Taxonomic sufficiency in a live–dead agreement study in a tropical setting

Paolo G. Albano a,*, Adam Tomašových b, Michael Stachowitsch c, Martin Zuschin a

a Department of Paleontology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria
b Earth Science Institute, Slovak Academy of Sciences, Dúbravská cesta 9, 84005 Bratislava, Slovak Republic
c Department of Limnology and Bio-Oceanography, Center of Ecology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria

Abstract

Community studies in paleontological research often rely on identification at taxonomic levels higher than species (mainly genus or family). Several studies have been conducted on paleo- and modern communities to identify the higher level of taxonomic identification that still depicts ecological patterns: genus- and, to a lesser extent, family-level identification are frequently sufficient. The use of higher taxonomic levels has not yet been explored in the context of studies comparing living and death assemblages (so-called “live–dead agreement studies”), notwithstanding their interest to quantify the fidelity of the fossil record and for environmental assessment. We conducted such exploration in a highly diverse tropical marine setting, targeting shelled molluscs. Our results suggest that the common practice of genus-level identification of paleontological samples allows for a proper reconstruction of the original biological community (e.g., variation in richness, ecological or taxonomic similarity) at the species level because (1) fidelity at the species and genus levels is very similar and (2) genera are sufficient to characterize between-habitat differences in composition and diversity. Live–dead agreement becomes even better at family and higher taxonomic levels for some metrics, but between-habitat differences in composition become weaker above the family level. However, at the genus and family levels, between-habitat differences are equally strong as at the species level. Genus-level identification may provide more robust results when one of the assemblages is dominated by a single species, because differences in abundance can be compensated by co-generic species. Moreover, in death and fossil assemblages, diagnostic characters get lost easily because of taphonomic processes such as abrasion, dissolution, and fragmentation; nonetheless, genus- and family-level identifications can still be reliable because the most conspicuous diagnostic characters which characterize higher taxa are more persistent after the death of the mollusc.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Community ecology studies in paleontological research often rely on identification at taxonomic levels higher than species (mainly genus or family) because (1) identification to generic level tends to be more robust when preservation of species-level morphological characters is variable or not optimal, or (2) genera are geologically long-lived and thus allow tracing community types through time (even when species identity changes) (e.g., Patzkowsky, 1995; Holland et al., 2001; Olszewski & Patzkowsky, 2001; Scarpone & Kowalewski, 2004; Redman et al., 2007). This approach has been tested: the differences in patterns between the species and the genus levels tend to have weak effects on paleoecological interpretations (Pandolfi, 2001), while family-level identification produced results similar to genus-level identification in only approximately 50% of cases (Forcino et al., 2012).

The fidelity of ecological information that can be preserved in fossil assemblages can be quantified by comparing living (LAs) and death assemblages (DAs), i.e. so-called “live–dead (LD) agreement studies”. This approach assumes that DAs are reasonable proxies for fully buried fossil assemblages (Kidwell, 2013). The use of supraspecific taxa has not been explored so far in this context. How does ecological or taxonomic similarity or variation in richness of living assemblages transfer to death assemblages at higher taxonomic levels? And, conversely, if the analysis of death and fossil assemblages is conducted at supraspecific levels, is it possible to assess the extent to which the results faithfully represent the original species-level biological community?

Furthermore, recent evidence suggests that average LD fidelity is significantly poorer in areas of human impact (Kidwell, 2007), opening interesting perspectives for the application of LD agreement studies for environmental assessment and identification of ecological baselines (Kidwell, 2009; Kidwell & Tomašových, 2013; Weber & Zuschin, 2012; Korpany & Kelley, 2014; Leshno et al., 2015; Zuschin & Ebner, 2015;
Albano et al., 2015b; Negri et al., 2015). In parallel with studies that addressed the comparability between species and higher taxa in fossil assemblages, multiple ecological studies tested whether higher taxa can detect pollution gradients in living assemblages. Ellis (1985) formalized the concept of "taxonomic sufficiency": "in any project, organisms must be identified to a level (species, genera, family, etc.) which balances the need to indicate the biology (including for example such matters as diversity) of organisms present with accuracy in making the identifications". Numerous studies reported that higher taxonomic levels tend to be sufficient to detect natural spatial patterns (e.g., Dethier & Schoch, 2006; Terlizzi et al., 2008; Beblicka et al., 2009) and disturbance to ecosystems due to anthropogenic pressures in marine ecosystems (e.g., Warwick, 1988a, 1988b; Ferraro & Cole, 1990, 1992, 1995; Vanderklift et al., 1996; Olsgard et al., 1998; Terlizzi et al., 2008). Similar studies assessed sufficiency in freshwater (e.g., Wright et al., 1995; Heino & Soininen, 2007; Jiang et al., 2013; Müller et al., 2013) and terrestrial habitats (e.g., Groom et al., 2010; Rosser & Eggleton, 2012; Timms et al., 2013). In most cases, family-level identification was sufficient, although results may vary among phyla (Heino & Soininen, 2007; Beblicka et al., 2009).

A high correlation is expected between species-level and supraspecific taxon indices or ordinations in datasets with low species/higher taxa (S/T) ratios. Therefore, datasets, or taxa, dominated by monotypic higher taxa are likely to produce a high level of taxonomic sufficiency (e.g., Olszewski & Patzkowsky, 2001). However, these ratios will tend to increase with the spatial and temporal extent of datasets and will be especially high in the tropics, where rates of evolutionary diversification are high (Krug et al., 2008). The ability of higher taxa to capture compositional and environmental gradients will also depend on whether species belonging to higher taxa respond to gradients similarly (ecological conservatism) and on their taphonomic similarity.

We test the hypothesis that within-phylum live–dead agreement follows the general patterns already described for paleocommunities and modern living assemblages, i.e., there is a good correlation in community patterns based on species-level and lower supraspecific taxa. Mismatch between LAs and DAs was assessed by measuring similarity in taxonomic composition, rank–order agreement in relative abundance of taxa, taxonomic richness, evenness, and by conducting multivariate analysis. Our case study is a highly diverse subtidal molluscan assemblage in the Persian (Arabian) Gulf. This dataset is characterized by high species richness (ca. 350 species) and by one bivalve species that dominated the DA but was absent from the LA, causing a lower-than-expected fidelity.

### 2. Materials and methods

#### 2.1. Dataset

Samples containing LAs and DAs were collected in 1999 off the United Arab Emirates, around two oil infrastructures in the Umm Al Dalkh field (UA, 25 km north-west of Abu Dhabi) and Zakum field (ZK, 84 km north-west of Abu Dhabi) (Fig. 1). Sediments showed very low contamination levels in both fields, and the LAs showed little disturbance (Albano et al., 2015a). Moreover, a comparison between LAs and DAs showed no evidence of community shifts (Albano et al., 2015b). Therefore, these samples can be regarded as coming from undisturbed environments. The seafloor at UA is a mixture of sand and shells, lies between 17 and 22 m depth, and is mostly inhabited by a small, sand-dwelling fauna. The seafloor at ZK consists of the so-called cap rock topped by a thin (ca. 1–5 cm) layer of sand, lying between 6 and 17 m depth. The fauna is characterized by less frequent sand-dwelling species and more frequent large-sized organisms, including scattered coral colonies.

At each location, a grid of sampling stations was designed over an area of ca. 45 km² in UA (13 stations) and ca. 240 km² in ZK (16 stations). Sampling for sediments and benthic organisms was carried out by divers with large aluminum scoops used in earlier benthic surveys in the area (Coles & McCain, 1990). The size fraction analyzed here was sieved with minimum 2 mm mesh size and maximum 5 mm mesh size, and then sorted under dissecting microscopes. The minimum mesh size helped avoid the collection of juvenile specimens which would have proved daunting to segregate and identify in this diverse and taxonomically poorly known tropical area. The maximum mesh size helped remove large but patchily distributed individuals which would be undersampled in the DAs due to the limited sampled area. Moreover, the species in this size range are expected to have more comparable “taphonomic half-lives” (Kidwell, 2002) and therefore reduce the effect of differential destruction rates among species.

All living molluscs were extracted from samples, while the extraction of empty shells from DAs was carried out until at least 1000 skeletal elements were collected at each station. In total, the study is based on 523 living individuals and 50,000 skeletal elements. To avoid counting a skeletal element (whole shells for gastropods and scaphopods, valves for bivalves and polyplacophorans) twice, only fragments larger than half a whole skeletal element were considered. Skeletal element counts were divided by the number of skeletal elements of each taxon.

---

![Fig. 1. The study area showing the two offshore oilfields (Umm Al Dalkh and Zakum) in which the platform structures were sampled.](image-url)
by performing Mantel tests with 999 permutations (Mantel, 1967). Anderson, 2001). Results at different taxonomic levels were compared by permutational multivariate analysis of variance (PERMANOVA, McArdle & Anderson, 2001; Oksanen et al., 2015). Analyses were computed at the field scale, pooling station abundances into field abundances, for univariate indices, and using stations as replicates in multivariate analyses. To account for sample-size differences between DAs and LAs, each DA was resampled without replacement to the sample size of the corresponding LA with 1000 iterations before computing univariate indices and conducting multivariate analyses. In the latter, the smaller sample size of LAs would have inevitably increased their dispersion (Tomašových & Kidwell, 2009).

In our dataset, LD agreement results at the species level were heavily influenced by a single bivalve species, Ervilia purpurea (Lamy, 1914), which was dominant in the DA (up to 90% at some sites) but totally absent from the LA. The lack of this species in living assemblages was related to spatial and temporal volatility in its abundance rather than to anthropogenic disturbance or to high preservation potential (Albano et al., 2015b). Therefore, we compared taxonomic sufficiency with and without Ervilia to explore the robustness of different taxonomic levels to such dominance patterns. Results without Ervilia are presented only when significantly different from the complete dataset.

All analyses, tests and plots were conducted in the statistical programing environment R, version 3.1.1 (R Development Core Team, 2009), and using the “vegan” package (Oksanen et al., 2013).

3. Results

3.1. Species richness and evenness

LAs had 42 and 38 species in UA and ZK, respectively. DAs were much more diverse: 253 species were found in UA and 160 in ZK, but much more diverse: 253 species were found in UA and 160 in ZK, but

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of taxa (S) and ratio between the number of species and of supra-specific taxa (S/T) in Umm Al Dalkh (UA) and Zakum (ZK) fields, for living (LA) and death (DA) assemblages (raw values (obs) and mean values of resampled datasets to standardize sample size (stan)).</td>
</tr>
<tr>
<td>UA</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Multivariate analysis was conducted with non-metric multidimensional scaling (NMDS) (Kruskal & Wish, 1978) using Bray–Curtis distances on square-root transformed proportional abundances. Differences in composition between LAs and DAs and between the two fields (separately for LAs and DAs) were assessed by permutational multivariate analysis of variance (PERMANOVA; McArdle & Anderson, 2001; Anderson, 2001). Results at different taxonomic levels were compared by performing Mantel tests with 999 permutations (Mantel, 1967). Guillo & Rousset (2013) suggested that Mantel tests should not be used when both variables are spatially autocorrelated. The Mantel-test correlations between Bray–Curtis distances and spatial Euclidean distances (separately for LAs and DAs) are low and insignificant.

Analyses were computed at the field scale, pooling station abundances into field abundances, for univariate indices, and using stations as replicates in multivariate analyses. To account for sample-size differences between DAs and LAs, each DA was resampled without replacement to the sample size of the corresponding LA with 1000 iterations before computing univariate indices and conducting multivariate analyses. In the latter, the smaller sample size of LAs would have inevitably increased their dispersion (Tomašových & Kidwell, 2009).

In our dataset, LD agreement results at the species level were heavily influenced by a single bivalve species, Ervilia purpurea (Lamy, 1914), which was dominant in the DA (up to 90% at some sites) but totally absent from the LA. The lack of this species in living assemblages was related to spatial and temporal volatility in its abundance rather than to anthropogenic disturbance or to high preservation potential (Albano et al., 2015b). Therefore, we compared taxonomic sufficiency with and without Ervilia to explore the robustness of different taxonomic levels to such dominance patterns. Results without Ervilia are presented only when significantly different from the complete dataset.

All analyses, tests and plots were conducted in the statistical programing environment R, version 3.1.1 (R Development Core Team, 2009), and using the “vegan” package (Oksanen et al., 2013).
became equally diverse when size-standardized (43 species in UA and 38 in ZK). Although the species/genus ratio (S/G) of raw DAs (1.29–1.48) exceeds S/G of LAs (~1.1), this ratio is close to 1.1 in standardized DAs (Table 1). The species/family ratio was also similar between LAs and standardized DAs. At the order and class levels, species were more evenly distributed among taxa.

Fidelity in species richness (ΔS) and evenness (ΔPIE) was comparable at all taxonomic levels (Fig. 2), with the single exception of a drop in fidelity at the order level in UA. As could be expected, datasets that did not contain the dominant *E. purpurea* in DAs had a higher fidelity in evenness (ΔPIE ~0 rather than −0.2).

3.2. Live–dead differences in composition

Jaccard and Bray–Curtis similarity between LAs and DAs increased from the species to the order level and peaked at the class level (Fig. 3). Results with the dataset without the dominant *E. purpurea* show almost identical trends and very similar values (not illustrated).

The rank–order correlation in abundance between LAs and DAs gradually increased with taxonomic level, irrespective of whether *E. purpurea* was retained or not in the dataset (Fig. 4). Significance values are, however, above 0.05, with the single exception of the order level in UA. The species and genus levels showed very similar but small (~0) values in the Zakum field.

The NMDS (Fig. 5) shows that LAs and DAs were well separated in two-dimensional space at the species level. At the genus level, segregation was also good. Differences between LAs and DAs within each field
4. Discussion

4.1. Are within-habitat properties preserved similarly in LAs and DAs at different taxonomic levels?

Most genus-level metrics of live–dead agreement in composition and in compositional gradients (i.e., fidelity of among-sample similarities) led to similar results as the species-level metrics. Moreover, distance matrices at the genus level were very similar to those at the species level. This implies that paleoecological information obtained from fossilized molluscan death assemblages identified at the genus level enable reconstructing the original biological signature in the same way as identification at the species level.

Moreover, the genus and family levels seem to be more robust to extreme dominance patterns such as those present in our dataset, where the bivalve *Ervilia purpurea* made up on average 30–50% of the DAs at the field scale whereas it was completely missing from the LAs (Albano et al., 2015b). For example, LD agreement in composition was much higher at the genus than at the species level in ZK, because living individuals of the co-generic *Ervilia scaliola* Issel, 1869 compensated for the lack of *E. purpurea* in the living assemblage (PERMANOVA, Fig. 5). However, this averaging out of within-genus variation in live–dead agreement in abundance at the species level can be expected only when genera are not monospecific and co-generic species have similar environmental preferences.

3.3. Live–dead differences in compositional gradients

NMDS ordinating both fields shows that differences between UA and ZK are preserved by DAs at the genus level (Fig. 5; PERMANOVA, p < 0.05, Fig. 9). The difference then declined progressively and classes no longer detected differences between the two fields. Results using datasets without the dominant *E. purpurea* were very similar (not illustrated). Although Mantel-test correlation between compositional distances among LAs and among DAs did not decline at higher taxonomic levels, between-field differences became smaller relative to within-field differences, and PERMANOVA thus becomes insignificant at higher levels.

Dissimilarity matrices at the species level were very similar to those at the genus level (Mantel test, mean r = 0.9, p < 0.05); similarity decreased at the family level (r = 0.8, p < 0.05) and further at the order and class levels (mean r ≤ 0.5, p < 0.05) (Fig. 10). Datasets without *E. purpurea* showed the same decline in similarity (Fig. 10).
Identification at the family level also provided results similar to the species level for several metrics and analyses. Among the exceptions, the rank–order correlation in abundance between LAs and DAs steadily increased with increasing taxonomic level; this result was expected because dominance patterns are preserved when pooling abundances at higher taxonomic levels. The reduction of the number of taxa also explains the pattern of the Bray–Curtis similarity index, which gradually increased with taxonomic level. These results imply that trends in rank–order of taxa abundance and in Bray–Curtis similarities in the original biological community at the species level cannot be inferred by a paleocommunity analysis conducted at the family (or higher) level. In contrast, some metrics (such as taxonomic richness and evenness) also gave results similar to the species level at high taxonomic levels, such as order and class.

### 4.2. Are between-habitat differences preserved similarly in LAs and DAs at different taxonomic levels?

Time-averaged molluscan DAs do capture compositional and environmental variations among samples of the living assemblage (e.g., Tomašových & Kidwell, 2009; Albano & Sabelli, 2011; Weber & Zuschin, 2012; Albano, 2014; Zuschin & Ebner, 2015). The two areas studied in our work are tens of kilometers apart. They share similar benthic ranges, but have different substrates. At this spatial scale, differences between areas detected by LAs were detected by DAs up to the order level (PERMANOVA, Fig. 6). These results are in accordance with other studies that found, for example, that family data still represent effective taxonomic surrogates in detecting β-diversity patterns in diverse marine settings (Terlizzi et al., 2008; Bett & Narayanaswamy, 2014).

### 4.3. Why do genera and families work?

The reason for the sufficiency of genus- and family-level identifications is partly related to the ratio between the number of species and genera/families: it was below 2 in our datasets, while species/order and species/classes ratios were above 2.5. Therefore, our results confirm that ratios between 2 and 2.5 may be thresholds beyond which multivariate analyses at coarser levels of resolution may not be well correlated with those performed at the species level (Timms et al., 2013).

Moreover, similar species/higher taxon ratios in corresponding LAs and DAs can contribute to the surrogacy of higher taxa in live–dead agreement comparisons, because they imply the same degree of information loss with increasing taxonomic level. Such similar ratios also depend on homogeneity of shell durability within higher taxa: if a taxon includes both fragile and robust taxa, then the former may be more easily lost in DAs, causing lower S/T ratios than in LAs.

Closely related organisms also tend to be ecologically similar (e.g., Cavender-Bares et al., 2009). Accordingly, environmental gradients (e.g., differences in composition and structure between different habitats) can be expected to be preserved at higher taxonomic levels. For example, invertebrate tolerance within orders is less variable than among orders in freshwater habitats: all species belonging to sensitive families (e.g., differences in composition and structure between different habitats) can be expected to be preserved at higher taxonomic levels.

![Fig. 8. Boxplots of Mantel-correlation tests between species and supraspecific level Bray–Curtis distances of LAs and DAs in Umm Al Dalikh (left) and Zakum (right) fields. Correlation is very high at the genus and family levels and is much lower at the order and class levels. Correlations are always significant (p < 0.05).](image1)

![Fig. 9. PERMANOVA R2 (solid circles) and F (grey triangles) values showing differences between UA and ZK in LAs (left) and DAs (right), with associated p-values. Differences between UA and ZK DAs are fully significant up to the family level. Grey line: 0.05 significance level.](image2)

![Fig. 10. Boxplots of Mantel-correlation tests between species and supraspecific level Bray–Curtis distances of LAs and DAs of Umm Al Dalikh (left) and Zakum (right) fields (see also NMDS in Fig. 5). Correlation is very high at the genus and family levels and is much lower at the order and class levels. Values are always significant (p < 0.05).](image3)
4.4. Taxonomic sufficiency helps when diagnostic characters get lost

The good sufficiency results at the genus and family levels have important operational consequences. The loss of taxonomically important diagnostic characters is common after the death of organisms due to abrasion, dissolution, fragmentation and other damaging processes acting on skeletal remains. The identification at the species level may also be inaccurate or imprecise if carried out by different operators. In contrast, genus- and especially family-level identification is certainly more accurate, even if conducted by non-specialists. Taxonomy at the family level is based on more relevant characters which can be more easily retained in specimens subject to taphonomic damage; it is also more stable in time and among taxonomic schools, allowing a better comparability of results among different works.

4.5. Extrapolation to other phyla

These results on mollusc assemblages are not necessarily transferable to other phyla because levels of taxonomic sufficiency of different organisms are variable. Firstly, different ratios between species and supraspecific taxa for different phyla may play a role (e.g., Heino & Soininen, 2007). Secondly, although families generally represent fairly homogeneous groups of organisms sharing basic functional traits (Warwick, 1993), the level of such intra-family functional redundancy could depend on the classification criteria adopted and on taxonomic traditions (Bertrand et al., 2006), which may not be consistent across phyla (Warwick & Somerfield, 2008).

5. Conclusions

The results of our case study suggest that the common practice of genus-level identification of paleontological samples allows for a proper reconstruction of the original biological community at the species level, because fidelity at the species and genus levels is very similar. Family-level identification may also be sufficient. Live–dead agreement becomes even better at family and higher taxonomic levels for some metrics, but between-habitat differences in composition become weaker above the family level. At the genus and family levels, however, between-habitat differences are equally strong as at the species level. The effect of some species that reduce live–dead agreement (i.e., overwhelmingly abundant in the DA and absent from the LA) is averaged out at the genus level by abundances of co-generic species that occur in the LA. Although caution should be exercised when extrapolating results to other phyla, it is likely that the same results can be obtained when analyzing communities with expected low (<2) ratios between species and higher taxon numbers, and with similar ratios in living and death assemblages.

Acknowledgments

Field work was conducted in the framework of the “Marine environmental monitoring programme” of Zakum and Umm al Dalkh oilfields by Texpior G.m.b.H. on behalf of the Zakum Development Company (ZADCO). Ernst Geutebrück, Texpior, kindly consented to the reanalysis of samples and the dissemination of the results. Reinhard Kikinger and M.S. took and initially processed the benthic samples on board. Antonio Bonfitto, Bruno Dell’Angelo, Manfred Herrmann, Graham P. Oliver, and Bruno Sabelli are thanked for support in species segregation and identification. Jan Steger, Mirko D. Gennaro and Nadezhda Filippova helped with sorting and identification of samples. P.G.A. was funded by a Marie Curie fellowship by the European Commission (PEF-GA-2012-328089). A.T. thanks the Slovakian Scientific Grant Agency for support (VEGA 2/0136/15).

References
