to identify mechanisms by which environmental change might interact with body size to affect speciation and extinction rates. Note, however, that it is unclear if the same environmental changes that led to anagenetic evolution and immigration of small-bodied species also generated species selection. Species selection also favoured large-bodied mammals immediately before PETM, and was nearly as strong in the intervals just prior to the PETM (middle to late Clarkforkian) as across the PETM (figure 3c).

Species selection in our dataset is not necessarily representative of species selection occurring under other conditions. Some previous theory and data suggest that species selection in mammals generally will favour smaller body sizes [37,38]. Yet species selection may also favour species with larger geographical ranges because they are less prone to extinction [1]. Body size and geographical range size are positively correlated in many groups, though not all [39], so species selection for larger geographical ranges might lead to correlated species selection for larger body size. Furthermore, recent evidence indicates that net diversification rates of animals generally are
independent of body size, with both speciation and extinction rates per generation decreasing with body size [40].

The large magnitude of species selection is somewhat surprising, given the few extinction and speciation events in the dataset. Across each of the intervals, only a small proportion of ancestral species became extinct without leaving any descendants, and only a minority left more than one descendant (electronic supplementary material, figure S1). The magnitude of species selection, however, does not scale in any simple way with the number of speciation and extinction events, just as the response to natural selection in microevolution does not scale in any simple way with the number of births and deaths. Rather, the magnitude of species selection increases with the steepness of the selection gradient and the variance in ancestral body size, and decreases with the number of descendant species deriving from the average ancestral species (see equation (2.2)). Species selection was strong across the PETM principally because there were sufficient extinction events, relative to speciation events, that the average ancestral species left relatively few descendants (electronic supplementary material, figure S1). This result may be general, with the effects of species selection on mean trait values large in magnitude when extinction rates are high relative to speciation rates. However, strong species selection does not arise solely from steep species selection gradients. It also can arise from low mean absolute ‘fitness’ of ancestral species, which magnifies interspecific variation in ‘fitness’.

(b) Strengths and limitations of the extended Price equation

Our results illustrate several strengths of the extended Price equation. Because the equation is retrospective, it provides a complete and exact partitioning of all possible causes of directional macroevolutionary change. Given data on species’ traits at two times the ancestor–descendant relationships linking those two times and the identity of immigrants, it specifies what must have happened in between those times to account for the observed change in mean trait value [41]. No contributing factor is omitted, nor does any unexplained ‘residual error’ remain. Here, the extended Price equation revealed that anagenetic change and immigration did not exhaust the possible drivers of macroevolutionary change; rather, species selection was also operating. In addition, the extended Price equation separated co-occurring effects acting in opposing directions and further quantified the effects of different mechanisms on a common scale. These features of the extended Price equation improve on previous studies that use separate tests for different macroevolutionary mechanisms.

The analysis does not separate the effects of species selection acting directly on body size from correlational selection arising from species selection on some other trait correlated with body size. However, given data on multiple characters, the Price equation can separate direct from correlational selection [10]. Moreover, as with any macroevolutionary study, our results reflect the available data. New fossil discoveries, phylogenetic revisions or the addition of more species might change estimated body sizes and fitnesses. Our findings are also limited by the temporal resolution of the fossil record across the PETM. Species selection, anagenesis and immigration across the Clarkforkian/Wasatchian boundary represent net effects integrating macroevolution during the PETM with macroevolution immediately before and after the event. Even the stratigraphic resolution within the Bighorn and Clarks Fork Basins of Wyoming is insufficient to isolate macroevolution during the PETM. That our analysis recovered several features of larger datasets (e.g. dwarfing across the PETM, immigration of small-bodied taxa) provides confidence that our results are robust.

An alternative approach to estimate trait change attributable to species selection and anagenesis would entail approximating the parameters of a stochastic macroevolutionary model from the data [42–45]. This approach is intriguing, but parameter estimates are sensitive to model choice [43,45], and even minor assumptions can lead to flawed interpretations [46]. Furthermore, no stochastic macroevolutionary model incorporates immigration. The extended Price equation does not make any substantive biological assumptions, and has the ability to integrate the influence of immigration on macroevolution.

(c) Conclusion and future directions

Our findings demonstrate that the dramatic decrease in mean mammalian body size within the Bighorn and Clarks Fork Basins during the latest Palaeocene and early Eocene principally occurred through immigration and anagenetic changes. In contrast to the directionality of these two mechanisms, however, species selection favoured larger-bodied resident taxa. Interestingly, body size variation among ancestral species did not entirely explain the number of descendants each ancestor produced, suggesting that there may be other species-level traits independent of body size that species selection acted upon. These results, in addition to other recent findings [9,10], indicate that species selection might be more influential and more extensive in macroevolution than previously recognized.

Addressing this possibility requires quantifying species selection in other systems. The extended Price equation can be applied to any macroevolutionary dataset for which character states and the evolutionary relationships of species can be assessed, including 750 000-year-old mammalian faunas from cave deposits in Colorado, USA [47], latest Cretaceous (approx. 70 Ma) dinosaurian lineages from the Western Interior of North America [48] and conodont assemblages across the Late Devonian mass extinction (approx. 375 Ma) in the Prototethys Ocean [49].

Rapid global climate change can lead to rapid evolutionary responses [36]. On short time scales of tens to hundreds of generations, responses to climate change are likely to be dominated by within-lineage evolution [50]. On longer time scales, our results illustrate that differential speciation and extinction can be just as important as within-lineage forces. Indeed, as current climate change continues to outpace within-lineage adaptation [51], evolutionary responses at short time scales may be governed by species selection in the form of differential extinction.

Data accessibility. Electronic supplementary material, datasets S1–S4, are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.1h0j6.

Authors’ contributions. J.W.F. designed the research, which was further developed by all the authors. All authors except J.W.F. compiled and processed the data. All authors analysed the data. B.D.R. and J.W.F. wrote the manuscript with input from all other authors.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by an NSERC Alexander Graham Bell Canadian Graduate Scholarship to B.D.R., NSERC Discovery Grants to J.W.F. and J.M.T., and a CONACYT graduate scholarship (no. 310423) to C.R.B.-O.


