TIME-AVERAGING AND FIDELITY OF MODERN DEATH ASSEMBLAGES: BUILDING A TAPHONOMIC FOUNDATION FOR CONSERVATION PALAEOBIOLOGY

by SUSAN M. KIDWELL

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

Typescript received 24 October 2012; accepted in revised form 5 March 2013

Abstract: Ecosystems today are under growing pressure, with human domination at many scales. It is difficult, however, to gauge what has changed or been lost – and why – in the absence of data from periods before human activities. Actualistic taphonomic studies, originally motivated to understand preservational controls on deep-time fossil records, are now providing insights into modern death assemblages as historical archives of present-day ecosystems, turning taphonomy on its head. This article reviews the past 20 years of work on the temporal resolution and ability of time-averaged skeletal assemblages to capture ecological information faithfully, focusing primarily on molluscs from soft-sediment seafloors. Two promising arenas for ‘applied taphonomy’ are then highlighted: (1) using live-dead mismatch – that is, observed discordance in the diversity, species composition, and distribution of living animals and co-occurring skeletal remains – to recognize recent anthropogenic change, and (2) using time-averaged death assemblages as windows into regional diversity and long-term baselines, as a supplement or substitute for conventional live-collected data. Meta-analysis and modelling find that, in unaltered habitats, live-dead differences in community-level attributes can be generated largely or entirely by time-averaging of natural spatial and temporal variability in living assemblages, on time frames consistent with the range of shell ages observed in death assemblages. Time-averaging coarsens the temporal and spatial resolution of biological information in predictable ways; by comparison, taphonomic bias of information arising from differential preservation, production and transport of shells is surprisingly modest. Several challenges remain for basic taphonomic research, such as empirical and analytical methods of refining the temporal resolution of death assemblages; assessing the fate of resolution and fidelity with progressive burial; and expanding our understanding of the dynamics of skeletal accumulation in other groups and settings. Rather than shunning human-impacted areas as inappropriate analogues of the deep past, we should capitalize on them to explore the fundamental controls on skeletal accumulation and to develop robust protocols for bringing time-averaged death assemblages into the toolkits of conservation biology and environmental management.

Key words: taphonomy, anthropogenic, ecosystems, meta-analysis, molluscs, mammals.

Death assemblages – the taxonomically identifiable, dead or discarded organic remains encountered in a landscape or seabed – reflect input from current and past generations of organisms that have lived in an area, either temporarily or permanently. Although the relative abundances of species at a particular location may be more or less altered by postmortem transport, incomplete preservation and differences in life span (taphonomic bias), and by the time-averaging of generations, these dead individuals are the direct empirical evidence of the former existence of populations on some spatial scale and within some past time frame. The critical question is how much we can infer about the ecology of these past communities from the mostly disarticulated, biomineralized or otherwise refractory tissues that survive under ordinary, well-oxygenated conditions.

Along modern coastlines, the formation of such ‘ordinary’ death assemblages (DAs) has been the subject of scientific analysis for the last c. 100 years, focusing initially on the behaviour of shells and carcasses in moving fluids, rates and processes or controls of disarticulation and disintegration, and their eventual incorporation into sedimentary deposits. These are the realms of biostratinomy and necrolysis sensu Seilacher (1970) and Behrensmeyer and Kidwell (1985). Regardless of your taxonomic or environmental focus, Schäfer’s (1972) tome is still well
worth reading. In the last 30 years, these studies have morphed into ‘taphofacies’ analysis, direct dating to establish scales of skeletal time-averaging, laboratory and field experiments on the effects of specific agents and tissue types on persistence and transport, and stratigraphic tests of evolutionary-scale ‘mega-bias’, inclusive of post-burial diagenesis; see comprehensive reviews by Martin (1999), Behrensmeyer et al. (2000), and Kidwell and Holland (2002) and edited volumes by Allison and Briggs (1991), Donovan (1991), Traverse (1994), Allison and Bottjer (2010), and McGowan and Smith (2011). Analyses of the ecological fidelity of multi-species assemblages were well underway by the 1950s, motivated by questions about the spatial resolution and environmental significance of ‘biofacies’ and the climatic resolution of pollen, foraminifera and other microfossil groups, and expanded into vertebrates via palaeoanthropology in the 1970s. These tests of compositional fidelity have relied primarily on various forms of ‘live-dead’ comparison in modern environments. By the late 1990s, there had been so many live-dead studies of shallow-marine, soft-sediment mol-luscs that formal meta-analysis (statistical synthesis) was possible, and by the late 2000s, the effects of time-averaging could be modelled using parameters set by real-world values, drawing on many numerical age determinations for shelly assemblages and live-dead comparisons supported by large live samples (a shortcoming of many early studies, see Kidwell and Bosence 1991).

Here, I review the temporal resolution (acuity) and ecological fidelity (faithfulness) of multi-species DAs in ordinary, well-oxygenated settings, using insights gained from new field studies on diverse animal groups, meta-analyses and modelling to update treatments by Kidwell and Bosence (1991), Kidwell and Flessa (1995) and Behrensmeyer et al. (2000). I also recommend that we evaluate DAs not simply as analogues or precursors of fully buried fossil assemblages, but as sources of much needed historical data on the status of modern-day communities and ecosystems, which are under increasing domination by humans (Vitousek et al. 1997; Millenium Ecosystem Assessment 2005; Watling 2005; Worm et al. 2006; Diaz and Rosenberg 2008; Halpern et al. 2008, 2012; Strong and Frank 2010; Puig et al. 2012). Even in marine systems, overexploitation, eutrophication, habitat conversion and species introduction can have deep, multi-millennial histories regionally, and all have intensified over the last few centuries, accelerating and becoming truly global in the last several decades (Jackson et al. 2001; Lütze et al. 2006; Orth et al. 2006; Waycott et al. 2009; Worm et al. 2009; Taylor 2010; Halpern et al. 2012). There is thus a strong, societally motivated scientific need for longer-term perspectives on the status of species, habitats and ecosystems. How are present-day species ranges, abundances, community compositions and trophic structures modified from recent, less perturbed states? What is the natural range of variability in past systems? How do populations and communities respond to stress? Are present-day rates of change in natural systems unprecedented? Direct observations on living systems are generally not sufficiently long running to evaluate many processes, commonly focus on only a few species of commercial or other interest, and monitoring usually does not start until human stresses are already underway or imminent.

There have now been many general calls for palaeontology, broadly defined, to become a standard research partner in ecology and conservation biology (i.e. inclusive of zooarchaeology and museum archives; Lyman 1996; Swetnam et al. 1999; Lyman and Cannon 2004; Hayashida 2005; National Research Council 2005; Willis and Birks 2006; Dietl and Plessa 2009, 2011; Smol 2010; Hoeksema et al. 2011; Williams 2011; Willis and MacDonald 2011; Brewer et al. 2012; Conservation Paleobiology Workshop 2012; Louys 2012).

DAs from ‘core-tops’ and other surficial deposits are underexploited sources of biological information, reflecting accumulation on the decadal to millennial scales that are most relevant to evaluating anthropogenic impacts and macroecological dynamics (see Kidwell and Tomsovyich 2013, for an overview of the biological applications of DAs). Because DAs comprise ancient populations of largely extant species and sidestep the uncertainties of later diagenesis, they also represent a smaller step than do fully buried fossil assemblages, both technically and conceptually, for integration with neontological data and models. Nonetheless, aside from biogeographers and macroecologists, most biologists remain sceptical of information that is not based on direct observation of living organisms. Among ecologists, expectations are even more stringent, with controlled manipulative experiments and quadrant-scale data gathered over a few seasons as the standard resolution. Genuine integration of palaeontology with ecology and conservation biology (a true discipline of conservation palaeobiology) will thus require effort by palaeontologists on many fronts, including direct collaborations with biologists and managers and cross-training of graduate students and postdocs; we will need a change in scientific culture (National Research Council 2005; Conservation Paleobiology Workshop 2012).

This review is intended for palaeontologists and others interested in the taphonomic underpinnings of multi-species, assemblage-level palaeo data, but has the further aim of turning actualistic taphonomy on its head. Rather than evaluating the preservation of organic remains in modern systems as a guide to biological analysis of deep-time fossil records, to what extent can time-averaged DAs be used to evaluate the millennial-to-decadal-scale changes and drivers in present-day species, communities and biomes? Human-impacted areas violate our usual conception of
an appropriate setting for uniformitarianism, but they can provide superb opportunities for rigorous, ‘unnatural’ experiments on: (1) the effects of environmental variation that is of known timing and quantity on rates and selectivity of preservation, and (2) lag times in registering history via a time-averaged record. I describe here two arenas where societal needs are acute and conditions are excellent for basic taphonomic research: (1) using live-dead mismatch — that is, observed discordance in the diversity, species composition and distribution of living communities and co-occurring time-averaged DAs — to recognize recent anthropogenic change in the natural ecological baseline of a system and (2) using time-averaged DAs to estimate diversity and variability in unimpacted areas, as a supplement to or substitute for conventional live-collected data.

The focus here is on shelled molluscs from shallow-marine, soft-sediment seafloors, which have seen the bulk of actualistic taphonomic work and constitute a large proportion of marine macrobenthic biomass and diversity, both today and in the post-Palaeozoic fossil record. Complementary information on other groups and settings will be mentioned when studies suggest notable differences or similarities. To visualize new taphonomic research opportunities, however, readers are encouraged to mentally substitute terms such as leaf litter, colonies, tests and bones for ‘shells’ in any section of this article. The findings for molluscs summarized here can be used as null hypotheses (a comparative benchmark) for the resolution, fidelity and preservational controls on any time-averaged DA. Taxonomic groups with significantly lower, higher or more heterogeneous intrinsic durability than molluscs will doubtlessly present different numbers, as will assemblages from different settings: such variation points the way to underexplored variables. Finding ways to compensate analytically for variation in data quality and quantity is also simply part of the business of developing practical protocols, both for near- and deep-time analysts.

**WHAT ARE DEATH ASSEMBLAGES?**

A death assemblage (DA) is the set of taxonomically identifiable, dead or discarded organic remains present in the surficial mixed layer of a landscape or seafloor. Examples include shelly debris on the beach, leaf litter on the forest floor, bones concentrated by predators and core-top assemblages of pollen and other microfossil groups. The DA reflects input from past generations of organisms that have lived in the area, either temporarily or permanently, and is typically time-averaged to some degree. The more inherently durable the tissues of the local living community, the more favourable the local environment to preservation and/or the slower the rate of permanent burial, then the longer the interval of time and the more generations that can be represented by DA accumulation.

Death assemblages can include transported remains and may even be dominated by them, such as pollen carried by wind or water to a pond or adjacent ocean and bones concentrated or dispersed by a predator. In principle, such postmortem movement can bias biological information. For example, a single nonindigenous species that is easily transported might come to dominate a local DA (although finding examples for specimens >5 mm is quite difficult), and in some settings, many or all species might be allochthonous, for example, in washover fan deposits, tidal inlets and river mouths (see Kidwell and Bosence 1991, table IX). However, most empirical tests, going back to the earliest actualistic tests of biofacies (Kindle 1916; Parker 1956; Ladd et al. 1957), find that postmortem transport does not homogenize macrobenthic species occurrences across seafloors and landscapes. Spatial fidelity was the primary focus of virtually all live-dead studies up through the 1970s and concern persists (Donovan 2002; Dominici and Zushcin 2005). Workers, nonetheless, continue to find remarkable facies-scale resolution of macrobenthic DAs to living assemblages (LAs) in diverse settings, consistent with the general conclusion of Kidwell and Bosence (1991). Such habitat-scale spatial resolution exists even in some high-energy estuarine settings (Poirier et al. 2010). Similar spatial fidelity is observed in terrestrial mammal DAs, where calving and wintering grounds can be distinguished (Behrensmeyer and Miller 2012; Miller 2012a).

The most consistent effect of postmortem transport is simply to coarsen spatial resolution of the locally sampled DA, that is, coarsen the grain or pixel size of the image. The most obvious example is the pollen assemblage on a pond floor. This entirely allochthonous DA is of course completely wrong about the presence of trees living in the pond, but can, nonetheless, preserve biological information (including the relative rarity of tree species) from the broader surrounding landscape; the area reflected by the DA is proportional to the surface area of the pond or lake (see the still excellent review by Jackson (1994); Burnham et al. 1992; Meldahl et al. 1995; Burnham 2006). Beach strandings provide remarkably good samples of open-marine mammal populations (Liebig et al. 2003, 2007; Pyenson 2010, 2011). Similarly, small-mammal DAs produced by raptors are allochthonous to the roost site, but can accurately reflect the relative abundances of species and habitat types in the predator’s home range (Terry 2010a). Macrobenthic in woodrat middens are reliable records of surrounding vegetation at a c. 100 m scale (Lesser and Jackson 2011), and mussels in archaeological middens are also locally sourced (Peacock et al. 2012). On highly patchy seafloors, for example, seagrass...
meadows with pockets of sand and sandy or gravelly seafloors with locally exposed bedrock, sedimentary samples yield molluscan DAs with a mix of local soft-sediment species and ‘exotic’ epiphytic, endolithic and encrusting species that are not sampled alive without specialized hand-collecting (Russell 1991; Albano and Sabelli 2011). Whether these dead-only species should be called allochthonous elements is semantic: DAs correctly identify spatial variation and the within-habitat mix of bottom types. Palaeoecologists of course also need to appreciate that time-averaging alone, even in the absence of post-mortem transport, tends to ‘spatially average’ ecological data, because of random colonization events, in situ ecological succession and, on longer timescales, the migration of habitats (see discussions below).

Molluscan DAs are the taxonomically identifiable empty shells that are collected from a standardized area or volume of the seabed, usually as part of an effort to sample the living fauna. Samples typically come from the uppermost c. 5–20 cm of the seabed, excavated by hand or using remotely deployed gear such as a Van Veen grab, and are usually processed using sieves with 0.5, 1 or 2 mm mesh openings; specialized sampling methods are typically required for hard substrata. Only dead shells that represent a unique individual are counted: bivalve shell fragments must retain more than half of the hingeline, and gastropod fragments must retain the apex, and most workers count each disarticulated bivalve specimen as an individual. Including all identifiable fragments, regardless of their completeness, is generally ill-advised because it skews counts of individuals towards species with especially distinctive ornamentation or other features. Some workers count only unbroken specimens, which much reduce sample size, and some divide their counts to compensate for a living individual having more than one skeletal element (for quantification of effects, see Grayson 1984; Kowalewski et al. 2003; Lyman 2008). On the seafloor, dead shells may be fully exposed, partly exposed or fully covered by sediment, but all still reside within the ‘mixed layer’ of the sedimentary column. This part of the seabed is subject to physical reworking, bioturbation and diverse processes such as bioerosion, dissolution, maceration and encrustation that can destroy shells or otherwise make them taxonomically unidentifiable. This mixed layer is also characterized by the more or less continuous addition of newly dead shells, which refresh and update the composition of the DA.

As a time-averaged accumulation, a DA thus contrasts with the local LA, which is the multi-species set of living organisms sampled on a landscape or seafloor. In actualistic taphonomic studies, information on the LA is almost always based on a single survey of standing populations, usually at the same time that the DA is sampled. For molluscs, the LA is sieved from the same set of sedimentary samples used to sample the DA. This approach (using a one-time sampling of the DA and LA to generate data) overwhelmingly dominates the live-dead literature for most groups. Axiomatically, it yields nonaveraged, temporally high-resolution data on LA species richness, composition and relative abundances, which are obviously of a different temporal scale than the DA. Exceptionally, several years or decades of quantitative LA data might be available for a site or region, for example from national parks and biomonitoring efforts, but these data usually target only a few species or focus on human-impacted areas. In a few cases, live-dead authors have repeat-sampled the LA themselves for a study area (Peterson 1976) or repeat-sampled both the LA and the DA (Staff et al. 1986; Terry 2010a, b). Species checklists compiled from many decades of systematic and biogeographical surveys in a region provide a closer approximation of time-averaged data, although sampling methods, sampling intervals and focal habitats usually vary over time. Comparing the DA from a single or limited number of sites against a regional checklist is, nonetheless, a useful and underexploited means of evaluating DAs as surrogates or estimators of regional diversity (Bouchet et al. 2002; Warwick and Light 2002; Zuschin and Oliver 2003; Smith 2008; Tomasovych and Kidwell 2010a, b).

Regardless of how they ultimately pool their LA survey data, biologists typically collect it via seasonal or other short-term ‘snapshots’ of standing populations. Such biological information is thus fundamentally different from time-averaged DA data. My favourite metaphor for a time-averaged DA is the floor of a teenager’s bedroom. At the end of the week, the accumulation of cast-off clothing (mixed by his/her occasional search for a T-shirt to wear again, augmented by items dropped by younger siblings and reduced by items dragged off or destroyed by the family pet) provides a picture of the local clothes-wearer over that interval of time. The time-averaged picture can do an excellent job of capturing the range of temperatures (and relative frequency of temperatures) during the week as well as activities, such as athletics and a formal dinner with grandparents. This kind of data contrasts fundamentally with the picture created by any random snapshot of the teenager running out the kitchen door; such a photograph has higher temporal and ecological resolution, but a far narrower temporal scope and may be quite atypical of the week. (Time-)averaged data are thus not of poorer quality than nonaveraged data. They are simply true at a coarser temporal scale and can be far more useful than nonaveraged data for many questions.

This contrast between time-averaged and nonaveraged data is analogous to the effects of spatial scaling on species richness, formalized as the famous species–area relationship (Preston 1960) and increasingly explored as the
species–time relationship (Rosenzweig 1995; Adler and Lauenroth 2003; Scheiner et al. 2011). Living organisms sampled from (summed over) a large area will almost always have a larger richness than a sample of a single point or small patch within that region because the larger area encompasses more random spatial variability and structured environmental heterogeneity. The larger number of individuals encountered is also a factor, although measured richness tends to increase with spatial scale even if sample size is held constant. Thus, neither local (alpha, point) nor regional (gamma) diversity is better or truer information about diversity; these are simply perceptions of diversity at different spatial scales.

**CALIBRATING TIME-AVERAGING**

Over the last 20 years, actualistic taphonomists have learned two key things from tackling this problem directly on soft-sediment molluscan DAs (for other reviews, see Kowalewski and Bambach 2003; Kosnik et al. 2009).

1. **Large windows of time per assemblage**

Grab samples of the upper 5–20 cm of seabed can yield bivalve shells that are thousands and even tens of thousands of years old, based on radiocarbon dating or, increasingly, radiocarbon-calibrated amino acid racemization dating (Miller and Clarke 2007), making it clear that even aragonitic, dissolution-prone shells can survive on this timescale. These specimens are unlikely to have spent their entire postmortem existence at the sediment–water interface, where destructive processes are most intense, but have, nonetheless, been durable to one or more cycles of burial and exhumation within the uppermost mixed layer of the sedimentary column over extended periods. Absolute age estimates of shells from intertidal or very shallow subtidal species (brackish-water oysters, reef-top coral species) sourced from modern-day shelf sediments have long been used by geologists to track the rise of sea level since the last glacial maximum. Shells from modern-day outer-shelf DAs yield maximum ages of 10–20 000 years, whereas shells from inner shelves and beaches typically yield maximum ages of a few thousand years (Table 1; and see compilation of scattered published dates by Flessa and Kowalewski 1994). Millennial-scale time-averaging is also inferred for lacustrine gastropod assemblages (Cohen 1989). The youngest shells in actively accumulating assemblages are nearly zero years old, because they co-occur with living populations.

These impressive age ranges within modern DAs imply a much greater durability of shells, especially of aragonitic shells, than would be extrapolated from short-term experiments where half of all shells may be lost within several months to years (e.g. the recapture experiments of Cummins et al. 1986a; Powell et al. 2006). As long appreciated (Driscoll 1970; reviewed by Kidwell and Bosence 1991), even temporary burial below the sediment–water interface provides shells a refuge from destruction. Persistence through episodic, small-scale burial–exhumation cycles is also implied by the complex internal stratigraphy of fossil shellbeds and their common association with discontinuity surfaces, signifying skeletal accumulation under conditions of low net sediment aggradation (Fürsich 1978; Kidwell and Aigner 1985; Kidwell 1991, 1993). Temporary burial within the mixed layer of the seabed is consequently a fundamental component of most models of DA accumulation, both qualitative (Seilacher 1985; Davies et al. 1989; Kidwell 1989) and quantitative (Sadler 1993; Oliszewski 2004; Tomasovych et al. 2012). Biological encrustation and early diagenetic conditioning of shells (Glover and Kidwell 1993; Walter et al. 1993; Burdige et al. 2010) almost certainly also reduce shell loss rates in both carbonate and siliciclastic settings (discussions in Kidwell 1989, 1991; Powell 1992; Best et al. 2007). Exceptionally, where the modern shoreline erodes Pleistocene and older strata, quite old shells can be injected into a modern DA (for reviews, see Kidwell 1991; Kidwell and Bosence 1991; Wehmiller et al. 1995; Kowalewski and Bambach 2003). Although such shells can preserve fine morphological features such as growth rings, they have probably been diagenetically stabilized, much as ‘prefossilization’ via mineral infilling and recrystallization favours the recycling of vertebrate material (Reif 1982; Martill 1991; Rogers and Kidwell 2000; Koenig et al. 2009).

2. **Dominance by recently dead individuals**

Shells might survive equally from all increments of time within the total window of time-averaging, producing a uniform frequency distribution of shell ages; this is the simplest possible model of time-averaging, with simple summing of all shells ever produced by a series of local LAs at a site. It is more likely, however, that, even given a constant rate of input over time, younger cohorts will dominate because some portion of older shells (perhaps the majority ever produced) will have been lost to general wear and tear during their residence in the upper mixed layer of the sedimentary column (Kidwell and Bosence 1991).

The application of relatively inexpensive amino acid racemization dating to shelly assemblages has, over the last decade, made it possible to test these disparate models of time-averaging and temporal resolution: dozens and even hundreds of shells can be dated with decadal resolution, allowing confident evaluation of the shape of shell-
age frequency distributions. These empirical tests find that actively accumulating DAs on intertidal and subtidal seafloors have strongly right-skewed, ‘hollow’ shell-age frequency distributions, with a long tail of very old shells up to thousands or tens of thousands of years, but with half or more of all shells dating to the most recent few decades (in lagoons) or centuries (on open continental shelves; Fig. 1; median values in Table 1). These frequency distributions are typically fitted with exponential functions and compared using the implicit ‘taphonomic half-life’ of the shells (cf. Cummins et al. 1986a), a useful heuristic for the dynamics of shell loss and for modelling the likely temporal resolution of random subsamples of a DA (Flessa et al. 1993; Olszewski 1999, 2012b). Conceptually, a given DA is dominated by recent input because, after death, shells ‘decay’ randomly at a constant albeit exponential rate; for a tabulation of molluscan shell half-lives, see Kosnik et al. (2009). Shell half-life is often used as a measure of the effective temporal resolution of a DA; alternative metrics are the median shell age (my preference; does not presume an exponential rate of shell loss) and the dispersion of shell ages (e.g. the interquartile

TABLE 1. The temporal scale (total range of shell ages) and effective resolution (median shell age) of bivalve death assemblages (DAs) that are actively accumulating, ordered by median age.

<table>
<thead>
<tr>
<th>Setting</th>
<th>Genus (N specimens)</th>
<th>Median age (years)</th>
<th>Total age range (years)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siliciclastic bay, 3 sites at 2 m (Baja California MX 27°N)</td>
<td>Chione (72)*</td>
<td>[0, 0, 0]</td>
<td>[1011, 1045, 360]</td>
<td>Meldahl et al. (1997)</td>
</tr>
<tr>
<td>Siliciclastic lagoon, 6 sites (Copano Bay TX 28°N)</td>
<td>Mulinia (100)</td>
<td>4</td>
<td>&lt;60, except eroded outliers</td>
<td>Olszewski (2012b)</td>
</tr>
<tr>
<td>Carbonate perireefal, 5 sites at 1–44 m (Caribbean Panama 9°N)</td>
<td>Pitar, Codakia (24)</td>
<td>72 (F)</td>
<td>3300</td>
<td>Kidwell et al. (2005)</td>
</tr>
<tr>
<td>Carbonate back reef, 2 sites at 7 m (GBR AU 18°S)</td>
<td>Scissulina (110)</td>
<td>102, 338</td>
<td>1481, 2096</td>
<td>Kosnik et al. (2012)</td>
</tr>
<tr>
<td>Carbonate back reef, 2 cores at one 7-m site (GBR AU 18°S)</td>
<td>Tellina (66, 245)</td>
<td>10.5 (top 15 cm), 141 (entire 120 cm)</td>
<td>2148, 4670</td>
<td>Kosnik et al. (2009)</td>
</tr>
<tr>
<td>Carbonate back reef, 2 sites at 7 m (GBR AU 18°S)</td>
<td>Pinguitellina (150)</td>
<td>370, 516</td>
<td>4674, 3877</td>
<td>Kosnik et al. (2012)</td>
</tr>
<tr>
<td>Siliciclastic shelf, 4 sites at 4–35 m (Caribbean Panama 9°N)</td>
<td>Pitar (15)</td>
<td>375 (F)</td>
<td>5400</td>
<td>Kidwell et al. (2005)</td>
</tr>
<tr>
<td>Siliciclastic intertidal, 2 sites (Sonora MX 31°N)</td>
<td>Chione (30)*</td>
<td>483, 427 [405, 312] (F)</td>
<td>3569, 1752</td>
<td>Flessa et al. (1993)</td>
</tr>
<tr>
<td>Siliciclastic shelf, 17 sites at 23–58 m (Southern CA 34°N)</td>
<td>Nuculana (232)</td>
<td>629</td>
<td>12 712</td>
<td>Kidwell et al. (2010)</td>
</tr>
<tr>
<td>Siliciclastic shelf, 2 sites at 10–30 m (Brazil 23°S)</td>
<td>Semele (75)</td>
<td>760, 964</td>
<td>4437 years at 10 m, 4271 years at 30 m</td>
<td>Krause et al. (2010)</td>
</tr>
<tr>
<td>Siliciclastic intertidal (North Sea 54°N)</td>
<td>Cerastoderma (9)</td>
<td>[5684] (F)</td>
<td>7355 [7839]</td>
<td>Flessa (1998)</td>
</tr>
<tr>
<td>Siliciclastic beach, 21 sites (New Jersey 40°N)</td>
<td>Mercenaria (c. 200; max age from 18 radiocarbon dates)</td>
<td>No data (F)</td>
<td>[7664], except 4 eroded outliers &gt;30 ka</td>
<td>Wehmiller et al. (1995)</td>
</tr>
</tbody>
</table>

Species with known or suspected recent change in living populations

<table>
<thead>
<tr>
<th>Setting</th>
<th>Genus (N specimens)</th>
<th>Median age (years)</th>
<th>Total age range (years)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siliciclastic shelf, 14 sites from 19–72 m (Southern CA 34°N)</td>
<td>Parvilucina (235)</td>
<td>36 (high in 20th century)</td>
<td>10 758</td>
<td>Kidwell et al. (2010)</td>
</tr>
<tr>
<td>Siliciclastic shelf, 2 sites at 38 and 89 m (Southern CA 34°N)</td>
<td>Cyclocardia bailyi (50)</td>
<td>2400, 11 500 (X in AD 1950s?)</td>
<td>5900, 17 500</td>
<td>Kidwell et al. (2010)</td>
</tr>
<tr>
<td>Subtropical chenier, 2 samples (Baja Calif MX 31°N)</td>
<td>Chione (20)</td>
<td>Mean 182, 134 (X in AD 1930s)</td>
<td>279, 193</td>
<td>Kowalewski et al. (2000)</td>
</tr>
</tbody>
</table>

All shells dated by radiocarbon-calibrated amino acid racemization except for early studies (denoted by *) that used only radiocarbon. Values in square brackets have been calculated using shell ages provided by Kosnik et al. (2009), who used the Marine04 curve released in 2004 to recalibrate the raw radiocarbon shell ages reported by original authors. Some median ages are overestimates of effective resolution, either because authors aimed to quantify total range and thus avoided dating fresh-looking shells (denoted with F) or because living individuals are rarely sampled in the study area, suggesting local extirpation and thus cessation of shell supply (denoted with X; last three rows).
On inspection, the shapes of many (most?) shell-age frequency distributions are ‘more hollow’ than a simple exponential curve; there are more shells in the youngest age bin, fewer in the intermediate age bins and a longer, flatter tail of old shells than would be produced by a constant, exponential rate of loss (Fig. 1). Based on hundreds of dated shells from aragonitic species on the continental shelf of southern California, Adam Tomasovych and I are finding that these rather L-shaped shell-age frequency distributions are best fit using two exponential rates: an initially very high loss rate of individuals, but then an order of magnitude or more slower rate that applies to survivors (Kidwell et al. 2010; Tomasovych et al. 2012; Fig. 1D). This shape implies that, like a newly purchased machine part, if a shell does not ‘fail’ soon after arrival in the DA (and most do disappear, hence the steep initial slope), the likelihood of its destruction drops very strongly, probably because it has been temporarily sequestered by burial within the mixed layer and/or has become less reactive because of early diagenetic conditioning. Shells that survive the perilous initial postmortem phase subsequently persist over long periods, defining the total range of time-averaging in that setting. To the extent that these long-persisting specimens are drawn randomly from each past cohort of newly dead shells, they should reflect, in damped numbers, the original numbers of individuals per species that died rather than being biased towards intrinsically durable body types. The portion of the DA constituting the long tail of old shells thus might retain a reliable ecological signal, albeit one of coarse temporal resolution; the presence of aragonitic species in this tail indicates that it is not simply a diagenetic residuum (note that all data sets summarized in Table 1 and Fig. 1 are based on aragonitic bivalves). We are currently modelling the consequences of this particular two-phase fossilization
dynamic for the fidelity and temporal resolution of biological information (Tomasovych et al., 2012).

Consequences for biological information

Regardless of whether time-averaging is accompanied by one or two phases of exponential shell loss from cohorts, the large proportion of very young shells in right-skewed shell-age frequency distributions arises from recent input of shells to the DA and thus reflects at least partial autocorrelation of the DA with the local LA: the LA constantly refreshes the DA drawing on whatever species currently dominate the LA. This dominance by recently dead shells almost certainly explains why molluscan DAs can do a fairly good job of reflecting the species that are most abundant in the current LA, and the long tail of older specimens no doubt contributes to the DA capturing many species that are present regionally but rarely encountered alive (Kidwell 2002a; Tomasovych and Kidwell 2011; Olszewski 2012a; Tomasovych et al., 2012). Studies with repeat sampling of a DA demonstrate the ability of the DA to change in response to a changing LA on subannual-to-decadal timescales (Cummins et al. 1986a; Perry 1996; Ferguson and Miller 2007; Western and Behrensmeyer 2009; Fig. 2). The species identity of these new cohorts can pull the average composition of the DA away from its previous position within multivariate space if the new cohorts are sufficiently numerous and taxonomically distinct from earlier ones. The composition of the DA lags behind these temporal shifts and volatility in the composition of the LAs that feed it, simply because the DA is a summed record of many preceding LAs. The likely damping effect of time-averaging on species composition is long recognized conceptually (Peterson 1977) and emerges consistently from new quantitative forward models (see section on ‘Estimating diversity’ below).

Viewed another way, the survival in a DA of at least a subset of shells from earlier stages of time-averaging provides a memory of past community states that might have differed strikingly from the present-day LA. The larger the survival of shells from earlier cohorts, the greater the ‘taphonomic inertia’ of the DA to change, and thus, the stronger its memory of earlier phases in its accumulation (term from Kidwell 2008). Such inertia in composition is one possible cause of live-dead discordance in species’ presence or abundance. Natural short-term variability in the LA means that a snapshot of the LA will usually diverge from the (time-)averaged norm, and if the environment (and LA) actually shifts directionally over time, then a snapshot of the LA late in that history might diverge significantly. Strong natural variability in LAs and natural environmental changes during the window of time-averaging have been invoked to explain live-dead mismatch by authors back to the 1950s at least, even when they had expected to discover postmortem bias in preservation (Ladd et al. 1957; Warme 1969), and several authors have suspected human-driven community change.

FIG. 2. Since the 1960s, the relative abundances of live-surveyed mammal species (dark bars in A, black icons in B) have changed significantly in Amboseli National Park, Kenya, both in the park overall (A) and within a series of habitats (B). Death assemblages (DAs) sampled in the same areas have changed in comparable ways, demonstrating their ability to respond rapidly to directional change in community composition, especially in this tropical setting that experiences fairly rapid postmortem loss of bones (total persistence <100 years; analysis excludes bones that are in very advanced states of weathering). A, Live-dead comparison of the proportion of individuals that belong to the functional group of browsers or mixed feeders, which reflects the amount of wooded habitat in the park. B, Live-dead comparison based on multivariate analysis of relative abundance data for all species; arrows denote shifts observed from 1970–1976 to 1999–2004. From Western and Behrensmeyer (2009).
as a cause of live-dead mismatch in their system (Greenstein et al. 1998; Staff and Powell 1999). Global meta-analysis finds that in fact, anthropogenic stresses in the local system, especially eutrophication, are the strongest correlate of live-dead discordance in molluscan DAs, suggesting that the inherent taphonomic inertia of time-averaged DAs can be used to recognize shifted baselines in regions where the historical pre-impact status of the system is unknown (Kidwell 2007, 2008, 2009; see section on ‘Recognizing anthropogenic change’).

**Time-averaging in other groups**

Direct-dating data for other groups with mineralized or refractory tissues are scarce, but convey the same message of substantial persistence and thus the potential for a significant memory of ecological history, despite high rates of loss observed in short-term experiments on postmortem disintegration (for more context, see Kidwell and Behrensmeyer 1993; Kidwell and Flessa 1995; Behrensmeyer et al. 2000). Direct dating of modern shelf brachiopod DAs, for example, indicates maximum shell ages of thousands of years, comparable with those of co-occurring aragonitic bivalves (Krause et al. 2010) and contrary to expectations that rapid disintegration at least of inarticulate and punctate brachiopods might create a secular Phanerozoic trend in bed-scale temporal resolution (Kidwell and Brenchley 1994). The problem of differential durability between organic-rich and organic-poor shell types presumed by that trend thus still needs consideration (Tomasovych and Rothfus 2005; Tomasovych and Zuschin 2009). In general, we are still only beginning to understand variation in time-averaging among benthic groups. Martin et al. (1996), for example, found comparable, millennial-scale time-averaging in co-occurring intertidal benthic foraminiferal and bivalve assemblages, whereas Kosnik et al. (2009) detected longer taphonomic half-lives in durable than in fragile molluscan species, more in line with intuitive expectations (Kowalewski 1997). A unique evaluation of corals from uplifted Pleistocene reefs finds total age ranges of 800–1060 years per horizon, notably less than the multiple thousands typically found for molluscan assemblages and quite small given the average 20- to 80-year lifespans of coral colonies (Edinger et al. 2007). Scarce radiocarbon data for mammal bones find total age ranges in the order of a few hundred years for raptor-concentrated small-mammal DAs (Utah steppe rock shelters; Terry 2008) and c. 100 years for large ungulate DAs in temperate settings (Miller 2011), but millennial-scale time-averaging in the Arctic (Sutcliffe and Blake 2000; Miller 2012b). These age ranges contrast strongly with the decadal-scale age ranges extrapolated from mammalian bone weathering in subtropical and tropical settings (Behrensmeyer 1978; see review by Ross and Cunningham 2011).

**Resolution of fossil assemblages**

Whether fully buried fossil assemblages, both in Holocene cores and in older stratigraphic records, exhibit the same temporal resolution as modern-day DAs is a critical issue for integrating data from multiple sources. The more closely that fossil assemblages resemble DAs in temporal resolution (i.e. the more true our assumption that DAs are good analogues of fossil assemblages), then the more seamless the eventual merging of palaeo- and neocological data. Do fossil assemblages exhibit the same right-skewed age-frequency distributions as modern DAs, implying a ‘freezing in’ of temporal resolution such as via sudden complete burial events (cf. Flessa et al. 1993)? Or, with the increasing isolation from a supply of new shells that comes with gradual burial, do age-frequency distributions become flatter because they quickly lose the youngest most reactive shells? Or do they become bell-shaped (normal, from upward smearing) or perhaps even left-skewed? Of these, only flattening requires a decrease in effective temporal resolution. All of these alternative shapes have been reported for DAs that have become cutoff from LAs: for example, shells from relict Holocene beach ridges (Kowalewski et al. 1998); land snails from buried aeolian records (Yanes et al. 2007); corals from an uplifted reef terrace (Edinger et al. 2007); and a downcore series of small-mammal assemblages (Terry 2008). Based on molluscan DAs on the southern California shelf, I suspect progressive flattening of shell-age frequency distributions with burial, owing to rapid decay of shells from the youngest age bins, which are no longer being replenished from the LA (Kidwell et al. 2010; Tomasovych et al. 2012); we are presently testing for change in the shape of shell-age frequency distributions with burial, using newly acquired 3-m-long cores that ensure penetration beyond the surficial mixed layer. Within the mixed layer, median ages and right-skewed age-frequency distributions are both expected and observed to be homogeneous (Kosnik et al. 2007, 2009).

**LIVE-DEAD TESTS OF FIDELITY**

Live-dead comparisons are the most common actualistic method of quantifying the fidelity (faithfulness) of the fossilization of ecological information. This uniformitarian extrapolation assumes that DAs are reasonable proxies for fully buried fossil assemblages, that is, that the assemblage undergoes no change in temporal resolution or differential loss of species as a result of the burial process.
and later diagenesis. These assumptions require explicit testing, as mentioned above. However, regardless of the outcome of such tests on the effects of permanent burial, live-dead studies are certainly useful for evaluating modern, mixed-layer DAs as decadal- to millennial-scale archives of present-day ecosystems.

In live-dead studies, the number of individuals (or species, functional groups or age classes) occurring alive (LA) is compared with the number occurring dead (DA). These comparisons are always made at a specified scale (e.g. for a given point in a habitat, for the habitat as an entity, for a multi-habitat landscape) and for a specified group (e.g. all macrobenthos, molluscs only, bivalves only). Many measures used to describe a living community can, with little or no modification, be used to describe a DA, for example, species richness, evenness or a listing of species present, perhaps ranked by relative abundance. Consequently, methods that are widely used to compare two samples of live-collected individuals can be used for live-dead comparison, for example, a Jaccard index of taxonomic similarity between living and dead lists or the coefficient of rank correlation in species abundances alive and dead. Multivariate methods are also possible. For example, do samples of the DA in a habitat plot inside, outside or partially overlapping the multi-dimensional space occupied by samples of the LA from the same habitat? Live-dead agreement can also be assessed for single species. For example, how closely does the body-size or age-class distribution of Species X observed in the time-averaged DA match that of a standing population (Cummins et al. 1986b, Tomasovych 2004)? Alternatively, the observed DA can be compared with the species-abundance, body-size or age-class distribution that should be produced (the ‘expected dead’), given some model of mortality acting on standing populations (Van Valen 1964; Behrensmeier et al. 1979; Kidwell and Rothfus 2010).

Live-dead agreement thus cannot be reduced to a single canonical value because so many different biological attributes can be considered. It varies among major groups owing to differences in intrinsic postmortem durability and can vary even within a single group depending on the collecting method (e.g. sieving effect recognized by Kidwell 2001, 2002b; Kowalewski and Hoffmeister 2003). It can also vary among environments, holding the group, sampling method and biological metric constant. For example, for molluscs, live-dead agreement is on average poorer on shelves than in estuaries and lagoons (Kidwell 2001, 2002b), and taphonomic processes other than within-habitat time-averaging (probably environmental condensation) are more important there (Tomasovych and Kidwell 2011). Live-dead agreement for molluscs is lower in reefs than in soft seafloors, owing in part to dead specimens becoming overgrown by living organisms (Zuschin et al. 2000; Zuschin and Oliver 2003; Zuschin and Stachowitsch 2007). Similar challenges confront live-dead studies of coral (Greenstein 2007).

In general, however, some optima for live-dead agreement are emerging. For example, in marine macrobenthic systems, fairly poor live-dead agreement results when the living and dead lists of all macrobenthos encountered in a habitat are compared. In such studies, polychaetes almost always constitute the majority of macrobenthic species and individuals, but leave no macroscopic biomineralized remains, thus ensuring that DAs are dominated by some other group, usually molluscs (Schopf 1978; Staff et al. 1986; Staff and Powell 1988, 1999). Biomass-based comparison of the same lists, in contrast, yield higher live-dead agreement because, by virtue of their relatively large body sizes, molluscs almost always dominate the LA as well as the DA (unique test by Staff et al. 1985). Live-dead agreement generally improves, up to a point, as the taxonomic or functional focus of the analysis is (1) narrowed and (2) shifted to the most intrinsically durable groups, such as shelled molluscs or corals. This narrowed focus reduces variation in body type or size among taxa in the analytical group and thus diminishes the opportunity for preservation bias of diversity and other ecological information. Live-dead agreement declines, however, with too narrow a focus. For example, live-dead comparison of the raw numerical or proportional abundance of a single species at a site (for example, via a chi-square test) commonly shows poorer agreement than is suggested by regression through a bivariate plot of living versus dead proportional abundances of multiple species in an assemblage. Similarly, a rank-order test of species’ living and dead abundances that focuses only on the few most abundant species commonly yields lower correlations than the same test applied to the entire assemblage inclusive of rare species. To avoid zeros in data matrices, some authors inappropriately eliminate species that occur either live-only or dead-only, which inflates live-dead agreement.

Based on meta-analyses of molluscan data sets from tropical and temperate nonreefal seabeds, live-dead agreement in community-level attributes becomes quite good when the entire abundance spectrum of the assemblage is used, when the question is posed at the spatial scale of a habitat or facies rather than for a single point and when the LA and DA are each characterized using 20 or more individuals, reflecting the pooling of two or more point samples. These findings can be used to guide the choice of study systems and sampling methods. For example, positive Spearman’s correlation coefficients (the commonest measure of live-dead agreement in species relative abundances) indicate that a DA, on average, captures whether species were among the most abundant or among the rarest in the source LA, contrary to concerns that such information would be badly biased by differen-
tial production and preservation of species (Kidwell 2001, 2007). At finer scales (e.g. comparing the LA and DA from a single sampled point, the equivalent of a single bed within an outcrop or core), live-dead agreement in this and other biological attributes is almost always poorer, even if the sample size (numbers of individuals) is comparable with that of a facies-level analysis. For tests of scaling effects on multiple biological measures, see section on 'Estimating diversity' below and Tomasovych and Kidwell (2009a).

Although not yet tested explicitly, these patterns should be general properties. For example, comparing the leaf litter in a forest patch with the standing populations of all plant species occurring there, from delicate herbs to trees with tough leaf cuticle, will almost certainly yield poorer agreement than if only woody species are considered, just as comparisons of lacustrine pollen assemblages with surrounding forests show poorer agreement if animal- as well as wind-pollinated species are considered as the two have strongly disparate potentials for pollen transport (Jackson 1994; Davis 2000); see Sims and Cassara (2009) for an analogous bias for seeds. Similarly, live-dead agreement in richness, species composition and relative abundances is quite good for small mammals (Terry 2010a) and for large mammals (Behrensmeyer and Boaz 1980; Miller 2011), but poorer for groups with more heterogeneous body sizes, such as birds (Behrensmeyer et al. 2003; Turvey and Blackburn 2011), which also have inherently lower preservation potential than mammals (Cruz 2008).

Finally, presence–absence tests of live-dead agreement can be applied to a broad array of groups and settings and thus promote cross-scale comparisons. They provide a relatively simple means of quantifying per-taxon 'fossilization rates', although none are true rates. For example, 'What percentage of species sampled alive are also encountered dead?' can be asked at any spatial scale: at a site, in a habitat, in a region (Kidwell and Bosence 1991; Kidwell and Flessa 1995; Johnson 2006; Kidwell and Tomasovych 2013) and across scales (e.g. 'What percentage of species in a regional checklist can be found among the dead at a single site?', discussed above). Such questions can also be posed at many taxonomic levels and can be expanded temporally, as in:

1. 'What percentage of species known dead or alive in a local area today is also present in the local Quaternary fossil record?' Examples of answers are as follows: 61 per cent of rocky shore mollusc species in temperate California (Russell 1991) and 30–60 per cent of mollusc species in fjords, although 100 per cent are present in the fossil record of the larger Arctic region (Gordillo and Aitken 2000);

2. 'What percentage of species living today in a province have a Pleistocene record somewhere within that province?' Examples of answers are as follows: 85 per cent of Californian bivalve species (Valentine 1989) and >88 per cent of mollusc species in temperate Chile, if corrected for the 'discoverable fraction' (Rivadeneira 2010; and see Cooper et al. 2006);

3. 'What percentage of genera known from the Recent (dead or alive, anywhere in the world) have been reported as fossils of any age, anywhere in the world?' Examples of answers are as follows: 76 per cent of 1292 bivalve genera (Valentine et al. 2006; see Harper 1998 and other authors in Donovan and Paul 1998).

For molluscs, these studies indicate that fossilization bias acts against genera with very small bodies, high-organic aragonitic shells and/or deep-water preferences (because few outcrops are available of such facies). Threar comparisons of living, death and Quaternary fossil assemblages from a small area are also useful. Although one might assume that live-dead agreement will be greater than living-fossil agreement, results vary considerably owing to various combinations of differential preservation and natural and anthropogenic environmental changes (Palmqvist 1993; Greenstein et al. 1998; Edinger et al. 2001; Greenstein 2007; Erthal et al. 2011; Yanes 2012; see section on 'Live-dead mismatch' below).

Such presence–absence measures of successful fossilization deserve more systematic empirical exploration and can be phrased in diverse hypothetic–deductive terms, as in Schopf's (1978) classic analysis of the Puget Sound's intertidal macrobenthos (repeated in Arctic waters by Aitken (1990) and for rocky shores by Johnson (2006)). Schopf asked: What percentage of genera listed in the regional checklist of live-collected invertebrates has a strong potential for fossilization? The answer was 30 per cent, because these genera have sturdy calcareous shells or tubes; 40 per cent of all genera have a poor potential owing to rapid postmortem disarticulation or largely uncalcified skeletons, and the remaining 30 per cent have no potential under ordinary conditions because they lack macroscopic hard parts. He also asked: What percentage of regional genera is known to have a fossil record somewhere in the world (40 per cent) and what are the expected biases among habitats (insignificant, somewhat surprisingly) and among trophic groups (considerable, with preservation favouring herbivores). 'Recent versus fossil' tests of species abundance information are subject to far more assumptions, such as comparison of modern abundances of breeding birds with occurrences in Quaternary natural and archaeological assemblages (Turvey and Blackburn 2011).

**RECOGNIZING ANTHROPOGENIC CHANGE**

The mismatch of paired living and DAs has usually been taken as a distressing indicator of a potentially unreliable
record. However, LAs and DAs might diverge in composition for many reasons:

1. **Differential preservation**, so that the DA is biased against species or ontogenetic age classes with low intrinsic durability;
2. **Postmortem transport** of specimens into or out of the area, for example, by currents, tides, rafting and predators;
3. **Differential turnover** of living populations, such that species with short-lived individuals are over-represented relative to long-lived individuals (lifespan bias). Clutch size, age of first reproduction and frequency of reproduction affect the size of the standing population as well as the numbers of individuals dying and so should not by themselves bias the DA;
4. **Natural temporal variability in the LA**, so that any one-time sampling of the LA will not resemble the local DA even in the absence of postmortem bias, given the contrast in scale between nonaveraged live data and time-averaged dead data, and even when live-dead differences in sample size are factored out; LA data based on a single census are an ‘inadequate’ characterization of the community;
5. **Environmental change within the window of time-averaging**, with the DA retaining a memory of former habitats or community states that existed, at least temporarily, at the accumulation site; palaeontological concept of faunal or environmental condensation.

Factors 1–3 constitute taphonomic bias *per se*: differences in the intrinsic durability and productivity of organisms, interacting with extrinsic environmental conditions at the accumulation site, cause the richness, structure, composition, spatial resolution or other attributes of the DA to diverge from the LA. Factors 4 and 5, on the other hand, are changes in DA attributes that arise from a contrast in the temporal scale. For example, even if all species have identical rates of population turnover and equal potential for preservation and postmortem transport so that no taphonomic bias can arise, a time-averaged DA will be richer than any single sample of the LA because the DA is a temporally inclusive sample of random, short-term variability in LA species composition (Factor 4). The longer the window of time-averaging, the more likely that the local environment and LA will shift beyond random, within-habitat levels of variability (Factor 5), much as broadening the spatial scale over which living individuals are sampled tends to increase the number of distinct environments and communities encompassed. Time-averaging, of course, also increases the window of opportunity for taphonomic bias to arise, but the effects of such bias need to be kept distinct from the effects of time-averaging, both conceptually and analytically (distinction in Tomasovych and Kidwell 2009a, b, 2010a, b, 2011; Tomasovych *et al.* 2012).

Evaluating these alternative explanations for live-dead differences, a global meta-analysis of molluscan habitat-scale data sets (38 from open shelves, 73 from estuaries and lagoons) found that the strongest correlate of significant mismatch between the taxonomic composition and rank abundance of living communities and co-occurring time-averaged DAs was the presence of strong recent human stress in the study area (Kidwell, 2007, 2008, 2009; Fig. 3). Not all areas with known anthropogenic stress exhibited poor live-dead agreement: c. 40 per cent of data sets from such areas displayed the same, relatively good, live-dead agreement as did data sets from areas with zero or very low human stress. However, data sets with poor live-dead agreement all came, with few exceptions, from areas with a long history of chronic or acute stress, such as intense agriculture in the watershed, strongly polluting coastal industry or urbanization, as established from independent governmental and scientific reports (see Table 2 for scoring system). The implication was that live-dead discordance arises because the DA retains a memory of the former abundance of species, prior to changes in the LA associated with anthropogenic stress (Factor 5).

To test this meta-analytical hypothesis, the compositions of the LA and DA must be evaluated closely, with special attention paid to the natural history of the species that are most responsible for live-dead mismatch; alternative hypotheses of time-averaging (Factor 4) and taphonomic bias of the DA (Factors 1–3) must be rejected (Table 3). (1) Which species are disproportionately abundant in the LA relative to the DA or perhaps occur ‘live-only’? The implication is that their populations are increasing or entirely new to the system, but this live-dead contrast might indicate instead that their shells are difficult to preserve in general or under local conditions or that the DA is undersampled. Although dead mollusc shells are usually far more abundant than living individuals (Kidwell 2002b), the opposite can be true, especially in coastal marsh, cold-water and freshwater settings. (2) Which species are disproportionately abundant in the DA or perhaps occur ‘dead-only’? One implication is that their local populations are waning or might have gone locally extinct, but the contrast might reflect strong postmortem import from other habitats or sampling bias against living specimens owing to sampling method. (3) Finally, even if taphonomic explanations and/or small or disparate sample sizes seem inadequate to explain observed live-dead differences, does the implicit historical change in the ecosystem make sense in the light of what is known about human stresses? In some areas, of course, the LA may be so degraded that all or almost all molluscan species occur dead-only.

For the data sets in my meta-analyses, I found that live-dead discordance overwhelmingly signified ecological
change within the window of time-averaging rather than taphonomic bias or inadequate sampling and was highly consistent with known human stress. To use geological terms, the observed live-dead differences made sense as ecological deformation (strain, biological response) driven by an independently known local external force (stress, human activities). For example, in open-shelf data sets, species that were dead-only or otherwise ‘overabundant’...
TABLE 2. Semi-quantitative scale used to score human stress on a study area at the time of sampling, as applied to anthropogenic eutrophication (AE) and bottom trawling (BT).

<table>
<thead>
<tr>
<th>Score</th>
<th>General meaning</th>
<th>Anthropogenic eutrophication</th>
<th>Bottom trawling</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Stress absent or negligible</td>
<td>AE0: no human settlements in area nor active land clearance</td>
<td>BT0: no exploitation or only artisanal harvesting with minimal habitat destruction</td>
</tr>
<tr>
<td>0.5</td>
<td>Stress possibly present</td>
<td>AE0.5</td>
<td>BT0.5</td>
</tr>
<tr>
<td>1</td>
<td>Stress definitely present</td>
<td>AE1: coastal development and/or major development in watershed</td>
<td>BT1: commercial harvesting, gear can dislodge benthos other than target species</td>
</tr>
<tr>
<td>1.5</td>
<td>Stress definite and possibly intense</td>
<td>AE1.5</td>
<td>BT1.5</td>
</tr>
<tr>
<td>2</td>
<td>Stress definitely present and intense</td>
<td>AE2: near to large point source, usually in addition to diffuse sources</td>
<td>BT2: especially intense commercial trawling, for example, more than once a year</td>
</tr>
</tbody>
</table>

This approach permits diverse kinds of information to be incorporated into a score, for example, government reports on trends in water quality or quantities of fish caught, academic research reports on pollution, date when local or state governments imposed regulations on pollution or fishing, the economic and cultural history of the region, historical insights from sedimentary cores on habitat decline and the expert but unpublished knowledge of scientists familiar with the area. Different kinds of data are available in different regions. The general categories can be applied to other human stressors (e.g. toxic pollution, thermal pollution, dredging and spoil dumping, introduced species) and scores could be added to produce a metric of ‘total human stress’. Reprinted from Kidwell (2009).

relative to their proportional abundance in the LA were not small, easily transported shells but a diverse array of fairly large-bodied, suspension-feeding bivalves preferring ordinary nutrient regimes, whereas species that were live-only or overabundant in the LA were organic-loving deposit feeders and chemosymbiotic species (Kidwell 2008). Live-dead discordance was higher on narrow steep shelves than on broad gentle shelves, suggesting that post-mortem transport might have been a contributing factor there, and some between-habitat transport of shells is likely. However, the presence of human stressors had higher explanatory value in multiple regression and partitioning, especially the presence of anthropogenic eutrophication (bottom trawling was also considered). In data sets from estuaries and lagoons, live-dead contrasts identified the loss of seagrass rather than increased organic input itself as the critical factor: species preferring or requiring seagrass were disproportionately abundant dead and some were dead-only, occurring in habitats that were described as unvegetated seafloor at the time of live-dead sampling (Kidwell 2007; Fig. 4A).

Detecting a signal of bottom trawling on molluscan communities requires focusing (1) on the subset of species known to be sensitive to physical disturbance, namely epifauna, nestlers and especially sedentary infauna and their commensals and (2) on the subset of bottom types where biologists find that trawling has the strongest negative consequences, namely seafloors containing shell and lithic gravel (Thrush et al. 1998; Collie et al. 2000; Thrush and Dayton 2002; Gray et al. 2006). Fauna on mobile sands are adapted to frequent disturbance, and faunas on muds respond to trawling by shifting to soft-bodied, non-molluscan assemblages. Although further data sets would be welcome, the results are encouraging. Live-dead data sets collected from areas with definite intense trawling, such as in the North Sea, English Channel and Gulf of Mexico, exhibit significantly poorer live-dead agreement than data sets from seafloors subject to moderate or no trawling (Kidwell 2009; Fig. 4B). This plot is also taphonomically interesting, because many of the species in the disturbance-sensitive group have calcitic shells and are thus generally expected to have a preservational advantage over exclusively aragonitic infaunal bivalves. Instead, no such preservational advantage is observed: the slope of the line of regression for data sets from nontrawled study areas is approximately one and the y-intercept is approximately zero. Thus, as a group, calcitic species are preserved in proportions comparable with their abundance alive under natural conditions.

Expanding this approach

Live-dead mismatch is thus a promising tool for retrospective evaluation of local and regional impacts, especially given the scarcity of historic survey and biomonitoring data even in highly developed areas. Explicit tests of this approach in other settings, with other groups and for additional human stressors are promising for conservation applications. For the purposes of this study, these also illustrate the potential for using human impacts to test basic taphonomic dynamics. Ferguson (2008), for example, found that molluscan DAs in Florida Bay could detect, with high spatial resolution and within a few years, the ecological impact of a small artificial line-source of nutrients, which favoured one type of seagrass over another. At a coarser spatial scale, Casey and Dietl (2010) found a gradient in molluscan live-dead rank-abundance agreement consistent with a known gradient of anthropogenic eutrophication within the relatively
TABLE 3. Working hypotheses (H1–5) to explain species that occur only in the living assemblage ('live-only') or nearly so (their proportional abundance alive is much greater than their abundance dead), and species that occur only in the time-averaged death assemblage ('dead-only') or nearly so (dead abundance ≫ living abundance).

<table>
<thead>
<tr>
<th>Observation</th>
<th>H1: Under-sampling bias</th>
<th>H2: Collection bias</th>
<th>H3: Time-averaging bias</th>
<th>H4: Taphonomic bias</th>
<th>H5: Ecological change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species is live-only and rare</td>
<td>Sample size DA small and &lt; LA</td>
<td>Unlikely</td>
<td>DA reflects very little time-averaging</td>
<td>Unlikely</td>
<td>Species has intrinsically low preservation potential</td>
</tr>
<tr>
<td>Species is live-only and is moderately or quite abundant</td>
<td>Unlikely</td>
<td>Unlikely</td>
<td>Relatively new, highly successful arrival</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species is dead-only and is rare</td>
<td>Sample size LA &lt; DA</td>
<td>Gear bias against LA</td>
<td>Most likely even if sample size LA is small</td>
<td>Postmortem exotic: most likely if consistently present in LA of adjacent habitat(s)</td>
<td>Waning; most likely if multiple dead-only species do not 'fit' with the LA</td>
</tr>
<tr>
<td>Species is dead-only and is moderately or quite abundant</td>
<td>Unlikely</td>
<td>Very strong gear bias</td>
<td>Unlikely</td>
<td>Postmortem exotic: most likely if abundant in LA of adjacent habitat(s)</td>
<td>Past dominant, now waning or extirpated</td>
</tr>
</tbody>
</table>

Hypotheses for an observed live-dead discordance should be evaluated in sequence from left to right, but are not mutually exclusive. LA, living assemblage; DA, death assemblage.

Rare species: represented by one or two specimens or at only few sites (low occupancy); the threshold percentage depends upon sample size but would typically be < or ≪ 1% of all individuals in an assemblage. Hypothesis 1 attributes strong live-dead discordance in a species solely to the small sample size of either the LA or the DA. H2 attributes discordance to methodologic bias against collection or detection of a species in the LA, for example owing to gear that cannot acquire deep-burrowing, cryptic or highly mobile individuals, or specimens attached to patches of hard substrata. H3 attributes discordance to the coarser temporal scale (time-averaging) of the DA compared to the LA. H4 attributes discordance to postmortem bias, for example long lifespan (few dead shells produced per unit time), low postmortem durability (e.g. small, thin and/or high-organic microstructure; prone to being overgrown) and/or high potential for transportation (e.g. low-mass shell, epifaunal life habit). H5 attributes discordance to a change in the LA within the window of time-averaging, resulting in species invasion, extirpation or significant increase or decrease in a species' population size; change may be natural or anthropogenic, and affect the biotic and/or abiotic environment. H5 is strengthened if the species that occurs live-only or dead-only does not ‘fit’ ecologically with species occurring both alive and dead, and if multiple species in the assemblage fit this description.

urban Long Island Sound (taxonomic similarity does not vary), and Michelson and Park (2013), evaluating ostracod assemblages in seven saline lakes from Bahamas, encountered poor live-dead agreement only in the lake with a known history of human use (plantation, penning of sea turtles). Focusing on raptor-concentrated small-mammal assemblages, Terry (2010b) encountered significantly poorer live-dead agreement in richness, evenness, taxonomic similarity and species rank-abundance in a steppe landscape with strong recent human disturbance (military training, invasive cheatgrass) than in a relatively undisturbed counterpart (Fig. 5A–B). Using a 10-year time series of abundance data encompassing the invasion of an alien predatory gastropod, Chiba and Sato (2012) found high live-dead agreement of species abundances in samples collected before invasion and poor agreement afterwards, reflecting decimation of the previously dominant bivalve (Fig. 6). Comparing bone assemblages with a decade of wildlife census data in Yellowstone National Park, Miller (2011) discovered high live-dead agreement in the rank-order abundances of ungulate species and significant offsets in proportional abundances only for species known to have undergone strong recent population changes (Fig. 7). Species that are disproportionately abundant in the DA compared with the present-day LA are ghosts of larger population sizes in previous decades.

Taphonomic research opportunities

Analysis of live-dead mismatch in human-stressed areas also provides excellent conditions for basic taphonomic research, using past intentional and unintentional manipulation of environments by humans as ‘unnatural’ experiments in fossilization (examples above, here, and in caveats section below). For example, areas with a known history of human-assisted invasion by one or more alien species provide opportunities to quantify taphonomic inertia and evaluate the dynamics of DA accumulation in general, with the alien species serving as tracers of live-dead equilibration and postmortem transport (Poirier et al. 2009, 2010). How long does it take for a new ecological dominant to actually dominate the DA, that is, to go through the taphonomic arc of being ‘live-only’ to
'more abundant alive than dead' to 'top taxon both alive and dead'? How many attempted invasions by a species does it take for always-rare-living species to register as such in the DA? Erthal et al. (2011), for example, found that an alien species dominated both the LA and the DA of a Brazilian river system, so that fully buried Holocene fossil assemblages were the best recourse for reconstructing the pre-impact community. Short scales of time-

**FIG. 4.** If live-dead discordance in a dataset genuinely arises from anthropogenic changes in the local living community, as suggested by meta-analyses (Fig. 3), then the species responsible for that mismatch should 'fit', i.e. make ecological sense, with the postulated stress. A, Molluscan datasets from estuaries and lagoons with no or only moderate anthropogenic eutrophication (scored as AE0 and AE1, Table 2) show quite good live-dead agreement in the proportional abundance of species that prefer seagrass meadows (regression line has slope c. 1 and y-intercept of c. 0), whereas these same species are disproportionately abundant in death assemblages, by c. 20%-points on average, relative to local living assemblages in areas with severe eutrophication (AE2; dotted regression line; adapted from Kidwell 2007). B, death assemblages from open shelf gravel habitats subject to chronic bottom-trawling (BT2) contain on average a significantly higher proportional richness of mollusc species sensitive to physical disturbance than do local living assemblages, indicating local extirpation of c. 30% of these species: the dashed regression line for BT2 datasets is parallel to but c. 30%-points higher than the c. 1:1 live-dead agreement observed in datasets from untrawled seafloors (BT0). Death assemblages from areas with intermediate trawling intensity (BT1) mostly fall between these two extremes (adapted from Kidwell 2009).

**FIG. 5.** Species richness (displayed as a rarefaction curve with 95 per cent confidence intervals, against number of individuals sampled) observed in modern small-mammal death assemblages collected under raptor roosts ('historic DA' collected from roost floor at start of study; solid line with dark grey field), compared with the live-trapped richness of species over the next 3 years ('current LA'; dashed line in white field) and skeletal remains produced at the roost during that same interval ('current dead assemblage'; solid black line in white field). Owls sample the richness of their home range very efficiently (compare current living and death assemblages) in both a relatively natural area (A) and a human-impacted study area (B; military activity, invasive cheatgrass). However, in the impacted area (B), the historic DA is far richer than the current living and death assemblages, suggesting that the living assemblage has historically been more diverse. C, Data from live trapping in the now-impacted region of B show that the small-mammal community was in fact much richer during the 1950s ('historical LA'; dashed line, light grey field), closely matching the richness of the historical DA. Time-averaged DAs thus provide good access to the richness of a study area and can detect recent community change when compared with the present-day LA. From Terry (2010b).
averaging in freshwater assemblages probably contribute to this live-dead equilibration. The known local extirpation of a key or particularly charismatic species also provides taphonomic opportunities. ‘Dead-only’ occurrences of a species are generally taken as prima facie evidence of decline in population size and/or geographical range (e.g. Nielsen et al. 2008; Simões et al. 2009; for vertebrate examples, see Lyman 1996; Lyman and Cannon 2004; Hardly and Barnosky 2009; Koch et al. 2009; Louys 2012). Knowledge of the time of extirpation or population crash thus provides a benchmark for assessing absolute or relative rates of postmortem taphonomic fading or burial on decadal and centennial scales that are far beyond the reach of ordinary manipulative experiments.

Caveats for historical ecology

The meta-analysis by Kidwell (2007) found that a strong live-dead mismatch between taxonomic composition and species’ rank abundances could be attributed to anthropogenic modification of the ecosystem and specifically anthropogenic eutrophication. Meta-analysis did not find the converse, that is, that live-dead discordance exists in all settings where stress from cultural nutrients is present. This ‘failure’ to recognize human stresses makes live-dead discordance a conservative method (it does not lead to false positives) and can have several explanations, as discussed by Kidwell (2007): (1) human impacts are so long-standing that the DA has equilibrated; this is espe-
cially likely if scales of time-averaging are relatively short, for example, are limited by shell durability; (2) sedimentation is so high that pre-impact DAs are already fully buried; this is especially likely in settings that are natural sediment sinks, for example, lagoons and harbours (Weber and Zuschin 2013); and/or (3) the composition of the postimpact community lies within the range of natural variability of the pre-impact community, that is, human activities did not have significant ecological consequences. In before–after studies, biologists commonly fail to detect significant change in benthic communities receiving cultural nutrients when the system is naturally meso- or eutrophic, much as bottom trawling only has significant consequences on seafloors that are not naturally disturbed. In general, areas with relatively mild, infrequent or spatially diffuse stress exhibit less ecological degradation.

Note that none of explanations (1–3) signifies an inability (i.e. a true failure) of DAs to capture information from local LAs. In situations (1) and (2), the DA that carries the pre-impact record is presumably down-core, and in situation (3), the DA correctly recognizes an absence of ecological strain despite human stress. The meta-analysis also found that DA sensitivity (memory) varies with the mesh size used to sieve specimens from seabed samples: clearer signals of baseline shift are captured by molluscan live-dead data sets collected using coarse mesh (>1 mm) than fine mesh (≤1 mm) because they focus exclusively on adult individuals (Kidwell 2007, 2008). Use of an overly coarse mesh (>5 mm), on the other hand, tends to decrease sample size, thus reducing statistical power and can omit key species or functional groups from analysis, such as small-bodied chemosymbiotic species and opportunists that arrive or bloom with cultural nutrients.

Such factors might explain the relatively good live-dead agreement observed in soft seafloor molluscan assemblages in areas with significant cultural histories. Note also that between-site variability in live-dead agreement within an area can be large. For example, in Figure 6, note the scatter 'after invasion’ site-level data points as opposed to the tight clustering of ‘before invasion’ data points; a comparable contrast in dispersion exists between AE ≥1 and AE0 habitat-level data points in the original meta-analysis of Figure 3. This between-site variability of live-dead values in impacted areas underscores the importance of sampling multiple sites in any study area. DAs are not so reliable that single grabs or cores are going to answer all questions, any more than single LA samples satisfy professionals biologist.

Studies in areas with important edible or otherwise commercial shell producers such as oysters and cockles demand caution: harvesting and other seafloor mining might have removed significant quantities of dead individuals and deposited them elsewhere (middens, heaps by canneries, dumped later on seafloor to seed new beds, converted to lime or aggregate), with potential to alter the dynamics of DA formation at a site or regional scale. See Powell et al. (2006, 2012), who took advantage of these factors to evaluate the dynamics of shell preservation. Intertidal and beach settings also require special caution given (1) the potential for natural erosion of Pleistocene and older specimens, commonly of locally extant species; (2) beach replenishment efforts, which commonly draw on offshore sands and require grey-literature analysis to document (discussed by Wehmler et al. 1995); and (3) the undocumented depletion of both LAs and DAs by public access for bait collecting, amateur shell collecting and interruption of long-shore drift, which all tend to be underestimated. Human activities can thus alter the composition of the DA as well as the LA, and some study areas may prove to be simply too altered on both sides of the live-dead equation for reasonable taphonomic and ecological analysis.

Using molluscan LA data from long-term biomonitoring efforts, Adam Tomasovych and I are finding strong signals in bivalve DAs in proximity to wastewater treatment plants, despite a steep high-energy shelf and fine-mesh (1 mm) data (Kidwell and Tomasovych 2011). Natural rates of sediment accumulation on the southern California shelf are lower than in estuaries, species diversities are quite high (increasing the power of standard metrics), there is no history of bottom trawling or shell removal, and nutrient stress has been both point sourced and historically acute within the twentieth century (peak sediment and nutrient loads in the 1970s, before enactment of the Clean Water Act; Stein and Cadien 2009).

**Using cores**

So far, few workers have attempted to bring sedimentary cores into molluscan-based analyses of human impacts. Most reviewers of grant proposals are convinced that too few shells will be recovered to support analysis and/or that bioturbation will have blurred temporal trends from the critical last century or so. However, using Pb210 and a known history of zinc pollution to establish an independent chronology, Edgar and Samson (2004; Edgar et al. 2005) found that the upper metre of push cores from New Zealand estuaries, acquired by divers, detected both: (1) a significant decline in the abundance of shells of a commercial scallop since the onset of trawling, consistent with known collapse of the fishery from overexploitation; and (2) an unappreciated and rather astounding fifty percent decrease in the sample-size-standardized richness of other molluscs. They were also able to detect a 20-year time lag in the onset of decline in an estuary where
commercial operations in fact began later, demonstrating the temporal and spatial resolution of ecological history possible using skeletal records (and see molluscan core-based detection of water management history in Florida Bay by Brewster-Wingard and Ishman 1999). In contrast, Armenteros et al. (2012) did not detect any significant up-section change in the composition of molluscan faunas in their <50-cm-long cores from the Gulf of Bat-abanó, nor any spatial patterns among cores, despite an array of suspected human stressors. Various possible explanations exist: (1) stress too slight and (2) core penetration too shallow, given deep bioturbation in the tropics. Their study, nonetheless, demonstrated, as do the other studies, that reasonable numbers of shells can be acquired per core increment, underscoring the general feasibility of mollusc-based historical analysis.

**ESTIMATING DIVERSITY**

Virtually any descriptive or analytical metric for living communities that is based on counts and/or body sizes of individuals can be applied to DAs; it is thus possible to evaluate the ability of DAs to preserve information on diversity and thus build long-term baselines. Examples include the following: live-dead agreement in species richness (Is the count of species occurring dead the same as, significantly greater than, or less than the count of species occurring alive?); taxonomic similarity (Do the living and dead species lists share many or few species in common?); species rank-abundance correlation (Are the species that are top-ranked by abundance in the LA also top-ranked in the DA or are they instead randomly distributed through an abundance-ranked list or perhaps exclusively rare?); and species proportional abundances or dominance (The same species may be top-ranked both live and dead but constitute 90 per cent of individuals in one assemblage and only 40 per cent in the other). Live-dead agreement can be assessed at any spatial scale (a single site, a habitat, a multi-habitat region, a province, the globe) and can be assessed across scales (e.g. the ability of the DA at a single site to capture or otherwise proxy for the regional species pool). We can also assess live-dead agreement in spatial variability (Are the species that are top-ranked by abundance in the LA also top-ranked in the DA or are they instead randomly distributed through an abundance-ranked list or perhaps exclusively rare?); and species proportional abundances or dominance (The same species may be top-ranked both live and dead but constitute 90 per cent of individuals in one assemblage and only 40 per cent in the other). Live-dead agreement can be assessed at any spatial scale (a single site, a habitat, a multi-habitat region, a province, the globe) and can be assessed across scales (e.g. the ability of the DA at a single site to capture or otherwise proxy for the regional species pool). We can also assess live-dead agreement in spatial variability (e.g. live-dead agreement in how species composition varies along an environmental gradient; live-dead agreement in the magnitude of taxonomic turnover between sites (beta-diversity); and live-dead agreement in the comparability of spatial and/or temporal variability (dispersion) among replicate samples). The detection of spatial and temporal variability in composition is as important as diversity itself in the search for pre-impact baselines. ‘Historical range of variability’ has also become a standard benchmark for scoring ecological strain (Morgan et al. 1994): Has the system been pushed outside its natural bounds and on what time frame is the present condition ‘unprecedented’?

Taphonomically, the quantitative utility of DAs depends on several things. These include their temporal resolution (see section on ‘Death assemblages’ above; time-averaging can be useful as in the teenager’s bedroom example), their spatial resolution (coarsens with time-averaging, even without postmortem transport; see below) and the extent to which DAs are only time-averaged and not also biased (see section on ‘Recognizing anthropogenic change’ above for differentiation of these factors). If DAs are only time-averaged, then we would only need to adjust for the coarser temporal and/or spatial scale of DA data relative to LA data and would not need to correct or otherwise compensate for taphonomic bias per se.

The critical question for basic taphonomic research thus becomes: in study areas with minimal human influence, how much of the observed live-dead difference in a community-level attribute can be explained simply in terms of random and other natural variability in the composition of the LA, that is, from within-habitat time-averaging (Factor 4 above)? Taphonomic bias from interspecies differences in shell preservation, production and postmortem transport (Factors 1–3 above) only needs to be invoked if the magnitude of observed live-dead differences exceeds that expected from the time-averaging of natural LA variability. Multiple authors (Pandolfi and Minchin 1996; Edinger et al. 2001; Zuschin and Oliver 2003; Tomasovych and Rothfus 2005) have evaluated this in single regions using spatially replicate samples of LAs to estimate the ‘cloud’ of LA variability in species’ relative abundances. This approach assumes that LA variation in space is comparable with LA variation in time. Peterson (1976, 1977) was the first to simulate the effects of time-averaging by pooling temporally replicate samples of LAs, focusing on the consequences for richness. Adam Tomasovych and I have taken this same basic approach to parameterize dynamic models, which assess the effects of time-averaging on a large number of biological attributes over time scales of several years to millennia (partly summarised in Table 4; see Kidwell and Tomasovych 2013). Spatially replicate point-scale samples and several time-series are used to estimate molluscan LA variability in regions. The death assemblages expected to arise from time-averaging are then compared to observed DAs in those same regions.

There are several key, first-order answers from genuinely natural systems, using molluscs in tropical and temperate soft sediments as the test system. The first, from meta-analytical synthesis of field studies, is that live-dead agreement in molluscan species composition (taxonomic similarity) and rank-abundance is in general quite good. In Figure 3, all except one pristine data set fall in the upper right corner (black icons; similar clustering is apparent in ‘before’ data points in the cross-plot of Fig. 6). Important
functional groups tested so far (seagrass dwellers and disturbance-sensitive species; Fig. 4) also show on average quite good (c. 1:1) live-dead agreement in proportional abundance and proportional richness, respectively, even though each group encompasses a range of body sizes, life habits and shell mineralogy. The overall strong signal (both in species’ presence–absence and relative abundance) is remarkably strong, especially given that each data point represents comparison of a temporally coarse DA with an instantaneous LA. Summing even a short time series of LA samples rapidly improves live-dead agreement in richness (Peterson 1977) and to some degree rank-abundance information, as in Kidwell (2001). Interestingly, land mammals return similarly strong fidelity in these same metrics (see Figs 2, 5, 7; Western and Behrensmeyer 2009; Terry 2010a, b; Miller 2011). In contrast, tropical-reef molluscs return good live-dead agreement if rare species are excluded and mixed to disappointing agreement if they are included; many dead specimens are lost due to postmortem overgrowth, transport by hermit crabs and disappearance into sediment-filled crevices, so that many rare species occur live-only (Zuschin et al. 2000; Zuschin and Oliver 2003; Zuschin and Stachowitsch 2007). Live-dead agreement among corals is reduced by disparate colony lifespans and breakdown rates, loss of fine morphologic detail needed for species-level identification and overgrowth (Edinger et al. 2001; Greenstein 2007).

A second key finding, from modelling, is that most of the molluscan live-dead differences observed in pristine settings can be explained entirely or largely by time-averaging (Table 4; summarizes the results of Tomasovych and Kidwell 2009a, b, 2010a, b, 2011). Time-averaging has predictable effects of the appropriate kind on the biological attributes that DAs are commonly observed to overestimate (e.g. site-level richness, evenness, proportion of rare species in a community) or underestimate (e.g. dominance by a single species, beta-diversity), and has little effect on attributes where live-dead agreement is observed to be fairly high (e.g. taxonomic similarity and log-transformed abundance, which approximates rank-abundance; see Olszewski 2012a on the modelled robustness of rank abundance). Our work shows that: (1) these effects on richness, etc., exist even when the DA is subsampled to match the LA and so are not artefacts of disparate sample sizes; (2) most effects of time-averaging are stronger (more pronounced) at the point scale than at the habitat scale, thus knowing the absolute scale of time-averaging in a DA is less important if you are comparing diversity among habitats rather than among single sites or beds; and (3) these effects can arise within the first few decades to centuries of time-averaging. Although time-averaging tends to damp (reduce) between-site and between-habitat differences in species composition (beta-diversity), DAs still detect spatial and environmental gradients in composition when such gradients exist among LAs and usually detect local species optima (segments of the gradient where a species has its highest abundance or occurrence alive). Much of beta-diversity is ‘transferred’ to the local (alpha) scale (i.e. the DA at a point has a diversity that is comparable with summing multiple points), thus time-averaged data from a single site or small series of sites provide a reasonably accurate estimate of the number and identity of species in the regional pool and, more surprisingly, their relative abundances, which is logistically extremely difficult to achieve via live collections alone. Note that we operationally defined regional diversity as the total species list in the live-dead study. Finally, as anticipated in part by conceptual models (Peterson 1977; Fürsich and Aberhan 1990), time-averaging alters the abundance structure of the assemblage. Random switching in the identity of the most abundant species in the LA over time decreases each of their proportional abundances in the time-averaged DA, thus reducing the initial steepness of the time-averaged rank-abundance distribution: time-averaging reveals multiple co-dominant species, with none of them constituting as large a proportion of individuals as when alive (dominance is reduced). The proportion of rare (single-individual) species in the assemblage also increases with progressive time-averaging because of the larger window for random colonization events by species that are everywhere rare. The rank-abundance distributions of time-averaged DAs thus tend to be flatter and have longer tails of rare species than any of the LAs they are drawn from. This flattening of the abundance structure with time-averaging contributes to the high yield of species from relatively small samples of DAs, making them an efficient means of estimating regional richness and species composition, albeit at a coarser temporal scale than live collections.

Development of a practical protocol

The good news is that DAs from unimpacted areas are, to a first-order, simply temporally and spatially coarse-resolution versions of LAs, with minor bias and/or condensation of natural environmental changes. Such scaling effects on species richness are familiar, at least in broad principle, to both applied and academic ecologists (e.g. species–area effect). Molluscan DAs have good resolution of habitats on length scales of tens to hundreds of metres (facies, biotopes) and a temporal window of centuries to millennia or more, depending on the setting, although most shells derive from the most recent few decades to centuries. Unfortunately, the gold standard of spatial resolution for most ecological analysis is a site (point, quadrat, plot), even for regional surveys, and the usual standard for temporal resolution is a single season, even for multi-year studies. DAs thus represent a significant...
**TABLE 4.** Fidelity of molluscan death assemblages (DAs) to living assemblages (LAs) for selected biological attributes, based on statistical synthesis of multiple data sets from subtidal seabeds in tropical and temperate areas.

<table>
<thead>
<tr>
<th>Biological attribute</th>
<th>Observed death assemblage (habitat scale)</th>
<th>Source</th>
<th>Variance and mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw species richness (presence-absence)</td>
<td>Larger count of species on average (median raw DL ratio of species = 2.6), owing to larger number of dead than living individuals in most sediment samples (median raw DL ratio of individuals = 8)</td>
<td>85 mostly pristine data sets (Kidwell 2002a)</td>
<td>Raw DL ratios of individuals and species are smaller in cold temperate and boreal provinces based on preliminary data; suggests less time-averaging (lower preservation rates), lower productivity and/or less beta-diversity in molluscan LAs (unpub. obs.)</td>
</tr>
<tr>
<td>Sample-size-standardized species richness</td>
<td>Larger count of species on average regardless of standardization method; median DL ratio of species richness is 1.22 based on subsampling with replacement, 1.52 based on subsampling without replacement (40 data sets); 1.3 at habitat scale and 1.8 at point scale (using 31 data sets)</td>
<td>85 mostly pristine data sets (Kidwell 2002a); 40 exclusively pristine data sets (Tomasovych and Kidwell 2010a)</td>
<td>Modelling shows that time-averaging alone can increase DA richness by 1.6 (2.1 at point scale): richness always increases with time-averaging, even holding sample size constant and with no change in environmental conditions, and the increase is always greater at point than at habitat scales (Tomasovych and Kidwell 2010a)</td>
</tr>
<tr>
<td>Shape of rank-abundance distribution (RAD)</td>
<td>RADs of DAs are on average flatter (lower dominance by a single species) and have longer tails of rare species (represented by singleton or doubleton individuals) than LAs, resulting in greater average richness and evenness in size-standardized samples. See same effect at point scale</td>
<td>31 data sets (Tomasovych and Kidwell 2010a); anticipated by conceptual models of Fürsich and Aberhan (1990)</td>
<td>Modelling of within-habitat metacommunity dynamics shows that (1) stochastic switching in the identity of the most abundant species in the LA over time (a few decades to centuries) decreases their proportional abundances in the DA, thereby reducing the initial steepness of the RAD, and (2) rare metacommunity species are temporally short lived in local communities, fostering accumulation of many rare species in the DA</td>
</tr>
<tr>
<td>Estimated size of regional species pool (gamma richness)</td>
<td>The DA sampled at a point in space or time (DA alpha richness) captures more of the total (gamma) richness of a region or time series than does the LA sampled at a point (DA alpha = median 0.8 of DA gamma versus LA alpha = median 0.6 of LA gamma)</td>
<td>9 regional data sets, each including several habitats, and 2 time series in a single habitat (Tomasovych and Kidwell 2009a)</td>
<td>Modelling indicates that DA alpha richness increases rapidly in the initial decades to centuries of time-averaging, owing to chance colonization by patchy and/or ephemeral species; this effect occurs even when samples are size standardized (Tomasovych and Kidwell 2010a)</td>
</tr>
<tr>
<td>Taxonomic composition Univariate similarity indices</td>
<td>Median similarity of DA and LA is 0.90 using Chao’s sample-size-corrected version of the Jaccard index; median 0.92 in estuaries and lagoons, 0.84 on shelves</td>
<td>18 shelf data sets (Kidwell 2008), 27 estuarine data sets (Kidwell 2007)</td>
<td>On hard substrata, DL differences in richness are larger owing to rare species that occur dead-only or live-only; capture of LA presence/absence is otherwise good (Zuschin et al. 2000)</td>
</tr>
<tr>
<td>Biological attribute</td>
<td>Observed death assemblage (habitat scale)</td>
<td>Source</td>
<td>Variance and mechanisms</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------------------------</td>
<td>--------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Multivariate analysis of taxonomic dissimilarity</td>
<td>Using presence–absence data, median DL difference is 0.4 (Jaccard dissimilarity, thus a similarity of 0.6); using relative abundance data, median DL difference is 0.35 (Horn-Morista dissimilarity)</td>
<td>31 data sets (Tomasovych and Kidwell 2010a)</td>
<td>Despite a shift from mean LA composition, the mean point-scale DA composition frequently lies within the cloud of replicate point-scale LA samples; such dissimilarity can be generated largely or entirely by time-averaging of random variability in LAs</td>
</tr>
<tr>
<td>Similarity in species rank-order abundance</td>
<td>Median Spearman’s rank correlation rho of raw DA and LA species lists is 0.58 in estuaries and lagoons (69% of DL correlations are significantly positive) and 0.38 on shelves (61% significantly positive)</td>
<td>18 shelf data sets (Kidwell 2008), 27 estuarine data sets (Kidwell 2007)</td>
<td>Similar results from 85 data sets from mostly pristine areas; rho is higher if exclude juvenile individuals (Kidwell 2001, 2002b). Rho &lt;1.0 in pristine settings arises from within-habitat time-averaging of natural LA variability, natural environmental condensation and taphonomic bias</td>
</tr>
<tr>
<td>Proportional abundances of species</td>
<td>A given species can be more or less abundant in the DA than in the LA (observed DL differences range from −0.51 to +0.47), but the median difference is quite small (−0.001; IQR is 0.02). In each data set, only a few species are responsible for DL differences in proportional abundance; most species are rare both alive and dead</td>
<td>193 species occurrences from 7 pristine habitats (Kidwell and Rothfus 2010); 31 data sets (Tomasovych and Kidwell 2011)</td>
<td>DL differences in proportional abundance are not significantly correlated with either lifespan or adult body size, because neither variable correlates with abundance in LA. Most DL differences arise from within-habitat time-averaging of LA variability; residual ‘unexplained’ DL differences in 25–65% of data sets must owe to taphonomic bias and/or environmental condensation (between-habitat time-averaging)</td>
</tr>
<tr>
<td>Estimating species identities and abundances in the species pool at broader spatial scales</td>
<td>The species composition and rank-abundance distribution of the DA at a point approaches that of the source metacommunity, given time-averaging of LAs on decadal to centennial scales, using neutral and non-neutral, dispersal-limited dynamics of species colonization</td>
<td>Model parameterized using 31 data sets and compared with data in 12 data sets (Tomasovych and Kidwell 2010a); direct tests of 2 time series with known time-averaging (Tomasovych and Kidwell 2010b)</td>
<td>DAs collected at a point (and especially at a habitat) level are an efficient means of generating a regional species list and estimating species' relative abundances at a regional scale, data that are difficult to acquire from live-sampling alone. Can assume that species do not differ in shell preservation and individual lifespan</td>
</tr>
<tr>
<td>Spatial patterns Beta-diversity</td>
<td>DAs show less turnover in species composition among points than do LAs (lower beta-diversity): in 9 of 11 data sets, between-point dissimilarity of DAs is positively correlated with that of LAs but is damped by c. 25% (fewer compositionally distinct communities). DAs thus tend to underestimate the true spatial variability of LAs at both point and habitat scales</td>
<td>9 regional data sets and 2 time series in a single habitat (Tomasovych and Kidwell 2009a); 31 data sets (Tomasovych and Kidwell 2010b)</td>
<td>Modelling shows that reduced beta-diversity can largely be explained by within-habitat time-averaging of natural temporal variability in colonization, without recourse to significant environmental condensation or postmortem spatial mixing. Reduced beta anticipated by early models (Miller and Cummins 1990). Temporal, down-core variability should also be damped (Tomasovych and Kidwell 2010b)</td>
</tr>
<tr>
<td>Biological attribute</td>
<td>Observed death assemblage (habitat scale)</td>
<td>Source</td>
<td>Variance and mechanisms</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Variation in community composition along environmental gradients</td>
<td>DAs detect a significant gradient where LAs detect a gradient in 6 of 7 data sets using presence–absence data (5/7 using proportional abundance data), and the DA gradient is as strong as or stronger than the LA gradient. Environment explains about the same proportion of between-point compositional variation among DAs as it does among LAs.</td>
<td>9 regional data sets (Tomasovych and Kidwell 2009b). And see powerful single-region tests by Miller (1988; Miller et al. 1992; Ferguson and Miller 2007)</td>
<td>Variation along gradients persists despite the potential for spatial mixing and is commonly strengthened by within-habitat time-averaging. The environmental resolution of DAs supports palaeoenvironmental analysis using molluscs as well as palaeoecological discrimination of species-sorting and mass-effects models of meta-community structure.</td>
</tr>
<tr>
<td>Habitat- (facies-) level preferences of species</td>
<td>Weighted meta-analytic mean 73 ± 3% (simple grand mean = 78%) of all individuals in the DA are drawn from species that are documented alive in that habitat (same set of samples)</td>
<td>85 data sets from mostly pristine settings (Kidwell 2002a)</td>
<td>Because LA information is based on a one-time survey and typically yields small numbers of individuals, this test provides a minimum estimate of the true spatial fidelity of DA individuals.</td>
</tr>
<tr>
<td>Variation in single-species abundance along environmental gradients</td>
<td>Species niche optima (positions of maximum living abundance and occupancy along environmental gradients) are detected in 7 of 9 data sets; DAs do not detect optima where none are evident in the LA.</td>
<td>9 regional data sets (Tomasovych and Kidwell 2009b)</td>
<td>Median rank correlation between-species optima in DAs and LAs are significantly positive regardless of data transformation; DAs have less ability to reflect niche breadth (blurred by time-averaging) and underestimate maximum abundance (carrying capacity).</td>
</tr>
</tbody>
</table>

Unless otherwise noted, all study areas had minimal human activities at the time of sampling. Only significant results are reported. The capacity of time-averaging to explain observed live-dead agreement was evaluated by dynamic modelling: LA data are used to produce an ‘expected’ DA assuming time-averaging, but no taphonomic bias, which can then be compared with the observed DA. Point scale = the set of living and dead individuals extracted from sediment taken at a single point or site on the seafloor. Habitat scale = specimens from two or more points within a relatively narrow depth zone, and bottom type are pooled before the living and dead lists are compared. Regional data set = LA and DA data from multiple habitats in a region are available. All analyses summarized here are based on habitat-level data sets and have been corrected for differences in living and dead sample sizes (numbers of individuals) within data sets, unless otherwise noted. DL = DA/LA ratio or DA − LA difference.
coarsening of temporal resolution for most biologists. Biologists (especially applied ecologists, biogeographers and macroecologists) are, nonetheless, increasingly open to the idea of using DAs, given the severe shortage of historical ecological data and the deep roots of some human activities. To paraphrase a paraphrase, if you are going into battle, a blunt battleaxe is better than no weapon at all (Smol 2007).

Winning over biologists to the value of DA data thus entails a substantive discussion of issues of scale, now that concerns with bias are being allayed. Eventually, however, it would also be useful to refine the temporal resolution of DAs, giving us a larger range of temporal scales for cross-comparison with LA data (biologists of course can also analytically coarsen the resolution of their point-scale time-series data, to meet us part way). The most obvious approach is to partition the DA according to the state of shell preservation, that is, to isolate and analyse only the subset of shells that are ‘fresh looking’. This operation should, in principle, eliminate some part of the long tail of the shell-age frequency distribution, as damage is irreversible and some field tests are promising (Yordanova and Hohenegger 2002). It presumes, however, that: (1) damage on a shell accrues with elapsed time, which is certainly true, assuming one can discriminate postmortem damage from ‘premortem’ fouling; only shell interiors should be used, as in Best and Kidwell (2000), to avoid ambiguity; (2) damage accrues at a reasonably constant rate among all shells in a given cohort, which is not necessarily true, given dynamism of the seabed (e.g. Driscoll 1970); and (3) all species in the LA accrue postmortem damage at the same rate, so that subsampling the fresh-looking ones would not bias against some groups, which is unlikely to be true in detail, but perhaps is true enough.

The correlation of shell damage with elapsed time since death (the idea of a taphonomic clock; presumption 2 above) has been tested in several systems, usually for only a single species in an assemblage (Powell and Davies 1990; Flessa et al. 1993; Kowalewski et al. 1994; Wehmiller et al. 1995; Martin et al. 1996; Meldahl et al. 1997; Kidwell et al. 2005; and for brachiopods, Carroll et al. 2003; Krause et al. 2010; see Kidwell 1998 for earlier review). Plots of shell age (usually amino acid dates) against a single aspect of shell condition (e.g. edge rounding or overall shell condition ‘taphonomic grade’) yield either a broad scatter of values, a weak indirect relationship (few truly old shells in pristine condition) or, in a minority of cases, a diffuse but significant positive relationship. Clearly, individual shells can start to acquire damage immediately postmortem or only after a significant delay, depending on postmortem residence at or near the sediment–water interface where many taphonomic processes are most diverse and intense. All authors infer this same underlying dynamic that depends on burial history, which has been supported by time-lapse field experiments. Shells tethered on strings for a few months or years at the sediment–water interface where they are free to undergo natural burial–exhumation cycles acquire varied degrees of damage, in contrast to the consistent levels and styles of damage acquired by shells that are held continuously either below or above the sediment–water interface (dissolution, maceration, bleaching or discolouration characterize the former; intense bioerosion and encrustation characterize the latter; Driscoll 1970; Best 2000; Best et al. 2007).

This erratic acquisition of damage by shells on annual (experimental) to Holocene (racemization) timescales reduces the precision of shell condition as a means of subdividing assemblages into younger and older subsets of shells and contrasts with the more monotonic weathering of bones in terrestrial settings (Behrensmeyer 1978; Cruz 2008; Miller 2011; Behrensmeyer and Miller 2012). However, despite the noisiness of the relationship between damage and shell postmortem age, I am unwilling to give up all hope, especially as we consider using DAs as tools for ecological analysis, ignoring concerns about their value as analogues for deep-time records. Shell colour and luster, for example, consistently emerge as useful variables for distinguishing Holocene shells from shells that have been reworked from older Pleistocene strata (citations above). The unlikely value of these variables for discriminating the relative ages of shells within pre-Quaternary fossil assemblages is thus irrelevant (Kolbe et al. 2011). Moreover, decadal deployment of shells in the subtropical shelf and slope experiments conducted by Powell et al. (2011a) have found trends in discolouration (initially from microbial colonists, leading eventually to bleaching), edge rounding and fine-scale surface texture (chalkiness from maceration and then softening and pitting from dissolution; see references therein for nonmoluscan groups). Ignoring noisy results over the first 2 years of deployment, they found, not unexpectedly, that rates of change in shell condition vary between environments. However, this variability would not in principle undermine using damage to age-subdivide DAs drawn from a single habitat. Powell et al. (2011a, b) found that between-species differences in rates of damage accrual (presumption 3 above) are less strong than between-habitat differences. This, too, is promising and consistent with many taphofacies studies based on undated DAs, although large samples are needed to resolve between-species variation within a habitat (see tests of variation with shell type and life habit by Best and Kidwell 2000; Lockwood and Work 2006). Powell et al. (2011b) also discovered that all shells end up at a similar, highly degraded state on a decadal timescale, despite different initial rates of alteration. That is, on a decadal timescale, shells of different
species within a postmortem cohort should ultimately resemble each other in a time-averaged DA.

We thus need to add taphonomic grading of shells to standard live-dead analysis to conduct a series of post hoc, analytical experiments that (1) are on timescales concordant with the full, centennial and longer scope of time-averaging in DAs, and (2) accommodate short-term burial-exhumation cycles as a steady-state condition in the mixed layer. These analyses should focus on new or archival live-dead data sets from impacted areas, so that there is an independently supported hypothesis of ecological change within the window of time-averaging. Large DA sample sizes are also needed, so that parsing shells into two, three or four taphonomic categories does not hinder statistical power. Does live-dead agreement in fact decrease as one compares the LA to subsets of the DA characterized by increasingly poor shell condition? Do live-dead differences in species composition and ranking among damage-graded sets of shells ‘fit’ with the known history of environmental change or are differences better explained by taphonomic bias (same procedure as in Table 3)? As an alternative for nonimpacted areas, one could test for consistency among damage subsets of the DA in, for example, richness, species composition and evenness, an approach used by Belanger (2011) to assess bias in fossil foraminiferal assemblages. Because live-dead comparison is not required in Belanger’s approach, archived DAs from taphofacies studies could be repurposed.

I suspect that some live-dead data sets produced by malacologists and fishery biologists probably did not count dead specimens below some cut-off in shell condition because live-dead comparison was not their aim. Applying some lower boundary for inclusion is a logical precaution to avoid taxonomic misidentification. Albano and Sabelli (2011), for example, graded the condition of all molluscan DA specimens and then limited their live-dead analysis to the top three preservational categories, motivated to evaluate the reliability of DAs that could be processed by a nonspecialist worker. Similarly, the large-mammal live-dead analysis in Figure 2 used data only from bones in the first two weathering stages to avoid preservational bias towards large-bodied species in the oldest part of the time-averaged DA. This intuitive approach (subdividing DAs into damage categories) could be a powerful means of increasing the temporal resolution (and decreasing the ecological memory) of DA data for historical ecological analysis, depending on the needs of the study. It presumes little between-species variation in rates of damage acquisition, however, or at the least a fairly rapid equilibration of damage levels (as in Powell et al. 2011b), and thus needs the next level of systematic testing as described here. We should thus not be discouraged by the mixed results of earlier tests of the taphonomic clock if our aim is to use DAs as archives of modern ecosystems.

**Modest bias of community-level data**

The capacity of within-habitat time-averaging to explain most or all live-dead disagreement in data sets from unimpacted study areas (Table 4) leaves little need to invoke postmortem bias from differential production, destruction and postmortem transport: summing of natural, within-habitat variability in the composition of LAs is almost always sufficient (Factor 4 in list above). Moreover, despite the clear potential for bias from each of many different processes capable of destroying or otherwise removing shells (e.g. dissolution, bioerosion, transport) and from interspecies differences in intrinsic durability and production rates, none of these biasing factors (alone or in combination) emerges as a primary control on the biological data archived by molluscan DAs, based on global meta-analyses of habitat-scale data sets. This conclusion has been consistent. The initial, informal analysis of 17 live-dead studies by Kidwell and Bosence (1991) concluded that the first-order explanation of observed live-dead differences was simply the inadequacy of most LA data to characterize the living community. Enlarging the database to 85 data sets (Kidwell 2001, 2002b) revealed a mesh-size effect on live-dead agreement (fine-mesh data sets include juveniles, which increase noise) and found no variation in live-dead agreement as a function of sediment grain size and only slight variation with bathymetry, with the poorest agreement on open shelves. Enlarging the database to increase power (101 exclusively subtidal data sets; Kidwell 2007, 2008) revealed human modification of LAs as the first-order control on live-dead agreement in richness, evenness, taxonomic composition and relative abundance, with a subsidiary mesh-size effect. In subsequent work with Adam Tomasovych focusing on approximately 10 relatively large regional data sets, each encompassing multiple habitats from an area with little or no human activities, it has been extremely difficult to find any significant increments of live-dead difference that must be explained by bias per se (Table 4; Tomasovych and Kidwell 2011): habitats with larger live-dead differences are usually also characterized by greater LA variability. Using a new, large, multi-habitat molluscan data set from the northern Adriatic Sea, Weber and Zuschin (2013) have tested and corroborated virtually all of the community-level attributes and mechanisms in Table 4, such as scaling effects and the ability of within-habitat time-averaging of LA variability to explain observed live-dead differences (see comparable results in freshwater molluscs by Tietze and de Francesco 2012).
The absence of strong between-habitat differences in molluscan DA fidelity so far, despite efforts to find it, and our ability to explain the between-habitat differences that exist in terms of LA attributes (higher random variability; anthropogenic shifts), is quite interesting. It seems to contradict the long scientific history of taphonomic experiments on the significant effects of shell types and particular processes, as well as observations that different taphonomic processes are important in different settings, as registered by taphofacies studies. The answer is probably that many taphonomic processes act in each habitat, each with potential to bias the composition of a multispecies assemblage in a different direction. As a result, the net effect on preservation and thus on biological information is effectively neutral, both in a given habitat and comparing between habitats. Even though many individual taphonomic processes vary in intensity with bathymetry, for example, none is so strong as to overwhelm other countervarying processes and trends in shell type. This deduction was forced by the first formal meta-analysis (Kidwell 2001). Further meta-analysis has revealed that human modification of LAs (globally) and trends in the variability of LAs (among unimpacted settings) are the strongest determinants of live-dead agreement (Kidwell 2007; Tomasovych and Kidwell 2009a, 2011).

This failure of taphonomic bias to emerge as a dominant factor in creating live-dead differences is not to say that taphonomic bias does not exist or rise to a problematic level in some data sets. For example, time-averaging alone can reduce live-dead taxonomic similarity by 0.1 or 0.2 units on a scale of 1.0, owing primarily to the larger number of rare species in the DA, many of which occur dead-only. Based on statistical partitioning, we found that, depending on whether presence-absence or proportional abundance data were used, 25–65 per cent of molluscan data sets in unimpacted settings had live-dead mismatch in species composition that could not be fully explained by within-habitat time-averaging; in multivariate space, the centroid of DA composition is shifted too far from the centroid of LA composition (Tomasovych and Kidwell 2010a; Tomasovych and Kidwell 2011). The importance of many dead-only species, each represented by just a few specimens, in creating live-dead disagreement in species richness and composition is a general phenomenon, such as on hard substrata where many rare species also occur live-only (Zuschin et al. 2000). Adam Tomasovych and I are presently investigating the accumulation in DAs of many rare, dead-only species as an epiphenomenon of the two phases of shell disintegration that are implicit in strongly right-skewed, L-shaped shellage frequency distributions.

Live-dead differences that cannot be explained by within-habitat time-averaging alone presumably reflect some combination of: (1) environmental condensation, that is, unsampled nonrandom LA variability such as from seasonal and interannual cycles or habitat migration, and (2) taphonomic bias from interspecies differences in shell production, preservation or postmortem transport. Albano and Sabelli (2011), for example, found that molluscan DAs were as effective as LAs in discriminating coralline algal and seagrass habitats, with good live-dead agreement in species richness and composition, but that species’ proportional abundances and trophic information were altered and probably genuinely biased by interspecies differences in lifespans. Molluscan carnivores and scavengers, with relatively long lifespans, dominated their LAs, whereas short-lived microherbivores dominate their DAs.

New research thus needs to dissect DAs for sources of residual, 'unexplained' live-dead offsets in composition. For modelling purposes, our definition of within-habitat variability (Table 4) has been very stringent. For example, seasonal changes in fauna are considered to be between-habitat variability, because they are driven by environmental change: only truly random demographic variability is considered to be within-habitat in origin. I thus tend to think that the 'live-dead differences unexplained by within-habitat variability' mentioned above are still likely to be scaling effects to some large extent, that is, products of within-habitat time-averaging sensu Kidwell and Bosence (1991), who conceived habitats as steady states of seasonal to interannual variation in salinity, temperature, etc. Larger-scale and directional changes were entailed in between-habitat time-averaging and environmental condensation, for example, habitat migration and ecological succession, including taphonomic feedback.

Any conclusion regarding mechanism of course does not erase the observed live-dead differences in composition, richness, etc. observed in fully natural systems. Those are real features. However, live-dead differences of such magnitude are simply part of our expectations about the coarse scale of time-averaged DAs, which are also understood to consistently damp spatial and temporal variability (Table 4). The question at hand determines whether this magnitude of offset in composition and damping of variability that arise with time-averaging are tolerable (Tomasovych and Kidwell 2010b).

Where does taphonomic bias dominate?

In modern environments, postmortem transport (of all species or of large numbers of one or two species) as well as reworking of significantly older shelly remains is most common in a few, sedimentologically distinctive settings, as summarized in Kidwell and Bosence (1991). If the stratigraphic context of your fossil assemblages suggests a washover fan or proximity to a tidal inlet or erosive
beach, caution is required, as it would also be in gravity-dominated sedimentary systems, although slumps can deliver allochthonous assemblages intact. Hurricanes can produce remarkably little between-habitat mixing of shells on shallow seafloors (Miller et al. 1992), and it is difficult to detect taphonomic smearing of assemblages between habitats even on steep continental shelves (Kidwell 2008). Bias in proportional abundance is most likely where one or two of the most abundant species in the LA have exceptionally low durability, especially shells that are extremely thin and/or dominated by a high-organic microstructure, and/or have exceptionally short lifespans (like some gastropods; Albano and Sabelli 2011). Small original sample size and severe analytical truncation of a data set can magnify the effects of such bias.

Deductively, however, a preservational weakness (biasing factor) should only modify species rank-abundance and abundance-based diversity measures if the weakness itself is correlated with abundance. For example, in our meta-analytical test for bias in species proportional abundances based on variation in lifespan, we found no consistent bias favouring short-lived, high-turnover species, contrary to intuition, and realized this is because lifespan varied randomly along the spectrum of species abundances in LAs (Kidwell and Rothlis 2010). If species abundances had been uniform (unlikely), or if the longest lifespans had been concentrated among the most abundant species (so that rare species in the LA became dominant in the DA), then variation in lifespan might create an ecologically misleading DA. The same principle is arguably true for other biasing factors. For example, small-bodied molluscs are less likely to be sampled as dead or fossil specimens (Valentine 1989; Cooper et al. 2006), but body size does not bias molluscan abundance data because of a weak relationship of size and abundance; both large- and small-bodied species can be rare or abundant. In contrast, body-size bias is quite important in land mammal and bird assemblages because large-bodied species are more consistently rare (Behrensmeier and Boaz 1980; Behrensmeier et al. 2003; Cruz 2008).

This overarching finding from meta-analysis and modelling (that the richness and composition of molluscan DAs overwhelmingly reflect the effects of time-averaging rather than postmortem bias) reveals the broader truth and robust logic of John Warme’s conclusions in his classic analysis of intertidal assemblages of Mugu Lagoon, California (Warme 1969). Contrary to expectations of strong bias from postmortem transport of dead shells across the flats and by a tidal channel, Warme (1969) found that live-dead differences in species occurrences and abundances were more readily explained (1) by the time-averaged nature of the DAs, which captured rare and patchy species not sampled alive during a one-time survey, and (2) by lateral shifting in habitats within the window of time-averaging. Shells from each successive habitat or community occupying a site enter the time-averaged DA, which is environmentally condensed but fundamentally true to the former existence of living populations at that site (‘faunal condensation’ sensu Fürsch 1978). Warme (1969) rejected bias from strong interspecies differences in preservation and postmortem transport as explanations: the most abundant ‘dead-only’ species did not have particularly robust shells, and small thin-shelled species in high-energy environments were not overly abundant as dead shells in nearby, low-energy ‘sink’ environments compared with their living populations there.

CONCLUSIONS AND FRONTIERS

Intense taphonomic analysis of modern DAs over the last 20 years, which has expanded to include meta-analysis and modelling, provides many new insights into the timescales and controls on the time-averaged, postmortem accumulation of skeletal remains on ordinary, well-oxygenated seafloors and landscapes and the implications for ecological information in multi-species assemblages. These findings are relevant to palaeoecological analysis in general, even though the potential for understanding present-day ecosystems has been stressed here. DAs have strong value for conservation biology and ecology as alternatives and complements to conventional high-resolution LA data, providing information on temporal and spatial scales where living organisms are impossible, logistically difficult or, in the case of rare and endangered species, unethical to sample. DAs also provide insights into anthropogenic changes in local systems and provide data on larger-scale, macroecological aspects of marine and terrestrial systems, such as regional distributions of body sizes, geographical range sizes and biomasses within food webs, all of which appear to be changing under human pressures (Tittensor et al. 2009; Fisher et al. 2010).

As taphonomists, directing actualistic effort towards conservation biology, ecology and environmental management forges new intellectual links. With time, it should create broader opportunities for research, employment and funding beyond geology and palaeontology programmes (e.g. to geography, landscape ecology, wildlife management) and beyond earth science agencies (e.g. wastewater, toxicology, coastal resources, national parks and reserves, environmental restoration, rapid biodiversity assessment). It also enlarges the aims of actualism (namely, towards the present and future rather than only the deep past) and recalibrates our judgment of ‘appropriate’ field areas. For example, rather than downplaying the human footprint or categorically rejecting impacted
areas, we should embrace local histories of human modification as unnatural manipulative experiments. Impacted areas are far more likely to have a well-documented environmental history and long-term biomonitoring data than their pristine counterparts. Unfortunately, from a societal perspective, the world is awash with regions where basic taphonomic variables have undergone human modification. These activities have created novel combinations and gradients of biological diversity, productivity, physical disturbance, sediment accumulation, bioturbation and toxicity that can be played off against natural systems, such as survive in other regions or, palaeontologically, down-core.

Many taphonomic challenges remain in the application of DAs and fully buried Holocene and older fossil assemblages to conservation biology, academic ecology and environmental management. Some have already been mentioned and others are only implicit, for example:

1. The overarching need to evaluate animal groups other than marine molluscs and land mammals. A substantial body of work, generated by a variety of authors, is also building for terrestrial and freshwater molluscs (Briggs et al. 1990; Cummins 1994; Rundell and Cowie 2004; Martello et al. 2006; Nielsen et al. 2008; Sólymos et al. 2009; Erthal et al. 2011; Yanes et al. 2011; Tietze and de Francesco 2012; Yanes 2012), but our understanding of corals and other nonmolluscan reef groups, marine vertebrates and birds is still nascent compared with their importance in conservation and the fossil record.

2. The question of possible down-core, postburial deterioration of temporal resolution and fidelity.

3. The question of latitudinal gradients in the nature of DAs. Boreal settings have received scant taphonomic attention for any animal group, and quantitative ecological information is also scarce there, placing a premium on developing DAs as data sources. Time-averaging of marine molluscs probably declines due to colder, undersaturated waters even at shelf depths, and shells are younger and subjected to erosion by ice. In contrast, time-averaging of mammalian bones appears to be higher in high latitudes due to lower microbial activity and less exposure to UV and scavengers (Miller 2012b).

4. Extracting the next level of taphonomic and historical ecologic information from archived DAs and from newly launched live-dead studies. For example, can temporal resolution be refined analytically, such as by using damage states or parsing intershell variability in age (Edinger et al. 2007; Simonson et al. 2012)? How reliable is biomass and body-size information? What are the consequences of time-averaging for sclerochronological, isotopic and other proxy data carried by skeletal remains?

Not least among the challenges facing the young field of conservation palaeobiology is overcoming resistance among biologists against drawing ecological inferences from nonliving materials and without manipulative experiments. Establishing the strengths and weaknesses of DAs under diverse conditions is thus critical. Equally important, however, will be worked examples of novel ecological insights produced by taphonomically astute analysis of palaeontological data. Such reports need to target journals outside of academic palaeontology and taphonomy to reach biologists and managers and preferably should reflect some degree of collaboration with them. Productive interaction with policymakers and the public also requires both a new vocabulary and a deeper mutual understanding of methods, goals and constraints (for excellent commentaries see Flessa 2009; Jackson and Hobbs 2009; Jackson et al. 2009; Hughes et al. 2010; Jackson et al. 2011; Jackson 2012). Palaeontologists have a rich set of roles to play as we meet the future.

Acknowledgements. Many thanks to Chicago colleagues for stimulating discussions about modern systems over the last decade, especially primary collaborator Adam Tomasovych along with Mairi Best, Yael Edelman-Furstenberg, Josh Miller, Tom Rothfus, Becca Terry, Kristen Jenkins Voorhies, others in the Death at Noon group and, for all of this plus helpful editing, David Jablonski. The manuscript was also improved by comments from G. P. Dietl, K. W. Flessa, P. J. Orr, S. Thomas, A. Tomasovych and K. J. Voorhies. Grateful thanks also to the many authors who have granted access to raw data sets for meta-analysis and to colleagues whose concern with ecosystem health and policy have motivated new research directions: K. W. Flessa, J. B. C. Jackson, S. T. Jackson and an extraordinary community of ben-thic biologists in southern California agencies. Research supported by NSF EAR-345897 and EAR-1124189 and by NOAA’s ‘Urban Oceans’ SeaGrant program administered by the University of Southern California.

Editor. Patrick Orr

REFERENCES


CHIBA, T. and SATO, S. 2013. Invasion of Lagopecula puchella (Gastropoda: Naticidae) and predator–prey interactions with bivalves on the Tona coast, Miyagi Prefecture, northern Japan. Biological Invasions. Published online September 2012, 15, 587–598.


886b. The size frequency distribution in paleoecology: effects of taphonomic processes during formation of molluscan death assemblages in Texas Bays. Palaeontol-ogy 29, 495–518.


PUIG, P., CANALS, M., COMPANY, J. B., MARTIN, J., PALAEONTOLOGY, VOLUME 56, 520

PRESTON, F. W. 1960. Time and space and the variation of species diversity in space and time.-


SADLER, P. M. 1993. Models of time-averaging as a matura-


Sadler, P. M. 1993. Models of time-averaging as a matura-


—-Stanton, R. J. Jr, Powell, E. N. and Cummins, H. 1986. Time averaging, taphonomy and their


