

References and Notes

1. Y. Aikawa, T. Umebayashi, T. Nakano, S. M. Miyama, *Astrophys. J.* **519**, 705 (1999).
2. D. Bockelée-Morvan *et al.*, *Astron. Astrophys.* **353**, 1101 (2000).
3. O. Mousis *et al.*, *Icarus* **148**, 513 (2000).
4. Y. Aikawa, E. Herbst, *Astrophys. J.* **526**, 314 (1999).
5. G. A. Blacke *et al.*, *Nature* **398**, 213 (1999); R. Meier *et al.*, *Science* **279**, 1707 (1998); R. Meier *et al.*, *Science* **279**, 842 (1998).
6. W. M. Irvine, F. P. Schloerb, J. Crovisier, B. Fegley Jr., M. J. Mumma, in *Protostars and Planets IV*, V. Mannings, A. P. Boss, S. S. Russell, Eds. (Univ. of Arizona Press, Tucson, AZ, 2000), pp. 1159–1200.
7. M. J. Mumma, M. A. Weaver, H. P. Larson, *Astron. Astrophys.* **187**, 419 (1987); M. Mumma, P. Weissman, S. A. Stern, in *Protostars and Planets III*, E. H. Levy, J. I. Lunine, Eds. (Univ. of Arizona Press, Tucson, AZ, 1993), pp. 1177–1252; J. Crovisier, in *Formulation and Evolution of Solids in Space*, J. M. Greenberg, A. Li, Eds. (Kluwer Academic, Dordrecht, Netherlands, 1999), pp. 389–426.
8. The OPR of H₂CO in comet C/1995O1(Hale-Bopp) was reported to be 1.5 ± 0.3 [M. Womack *et al.*, *IAU Circ.* **7474** (1997)]. The spin temperature is ~ 10 K and is inconsistent with other temperature range indicated by the OPR of H₂O and by the D/H ratio of H₂O and HCN and the depletion of neon in comet Hale-Bopp. We think that the OPR of H₂CO should be investigated carefully because its estimation depends on the assumed population in unobserved states.
9. H. Kawakita, J. Watanabe, *Astrophys. J.* **495**, 946 (1998); S. Tegler *et al.*, *Astrophys. J.* **384**, 292 (1992); U. Fink, M. R. Combi, M. A. DiSanti, *Astrophys. J.* **383**, 356 (1991).
10. Most of NH₂ molecules in comets are thought to be photodissociation products of ammonia. Another possible parent of NH₂ is NH₂CHO, which was detected in C/1995O1(Hale-Bopp) for the first time, but its abundance was less than 1/50 that of ammonia (2). Thus, we assume that all NH₂ is produced from ammonia through photodissociation.
11. M. K. Bird *et al.*, *Earth Moon Planets* **78**, 21 (1997); M. K. Bird *et al.*, *Astron. Astrophys.* **325**, L5 (1997).
12. K. Noguchi *et al.*, *SPIE* **3355**, 354 (1999).
13. H. A. Weaver *et al.*, *Science* **292**, 1329 (2001).
14. We used the "IRAF" software package provided by National Optical Astronomy Observatories (NOAO) with the standard reduction procedures for an echelle spectrum. The sensitivity calibration for instrument was performed as follows. The cometary spectrum was first normalized by the continuum (which is the reflected sunlight by cometary dust grains). Then the normalized spectrum was divided by a normalized solar spectrum (25). Finally, it was multiplied by the product of the solar spectrum and the albedo of dust grain determined from the low-resolution spectra of comet LINEAR (26). Thus, we obtained the final cometary spectrum in which the sensitivity of the instrument was calibrated relatively.
15. H. Kawakita *et al.*, *Publ. Astron. Soc. Jpn.* **53**, L5 (2001); H. Kawakita, K. Ayani, T. Kawabata, *Publ. Astron. Soc. Jpn.* **52**, 925 (2000).
16. The included transitions in our model are as follows (27): (i) the ro-vibronic transitions, $\tilde{A}(0, v_2', 0) - \tilde{X}(0, v_2'', 0)$, $18 \geq v_2' \geq 1, v_2'' = 0$ and 1; (ii) the ro-vibrational transitions, $\tilde{X}(0, v_2', 0) - \tilde{X}(0, v_2'', 0)$, $13 \geq v_2' \geq 8, v_2'' = 0$ and 1; (iii) the ro-vibrational transitions, $\tilde{X}(1, 0, 0) - \tilde{X}(0, 0, 0)$, $\tilde{X}(0, 1, 0) - \tilde{X}(0, 0, 0)$, $\tilde{X}(0, 0, 1) - \tilde{X}(0, 0, 0)$; and (iv) the pure rotational transitions in $\tilde{X}(0, 0, 0)$. The vibrational transition moment for (iii) and the permanent dipole moment were determined by ab initio calculation. The rotational part of transition moment was calculated by ASYROT program (28).
17. The fluorescence equilibrium was assumed for calculating population distribution of NH₂ [M. F. A'Hearn, *Astrophys. J.* **219**, 768 (1978); S. C. Tegler, S. Wyckoff, *Astrophys. J.* **343**, 445 (1989)]. Although calculations based on our model can reproduce the observation, we cannot avoid the collisional excitation near nucleus region. The influence of collisional excitation seems to be negligible for the present case.

18. M. E. Brown, A. H. Bouchez, A. H. Spinrad, C. M. Johns-Krull, *Astron. J.* **112**, 1197 (1996); H. W. Zhang, G. Zhao, J. Y. Hu, *Astron. Astrophys.* **367**, 1049 (2001).
19. M. Quack, *Mol. Phys.* **34**, 477 (1977).
20. S. Takano, N. Nakai, K. Kawaguchi, T. Takano, *Publ. Astron. Soc. Jpn.* **52**, L67 (2000).
21. K. Hiraoka *et al.*, *Astrophys. J.* **443**, 363 (1995).
22. K. Willacy, H. H. Klahr, T. J. Millar, Th. Henning, *Astron. Astrophys.* **338**, 995 (1998).
23. S. A. Stern *et al.*, *Astrophys. J.* **544**, L169 (2000).
24. M. J. Mumma *et al.*, *Science* **292**, 1334 (2001); D. Bockelée-Morvan *et al.*, *Science* **292**, 1339 (2001).
25. R. L. Kurucz, I. Furenlied, J. Brault, L. Testerman, *National Solar Observatory Atlas No. 1* (Harvard University, Cambridge, MA, 1984); M. P. Thekaekara, *Appl. Opt.* **13**, 518 (1974).
26. M. Fujii, unpublished data. The spectroscopic observation of C/1999S4 (LINEAR) by the low-dispersion

spectrograph was carried out on 6 July 2000 (UT). We can find the color of dust grains being 15%/100 nm for the comet.

27. I. H. Bachir, T. R. Huet, J.-L. Destombes, M. Vervloet, *J. Mol. Spectrosc.* **193**, 326 (1999), and references therein; R. J. Buenker, M. Peric, S. D. Peyerimhoff, R. Marian, *Mol. Phys.* **43**, 987 (1981).
28. F. W. Birss, D. A. Ramsay, *Comp. Phys. Com.* **38**, 83 (1984).
29. We thank S. Saito for a fruitful discussion on the OPR in the NH₂ generated from the ammonia through the photodissociation and D. C. Boice for improving our manuscript. NSO/Kitt Peak Fourier transform spectrometer data used here were produced by NSF/NOAO. This report is based on data collected at Subaru Telescope, which is operated by the National Astronomical Observatory of Japan.

12 July 2001; accepted 24 September 2001

Preservation of Species Abundance in Marine Death Assemblages

Susan M. Kidwell

Fossil assemblages of skeletal material are thought to differ from their source live communities, particularly in relative abundance of species, owing to potential bias from postmortem transport and time-averaging of multiple generations. However, statistical meta-analysis of 85 marine molluscan data sets indicates that, although sensitive to sieve mesh-size and environment, time-averaged death assemblages retain a strong signal of species' original rank orders. Naturally accumulated death assemblages thus provide a reliable means of acquiring the abundance data that are key to a new generation of paleobiologic and macroecologic questions and to extending ecological time-series via sedimentary cores.

The fidelity of fossil assemblages to their source communities has haunted paleontologists for decades (1–4), and this issue has become especially acute with the growing realization that relative abundance data are required to address such dynamical problems as taxonomic survivorship, clade interactions, and ecological structure over evolutionary time (5–7). The potential reliability of naturally accumulated death assemblages is also important to ecology, where longer temporal perspectives on community composition are needed to discriminate natural and anthropogenic factors in ecosystem change (8, 9).

Paleoecological reliability has been estimated primarily via field tests comparing the live community with associated assemblages of dead remains in modern environments [reviewed by (10)]. Strong quantitative guidelines have been developed via such live-dead studies for some groups, most notably pollen and macroflora, permitting modern and prehistoric records of continental ecosystem change to be integrated (11). In contrast, the

many tests conducted in marine molluscan communities, which dominate post-Paleozoic to Recent sedimentary seafloors, have yielded substantial variation in live-dead fidelity, especially for species abundance (12), and it is unclear how this variation is partitioned among methodological artifacts and true preservational bias.

To acquire a robust estimate of taphonomic (postmortem) bias, I reanalyzed numerical abundance at the species level from 85 molluscan data sets (bivalves and gastropods) to standardize the metric of live-dead agreement (Spearman rank-order correlation coefficient). The 85 data sets span fine- to coarse-grained seafloors (no reefs or hardgrounds) from marsh to middle shelf settings and are drawn from 19 independent studies by other authors in low and middle latitudes (0° to 54°N, median 34°) (13). Although individual sample size, number of stations per habitat, and species richness vary among data sets, all reflect quantitative benthic sampling of the uppermost 10 or 20 cm of the sedimentary column. In the majority of studies, live data reflect a single census, thus providing a very conservative estimate of true live diversity. In 12 of the 85 habitats,

Department of Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, IL 60637 USA. E-mail: skidwell@midway.uchicago.edu

REPORTS

multiseason time-series of the live fauna are available (3, 14–17) (1.75 years maximum duration of sampling program). In the present analysis, these 12 data sets are each represented by only one single-census value (census with maximum number of live individuals) and one multicensus value (total pooled live data). Variation among data sets in significance (P -levels of individual Spearman rank-order tests) was analyzed conventionally via scatterplots. Individual correlation coefficients (Spearman r -values) were weighted by data set size (number of species), and average values were calculated overall and for various subgroups to isolate methodological and natural effects (18).

Individual Spearman tests yield a wide scatter in P -levels that can largely be explained by sample size and mesh-size effects (Fig. 1). Data sets composed of specimens exclusively from meshes >1 mm yield more significant live-dead agreements ($P < 0.05$) than those generated using fine (≤ 1 mm) mesh, especially if data sets based on extremely small numbers of specimens (<100 live individuals) are excluded; only 60% of fine-mesh data sets with >100 live individuals show a significant correlation between live and dead rank-orders (23/38 Spearman rank tests, individual $P < 0.05$; Fig. 1A). In contrast, live and dead species' rank orders are significantly correlated in 92% of coarse-mesh data sets exceeding a threshold size of 100 live individuals (22/24 Spearman rank tests, individual $P < 0.05$; Fig. 1B) (19). With increasing quantities of live data, live-dead correlations in coarse-mesh data sets become more significant, whether data are increased by larger single-census collections or by pooling time-series data (Fig. 1B). Agreement in fine-mesh data sets does not increase with increasing single-census data, but only in response to pooling of time-replicate samples (Fig. 1A). These increases are also less marked than in coarse-mesh data sets, indicating that discrepancies between live and dead species' rank-order in fine-mesh data sets reflect a more fundamental methodological or taphonomic problem than inadequate sample size.

The rank order of dead species may be opposite, random, or consistent with live species (i.e., Spearman r may range from -1 to $+1$), and so it is meaningful that r -values are positive in 92% of the total 85 data sets, and in 97% of the 62 data sets with >100 live individuals. Thus, although not all comparisons are significant at $P < 0.05$, coarse- and fine-mesh death assemblages are qualitatively consistent: species that are numerically dominant in a single census of the live fauna tend to be among the most abundant dead in the same set of pooled samples, and species that are rare or absent alive are usually rare among the dead.

Meta-analytic synthesis of individual Spearman r -values quantifies the strength of this positive result, yielding an average weighted r of 0.45 ± 0.03 for the 85 data sets considered as a single group (18). Eliminating extremely small data sets (<100 live individuals) increases the estimate slightly (average weighted r of 0.47 ± 0.04 for the 62 large data sets). However, segregating data sets by sieve mesh-size has a larger, previously unrecognized effect: for data sets with >100 live individuals, the average weighted r is 0.54 ± 0.05 for coarse-mesh death assemblages (24 data sets) versus 0.38 ± 0.06 for fine-mesh (38 data sets).

Fine-mesh data sets show poorer live-dead agreement in species' rank-order for both ecological and taphonomic reasons. Ecologically, any census of mollusks based on ≤ 1 mm sieves will capture, and often be numerically dominated by, larvae and newly settled juveniles; because settlement is pulsed, and most recruits die before reaching adult sizes of $\geq \sim 2$ mm, the species composition and abundance of live data generated by fine sieves are especially sensitive to the timing of the live-census [e.g., multiseason fine-mesh data sets of (3) and (17)]. The very small-bodied adults (<2 mm) of some opportunistic (weedy) species have similarly erratic presence alive. Pooling of multiple seasons of live data should thus damp the volatility of species dominance (e.g., Fig. 1A). However, in fine-mesh data sets, ecologically transient populations of larvae and early juveniles are still likely to numerically overwhelm the late-juvenile and adult component (LJA) of the local community, even though these larger specimens dominate in terms of biomass (20, 21). Consequently, larval and early juvenile

rather than LJA dynamics will determine the ranking of live species in fine-mesh data sets. Taphonomically, shells ≤ 1 mm are especially prone to out-of-habitat transport [e.g., (22)]. They furthermore have extremely brief persistence in the surficial mixed zone of the seafloor [e.g., half-lives <100 days, versus immeasurably long persistence of specimens >1.8 mm (23)], owing to greater vulnerability to chemical dissolution during early burial and ingestion by predators. Like live fauna, the composition and abundance of species collected dead on ≤ 1 mm sieves is thus volatile, depending on the timing of sampling relative to pulses of juvenile death [e.g., (3)], and relative to peak-seasons of shell transport and dissolution [e.g., (17, 24)].

In contrast, the stronger and more significant live-dead agreement in rank-order found among coarse-mesh data sets (>1 mm, and especially ≥ 2 mm) reflects the relative ecological stability of the LJA populations that constitute the bulk of molluscan biomass, and probably a tendency for the most abundant species to have higher turnover rates [shorter life-spans, improving the likelihood that their dead shells remain among the most abundant (10)]. This live-dead agreement in rank-order also argues that dead shells >1 mm are taphonomically more durable and resistant to transport than their ≤ 1 mm counterparts, a relationship supported by a wide range of experimental and field observations [reviewed by (12)]. These meta-analytic results place into context the only biomass-based live-dead analysis of mollusks, where species from 0.5 mm mesh ranked by biomass (dominated by adults) exhibited qualitatively better live-dead agreement than when ranked by numerical abundance [dominated

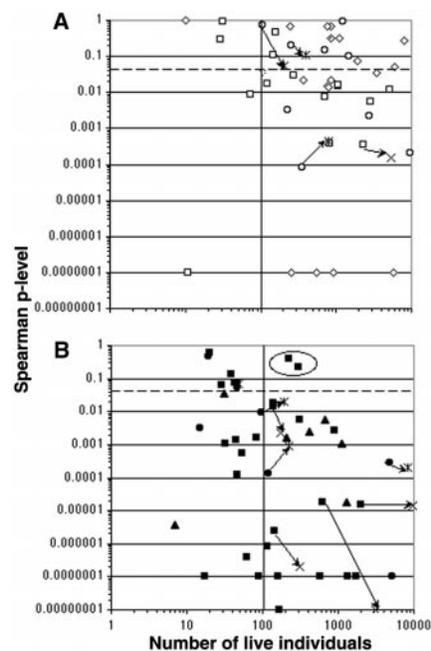


Fig. 1. Significance of rank-order agreement (Spearman test P -levels) as a function of sample size (number of live individuals) and inclusion of larvae and early juveniles (sieve mesh-size). Points lying below the dashed horizontal line are live-dead comparisons where rank-orders of live and dead species are significantly correlated at $P < 0.05$. (A) Results from all datasets ($n = 43$) generated using ≤ 1 mm mesh-size (open symbols). (B) Results from all datasets ($n = 42$) generated using >1 mm mesh-size (solid symbols); the two circled symbols are exceptional coarse-mesh datasets where live and dead species' rank orders are not significantly correlated despite information based on >100 live individuals. Both diagrams: Polygonal symbols indicate that dead species' rank order was compared with a single census of the local live fauna; diamond = marsh and creek, triangle = intertidal, square = coastal embayment, circle = shelf. x and * = dead species were compared with pooled time-series data on the local live fauna, from coastal embayment and shelf environments respectively. Arrows connect single-census and time-series results that are from the same place.

REPORTS

by juveniles; 2 lagoonal habitats (3, 25)].

The capture of species dominance by coarse-mesh, i.e. late juvenile- and adult-dominated, molluscan death assemblages in sedimentary substrata is corroborated by other indices of live-dead agreement. For example, virtually all ($95 \pm 5\%$) species sampled alive also occur dead in the same habitat (24 single-census >1 -mm-mesh data sets where the dead species list is based on at least 100 individuals; 95% confidence interval). In addition, most ($82 \pm 5\%$) individuals in these death assemblages are from species also censused alive. Thus, species encountered alive as "adults" (>1 mm body size) in any given census, which are arguably the ecological dominants for that molluscan community in terms of biomass and temporal persistence, are almost all represented among the dead in the same habitat, and the death assemblage is usually overwhelmingly composed of individuals from these species (26).

Meta-analysis also reveals clear differences in live-dead rank-order correlation as a function of substrate type. For both coarse- and fine-mesh death assemblages, species rank-order correlations are significantly higher in muddy habitats (seafloors of soft or firm mud, sandy mud, muddy sand) than in sandy habitats (well-sorted sands, lithic gravels and shell gravels; Fig. 2). Although the database is too small to partition variation more finely with confidence, rank-order agreement in muddy habitats appears to decline offshore (Fig. 2). In contrast, no significant trends are apparent across major environments among

sand and gravel substrata (Fig. 2).

The high rank-order agreement among coarse-mesh death assemblages from muddy substrata (average weighted r of 0.61 ± 0.09 ; Fig. 2) is very encouraging. Live-dead agreement generally prevails despite the high rates of carbonate dissolution documented in many marine muds [linked to microbial decomposition of organic matter (24, 27)] and the potential for exotic (out-of-habitat) shells to be introduced via rafting vegetation and storms [reviewed in (12)]. Although still strong, the lower rank-order agreement in sand- and gravel-hosted death assemblages (average weighted r of 0.50 ± 0.06) is expected and probably reflects multiple taphonomic factors, including higher rates of physical destruction and bioerosion, greater hydraulic import/export of shells, and greater time-averaging [especially in "relict" shell gravels and amalgamated sand bodies (28)]. Notwithstanding these genuine taphonomic biases, live-dead agreement in species rank-order among late-juvenile and adult-dominated death assemblages is still quite strong in absolute terms, and in both muddy and sand/gravel substrata exceeds the average agreement attained by larval- and early juvenile-dominated death assemblages (Fig. 2).

The pervasive impact of mesh-size on the fidelity of species dominance is crucial, both for understanding the nature of bias [focused on an ontogenetic subset of specimens (29)] and for guiding future applications of molluscan shell data. The mesh-size effect suggests a simple protocol for acquiring prehistoric

time-series data from molluscan death assemblages in sedimentary cores: by using only individuals ≥ 1.5 or ≥ 2 mm to characterize the live community, rather than the much finer mesh sizes (e.g., 0.3, 0.5, 1 mm) that typify marine macrobenthic surveys, live data could be directly compared with the most reliable down-core dead data. For comparisons entirely within the fossil record, a coarse-mesh protocol is already widespread among mollusk workers because ~ 1.5 to 2 mm is the usual lower size limit for taxonomically identifiable fossils (larval shells are rarely preserved except as part of a juvenile or adult). Thus, despite variation in the mesh-size used (2 mm, 3 mm, 5 mm, etc.), paleontologists have by default been focusing on the most reliable portion of the molluscan fossil record (30). A postjuvenile focus might also permit high-fidelity data to be isolated from the skeletal records of other marine and terrestrial metazoan groups.

This meta-analysis of a stringently controlled database demonstrates and quantifies the overall reliability of species dominance information from molluscan death assemblages, especially the portion composed of late juveniles and adults. Integrating modern and prehistoric records of benthic community change should thus not be fundamentally limited by natural taphonomic bias, and well-designed paleobiologic and macroecologic studies requiring abundance information should be justifiable, ranging from the extinction probabilities of rare versus abundant species, to long-term changes in how clades and trophic groups are packed into local communities.

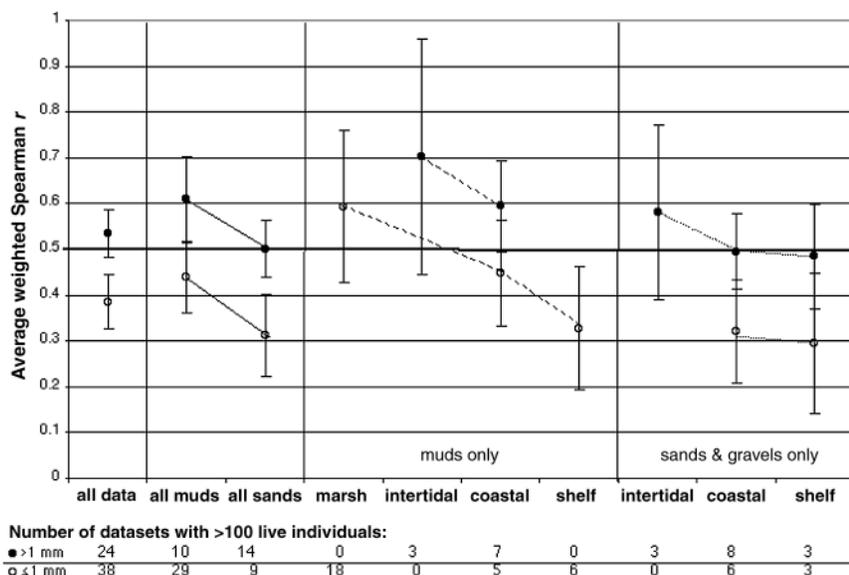


Fig. 2. Strength of rank-order agreement (average weighted Spearman r -values) is significantly higher among coarse-mesh data sets (individuals all >1 mm; solid symbols) than fine-mesh data sets (includes specimens ≤ 1 mm; open symbols), and in muddy substrata than in sands and gravels (points connected by solid lines; 95% confidence intervals). Environmental setting [marsh, intertidal, coastal embayment, shelf (13)] has a slight effect on live-dead rank-order agreement in data sets from muddy substrata (points connected by dashed lines), and no apparent effect among sands and gravels (points connected by dotted lines).

References and Notes

1. R. G. Johnson, *J. Paleontol.* **39**, 80 (1965).
2. J. E. Warme, A. A. Ekdale, S. F. Ekdale, C. H. Peterson, in *Structure and Classification of Paleocommunities*, R. W. Scott, R. R. West, Eds. (Dowden, Hutchinson & Ross, Stroudsburg, PA, 1976), pp. 143–169.
3. G. M. Staff, R. J. Stanton Jr., E. N. Powell, H. Cummins, *Geol. Soc. Am. Bull.* **97**, 428 (1986).
4. B. J. Greenstein, J. M. Pandolfi, *Bull. Mar. Sci.* **61**, 431 (1997).
5. R. K. Bambach, J. B. Bennington, in *Evolutionary Paleobiology*, D. Jablonski, D. H. Erwin, J. H. Lipps, Eds. (University of Chicago Press, Chicago, 1996), pp. 123–160.
6. M. L. McKinney, in *The Biology of Rarity*, W. E. Kunin, K. J. Gaston, Eds. (Chapman & Hall, London, 1997), pp. 110–129.
7. F. K. McKinney, S. Lidgard, J. J. Sepkoski Jr., P. D. Taylor, *Science* **281**, 807 (1998).
8. J. M. Pandolfi, *Am. Zool.* **39**, 113 (1999).
9. J. B. C. Jackson et al., *Science* **293**, 629 (2001).
10. S. M. Kidwell, K. W. Fleasa, *Annu. Rev. Ecol. Syst.* **26**, 269 (1995).
11. M. B. Davis, *Annu. Rev. Earth Planet. Sci.* **28**, 1 (2000).
12. S. M. Kidwell, D. W. J. Bosence, in *Taphonomy*, P. A. Allison, D. E. G. Briggs, Eds. (Plenum Press, New York, 1991), pp. 115–209.
13. Studies were included only if numerical abundance data were available for the complete molluscan fauna, including rare species, and if live and dead data were generated using a single known mesh-size (varies among studies from 0.3 to 5 mm; 1 mm median) [see supplementary materials, Part 1 (37)]. Twenty-six additional studies, including many classics in the paleoecological literature, did not meet these criteria

and are excluded [see supplementary materials, Part 2 (37)]. For each study area, spatially replicate samples were pooled into habitat-level datasets on the basis of sedimentary grain size, seafloor features (bedforms, vegetation, mass properties), and salinity. These habitats are comparable in scale and distinctiveness to sedimentary facies in the stratigraphic record, and were defined independently of faunal data. Habitats are grouped into four broad environments: salt marsh and tidal creek; intertidal flats and channels; coastal embayment (lagoons, estuaries, rias, and other semi-enclosed coastal bays where water energy, salinity, and/or oxygen level are reduced); and shelf (includes shoreface sands above fairweather wavebase, and an array of shallow and deep-water muds, muddy sands, and actively-building and relict shell gravels).

14. C. H. Peterson, Ph.D. thesis, University of California Santa Barbara, Santa Barbara, CA (1972).
15. A. I. Miller, M.S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA (1981).
16. R. Carthew, D. Bosence, *Palaeontology* **29**, 243 (1986).
17. J. Y. Aller, I. Stupakoff, *Cont. Shelf Res.* **16**, 717 (1996).
18. The 85 Spearman r -values were not normally distributed by a chi-square test ($P = 0.014$), suggesting that they should be transformed using Fisher's z before weighting (by $N-3$) and averaging [L. V. Hedges, I. Olkin, *Statistical Methods for Meta-Analysis* (Academic Press, New York, 1985)]. However, all stratified subsets of the raw r -values were normally distributed, and overall results of meta-analytically weighting and averaging z -transformed r -values do not differ significantly from those using untransformed r -values. The formal procedure of combining weighted results from many studies of disparate size and treatment (meta-analysis) has become a standard method in ecology, medicine, and the social and cognitive sciences where effect sizes are commonly small and diffuse, but is applied here to paleoecology for the first time.
19. Only 21% of ≤ 1 mm mesh comparisons show a significant correlation by a Sequential Bonferroni test [per W. R. Rice, *Evolution* **43**, 223 (1989)], whereas 84% of > 1 mm coarse-mesh comparisons are significantly correlated after correction. If the threshold mesh is set at ≥ 2 mm, then 100% of datasets show a significant correlation, including after correction.
20. D. J. Reish, *Ecology* **40**, 307 (1959).
21. G. Bachelet, *Mar. Environ. Res.* **30**, 21 (1990).
22. J. Y. Aller, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **118**, 181 (1995).
23. H. Cummins, E. N. Powell, R. J. Stanton Jr., G. Staff, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **52**, 291 (1986).
24. M. A. Green, R. C. Aller, J. Y. Aller, *Limnol. Oceanogr.* **38**, 331 (1993).
25. E. N. Powell, R. J. Stanton Jr., A. Logan, M. A. Craig, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **95**, 209 (1992).
26. Fine-mesh death assemblages have slightly lower representation of live species ($87 \pm 6\%$ found dead; 38 data sets with > 100 dead individuals), and lower agreement in species dominance ($68 \pm 6\%$ of dead individuals are from species censused alive). An earlier synthesis using a methodologically variable set of live-dead studies also yielded lower estimates than the present analysis [S. M. Kidwell, in *1998 Belle Baruch Conference on Organism-Sediment Interactions*, J. Y. Aller, S. Woodin, R. C. Aller, Eds. (Univ. of South Carolina Press, Columbia, in press); see supplemental materials for list of studies used (37)].
27. L. M. Walter, E. A. Burton, *Am. J. Sci.* **290**, 601 (1990).
28. Radiocarbon dating of shells in comparable marine habitats (bioturbated sediments in fully to seasonally aerated, level-bottom sedimentary seafloors) indicates that, except for open shelf shell-gravels where input can be summed over a few tens of thousands of years, time-averaging typically ranges from decades to centuries or a few thousand years [K. W. Flessa, M. Kowalewski, *Lethaia* **27**, 153 (1994)].
29. In general, strong rank-order agreement between a time-averaged death assemblage and a single-cen-

sus of live fauna suggests that either (i) rank-order of the standing live fauna does not change significantly over the duration of time-averaged input, or (ii) community composition does change, but the death assemblage is numerically dominated by the most recent cohort(s) of dead input. The latter scenario is ecologically more likely, and is also consistent with strongly right-skewed frequency distributions of shell age-since-death [e.g., K. H. Meldahl, G. A. Goodfriend, K. W. Flessa, *Paleobiology* **24**, 287 (1998)]. This scenario implies that, although species richness values may well reflect input from the entire duration of time-averaging [e.g., conclusions of (1-3, 10, 12, 14-16, 22)], dominance information might reflect only a final short segment of total elapsed time and thus have higher time-resolution.

30. Additional bias does accrue with lithification and subaerial emergence of marine sedimentary records, even where aragonitic shells persist, but species pres-

ervation is still very high. For example, J. W. Valentine [*Paleobiology* **15**, 83 (1989)] found that 77% of species living today in the Californian Province are preserved in Pleistocene terrace deposits.

31. Supplementary materials are available at www.sciencemag.org/cgi/content/full/294/5544/1091/DC1.
32. I thank original authors for discussion and permission to reanalyze their raw data, L.V. Hedges and C. W. Osenberg for meta-analytic statistical advice via the U.S. National Center for Ecological Analysis and Synthesis, M. Foote and D. Jablonski for early reviews, and the many individuals who assisted my search for datasets, especially associates of the U.S. National Museum of Natural History (Smithsonian Institution), Natural History Museum (London), Texas Bureau of Economic Geology, and California Academy of Sciences.

18 July 2001; accepted 11 September 2001

The Origin and Evolution of the Woolly Mammoth

Adrian M. Lister^{1*} and Andrei V. Sher²

The mammoth lineage provides an example of rapid adaptive evolution in response to the changing environments of the Pleistocene. Using well-dated samples from across the mammoth's Eurasian range, we document geographical and chronological variation in adaptive morphology. This work illustrates an incremental (if mosaic) evolutionary sequence but also reveals a complex interplay of local morphological innovation, migration, and extirpation in the origin and evolution of a mammalian species. In particular, northeastern Siberia is identified as an area of successive allopatric innovations that apparently spread to Europe, where they contributed to a complex pattern of stasis, replacement, and transformation.

Testing among models of species-level evolution in the fossil record ideally requires abundant samples that are finely stratified, accurately dated, and correlated across a broad geographical area (1). Most previous studies of fossil mammals have lacked the resolution to identify lineage splitting in contrast to phyletic change, nor have they offered sufficient geographical spread to distinguish in situ transformation from immigration (2, 3). Among large mammals, the mammoth lineage has one of the most complete records as well as pronounced adaptive morphological evolution through a time of well-studied environmental change. It also allows us to address the issue of geographical variation by sampling correlated sequences in both the European and Siberian parts of the mammoth's Eurasian range.

European mammoths (*Mammuthus*) have conventionally been divided into three chronospecies: the Early Pleistocene *M. meridi-*

onalis [recorded about 2.6 to 0.7 million years ago (Ma)], the early Middle Pleistocene *M. trogontherii* (~0.7 to 0.5 Ma), and the woolly mammoth *M. primigenius* of the late Middle and Late Pleistocene (~0.35 to 0.01 Ma). Important changes through this sequence include shortening and heightening of the cranium and mandible, increase in the height of the molar crown (hypsodonty), increase in the number of enamel bands (plates) in the molars, and thinning of the enamel (4-6) (Fig. 1). The dental changes resulted in increased resistance to abrasion, which is believed to correlate with a shift from woodland browsing to grazing in the open grassy habitats of the Pleistocene.

Critical to our study is the selection of samples that are chronologically restricted and independently dated (7). Dating methods for source deposits include radiometry (e.g., K/Ar or ¹⁴C), electron spin resonance/thermoluminescence, paleomagnetism, amino acid epimerization, first- and last-appearance datum of marine microfossils, and associated mammalian fauna. Samples from ~500,000 years ago (500 ka) onward can be tentatively correlated with marine isotope stages (MIS) (8).

The variable most frequently used in tracing elephantid evolution is lamellar frequency of

¹Department of Biology, University College London, London WC1E 6BT, UK. ²Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow 117071, Russia.

*To whom correspondence should be addressed. E-mail: a.lister@ucl.ac.uk