Paleobiology, 27(1), 2001, pp. 79–83

Completeness of the mammalian fossil record in the Iberian Neogene

David M. Alba, Jordi Agustí, and Salvador Moyà-Solà

Abstract.—We infer the absolute completeness of the mammalian fossil record of the Iberian Neogene, from a compiled database of preserved stratigraphic ranges for both species and genera and by means of the analytical tools developed by Foote and Raup (1996). We conclude that the mammalian fossil record from the Neogene of the Iberian Peninsula is very complete (more than 75% at the specific level, and more than 90% at the generic one), being a good indicator of how complete the record of terrestrial organisms can be at least under certain favorable conditions. Comparison with previously published results for well-known marine invertebrates indicates that the continental Iberian Neogene record is not significantly less complete, thus raising doubts about the importance of more episodic deposition in continental environments. It remains to be seen if continental faunas in general are as complete as the marine invertebrate record.

David M. Alba, Jordi Agustí, and Salvador Moyà-Solà. Institut de Paleontologia M. Crusafont (DB-Unidad Asociada CSIC), Escola Industrial 23, 08201 Sabadell (Barcelona), Spain. E-mail: dalba@porthos.bio.ub.es

Accepted: 21 June 2000

Introduction

''Determining the completeness of the fossil record is of crucial importance to palaeobiologists'' (Benton 1994: p. 185), since it is customary for both paleontologists and neontologists to complain about the incompleteness of the fossil record. Although this complaint can be followed back to Darwin's *Origin of Species* (Darwin 1859), where it was used to explain the absence of transitional forms expected by the theory of evolution, numerical analyses are rarely provided. The fossil record is obviously incomplete for several reasons (Benton 1994), and the degree of incompleteness can vary dramatically from group to group, but even in groups frequently found as fossils, only a proportion of the extinct species are preserved. Incompleteness, however, should not be confused with inadequacy, which depends on the aims of any particular study (Paul 1991), and will not be examined here.

Completeness can be defined and estimated in different ways (Paul 1991), although it generally refers to the completeness of the fossil list, which can be defined as ''the proportion of taxa that have left some fossil record'' (Foote and Raup 1996: p. 121; cf. Valentine 1986). To infer the proportion of past species known as fossils, a first (although too coarse) approximation would be estimating incompleteness as the proportion of stratigraphic ranges that are occupied by gaps in the fossil record for any given species (Paul 1991). A similar, more refined methodology can be used at higher taxonomic levels, like Benton and Storrs's (1994) estimates of the so-called relative completeness index (RCI), based on the calculation of the minimum implied gaps for any given cladogram and the range length of each taxon (see also Benton 1994). However, this kind of analysis is highly dependent on the reliability of phylogenetic (cladistic) hypotheses, because low values of completeness can be due not only to a small portion of taxa preserved (or inadequate sampling), but also to incorrect cladograms (due to homoplasy) (Wagner 1995).

Good estimates of completeness for some groups at a given time interval can be obtained by means of the more powerful, phylogeny-free analytical tools recently developed (Foote and Raup 1996; Foote 1997). This enables us to make much more reliable estimates of completeness on the basis of data from stratigraphic ranges alone. These inferences are derived from the computation of the probability of taxon preservation per stratigraphic interval (see Material and Methods). This has been shown, at least at the generic

0094-8373/01/2701-0006/$1.00

TABLE 1. Summary of main basins covering the Iberian Neogene.

Basin	Time interval	Reference
Ebro	Oligocene to early late Miocene	Krijgsman et al. 1996
Duero	Middle to early late Miocene	Krijgsman et al. 1996
Tagus and Loranca	Early to middle Miocene	Krijgsman et al. 1996
Vallès-Penedès	Early to late Miocene	Agustí et al. 1997, 1999
Calatayud-Daroca	Early to late Miocene	Krijgsman et al. 1996
Teruel	Late Miocene to Pliocene	Opdyke et al. 1997
Fortuna	Late Miocene to Pliocene	Garcés et al. 1998
Júcar	Pliocene	Opdyke et al. 1997
Granada	Late Miocene to Pleistocene	Agustí et al. 1990
Guadix-Baza	Pliocene to Pleistocene	Agustí and Moyà-Solà 1992

level, to be well correlated with another independent measure of completeness, the proportion of living families that have left some fossil record (Foote and Sepkoski 1999), thus strengthening the reliability of the inferences of completeness derived by means of this method.

Studies based on the RCI have shown that incompleteness is not as problematic as previously thought. Thus, even in tetrapods a good agreement exists in many instances between the stratigraphic order in which groups appear in the fossil record and the sequences of lineage divergence predicted by cladistic analyses (Norell and Novacek 1992; Benton and Storrs 1994). Both common sense and traditional knowledge suggest that the completeness of the marine fossil record (at least for benthic organisms from shallow environments) is better than that of terrestrial ones (e.g., Valentine 1969). Most of the deficiencies of the fossil record of durably skeletonized forms have been attributed to tectonic destruction or burial of fossiliferous sediments rather than to the lack of accumulation (Valentine 1989). Although this may be true for marine organisms, this is certainly not the case for the much more patchy continental fossil record, where accumulation is more irregular (episodic) and fossils are generally much more sparse. Many gaps in the continental record could result from original bad preservation rather than late deterioration, leading to greater incompleteness of the terrestrial fossil record.

Surprisingly, the comparison of the fossil records of tetrapods and echinoderms at the family level indicates that this is not necessarily true (Benton and Simms 1995). Thus, both the match between the stratigraphic and the cladistic order of group appearances and the computation of the RCI of Benton and Storrs (1994) indicate that the fossil record of tetrapods is at least as good as or even better than that of echinoderms (Benton and Simms 1995). The latter authors suggest that this apparent paradox could be due to more intensive study of the vertebrate fossil record, i.e., that their results might reflect the current state of knowledge rather than the record itself. To test these results more accurately, we infer the completeness of the mammalian fossil record of the Iberian Neogene at the genus and species level and compare it with published results on marine invertebrates. The method employed (Foote and Raup 1996; Foote 1997) has the advantages that it is performed at a lower taxonomic level and is independent of inferred phylogenies.

Our aim here is to test if (at least under certain circumstances) the record of terrestrial organisms can be as good as that of marine ones. The mammalian record of the Iberian Neogene provides a good opportunity to do so, because it is thoroughly sampled, comprising a number of endorheic and intramontane basins that cover most of the Neogene (see Table 1 and references therein). Most of the early and middle Miocene is well represented by four large basins (Ebro, Duero, Tagus, and Loranca), with special relevance of some marginal and intramontane basins (Vallès-Penedès and Calatayud-Daroca), whereas the late Miocene to Pliocene succession is best represented by Teruel and small basins of the Betic Cordillera (Fortuna, Júcar, Granada and

TABLE 2. Preserved range frequencies for species and genera.

	Frequency													
	f(1)	f(2)	f(3)	f(4)	f(5)	f(6)	f(7)	f(8)	f(9)	f(10)	f(11)	f(12)	f(13)	Total
Species	292	89	41	21	11	3	1	0	0	0	0	0	0	458
Genera	108	52	33	30	20	12	6	7	0	2	0	0	1	271

Guadix-Baza), with the last one including a unique mammalian succession that covers the whole Plio-Pleistocene.

Material and Methods

We compiled the stratigraphic ranges of species and genera of fossil mammals from the Neogene of Spain and Portugal using information from a database of paleontological literature. This database was assembled at the Institut de Paleontologia Miquel Crusafont in Sabadell (Barcelona, Spain), under the guidance of one of the authors (Agustí), as a result of the European Science Foundation network ''Hominoid Evolution and Environmental Change in the Neogene of Europe'' (Agustí et al. 1998). We followed the MN (Mammal Neogene) biochronological units convention (see reviews in Mein 1999 and Steininger 1999), from MN2 to MN17 (with MN7/8 combined). Identifications ''cf.'' were used to compute ranges, whereas identifications ''sp.'' were omitted for the analysis at the specific level unless no other identified species of the genus at issue was present.

To compute completeness, we followed the discrete time interval and infinite window devised by Foote and Raup (1996), which gives results similar to those produced by later refinements of the method (Foote 1997). According to the procedure employed, the range of a taxon is the number of intervals (in this case, MN units) from its first to last occurrence, inclusive (so that single occurrences, called single hits, have a range of one). Details of the method are described by Foote and Raup (1996) and Foote (1997). The first step was to compute the frequencies of taxa with preserved stratigraphic range t: $f(t)$. These data were used to calculate the preservation probability R (probability that a taxon is preserved at least once in one interval), which is estimated, assuming stochastically constant preservation probability, by the range-frequency ratio *FreqRat*, calculated as follows:

$$R \approx FreqRat = f(2)^2/[f(1) \cdot f(3)]. \qquad (1)$$

The probability $P_1(T)$ that a taxon is preserved at least once (range equal to or greater than 1), given an original duration T, is calculated as

$$P_1(T) = 1 - (1 - R)^T. \qquad (2)$$

The estimated preservation probability R (probability of preservation per taxon per interval) is not enough to infer the quality of the fossil record, since it should be considered with respect to the distribution of true durations (Foote and Raup 1996). The original durations $h(T)$ can be inferred assuming an exponential distribution of original durations; this means that extinction rate q is constant and can be calculated from the ln-slope of range frequencies (except that of single hits), as the slope with the sign reversed. The probability that the original duration is equal to T can be therefore computed as follows:

$$h(T) = e^{-q(T-1)} - e^{-qT}. \qquad (3)$$

Completeness P_p can therefore be calculated as the sum of proportions of taxa preserved for each original duration, i.e., as the sum, for all original durations, of the product between the probability $h(T)$ of having an original duration T and the probability $P_1(T)$ of being preserved at least once in an interval given an original duration of T:

$$P_P = \sum_{T=1}^{\infty} h(T) \cdot P_1(T). \qquad (4)$$

Results

The preserved range frequencies for species and genera have been reported in Table 2. When single hits are excluded, these range frequencies were found to fit very well an exponential distribution: $r = 0.990$ ($p < 0.001$)

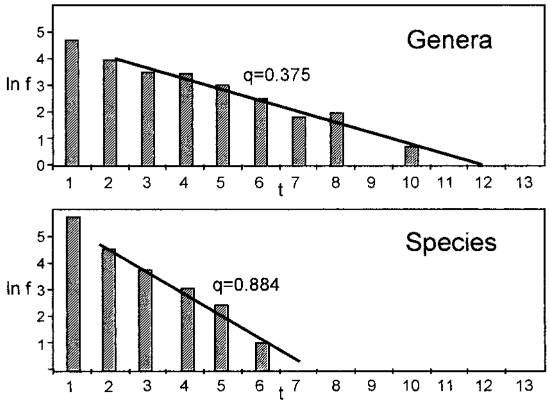

FIGURE 1. Logarithmic plot of range frequencies (natural logs) for genera and species and least-squares regression line. f = frequency; t = actual range; q = extinction rate (the slope with reversed sign). Single hits $f(1)$ and frequencies of zero have been excluded from the computation of the regression.

for species and $r = 0.989$ ($p < 0.001$) for genera (Fig. 1). We therefore estimated the range-frequency ratio by using raw range frequencies (R). The values obtained were $R = 0.662$ for species and 0.759 for genera. The constant extinction rate per interval (as estimated from the slope of the ln-range frequencies regression) was 0.884 for species and 0.375 for genera. These values provide estimates of completeness of 77% ($P_p = 0.769$) at the specific level and 91% ($P_p = 0.910$) at the generic level.

Discussion

From the above-mentioned results, we can conclude that more than three/fourths of the mammalian species that existed during the Miocene and the Pliocene of the Iberian Peninsula have been preserved in the fossil record. Completeness is higher for genera, which seems logical since genera usually include more than one species and have longer ranges. Although different, results at both taxonomic levels are consistent with one another, suggesting that data at the generic level are enough to obtain rough estimates of the proportion of past diversity preserved in the fossil record.

This figure of completeness at the species level indicates a very good fossil record, especially if we take into account that future refinement of the database will tend to increase, rather than decrease, this figure (by extending preserved known ranges and including previously excluded taxa). According to these results, the Iberian fossil record is more complete (77%) than the extensively studied record of Cenozoic mammalian species from North America (58% complete, as calculated by Foote and Raup [1996], on the basis of the compilation of stratigraphic ranges by Alroy [1992, 1994] for 2941 Cenozoic species and 0.7-m.y. intervals). The results are not fully comparable, however, for several reasons, including longer time span, much more extensive area, and shorter stratigraphic intervals in the latter case.

It remains to be seen whether our results merely reflect the restricted time span and/or geographical area analyzed, or are more representative of the completeness of the terrestrial fossil record as a whole. The computation of completeness in other similarly restricted geographical areas from Eurasia and in all of them combined would allow us to test our results, but this would require a more extensive database and is outside the scope of this paper. We therefore take our results from the Iberian Neogene as tentatively indicating how complete the terrestrial fossil record can be.

Interestingly, the comparison with previously published figures indicates that the completeness of terrestrial mammals is not so different from that of marine invertebrates, thus partially confirming the results of Benton and Simms (1995) at the family level. For example, our figures at the species level (77%) are lower than those for Upper Cambrian-Lower Ordovician trilobites from Oklahoma and European Jurassic bivalves (90% [Foote and Raup 1996]), but higher at the generic level (91%) than those for Ordovician–Devonian crinoids (70% [Foote and Raup 1996]). This suggests that, at least locally, the terrestrial fossil record can be as good as the marine one. The unexpected completeness of the terrestrial record may indicate that hiatuses in deposition are fewer than expected, or alternatively that there is considerable redundancy among continental localities that compensates for less-continuous accumulation.

Conclusions

The record of Iberian Neogene mammals seems to be a good indicator of how complete the record of terrestrial organisms can be, at the genus and species levels, under conditions favorable for preservation and with extensive sampling. The comparison with values of completeness for well-known marine invertebrate groups with a high fossilization potential indicates that the continental fossil record can be approximately as good as the marine one, thus suggesting that depositional differences are perhaps less significant than previously thought for the reliability of the continental fossil record. These results should be refined by improving our database and extending it to larger geographic areas. Above all, completeness estimates from other fossil records (different taxa, geographic areas, and/or periods of time) are required for more extensive comparisons. The record from the Iberian Neogene is one of the most extensively studied and best documented for mammals, so that more information is needed to see whether the results reported here represent the rule, rather than the exception, for the continental fossil record.

Acknowledgments

We are indebted to the European Science Foundation for providing funds for the elaboration of the database used in this study. We are also grateful to M. Foote for insightful comments and suggestions during the review process. This paper was possible thanks to the project DGICYT PB97-0157 of the Spanish Ministry of Education and Culture, and D. M. A. has been supported by a predoctoral fellowship (1999FI 00765) from the Generalitat de Catalunya.

Literature Cited

Agustí, J., and S. Moyà-Solà. 1992. Mammalian dispersal events in the Spanish Pleistocene. Courier Forschungs-Institut Senckenberg 153:69–77.

Agustí, J., C. Castillo, M. Freudenthal, E. Martín-Suárez, and J. A. Peña. 1990. On the presence of *Allophaiomys* (Rodentia) in the Granada basin (Spain). Scripta Geologica 93:35–40.

Agustí, J., L. Cabrera, M. Garcés, and J. M. Parés. 1997. The Vallesian mammal succession in the Vallès-Penedès basin (northeast Spain): paleomagnetic calibration and correlation with global events. Palaeogeography, Palaeoclimatology, Palaeoecology 133:149–180.

Agustí, J., P. Andrews, M. Fortelius, and L. Rook. 1998. Hominoid evolution and environmental change in the Neogene of Europe. Journal of Human Evolution 34:103–107.

Agustí, J., L. Cabrera, M. Garcés, and M. Llenas. 1999. Mammal turnover and global climate change in the late Miocene terrestrial record of the Vallès-Penedès basin (NE Spain). Pp. 397–412 *in* J. Agustí, P. Andrews, and L. Rook, eds. Hominoid evolution and climatic change in Europe, Vol. 1. The evolution of Neogene terrestrial ecosystems in Europe. Cambridge University Press, Cambridge.

Alroy, J. 1992. Conjunction among taxonomic distributions and the Miocene mammalian biochronology of the Great Plains. Paleobiology 18:326–343.

———. 1994. Quantitative mammalian biochronology and biogeography of North America. Ph.D. dissertation. University of Chicago, Chicago.

Benton, M. J. 1994. Paleontological data and identifying mass extinctions. Trends in Ecology and Evolution 9:181–185.

Benton, M. J., and M. J. Simms. 1995. Testing the marine and continental fossil records. Geology 23:601–604.

Benton, M. J., and G. W. Storrs. 1994. Testing the quality of the fossil record: paleontological knowledge is improving. Geology 22:111–114.

Darwin, C. 1859. On the origin of species. A facsimile of the first edition with an introduction by Ernst Mayr, 1964, Harvard University Press, Cambridge.

Foote, M. 1997. Estimating taxonomic durations and preservation probability. Paleobiology 23:278–300.

Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22:121–140.

Foote, M., and J. J. Sepkoski Jr. 1999. Absolute measures of the completeness of the fossil record. Nature 398:415–417.

Garcés, M., W. Krijgsman, and J. Agustí. 1998. Chronology of the late Turolian of the Fortuna basin (SE Spain): implications for the Messinian evolution of the eastern Betics. Earth and Planetary Science Letters 163:69–81.

Krigjsman, M., M. Garcés, C. G. Langereis, R. Daams, J. van Dam, A. van der Meulen, J. Agustí, and L. Cabrera. 1996. A new chronology for the middle to late Miocene continental record in Spain. Earth and Planetary Science Letters 142:367–380.

Mein, P. 1999. European Miocene mammal biochronology. Pp. 25–38 *in* G. E. Rössner and K. Heissing, eds. The Miocene land mammals of Europe. Fritz Pfeil, Munich.

Norell, M. A., and M. J. Novacek. 1992. The fossil record and evolution: comparing cladistic and paleontologic evidence for vertebrate history. Science 255:1690–1693.

Opdyke, N., P. Mein, E. Lindsay, A. Perez-Gonzalez, E. Moissenet, and V. L. Norton. 1997. Continental deposits, magnetostratigraphy and vertebrate paleontology, late Neogene of Eastern Spain. Palaeogeography, Palaeoclimatology, Palaeoecology 133:129–148.

Paul, C. R. C. 1991. Completeness of the fossil record. Pp. 298–303 *in* D. E. G. Briggs and P. R. Crowther, eds. Palaeobiology: a synthesis. Blackwell Scientific, Oxford.

Steininger, F. F. 1999. Chronostratigraphy, geochronology and biochronology of the Miocene ''European Land Mammal Mega-Zones'' (ELMMZ) and the Miocene ''Mammal-Zones (MN-Zones).'' Pp. 9–24 *in* G. E. Rössner and K. Heissing, eds. The Miocene land mammals of Europe. Fritz Pfeil, Munich.

Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. Palaeontology 12:684–709.

———. 1989. How good was the fossil record? Clues from the Californian Pleistocene. Paleobiology 15:83–94.

Wagner, P. J. 1995. Stratigraphic tests of cladistic hypotheses. Paleobiology 21:153–178.