Do species populations really start small? New perspectives from the Late Neogene fossil record of African mammals

E. S. Vrba[1*] and D. DeGusta[2]

[1]*Department of Geology and Geophysics, Yale University, PO Box 208109, New Haven, CT 06520-8109, USA*
[2]*Department of Anthropological Sciences, Stanford University, Stanford, CA 94305-2117, USA (degusta@stanford.edu)*

This analysis of all known African larger mammals of the past 10 Myr offers new perspectives on the geographical circumstances of speciation. Our central question is: does the fossil evidence support the idea that most new species start as small populations and, if true, how long is the average growth interval until species are established at their mean later size? This simple question is important to unravelling the competing claims of rival models of speciation. We approached it by direct use of fossil data, which, to our knowledge, has not been done previously. We compared the numbers of fossil site records, as a proxy for magnitude of geographical spread, between survivorship intervals across all species. The results show that the average mammal species has indeed started its life in a relatively small population, and thereafter increased rapidly in geographical spread to reach its long-term equilibrium abundance by about 1 million years after origin. Some theoretical implications of these results are discussed.

Keywords: speciation; vicariance; isolated small populations; physical change; African mammals; Late Neogene

1. INTRODUCTION

In the mid- to late-nineteenth century John Thomas Gulick, a missionary and naturalist, was studying the speciose land snail fauna on the volcanic slopes of the Hawaiian Archipelago (Gulick 1872, 1905). He recorded that lava flows typically separate closely related species of very small total geographical distributions. He noted that these sister-taxa, despite living in very similar environments, were strongly diverged in colour, shape, and in feeding and other behaviours. Gulick argued that the cause of speciation in these cases is not well explained by selection among competitors, namely, by the 'survival of the fittest' (Gulick 1872, p. 224): 'The conditions under which they live are so completely similar, that ... it does not follow that the 'Survival of the Fittest' ... is the determining cause ...'. Instead, Gulick believed that the separation of a small population from its parent species (or vicariance as we now term it) was seminal in initiating speciation. He reasoned that since vicariance results in the chance apportionment of different genetic variants among small isolated populations, it can thereby lead to speciation.

Gulick's early focus on population dynamics differed from Darwin's view of speciation which focused on competition: '... each new species is produced ... by having some advantage over those with which it comes into competition; and the consequent extinction of the less-favoured forms almost inevitably follows' (Darwin 1859,

p. 320). Darwin acknowledged the influence of climate but he stressed its effects on competition rather than on population structure: 'in so far as climate chiefly acts in reducing food, it brings on the most severe struggle between the individuals ... which subsist on the same food' (Darwin 1859, p. 68).

This early polarity between competition and vicariance has persisted in the speciation debate, and is reflected by current models of the geographical circumstances of speciation. These range from a strong version of the sympatric model of speciation (complete integration and required competitive interactions of the speciating organisms with those of the parent species in large populations) to the classic allopatric model of decisive and seminal separation of a small, possibly extremely small, founder population. Reviews that have variously discussed the biogeographic, genetic, reproductive behavioural and palaeontological implications of the two models can be found in Mayr (1963, 1982), Paterson (1981, 1982, 1985), Lande (1979), Futuyma & Mayer (1980), Templeton (1981, 1989), Carson (1982), Butlin (1989) and Vrba (1995*a*).

In this contribution we offer new perspectives on the geographical circumstances of speciation, based on analysis of the African fossil mammals of the past 10 Myr. The term 'speciation' refers to sexually reproducing lineages (Mayr 1942; Carson 1982) and means the process of genotypic and phenotypic divergence in which a subpart of the parent species evolves a new fertilization system that is incompatible with the parental system (Paterson 1981). (The fertilization system is interpreted broadly to include all elements of reproduction that culminate in zygote formation: from meiosis, through the processes by which mating partners find, recognize and mate with each other,

*Author and address for correspondence: Kline Geology Laboratory, 210 Whitney Avenue, New Haven, CT 06511, USA (elisabeth.vrba@yale.edu).

Phil. Trans. R. Soc. Lond. B (2004) **359**, 285–293
DOI 10.1098/rstb.2003.1397

285

to fertilization.) The phylogenetic pattern of speciation can either show irreversible lineage branching (if both parent and daughter populations persist) or a single unbranched lineage if the parent species becomes extinct. Our central question is: does the fossil evidence support the view that most new species of large mammals start as small populations and, if true, how long is the average growth interval until species are established at their mean later size?

The answer has several theoretical implications. If speciation commonly occurs in small populations, then there is a predominant causal role for agencies that split apart previously continuous geographical distributions of species. The overwhelmingly most important causal agencies of this kind are physical changes on the Earth's surface, the most common and regular among which are the global and local effects of astronomically initiated climatic oscillations. Physical changes also include the results of dynamics in the Earth's crust and deeper layers as manifested, for instance, by rifting, uplift, sea-level change and the lava flows that Gulick regarded as important to speciation in Hawaiian snails. If physical change does indeed have a seminal role in the origin of species, then the regularities of physical change may also lead to consequent regularities in biotic evolution, regularities that could produce consistent macroevolutionary patterns among diverse lineages in relation to environmental change. It would imply much closer linkages between the physical and biotic dynamics on Earth than has traditionally been acknowledged. It would also imply that the causal focus of speciation theory cannot be restricted to the level of organismal interactions and selection within populations, but needs to include the population structures and dynamics at the species level.

In summary, although the question 'do species start small?' is simple, it is nevertheless centrally important to understanding the processes by which species originate. We have approached this question by direct use of fossil data, which, to our knowledge, has not been done previously.

2. MATERIAL AND METHODS

One of us (E.S.V.) assembled a database of the fossil records of all the African larger mammal species known over the past 10 Myr. Larger mammals include all taxa other than Insectivora, Chiroptera and Rodentia smaller than Hystricidae (see taxon list and literature sources in Vrba (2000)). Only taxa, the initial description of which as separate species appears to have survived subsequent systematic revision, were included. Time was divided into 20 intervals of 0.5 Myr each ($10 > t \geq 9.5$ Ma, ... $0.5 > t \geq 0$ Ma). For a given species during a given 0.5 Myr interval, the number of fossil site records (we use the abbreviation sr for this number), distinct in geographical location and/or time, from which that species is known serves as a proxy for the magnitude of its geographical spread and abundance during life in that interval.

The sr for each species in each 0.5 Myr interval was counted. The notation f, $s1...sn$ refers to survivorship intervals and survivorship categories of species: the earliest time interval in which a species appears in the fossil record is denoted as the f interval (of first appearance), the immediately following one is the $s1$ interval, the next the $s2$ interval, and so forth, until the last one

from which the species is known, the sn interval. For example, if fossils of species X are known from 4.6 Ma to 3.1 Ma, the f interval for species X would be 4.5–5.0 Ma, its $s1$ interval 4.0–4.5 Ma, its $s2$ interval 3.5–4.0 Ma, and its $s3$ interval 3.0–3.5 Ma. It would not have an $s4$ interval. The presence of species X in the 4.5–5.0 Ma interval would be one of 259 species records in the f survivorship category (table 1), and so on. For each of these intervals (f, $s1$, $s2$, $s3$), species X would have a certain number of site records (sr). Therefore there is the f sr, $s1$ sr, $s2$ sr and $s3$ sr for species X. For the entire dataset, then, we can investigate the average sr for all f (regardless of when f occurs in absolute time), namely the average sr for the f survivorship category, as compared with the average sr for the $s1$, $s2$, etc., survivorship categories.

As our survivorship intervals are based on a division of the past 10 Myr into intervals of 0.5 Myr each, it is important to note that the chronology for these African fossil site records is relatively good. It is based on numerous independent studies (see literature sources in Vrba (2000)). More than 70% of these fossil site units have been dated by radiometric and/or palaeomagnetic methods. Typical error limits can be illustrated by the contributions in Vrba *et al.* (1995). (For example, fig. 23.2 in Brown (1995) lists radiometric dates for tephra and basalts associated with some of the Kenyan and Ethiopian fossil site units included in our analysis: the 21 dates range from 4.35 Ma to 0.74 Ma, with errors ranging from ± 0.01 Ma to ± 0.08 Ma and mean error ± 0.04 Ma.) Because of this comparatively rich framework of physical dates for the African fossil sites, the biochronological estimates for the remaining site units (which are arrived at by faunal comparisons with the physically dated assemblages) are generally good as well. Nevertheless, there are likely to be fossil site records in our data that are not placed into their correct 0.5 Myr time intervals, but instead into a preceding or succeeding interval. But we suggest that the proportion of such errors is small, and that they are unlikely to result in erroneous rejection of the null hypothesis by our comparative tests among survivorship categories f, $s1...sn$ (regardless of when each occurs in absolute time).

The initial data were reduced by removing all records of last appearances (sn records), because we expect the geographical spread of species to decline towards extinction and we are interested in the transition from speciation to later equilibrium geographical distribution. This includes elimination of species with first and last appearance records in the same 0.5 Myr time interval. We also excluded records for higher taxa unidentified to species level, and those for species with first appearances previous to 10 Ma.

In our calculations of average sr for survivorship categories we included only records of $sr \geq 1$. For instance, let us suppose that our hypothetical species X has no fossil record in its $s2$ interval 3.5–4.0 Ma ($s2$ $sr = 0$), although one can infer from its presence in the earlier and later record that it was surviving in Africa throughout the 3.5–4.0 Myr interval. In this case we would not count species X in the species total used to arrive at the sr average for the $s2$ survivorship category. We did this because our main aim was to achieve the best possible comparison of the f with the later $s1...sn$ categories, and only positive records can be used for this comparison; although we have totals of instances of $sr = 0$ for surviving species, we do not have the comparable information for the f category which, by definition, has $sr \geq 1$.

The null hypothesis we tested is that the geographical spread of species remains approximately constant from their origin onwards. In that case, successive survivorship categories f, $s1$,

Table 1. Representation in the African fossil record, 10 Ma to Recent, of larger mammal species in survivorship categories. (The columns represent numbers of site records ($sr = x$) per species per 0.5 Myr-long intervals ($10 > t \geqslant 9.5$ Myr, ...$0.5 > t \geqslant 0$ Myr). The entries in the sr table are total numbers of instances of species with $sr = x$, across all species recorded in each survivorship category: f, first appearance in a 0.5 Myr-long interval; $s1$, survivor recorded in the interval immediately succeeding the f-interval, and so on; average sr, average site record number; n, number of 'species instances' in site record-survivorship categories (note that the same species commonly occurs in more than one survivorship category). Only positive records were included in the estimates in the last two columns.)

survivorship category	site record distribution								
	$sr = 1$	$sr = 2$	$sr = 3$	$sr = 4$	$sr = 5$	$sr \geqslant 6$	Σsr	n	average sr
f	138	67	22	14	12	6	495	259	1.91
$s1$	48	29	12	15	9	24	435	137	3.18
$s2$	23	17	9	12	9	22	364	92	3.96
$s3$	17	15	13	12	6	14	277	77	3.60
$s4$	9	15	7	7	3	7	156	48	3.25
$s5$	13	10	6	2	3	8	146	42	3.48
$s6$	9	5	3	1	3	6	99	27	3.67
$s7$	2	3	1	3	0	2	41	11	3.73
$s \geqslant 8$	6	6	2	2	1	1	45	18	2.50
Σ	265	167	75	68	46	90	2058	711	

$s2$, ...$s7$ (table 1) should not differ significantly in sr. The first part of our analysis tests this by using the χ^2-test.

The second part of our analysis is an attempt to grapple with the problem of biased recovery in the fossil record, and to obtain more realistic estimates of the totals of species instances and average sr, in the 'true' survivorship categories in the living biota. For instance, the total of species seen in a given interval, t, of the fossil record in the f category (first appearances of species) includes an unknown proportion that in the real world originated in previous intervals but were not detected in the fossil record before t. The same is true for the instances of species that appear in the fossil record to be in their second, third, etc., intervals of survivorship, but which in the living world had actually been in existence for a longer time. In other words, the observed totals of fossilized species in each survivorship category are inflated by excess erroneous data as a result of the vagaries of fossil preservation, a bias to which we will refer to as 'contamination'. We used an updated version of a model first described in Vrba (2000) to obtain rough estimates of the proportions of first appearances of species in the fossil record, in particular intervals that are 'true' first appearances, and thus allow us to correct for such 'contamination'. The model is presented as equation (A 1) and explained in Appendix A. A similar equation was used to estimate the relative proportions of true and contaminating components of the observed second and later appearances of species, namely in the interval immediately after their first records, and in later intervals.

In summary, first, we statistically tested the null hypothesis that the average number of sr do not differ significantly between survivorship intervals as directly found in the fossil record. Despite the above-described contamination of these raw data, this is worth doing. For instance, we know that the category of observed first appearances in the fossil record includes at least some species that originated in those same intervals in the real world; and the sr of these true originations, provided species are really significantly more rare near origin than later, might depress the total sr for this category sufficiently to be statistically detectable despite contamination. Of course, if no significant difference in sr is found between the f and $s1$ categories as recovered directly from the fossil record, their counterparts in

the living biota may still have differed in geographical extent and with the difference obscured by contamination. The contamination problem, in other words, biases the analysis in favour of the null hypothesis. We attempted to correct for the contamination in the second part of the analysis, despite the fact that it provides rough estimates unaccompanied by confidence intervals, so as to provide a more complete appreciation of actual events in the living biota and to aid in further discussion of this problem.

3. RESULTS

Our final dataset yielded a total of 2058 site records for 259 species distributed across the various survivorship categories (table 1). The average number of site records (sr) increases dramatically from the time interval of first appearance (f, mean $sr = 1.9$) to the following time interval ($s1$, mean $sr = 3.2$), as shown in figure 1. The χ^2 results (table 2) demonstrate that this increase in mean sr from f to $s1$ is statistically significant ($p < 0.001$), as is the increase in mean sr from $s1$ to $s2$–$s7$ (using only $sr > 1$ and $sr = 1$, $p < 0.05$). No significant changes are seen in mean sr across $s2$ to $s7$. Thus, from this first analysis, it appears that there is typically a sharp increase in the geographical spread of a species from the time of its origin (f) over the next *ca.* 1.0 Myr ($s1$, $s2$). This is the case despite the 'contamination' problem, described above, which biases the analysis against detecting such an increase.

Because there are no significant, or even notable, differences in mean sr between $s2$–$s7$ (table 1), we conclude that, although the $s2$–$s7$ survivorship categories are each contaminated by species records, which in the living world actually belonged to later survivorship intervals, this does not affect the estimates of sr averages in those intervals. Only the contamination of the f and $s1$ intervals, which are our main focus in any case, are expected to result in inflation of their sr averages. Thus we accepted the observed mean sr of 3.65 per 0.5 Myr interval for $s2$–$s7$ species, and used it together with the model in Appendix
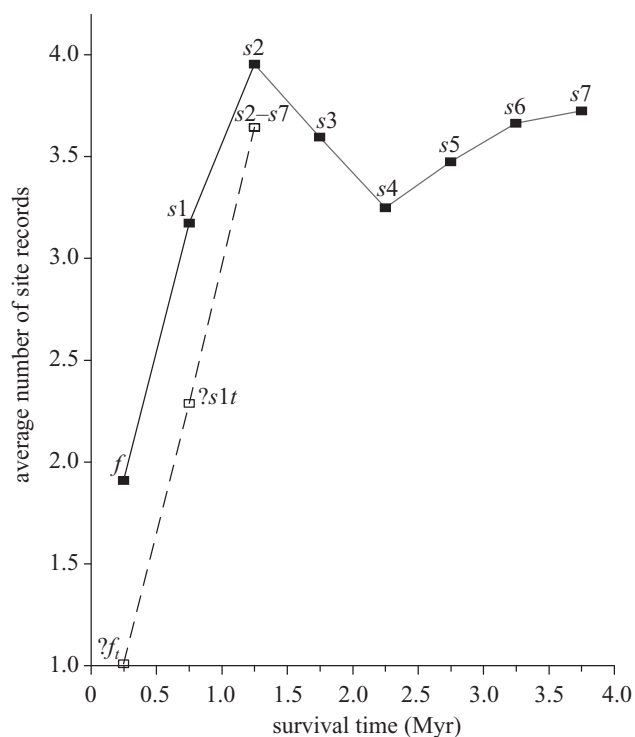
Figure 1. Solid line and symbols, change in average number of fossil site records by which African large mammal species are known in successive survivorship categories: *f*, *s*1, ...*s*7, over the past 10 Myr; *f*, first appearance in a 0.5 Myr-long interval; *s*1, 'first survivor', recorded in the interval immediately succeeding the *f*-interval of the same species, and so on. Dashed line and open symbols, estimates of the average site record numbers of: (i) 'true *f*' (?*f*ₜ, the species among the observed *f* species records which in the real world speciated during the same interval as their *f* records); (ii) 'true *s*1' (?*s*1*t*, the species among the observed *s*1 species which in the real world were also first survivors because they speciated during the immediately preceding interval), in relation to (iii) the mean number of site records for survivorship categories *s*2–*s*7. For derivation of the estimates, see Appendix A and table 3.

A to achieve more refined estimates of average *sr* for the true *f* and true *s*1 sets of species (table 3). These estimates confirm and magnify the results of the uncorrected analysis. The corrected average *sr* for *f* is estimated to be close to 1, the minimum, and roughly one-quarter that of the pooled *s*2–*s*7 intervals (figure 1, table 3). Correcting the estimate of mean *sr* for *s*1 decreases it to less than two-thirds of the value for later survivorship categories (figure 1, table 3).

4. DISCUSSION

We want to know whether species in life are more rare soon after origin than during their later histories. There are two aspects of species rarity that concern us. One is geographical rarity, as reflected by the size of the geographical area in which a species is found, and/or its population density in that area. The other is temporal rarity, or the length of time for which a species existed in a sampled time interval.

Unfortunately the measure we are using, change in number of distinct fossil site records from which species

are known in successive time intervals, is only a distant proxy for rarity. It is nevertheless expected to represent approximately the relative sizes of geographical distributions of species in different survivorship categories. By definition, *sr* = 1 signifies an observed distribution in only a single locality, a geographically very restricted area. Although we have not conducted a formal geographical analysis, the geographical span of a species typically increases rapidly as site records are added. An *sr* > 2 for a species in an interval usually indicates its presence in more than one major African area, and many instances of higher *sr* reflect those species' presence in several major regions, such as northern, eastern and southern Africa. Because the present survivorship samples are large, one might expect any real average differences in geographical spread to be evident in the *sr* comparisons.

There is also the question of how taphonomic biases might have affected the samples and our results. The large taxonomic, geographical and temporal scope of the present analysis holds out hope that the differences in fossil preservation potential between taxa, areas and time intervals will be distributed among survivorship categories in a manner that does not obliterate any large differences in rarity among them. More importantly, we are comparing the manner in which the abundance of a set of species changes over time, rather than considering the abundance of one species relative to another. As such, the taphonomic bias inherent to particular species (e.g. small-bodied taxa versus larger taxa) should not bias the comparison of their abundance at appearance relative to their later abundance (so long as that taphonomic bias remains relatively constant over time, as is expected). The same applies to fossil recognition and recovery bias.

Since time is our primary comparative axis, temporal issues have a greater potential to bias our analysis. For instance, a species that originated in the interval 5.5–5.0 Ma is typically present in that interval for less time (unless it originated right at 5.5 Ma) than species that originated in a previous interval and then survived through the 5.5–5.0 Myr period. On average, species are present in their appearance interval (*f*) for only half as much time as survivor species in that interval, and thus have only half the chance of being detected in the fossil record. We will refer to this as the 'half-life' effect. We interpret our results as a whole (tables 1–3; figure 1) to mean that the half-life effect of temporal rarity, in the interval of true species origination, cannot, on its own, account for the observed differences in average *sr* between the first appearance of species (*f*) and their immediately following two appearances (*s*1 and *s*2) in the fossil record (figure 1). Relative geographical rarity in the origination interval also played a part. First, in the analysis of the directly observed data (tables 1 and 2), the average *sr* for *f*, even in its erroneously inflated state, is approximately one-half of that for the established later categories. This alone suggests that removal of the contaminating component will depress the value well below the half-life expectation. Also, the *sr* of *s*1 is significantly lower than is that of the pooled *s*2–*s*7 categories (table 2), although no half-life effect is present in this comparison. Second, in the modified results, after correction for contamination (table 3), average *sr* for *f* is close to 1, about one-quarter of that for the later, established *s*2–*s*7 categories. This is a greater difference than

Table 2. Results of χ^2-tests of the null hypothesis H_0 that successive survivorship categories *f*, *s*1, *s*2, ...*s*7 (see table 1) do not differ significantly in numbers of site records in the fossil record.

survivorship categories level compared	χ^2 critical value	χ^2 significance level	χ^2 d.f.	χ^2 obtained in tests
f, *s*1	20.51	0.001	5	38.90***
*s*1, *s*2–*s*7	3.84	0.05	1	5.10*
*s*1, *s*2	11.07	0.05	5	4.10 n.s.
*s*2, *s*3	11.07	0.05	5	2.82 n.s.
*s*3, *s*4	9.45	0.05	4	2.31 n.s.
*s*4, *s*5	7.82	0.05	3	1.95 n.s.
*s*5, *s*6	7.82	0.05	3	0.32 n.s.
*s*6, *s*7	3.84	0.05	1	0.13 n.s.

* Significant at 0.05; ** significant at 0.01; *** significant at 0.001; n.s., not significant (considered but not found).

Table 3. Estimated total numbers of species, *n*, and average numbers of site records, average *sr*, in survivorship categories. (Estimates include numbers of species in the *f* survivorship category (first species' appearances per 0.5 Myr-long interval), which in the real world during those same intervals represent: (i) first species' appearances (or 'true *f*': f_t); (ii) *s*1 survivors (or 'false *f*' that in reality speciated in the previous interval: $f_{(s1)}$); and (iii) $s \geq 2$ survivors (or 'false *f*' that really belong to the $s \geq 2$ category: $f_{(2 \leq s \leq 7)}$. The estimates are from application of equation (A 1) (in Appendix A) to the African fossil data. A similar equation was used to estimate the proportions of *s*1 which in the real world during that interval represent (iv) *s*1 survivors (or 'true *s*1' = $s1_t$), and (v) *s*2...*s*7 survivors (or 'false *s*1' that really belong to $2 \leq s \leq 7$: $s1_{(2 \leq s \leq 7)}$). The sequence in the table reflects the sequence in which the estimation proceeds. Results marked ^ are estimated directly from the model applications; others are either observations in the fossil record (see table 1) or calculated from these together with the ^estimates.)

survivorship category	*n*	Σsr	average *sr*
$2 \leq s \leq 7$	315	1128	3.65
*s*1	137	435	3.18
f	259	495	1.91
(v) $s1_{(2 \leq s \leq 7)}$	89^	325	3.65
(iv) $s1_t$	48^	110	2.29
(iii) $f_{(2 \leq s \leq 7)}$	52^	190	3.65
(ii) $f_{(s1)}$	83^	190	2.29
(i) f_t	124^	115	0.93 ≈ 1

expected if the half-life effect were alone responsible. Also in this result (table 3), average *sr* for *s*1 has decreased to less than two-thirds of that for *s*2–*s*7.

As noted earlier, the taxa included in this analysis have generally been accepted as separate species by workers in the relevant systematic subfields. Although a proportion may have been wrongly classified as separate species and in reality represent intraspecific variants (either genetically based, such as subspecies or ecophenotypes), we suggest that most really are full species. First, *ca.* 30% of the taxa in our sample are extant, and are today recognized as specifically distinct from their closest relatives based on the broader range of neontological data, without hybrid intermediates evident today or in the fossil record. Second, the average persistence of larger African mammals over the past 20 Myr is 2.33 Myr (Vrba (2000), this includes all taxa that survived to at least one additional interval after that of first appearance), which is equivalent to survivorship interval *s*4. The species used are considered distinct; and we do not generally know of reported intermediate fossil forms between pairs of species, as would be expected if a single real species had been mistakenly identified as two species. There is also a strong pattern across the 259 included taxa (table 1) of restricted geographical spread in the interval of first appearance,

rapid increase over approximately the next 0.5 Myr by the *s*1 interval, a lesser additional increase by the *s*2 interval, and no change subsequently. This strong pattern is more consistent with the notion that most of the included taxa are indeed separate species than with one that they are subspecies. In general, subspecies would be expected to show more labile and reversible changes in the fossil record, both in phenotypic variation and in the geographical distribution of variants, and one might expect intermediate forms between the taxa in question (see Vrba (1998) for fossil hartebeests, Bovidae). Infraspecific variants have been found to extend their geographical ranges very fast both after arising in historical times (Mayr 1963) and in the fossil record on the glacial–interglacial timescale (see Sutcliffe (1985) for mammalian examples).

We conclude that the average large African mammal species has indeed started its life in a relatively small population, and thereafter increased in geographical spread to reach its long-term equilibrium abundance by about 1 Myr after origin. The question arises as to why the average time of increase in abundance and geographical distribution took so long. Kurten (1957) studied Pliocene–Recent dispersal episodes, at times when barriers to dispersal were removed by environmental changes, in a large sample of Eurasian mammalian species. He concluded

that 'an unchecked spread of some 1000 km in a century would seem a moderate estimate for most larger mammals' (p. 217). Kurten (1957) used a different sample from ours (Eurasian versus African) and different methods. This makes it difficult to interpret the significance of the apparent difference in rate estimates. If we assume that his conclusions on dispersal rate broadly apply to African mammals, then a comparison of the two sets of observations suggests that they may, in fact, be compatible. Kurten's (1957) estimate may apply to established species, whereas the rate of geographical range extension for young species, especially in the interval after first appearance, is much lower. Substantial increase of the geographical distribution of a species minimally requires: (i) a large net increase in numbers of organisms; and (ii) large new areas of suitable habitat that are accessible to them. Given the dynamism of environments in Africa since the Late Miocene, with repeated large-scale shrinking and spreading, vicariance and remerging, of alternative biomes (contributions in Vrba *et al.* (1995)), option (ii) looks less likely to be the factor that limits dispersal in early species than does option (i). We suggest that young species commonly have more stringent limits on their rates of increase in abundance than have established species. New species are subject to the population genetic effects of small numbers, and have new habitat specificities and ecologies. They remain in a destabilized state, still caught up in the aftermath of recent 'genetic revolution' (Mayr 1963) and also ecological revolution. These circumstances impose constraints, in recently diverged species, on birth rates greatly exceeding death rates, as needed for the examples of dramatic range expansion in the presumably older species documented in Kurten (1957). If it is true that the Eurasian species in Kurten's sample really were averagely older at times of their major dispersals than were the African species in our *f–s*1 categories, that raises a question about the second difference between the two sets of findings; why were the Eurasian established species capable of such rapid and large-scale range extension, whereas the African species in comparable survivorship categories (*s*2–*s*7) show approximated equilibrium at larger geographical ranges? We believe that the difference may be accounted for by the differing time-scales examined. Although our conclusions on rate of geographical spread are net estimates for 0.5 Myr-long intervals, in which the more rapid excursions of the astronomical cycles within each examined interval were averaged, Kurten's (1957) were based on individual episodes occurring at the shorter time-scale of those climatic cycles.

5. CONCLUSION

We conclude that the average large African mammal species has originated in a relatively small population, and thereafter increased in geographical spread to reach its long-term equilibrium abundance by *ca.* 1.0 Myr later. This result generally supports hypotheses of speciation that accord a major initiating role to reduced population size, such as through vicariance due to environmental change, and call into question those hypotheses that diminish the importance of small population size.

Our results cast doubt on the generality of speciation

models under which vicariance to small population size is mostly irrelevant, or at least not needed for speciation to occur. Many traditional neo-Darwinian discussions of speciation belong in this category. They share the view that speciation is initiated by many diverse factors, the most important being organismal interactions, such as competition, disease, predation, mutualism, etc. These factors can act alone to drive speciation (e.g. Van Valen's (1973) Red Queen's hypothesis and Rosenzweig's (1978) model of competitive speciation) or combine with environmental change to cause speciation. Under this view, which has been termed the competition paradigm (Vrba 1992), the causes of species origination differ between taxonomic groups and from one local area to the next. The implications of this view include that each evolutionary shift is largely an independent piece of history, and that we should therefore not expect evidence of common rules that act across species or coherent patterns among speciation events in relation to physical change (see examples of such arguments in Hoffman (1989), McKee (1993) and Foley (1994)). Another implication of such views is that the theory of macroevolution is synonymous with traditional Darwinian theory, namely that macro-patterns emerge as epiphenomena of natural selection and other processes within populations.

The implications of our results for African mammals contrast with those of such competition-based concepts in several respects. First, small population size caused by vicariance is crucial to speciation. Second, physical change is therefore indicated as a predominant precursor of speciation. In concluding this, we do not deny the importance of competition and other proximal biotic interactions. For example, a speciation event in grassland herbivores may be proximally caused by the disappearance of grassland and the appearance of shrubland bringing with it new parasites, diseases and competitors. But the question would arise, what physical changes may have caused the biotic changes in the first place? Finally, the predominant initiation of speciation by physical change appears to imply common rules for diverse lineages, by which some coherence is conferred on their patterns of speciation. This accounts for the relatively consistent pattern we observe among mammalian speciation events, a consistency not predicted by competition-based models.

By contrast, our results generally support hypotheses of speciation that accord a major initiating role to vicariance owing to physical change. The oldest such hypothesis is the allopatric speciation model. Our findings agree with theoretical formulations of allopatric speciation which call for separation, either by vicariance or by dispersal, of a small founder population as a necessary circumstance of most speciation events. Examples of such formulations range from the early argument of Gulick (1872), through those dating to the neo-Darwinian synthesis by Wright (1932, 1967) and Mayr (1942, 1963), to the strong and detailed arguments under the 'specific-mate recognition' model of Paterson (1978, 1981, 1985) and the 'flush–crash–founder' model of Carson (1957, 1971, 1982). Most of the cited papers, especially those of Wright, Paterson and Carson, as well as Lande (1979), Templeton (1981, 1989) and others, contain arguments from a genetic perspective that small populations are the most likely context for the initiation of speciation. Reasons include

that size bottlenecks associated with decreasing populations result in random genetic sampling, increased genetic diversity among the vicariant populations, and altered selective environments. Thus, our study is a case in which direct observations on the fossil record broadly confirm what population geneticists have concluded from neontological genetic observations and models.

We have not directly tested the model of punctuated equilibria (Eldredge & Gould 1972; Gould & Eldredge 1993), but the patterns that we found are consistent with that model's proposals of a high rate of phenotypic change in small isolated populations at species' origin followed by net stability in established species. As outlined above, the average duration of species over the past 20 Myr has been 2.33 Myr, with the longest duration being 14.8 Myr. These durations imply net long-term phenotypic equilibrium. Furthermore, there is also equilibrium in terms of the average geographical representation in larger African mammal species after about the first million years of their existence.

Another hypothesis based on physical change which is consistent with our results is the turnover pulse hypothesis (Vrba 1985, 1995*b*). Its essential point is that turnover (speciation and extinction) requires vicariance initiated by changes in the physical environment. The larger prediction, stripped to its bare essentials, is that most lineage turnover has occurred in pulses, varying from tiny to massive in scale, across disparate groups of organisms, and in predictable temporal association with changes in the physical environment. Subsidiary predictions are that species start in populations of reduced size and that the subsequent equilibrium phase of species' durations is marked by relatively larger geographical distributions.

The dataset of the African larger mammal fossil record assembled here can be used to investigate a variety of issues, and additional analyses are certainly needed. Even so, it is clear that the fossil record can be involved productively in testing various macroevolutionary hypotheses. In the present case, analysis of this fossil dataset indicates that species do, in fact, start small, a finding of some consequence for models of speciation and macroevolution.

APPENDIX A

Equation (A 1) estimates that $E(\Phi_i)$ is the expected number of species in the f survivorship category (of first global appearances, see text) in the fossil record in t_i, the ith time interval. n_i is the total number of species in the fossil record in t_i; N_i is the total number of species in the living biota in t_i; $\alpha_i = F_i/N_i$, where F_i is the number of species' first appearances in the living biota in t_i; β_j is the probability that species survive the jth survivorship interval, where β_0 refers to the interval of species origin, β_1 to the immediately subsequent interval after that of origin, and so on.

$$E(\Phi_i) = \frac{\alpha_i}{2}n_i + \alpha_{i-1}\beta_0\left(1 - \frac{n_{i-1}}{2N_{i-1}}\right)n_i$$
$$+ \alpha_{i-2}\beta_0\beta_1\left(1 - \frac{n_{i-2}}{2N_{i-2}}\right)\left(1 - \frac{n_{i-1}}{N_{i-1}}\right)n_i$$
$$+ \alpha_{i-3}\beta_0\beta_1\beta_2\left(1 - \frac{n_{i-3}}{2N_{i-3}}\right)\left(1 - \frac{n_{i-2}}{N_{i-2}}\right)\left(1 - \frac{n_{i-1}}{N_{i-1}}\right)n_i$$
$$+ \alpha_{i-r}\prod_{k=0}^{r-1}(\beta_k)\left(1 - \frac{n_{i-r}}{2N_{i-r}}\right)\prod_{k=0}^{r-2}\left(1 - \frac{n_{i-1-k}}{N_{i-1-k}}\right)n_i.$$

$$(A\ 1)$$

We applied equation (A 1) to estimate the relative proportions of 'true' and 'false' components of the f survivorship category (the f_t and $f_{(s1)}$ categories in table 3, the numbers of species in which, Φ_t and $\Phi_{(s1)}$, are respectively represented by the first term and the summed subsequent terms in equation (A 1)). The estimates were based on the summed totals, across the entire 10 Myr dataset, of the $\Phi_t, \Phi_{(s1)}, ...\Phi_{(sr)}$ components obtained by equation (A 1). These results were then used to estimate the average numbers of fossil site records of, and to compare the averages among, the survivorship categories (table 3).

The n_i in equation (A 1) are the observed totals of fossil species in 0.5 Myr-long intervals (Vrba 2000), reduced by omitting certain categories of species instances as outlined in the text. For the present study we assumed that N_i can be treated as constant N. The arguments as to why this is reasonable in the case of this all-African, 10 Myr record are given in Vrba (2000). We also treated $\alpha = F/N$ as a constant in this analysis because, although the α_i were found to vary through time in African mammal evolution (Vrba 2000), the use of a constant α is unlikely to distort the summed values across the entire dataset, which we estimate in the present study. We followed the method in Vrba (2000) for inferring a mean value of N per 0.5 Myr interval, using the fossil record of the past 0.5 Myr together with the living species' total, and reduced this value proportionally to reflect the present omission of certain categories of species instances. The β_j were estimated from a survivorship analysis of the observed fossil record.

Equation (A 1) was applied to each of the 20 intervals of equal 0.5 Myr duration in the African larger mammal record of the past 10 Myr. The results were summed across all intervals to give estimates of $\Sigma E(\Phi_i) = \Sigma E(\Phi_t) + \Sigma E(\Phi_{(s1)}) + ... + \Sigma E(\Phi_{(sr)})$. The sums of all expected and observed first appearances over the 20 intervals of the past 10 Myr are, respectively, $\Sigma E(\Phi_i)$ and $\Sigma O(\Phi_i)$. The 20 applications of equation (A 1), to the 20 time intervals, were solved simultaneously with the constraint that $E(\Phi_i) = O(\Phi_i)$, thus providing solved estimates for the constant α and for the expected individual and summed values in survivorship categories.

A similar equation was used to estimate the relative proportions of true and false components of the observed second appearances (the $s1$ category of species' presence, namely in the interval immediately after their first records).

REFERENCES

Brown, F. H. 1995 The potential of the Turkana Basin for paleoclimatic reconstruction in East Africa. In *Paleoclimate and evolution with emphasis on human origins* (ed. E. S. Vrba,

G. H. Denton, T. C. Partridge & L. H. Burckle), pp. 319–330. New Haven, CT: Yale University Press.

Butlin, R. 1989 Reinforcement of premating isolation. In *Speciation and its consequences* (ed. D. Otte & J. A. Endler), pp. 158–179. Sunderland, MA: Sinauer.

Carson, H. L. 1957 The species as a field for gene recombination. In *The species problem* (ed. E. Mayr), pp. 23–38. Washington, DC: American Association for the Advancement of Science Publication 50.

Carson, H. L. 1971 Speciation and the founder principle. *Stadler Genet. Symp.* **3**, 51–70.

Carson, H. L. 1982 Speciation as a major reorganization of polygenic balances. In *Mechanisms of speciation* (ed. C. Barigozzi), pp. 411–433. New York: Alan R. Liss.

Darwin, C. 1859 *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London: John Murray.

Eldredge, N. & Gould, S. J. 1972 Punctuated equilibria: an alternative to phyletic gradualism. In *Models in paleobiology* (ed. T. J. M. Schopf), pp. 82–115. San Francisco, CA: Freeman.

Foley, R. A. 1994 Speciation, extinction and climatic change in hominid evolution. *J. Hum. Evol.* **26**, 275–289.

Futuyma, D. J. & Mayer, G. C. 1980 Non-allopatric speciation in animals. *Syst. Zool.* **29**, 254–271.

Gould, S. J. & Eldredge, N. 1993 Punctuated equilibrium comes of age. *Nature* **366**, 223–227.

Gulick, J. T. 1872 On the variation of species as related to their geographical distribution, illustrated by the Achatinellidae. *Nature* **6**, 222–224.

Gulick, J. T. 1905 *Evolution, racial and habitudinal.* Publication no. 25. Washington, DC: Carnegie Institution of Washington.

Hoffman, A. 1989 *Arguments on evolution.* Oxford University Press.

Kurten, B. 1957 Mammal migrations. Cenozoic stratigraphy, and the age of Peking man and the australopithecines. *J. Paleontol.* **31**, 215–227.

Lande, R. 1979 Effective deme sizes during long term evolution estimated from rates of chromosomal rearrangement. *Evolution* **33**, 234–251.

McKee, J. K. 1993 Formation and geomorphology of caves in calcareous tufas and implications for the study of the Taung fossil deposits. *Trans. R. Soc. S. Afr.* **48**, 307–322.

Mayr, E. 1942 *Systematics and the origin of species.* New York: Columbia University Press.

Mayr, E. 1963 *Animal species and evolution.* Cambridge, MA: Harvard University Press.

Mayr, E. 1982 Processes of speciation in animals. In *Mechanisms of speciation* (ed. C. Barigozzi), pp. 1–19. New York: Alan R. Liss.

Paterson, H. E. H. 1978 More evidence against speciation by reinforcement. *S. Afr. J. Sci.* **74**, 369–371.

Paterson, H. E. H. 1981 The continuing search for the unknown and the unknowable: a critique of contemporary ideas on speciation. *S. Afr. J. Sci.* **77**, 113–119.

Paterson, H. E. H. 1982 Perspective on speciation by reinforcement. *S. Afr. J. Sci.* **78**, 53–57.

Paterson, H. E. H. 1985 The recognition concept of species. In *Species and speciation* (ed. E. S. Vrba), pp. 21–23. Pretoria, South Africa: Transvaal Museum Monographs No. 4.

Rosenzweig, M. L. 1978 Competitive speciation. *Biol. J. Linn. Soc.* **10**, 275–289.

Sutcliffe, A. J. 1985 *On the track of ice age mammals.* Cambridge, MA: Harvard University Press.

Templeton, A. R. 1981 Mechanisms of speciation: a population genetic approach. *A. Rev. Ecol. Syst.* **12**, 23–48.

Templeton, A. R. 1989 The meaning of species and speciation: a genetic perspective. In *Speciation and its consequences*

(ed. D. Otte & J. A. Endler), pp. 3–27. Sunderland, MA: Sinauer.

Van Valen, L. 1973 A new evolutionary law. *Evol. Theor.* **1**, 1–30.

Vrba, E. S. 1985 Environment and evolution: alternative causes of the temporal distribution of evolutionary events. *S. Afr. J. Sci.* **81**, 229–236.

Vrba, E. S. 1992 Mammals as a key to evolutionary theory. *J. Mammal.* **73**, 1–28.

Vrba, E. S. 1995a Species as habitat specific, complex systems. In *The recognition of species: speciation and the recognition concept* (ed. J. C. Masters, D. M. Lambert & H. G. Spencer), pp. 3–44. Baltimore, MD: Johns Hopkins University Press.

Vrba, E. S. 1995b On the connections between paleoclimate and evolution. In *Paleoclimate and evolution with emphasis on human origins* (ed. E. S. Vrba, G. H. Denton, T. C. Partridge & L. H. Burckle), pp. 24–45. New Haven, CT: Yale University Press.

Vrba, E. S. 1998 New fossils of Alcelaphini and Caprinae (Bovidae, Mammalia) from Awash, Ethiopia, and phylogenetic analysis of Alcelaphini. *Palaeontologia Africana* **34**, 127–198.

Vrba, E. S. 2000 Major features of neogene mammalian evolution in Africa. In *Cenozoic geology of southern Africa* (ed. T. C. Partridge & R. Maud), pp. 277–304. Oxford University Press.

Vrba, E. S., Denton, G. H., Partridge, T. C. & Burckle, L. H. 1995 *Paleoclimate and evolution with emphasis on human origins.* New Haven, CT: Yale University Press.

Wright, S. 1932 The roles of mutation, inbreeding, crossbreeding, and a selection in evolution. *Proc. VIth Int. Congress Genet.* **1** (ed. D. F. Jones), pp. 356–366. Brooklyn, New York: Brooklyn Botanic Garden.

Wright, S. 1967 Comments on the preliminary working papers of Eden and Waddington. In *Mathematical challenges to the neo-Darwinian theory of evolution* (ed. P. S. Moorehead & M. M. Kaplan), pp. 117–120. Philadelphia, PA: The Wistar Institute Symposium No. 5.

Discussion

B. Odgaard (*Institute of Earth Sciences, University of Aarhus, Aarhus, Denmark*). I may not have fully understood how the possible biases in the fossil record were taken into consideration, so this is maybe a question of clarification. One obvious bias in the fossil record is that common species tend to produce more fossil finds than rare species. Obviously, what happens at an event or period of climate change is that species that previously were common will now become rarer because they are no longer well-adapted to the climate of the area. Similarly, rare species may now become common. Such changes in the fossil record may be interpreted as events of extinctions and speciations but are really changes in frequency after climate change. How was this possible bias taken into consideration in the analysis of speciation rates based on the fossil record?

E. Vrba. We now have good evidence, at least for the Pliocene–Pleistocene, that most mammalian species responded to each excursion of the astronomical (or Milankovitch) climatic cycles by changes in geographical distributions, including alternate shrinkage and expansion and latitudinal displacements of geographical distributions. Thus I agree on this point. I also agree that species that are rare during an interval are often not seen in the fossil record for that interval, so that the earliest fossil appearance of a given species may postdate its origination

in the living world (and its last appearance pre-date its actual extinction). It is precisely to correct for this bias that I developed and applied the statistical model that I mentioned in my talk. The model uses the observed changes in fossil preservation through time to estimate the expected numbers of fossil first and last appearances in each interval under a null hypothesis of random variation in speciation and extinction rate over time (i.e., random variation about a constant probability of turnover). That is, this approach acknowledges that a proportion of the first appearances of species in an interval in reality are for species that originated earlier but were not preserved in the fossil record for those earlier times; and the model is designed to estimate that proportion under the null hypothesis. Then a statistical test is done of whether the values expected under the null hypothesis differ significantly from those observed in the fossil record. Also, bear in mind the large time-scale of the analyses. For example, whereas forest/woodland species X may disappear from the fossil record during a given glacial phase Y (when Africa was averagely more arid and open), provided that species survived it is expected to increase again in abundance and geographical range during subsequent interglacial climatic phases, with increased probability of fossil presence. Thus the palaeontologist would not conclude from the absence of species X during glacial Y that it became extinct then, because the larger pattern through time attests to X's survival through phase Y.

E. Verheyen (*Royal Belgian Institute of Natural Sciences, Brussels, Belgium*). Your talk nicely demonstrates that speciation rates—as inferred from fossil data—indeed appear to be highest when the environment changes. I wonder if recent molecular phylogenies are useful in showing (and already have shown) that speciation has indeed occurred in bursts?

E. Vrba. Years ago I suggested that the best way to test the turnover pulse hypothesis might be by use of molecular data. (The turnover pulse hypothesis is that speciation and extinction require vicariance initiated by changes in the physical environment, with the result that most lineage turnover has occurred in pulses in predictable temporal association with changes in the physical environment.) Provided that a molecular clock is approximately working in a large group of species, and even without knowing anything about the fossil record, one should detect speciation 'pulses' in 'relative molecular time' if the hypothesis is correct. Well, over the years (and especially recently) I have been contacted by numerous molecular systematists who informed me that speciation pulses are evident in their phylogenies and, with time calibration, can also be related to independent evidence of physical change. Collaborators and I have analysed the molecular phylogenies of Bovidae and other artiodactyls in great detail, and found a lot of 'star-phylogenies': namely, unresolvable polytomies of related species. We interpret the times of such polytomous origins to be times of bursts of rapid speciation, brought about by rapidly repeated episodes of vicariance induced by physical change. So, to answer you: molecular phylogenies can indeed be 'useful in showing (and already have shown) that speciation has indeed occurred in bursts'. I do not know of any publications that have directly, intensively and competently analysed this for groups of organisms. But I believe it should be done, and that we will see more and more such papers in future.