Vol. 491: 77–90, 2013
doi: 10.3354/meps10475

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published October 2

Evaluation of the taxonomic sufficiency approach for ichthyoplankton community analysis

Frank J. Hernandez Jr.[1,4,*], Laure Carassou[1,5], William M. Graham[2], Sean P. Powers[3,1]

[1]Dauphin Island Sea Lab, Dauphin Island, Alabama 36528, USA
[2]Department of Marine Science, The University of Southern Mississippi, Stennis Space Center, Mississippi 39529, USA
[3]Department of Marine Sciences, University of South Alabama, Mobile, Alabama 36688, USA

[4]*Present address:* Department of Coastal Sciences, The University of Southern Mississippi, Ocean Springs, Mississippi 39564, USA
[5]*Present address:* Rhodes University, Department of Zoology and Entomology, PO Box 94, Grahamstown 6140, South Africa

ABSTRACT: Ichthyoplankton identification is a time-consuming task, and often larvae cannot be identified to species due to a lack of adequate early life history descriptions. As a result, ichthyoplankton assemblage data are often analyzed at the family level, which results in a loss of taxonomic resolution, or at mixed taxonomic levels (e.g. family, genus and species combined), which can lead to difficulties in interpretation of results when a single species is included in multiple taxonomic groupings. The taxonomic sufficiency (TS) approach has been used extensively in other disciplines (e.g. benthic marine macrofauna) to address similar analytical constraints, but to date this method has not been rigorously examined for ichthyoplankton studies. In this study, an ichthyoplankton data set collected in the northern Gulf of Mexico was proportioned into 3 data subsets with varying levels of taxonomic resolution: (1) species level only; (2) species, genus and family levels; and (3) combined taxonomic levels. Comparisons were made for assemblage metrics (larval density, richness and diversity) calculated for each taxonomic subset, as well as multivariate analyses of temporal variations characterizing ichthyoplankton assemblages. Genus- and species-level similarity matrices were highly correlated, which suggests analyses at the genus level could serve as a good proxy for species when examining assemblage diversity. Multivariate results for seasonal patterns were similar among family-, genus- and species-level analyses. The common approach of analyzing ichthyoplankton assemblages at mixed taxonomic levels, however, is not as statistically rigorous as single taxonomic-level analyses.

KEY WORDS: Taxonomic resolution · Larval fish · Seasonality · Gulf of Mexico · Multivariate community analyses

INTRODUCTION

A pervasive problem in ichthyoplankton research is the lack of taxonomic information available to identify many early life stages to the species level. Because of the rapid morphological changes that occur during early ontogeny, a series of descriptions is required to document what Richards (2005) called a 'dynamic process with many changes appearing at different times and in different places'. In addition, the number of larval fish descriptions for a given region is often inversely related to faunal diversity (Fahay 2007). For example, larval fish descriptions for regions with relatively high taxonomic diversity, such as the tropical Indo-Pacific, Northwest Pacific and Western Central Atlantic (which includes the Gulf of Mexico), are available for only 10%, 34% and 40% of the known marine fish species in each region,

*Email: frank.hernandez@usm.edu

respectively (Fahay 2007). In the Western Central Atlantic, many of the unknown larval stages belong to commercially and recreationally important groups, such as serranids (Richards et al. 2005), lutjanids (Drass et al. 2000, Lindeman et al. 2005) and sparids (Powell & Greene 2005), which limits the utility of ichthyoplankton data for stock assessments and fisheries management. The use of molecular techniques to identify fishes (including larval stages) has increased in recent years, and provides a nice compliment to traditional identification based on morphometrics. To date, these efforts tend to focus on relatively few species within a given taxonomic group, such as individual families (Luthy et al. 2005, Vandersea et al. 2008, D'Alessandro et al. 2010, Marancik et al. 2010). Although molecular methods are available for examining complete larval assemblages (e.g. barcoding, microarrays), the expense of identifying large numbers of specimens from multiple species in regions of high diversity remains a limiting factor.

The inability to identify larval fishes to species levels often complicates statistical analyses and the resulting data interpretations. Many studies examining ichthyoplankton assemblage dynamics, for example, have utilized a suite of metrics and analyses (e.g. diversity, richness, cluster analysis, MDS, ANOSIM) on data sets with larvae identified at mixed levels of taxonomic resolution, i.e. combinations of family, genus and species identifications (Tolan et al. 1997, Marancik et al. 2005, Boeing & Duffy-Anderson 2008, Muhling et al. 2008), or simply grouped to family level (Vásquez-Yeomans 2000, Quattrini et al. 2005, Suthers et al. 2006, Syahailatua et al. 2011, Carassou et al. 2012, Muhling et al. 2012) when species identifications are problematic. Even closely related species may differ in their early life history traits (Richardson et al. 2007), thus analyzing data based on multi-level classifications of fish larvae may confound interpretation of results, as the same species may be partitioned into multiple taxonomic units, or lumped into one taxonomic unit with conspecifics. For example, identifiable bay whiff *Citharichthys spilopterus* larvae were included as a taxonomic unit in a detailed examination of larval fish seasonality and abundance in the Gulf of Mexico by Hernandez et al. (2010a). However, *C. spilopterus* specimens unidentifiable to species for various reasons (e.g. due to damaged condition or small size), were likely included in 2 other taxonomic groupings (*Citharichthys* spp. and unidentified Paralichthyidae). Also, the *Citharichthys* spp. taxonomic unit may contain as many as 5 species (*C. arctifrons*, *C. cornotus*, *C. gymnorhinus*, *C. macrops* and *C. spilopterus*) (Lyczkowski-Shultz & Bond 2005). While these taxonomic groupings are often an undesirable necessity, to date there has been no examination of the consequences of using multi-level and higher-level classifications of ichthyoplankton in community analyses.

As with ichthyoplankton, the identification of zooplankton, benthic meiofauna and macrofauna, planktonic ciliates and other relatively small marine organisms is time-consuming, and thus researchers are tasked with pragmatic decisions that weigh costs of resources (e.g. need for taxonomic expertise, expediency of sample processing) against the benefits of taxonomic resolution needed to address scientific objectives (Ellis 1985, Warwick 1988, Bertasi et al. 2009, Xu et al. 2011). As a result, there is a large and growing body of literature that addresses the taxonomic level of identification required to analyze invertebrate assemblage dynamics, particularly in response to anthropogenic disturbances (Ellis 1985, Warwick 1988, Mendes et al. 2007, Puente & Juanes 2008, Jimenez et al. 2010, Xu et al. 2011). Numerous studies (primarily related to pollution effects) have validated the 'taxonomic sufficiency' (TS) approach as a means of using coarse, higher-level identifications without substantial loss of information (e.g. Olsgard et al. 1998, Karakassis & Hatziyanni 2000, Gomez Gesteira et al. 2003). Ferraro & Cole (1990), for example, demonstrated that family-level identifications were adequate for detecting intermediate and large changes in community metrics for macrobenthos collected in the Southern California Bight (USA). Similarly, Xu et al. (2011) succeeded in using genus- and family-level identifications without significant loss of information while examining planktonic ciliate communities in Jiaozhou Bay (China). Most examinations of taxonomic sufficiency have focused on macrobenthos or meiofauna, but the approach has also been extended to examine terrestrial invertebrates (Pik et al. 1999, Cardoso et al. 2004), phytoplankton (Carneiro et al. 2010), macroalgae (Smale et al. 2010), and terrestrial mammals (Grelle 2002), among other taxa. To date, however, the taxonomic sufficiency approach has not been tested or utilized for ichthyoplankton studies, despite similarities in methodological and taxonomic constraints.

The main objective of this study was to examine the utility of the taxonomic sufficiency approach using ichthyoplankton data collected in the northern Gulf of Mexico. Comparisons were made for assemblage metrics (larval density, richness and diversity) determined from a single data set proportioned into 3 data subsets with varying levels of taxonomic resolution: (1) species-level only; (2) species, genus and

family levels; and (3) combined taxonomic levels. Our goal was to assess the validity of the commonly used practice of analyzing larval fish data at multiple levels of taxonomic resolution, and offer recommendations for future analyses of ichthyoplankton community dynamics, including ecological assessments of anthropogenic disturbances such as the Deepwater Horizon oil spill.

MATERIALS AND METHODS

Ichthyoplankton sampling

Ichthyoplankton samples were collected monthly from October 2004 to October 2006 at a single study location (water depth of 20 m) approximately 18 km south of Dauphin Island, Alabama, USA (30° 05′ 25″ N, 88° 12′ 42″ W). Sampling methodologies were detailed in Hernandez et al. (2010a,b). In brief, depth-discrete plankton samples were collected using a Bedford Institute of Oceanography Net Environmental Sampling System (BIONESS, Open Seas Instrumentation) with a 0.25 m^2 mouth opening fitted with 202 µm mesh nets. Towed times for each sample were generally short (mean = 2.5 min), with an overall mean filtered volume of 40.6 m^3 (SD = 7.0 m^3). Net contents were rinsed with seawater, sieved (149 µm mesh), and preserved in 4% formalin for 48 h before being transferred to 70% ethanol. A flowmeter (General Oceanics) mounted within the BIONESS frame estimated the volume of water filtered for each sample. Ichthyoplankton samples were sorted and larval fish were identified to the lowest possible taxonomic level at the Plankton Sorting and Identification Center (Szczecin, Poland) and at the Dauphin Island Sea Laboratory (Dauphin Island, Alabama). Most identifications (52%) were at the family level, followed by species (22%), order (14%), and genus (7%) levels (Hernandez et al. 2010a). Larval specimens were represented by 15 orders, 59 families, 81 genera and 64 species of fishes. In all, 1634 ichthyoplankton samples were processed and available for analyses (Table 1).

Table 1. Number of ichthyoplankton samples collected (Oct 2004 to Oct 2006) on the Alabama inner shelf and retained for multivariate analyses. For each analytical stage (I, II, III; see 'Materials and methods: Data analysis' and Fig. 1), samples retained are those in which fish larvae from at least 2 taxa were collected

Month	Sampling dates	No. of samples			
		Collected	Retained		
			I	II	III
Oct 2004	22	52	45	45	51
Nov 2004	16, 17, 29	88	73	78	87
Dec 2004	8	47	18	18	30
Jan 2005	6, 18, 19, 20, 21	124	39	52	86
Feb 2005	16	50	6	6	17
Mar 2005	29	23	33	20	23
Apr 2005	19	65	23	39	62
May 2005	9, 11, 12, 13, 17	120	65	66	118
Jun 2005	9	47	40	41	47
Jul 2005	13	48	25	25	47
Aug 2005	9	46	46	46	46
Sep 2005	14, 27, 28, 29	120	69	85	113
Oct 2005	11	31	22	30	31
Nov 2005	9, 29, 30	103	25	59	84
Dec 2005	1, 2, 16	40	45	21	31
Jan 2006	12	44	5	6	16
Feb 2006	7, 8, 9, 10, 17	103	55	57	80
Mar 2006	16	39	20	21	38
Apr 2006	12, 13	38	1	6	34
May 2006	1, 2, 3, 4, 17	113	64	66	111
Jun 2006	15	42	37	38	42
Jul 2006	5	46	18	19	40
Aug 2006	10	46	46	46	45
Sep 2006	8, 19, 20, 21, 22	112	94	99	112
Oct 2006	12	47	46	47	47
Total		1634	960	1036	1438

Data analysis

Analyses were parsed into 3 stages (I, II, III), each representing different levels of taxonomic data inclusion, to examine the effect of taxonomic resolution on the description and comparison of larval fish assemblages (Table 1, Fig. 1). The data subset for Stage I analyses included only fish larvae identified to the species level, from which 3 matrices were created by grouping larvae at the species, genus and family levels (Fig. 1). Because Stage I matrices were comprised of only larvae identified to the lowest possible taxonomic level (species), each matrix (family, genus and species) contained the same overall number of fish larvae. Similarly, Stage II included 3 matrices (fish larvae grouped at the family, genus and species levels), but included all fish larvae identified at least to the family level (Fig. 1). Because Stage II matrices were comprised of fish larvae from different taxonomic levels (family, genus and species), each matrix (family, genus and species) contained different numbers of fish larvae. Stage III included one matrix designed to examine the effect of analyzing larval fish data by combining multiple taxonomic levels (Fig. 1). This matrix was comprised of fish larvae from a mixture of taxonomic levels that included family, genus and species identifications. The progression of

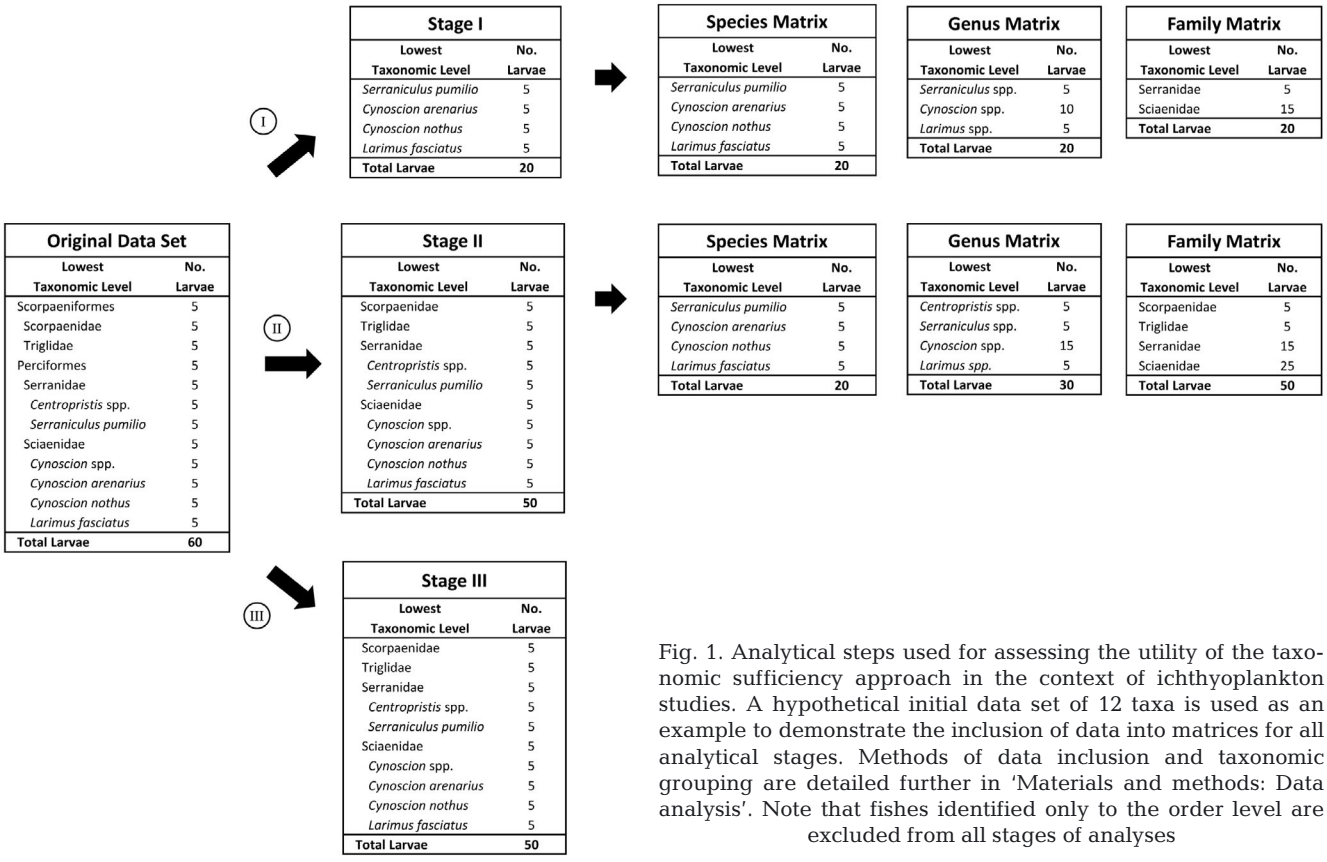

Fig. 1. Analytical steps used for assessing the utility of the taxonomic sufficiency approach in the context of ichthyoplankton studies. A hypothetical initial data set of 12 taxa is used as an example to demonstrate the inclusion of data into matrices for all analytical stages. Methods of data inclusion and taxonomic grouping are detailed further in 'Materials and methods: Data analysis'. Note that fishes identified only to the order level are excluded from all stages of analyses

this statistical design ranged from a relatively conservative approach (e.g. Stage I species-level matrix) that used a relatively small portion of the overall available data, but with high taxonomic resolution, to a more liberal inclusion of larval fish observations, but with mixed and overlapping taxonomic resolution (e.g. Stage III). In every case, order-level taxa were removed from the analyses.

For each data matrix listed above (n = 7), the abundances of fish larvae from each sample were standardized by the volume filtered and used to calculate mean density estimates (no. m^{-3}) for each month. Taxonomic richness and diversity for each sample were calculated based on the count of taxa and the exponential Shannon entropy index (exp H'), respectively (Jost 2006). Both metrics were averaged within each sampling month.

Each data matrix was also used in multivariate analyses of temporal variations characterizing ichthyoplankton assemblages. To improve graphical interpretations, samples without fish larvae or with only 1 taxon were excluded. The list of samples retained was kept constant from one matrix to the other within a given analytical stage (see corresponding number of retained samples in Table 1). To examine and compare the effect of commonly used data transformations, all multivariate analyses were done using raw, log(x + 1), presence/absence, square-root and fourth-root-transformed data. For each matrix, analysis of similarity (ANOSIM) was used to test for significant ($p < 0.05$) variations between years and between months in assemblage structure. Values of R statistics were used to compare the degree ('strength') of these variations from one taxonomic level or data combination to the other, on a scale of 0 (indistinguishable) to 1 (all similarities within months or years are less than any similarity between months or years) (Clarke 1993).

Two-dimensional graphical visualizations of seasonal patterns derived from each matrix were then generated using multi-dimensional scaling (MDS) based on Bray-Curtis similarities between samples. Finally, correlations between the species-, genus- and family-level matrices from Stages I and II were calculated using Spearman rank correlations based on the comparison of Bray-Curtis similarities between matrices (RELATE procedure, Primer) (Clarke & Ainsworth 1993).

RESULTS

Larval density, richness and diversity

Total larval fish density was characterized by a peak in August regardless of the taxonomic resolution considered (Fig. 2a–c), although the magnitude of this peak varied, increasing between Stage I vs. Stages II and III, coincident with the inclusion of additional taxa. Additionally, minor peaks in density (e.g. March, May and June) were not evident in the Stage I data analysis, but easily discernible in Stage II and Stage III analyses. In the Stage II analysis, monthly patterns in larval density were similar between taxonomic levels; however, significantly higher larval fish densities were observed during March through October for the family-level analysis (Fig. 2b).

The monthly pattern in larval fish richness (count of taxa) was similar for all stages, though again the magnitude varied (Fig. 2d–f). In general, richness peaked in June, August and October over the year. Results for the Stage I analysis were largely similar, although richness was relatively higher in the genus- and species-level analyses from August through November (Fig. 2d). In the Stage II analysis, larval fish richness was higher during most months (February through November) for the family-level analysis, relative to the genus- and species-level analyses (Fig. 2e). The overall pattern in taxonomic richness for the Stage III analysis was similar to that of the Stage II family analysis (Fig. 2e,f).

Larval diversity (exp H') remained relatively constant throughout the study period, with the exception of a decrease in October, regardless of the taxonomic resolution considered (Fig. 2g–i). However, exp H'

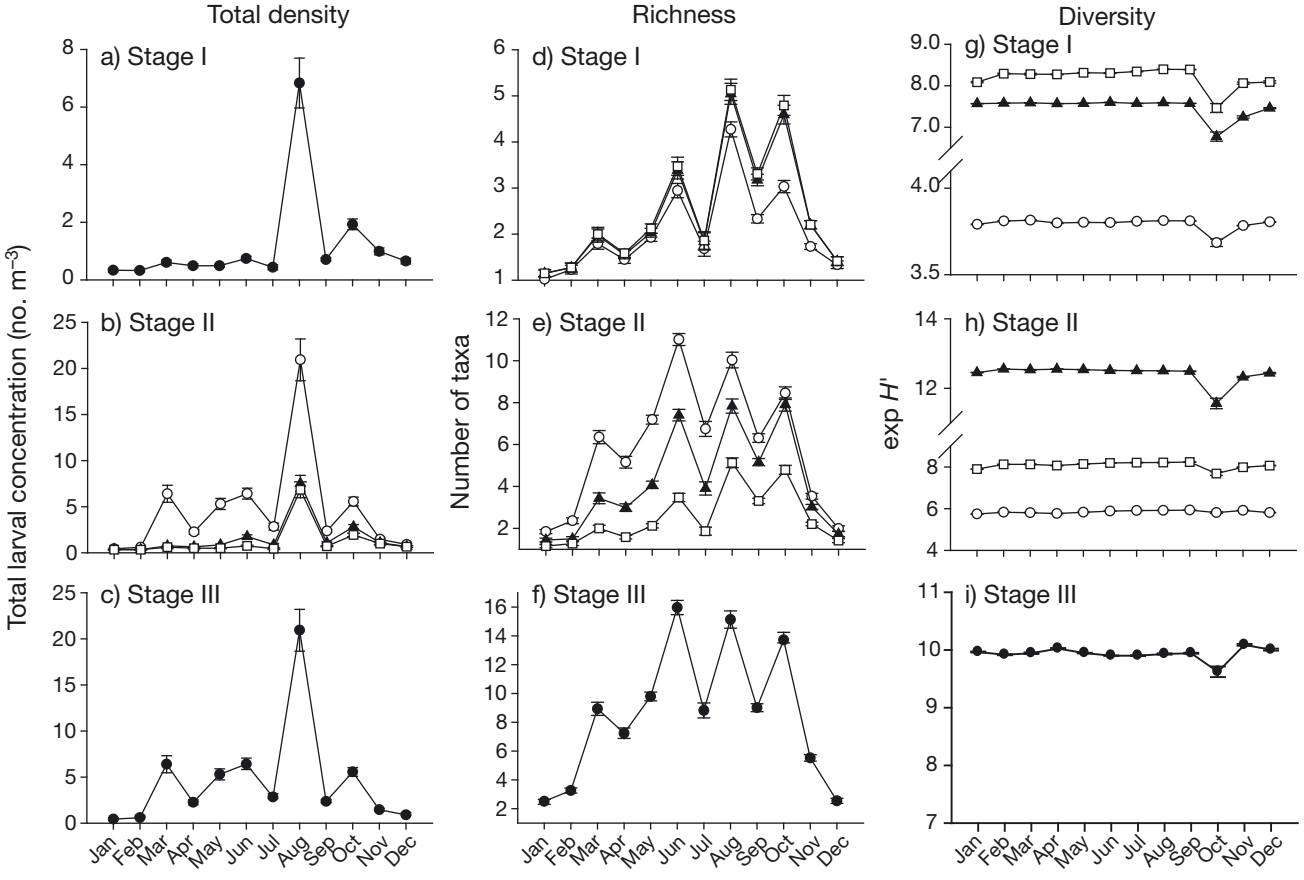

Fig. 2. Monthly patterns in total larval density (no. m^{-3}), richness (count of taxa), and exponential Shannon entropy diversity index (exp H') obtained when including only larvae identified to species and analyzed at different taxonomic levels (Stage I: a,d,g), when including all larvae identified to families, genus and species, and analyzed at different taxonomic levels (Stage II: b,e,h), and when including all larvae identified to families, genus and species and by combining different taxonomic levels (Stage III: c,f,i). When multiple curves are presented in a panel (b,d,e,g,h), open circles represent family-level data, filled triangles genus-level data, and open squares species-level data. Filled circles are used when only one curve is shown (c,f,i), or when data from the 3 taxonomic levels perfectly superimpose (a)

values showed different patterns depending on data selection. When only larvae identified to species were considered (Stage I), exp H' was consistently higher at the species level, followed by genus level and family level (Fig. 2g). Conversely, when all larvae identified to families were included (Stage II), exp H' was consistently higher at the genus level, followed by species level and family level (Fig. 2h). Similar to richness, the overall pattern in diversity for the Stage III analysis was similar to that of the Stage I and II analyses (Fig. 2e,f).

Detection of temporal variations in larval assemblages

The 3 matrices for Stage I used in multivariate analyses included a total of 67 503 larvae, representing 65 species, 55 genera and 26 families. The 3 matrices retained for Stage II included a total of 67 503 larvae identified at the species level (65 species), 87 875 larvae identified at the genus level (82 genera), and 250 747 larvae identified at the family level (57 families). The matrix retained for Stage III included a total of 250 752 larvae representing 152 taxa (see Table 1 for corresponding numbers of retained samples).

Significant variations in larval assemblages between years and months were detected for all taxonomic levels and data combinations ($p \leq 0.01$; Table 2). However, the strength of those variations differed among taxonomic levels and analytical stages (see R values, Table 2). In general, the relative strength of inter-annual and monthly variations in larval assemblages were stronger at the genus and species levels than at the family level, regardless of the data set used (Stage I or II) or data transformation. Genus- and species-level data provided very close results in all cases. When multiple taxonomic levels were combined (Stage III), the strength of variations detected were generally lower than those for genus- and species-level data and higher than those for family-level data (Table 2). Overall, the R statistics for ANOSIM results differed little between data transformations, though in most cases, the square-root or fourth-root transformations resulted in the highest R values (Table 2).

Multivariate pictures of monthly patterns

The seasonal structure (monthly patterns) of larval fish assemblages derived from Stage I data (Fig. 3) was more clearly visualized at the species level, followed by genus and family level, as supported by increasing stress values. The differences in statistical scores between species- and genus-level analyses remained minor, and species-level and genus-level plots showed similar patterns in seasonal structure (i.e. the same months were similarly well distinguished from one another and similarly organized along the 2D plots). Family-level plots were gener-

Table 2. R statistics from ANOSIM testing for differences between years and months in larval fish concentrations at different taxonomic levels analyzed separately or combined together, and with different data transformations. $^{**}p \leq 0.001$, $^{*}p \leq 0.01$

Step-data set	Year				Month			
Transformation	Family	Genus	Species	Combined	Family	Genus	Species	Combined
I — Larvae identified to species only								
None	0.023*	0.154**	0.133**	–	0.295**	0.536**	0.566**	–
Square-root	0.034**	0.158**	0.137**	–	0.334**	0.568**	0.589**	–
Fourth-root	0.043**	0.158**	0.138**	–	0.341**	0.563**	0.582**	–
Log(x + 1)	0.025*	0.155**	0.134**	–	0.310**	0.564**	0.574**	–
Presence/absence	0.050**	0.152**	0.134**	–	0.331**	0.526**	0.548**	–
II — All larvae, distinct taxonomic levels								
None	0.052**	0.114**	0.110**	–	0.380**	0.515**	0.527**	–
Square-root	0.077**	0.123**	0.114**	–	0.453**	0.562**	0.559**	–
Fourth-root	0.081**	0.124**	0.116**	–	0.464**	0.570**	0.563**	–
Log(x + 1)	0.067**	0.117**	0.111**	–	0.401**	0.527**	0.536**	–
Presence/absence	0.071**	0.120**	0.115**	–	0.423**	0.552**	0.541**	–
III — All larvae, combined taxonomic levels								
None	–	–	–	0.096**	–	–	–	0.463**
Square-root	–	–	–	0.104**	–	–	–	0.493**
Fourth-root	–	–	–	0.104**	–	–	–	0.496**
Log(x + 1)	–	–	–	0.103**	–	–	–	0.471**
Presence/absence	–	–	–	0.100**	–	–	–	0.489**

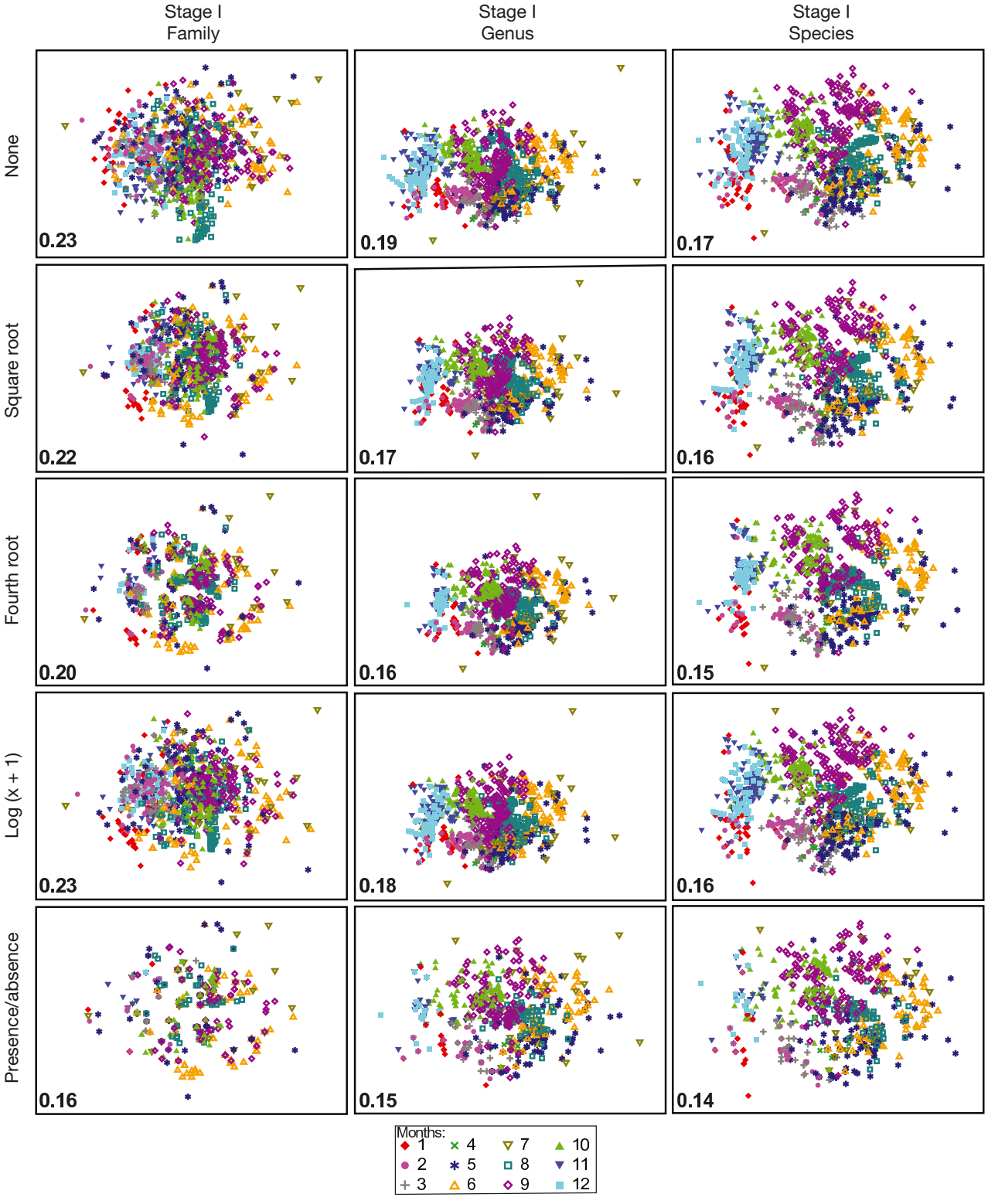

Fig. 3. Multidimensional scaling based on Bray-Curtis similarities between samples, showing the seasonal structure (Months 1–12 = Jan–Dec) of larval densities observed at different taxonomic levels and with different data transformations, based on 3 matrices including only larvae identified to species (Stage I). Bolded labels indicate stress values for each analysis. See Table 1 for the number of samples included in the analyses. For visual clarity, 1 and 6 samples were removed from the family and species plots, respectively (this did not affect the corresponding stress values)

ally characterized by low statistical performances (stress ≥0.20), except for presence/absence data, which, however, provided a poor resolution of monthly patterns in larval density on the MDS plot (Fig. 3).

The seasonal structure of larval fish assemblages derived from Stage II data was more clearly visualized at the species level, followed by genus and family level, as supported by increasing stress values (Fig. 4). The analyses of species-level data provided the best statistical performances for all data transformations applied (stress ≤ 0.17). Genus-level data performed better when transformed as fourth root and presence/absence (stress < 0.20), while family-level data always provided relatively high stress values (stress > 0.20). However, graphical representations of monthly patterns in larval densities were generally informative at all taxonomic levels, with months being similarly well distinguished and organized along the 2D plots (Fig. 4). Stage III data were not useful in analyzing the seasonal structure of larval fish assemblages, with stress values exceeding 0.22 whatever the transformation applied (Fig. 5).

Correlations between taxonomic levels

The close similarity between genus-level and species-level data revealed by the ANOSIM and MDS analyses was confirmed by the high correlation coefficients linking both matrices at each analytical stage (r ≥0.767; Table 3). Correlations between family-level and genus-level matrices, and between family-level and species-level matrices were lower, although reaching relatively high values for particular data transformations (e.g. r = 0.659 for fourth-root-transformed data; Table 3).

DISCUSSION

To our knowledge, there has been no previous examination of the efficacy of the taxonomic sufficiency approach for ichthyoplankton data, despite the methodology having been used for larval fish studies in one form or another (i.e. analyses at higher- or multi-level taxonomic groupings) (Rakocinski et al. 1996, Espinosa-Fuentes & Flores-Coto 2004, Marancik et al. 2005, Duffy-Anderson et al. 2006, Faria et al. 2006, Carassou & Ponton 2007, Brodeur et al. 2008, Carassou et al. 2012). Historically, the TS approach has been applied to invertebrate organisms that are closely related in their morphology, and requires considerable taxonomic expertise, time and resources for identification. The larval stages of many marine fishes, however, remain undescribed at the species level, so the limitations lie not only with the expertise of individual taxonomists, but with the overall lack of early life stage descriptions. Unless molecular methods are employed (often a prohibitively expensive alternative), researchers are left with the option of higher- or mixed-level taxonomic analyses, or the removal of higher-level taxa from their analyses. So while the TS approach for ichthyoplankton analyses has not been critically examined until now, the method has been used historically out of necessity.

Given the numerous pragmatic concerns that are common to the identification of ichthyoplankton and other marine organisms where the TS approach has been utilized, the TS approach is seemingly well-suited for multi-taxa ichthyoplankton data analysis. First, many marine fish larvae collected in the northern Gulf of Mexico (and other marine systems) can often be identified to the family level only, which limits the scope of concerns regarding taxonomic relatedness that are debated in TS analyses of invertebrate fauna which often include order, class, phylum and functional group levels (Bates et al. 2007). In our study for example, over 80% of the larvae collected were identified to at least the family level; of the order-level taxa (which constituted 14% of the total), the overwhelming majority (83%) were from a single order (Clupeiformes). Further, in a comprehensive literature review of taxonomic sufficiency studies Bevilacqua et al. (2012) determined that the effectiveness of using higher taxonomic levels was generally high when the higher taxa to species ratio (e.g. family:species or genus:species ratios) exceeded a value of 0.4. Based on data compiled for a recent biodiversity assessment for the Gulf of Mexico (McEachran 2009), the genus:species ratio for bony fishes is approximately 0.47, which suggests that genus-level identifications may provide an adequate surrogate for species in the northern Gulf of Mexico. This is consistent with our results, which showed that genus-level data were sufficient for obtaining statistically similar descriptions of temporal patterns in ichthyoplankton assemblages, as compared to species-level data (Table 2, Figs. 3 & 4).

However, our results also suggest that there are trade-offs to consider when deciding which taxonomic level may be sufficient for examining ichthyoplankton assemblage metrics. Estimates for the simplest metrics (total density and richness) varied with the level of data inclusion, as expected. The rela-

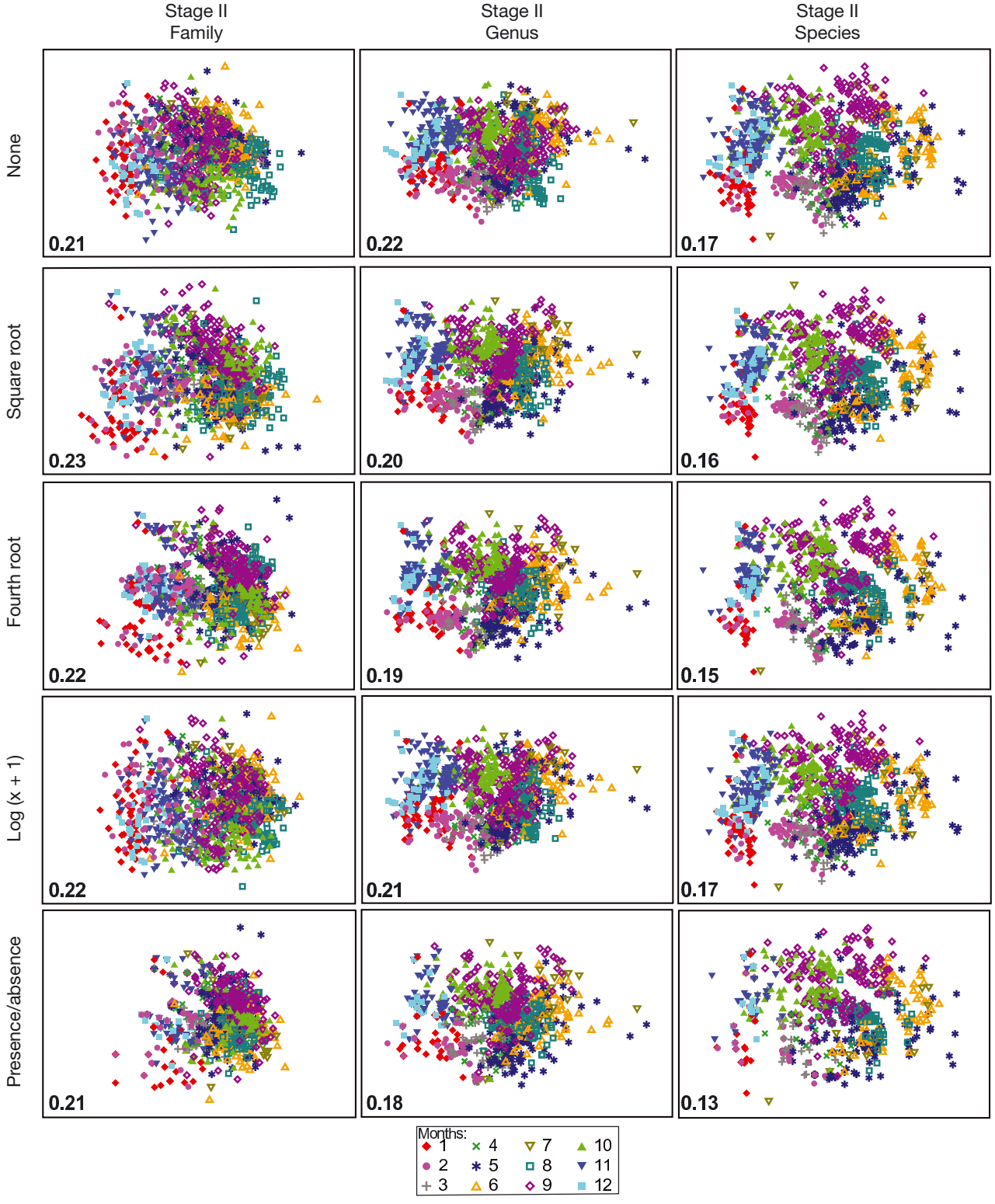

Fig. 4. Multidimensional scaling based on Bray-Curtis similarities between samples, showing the seasonal structure (Months 1–12 = Jan–Dec) of larval densities observed at different taxonomic levels and with different data transformations, based on 3 matrices including all larvae identified to families, genus and species (Stage II). Bolded labels indicate stress values for each analysis. See Table 1 for the number of samples included in the analyses. For visual clarity, 6 samples were removed from the species plots (this did not affect the corresponding stress value)

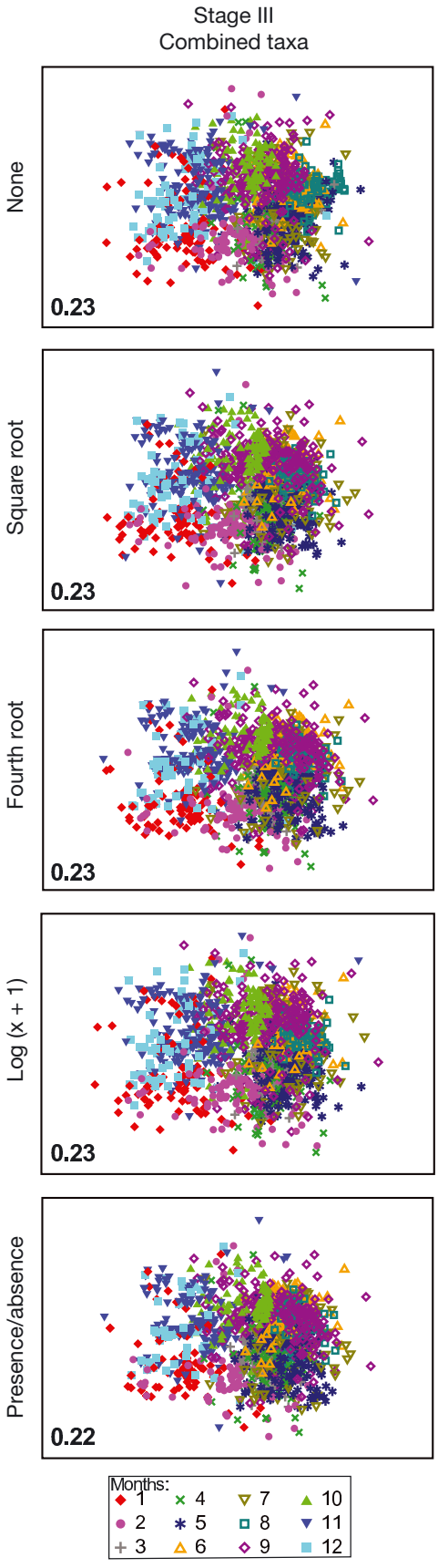

Fig. 5. Multidimensional scaling based on Bray-Curtis similarities between samples, showing the seasonal structure of larval densities observed when combining different taxonomic levels and applying different data transformations, based on a matrix including all larvae identified to families, genus and species (Stage III). Bolded labels indicate stress values for each analysis. See Table 1 for the number of samples included in the analyses

Table 3. Spearman rank correlations (r_S) characterizing pairs of similarity matrices based on larval concentrations analyzed at different taxonomic resolutions, and with different data transformations. **p = 0.001

Step-data set Transformation	Family × Genus	Family × Species	Genus × Species
I — Larvae identified to species only			
None	0.455**	0.431**	0.924**
Square root	0.523**	0.497**	0.929**
Fourth root	0.574**	0.540**	0.929**
Log(x + 1)	0.452**	0.434**	0.927**
Presence/absence	0.605**	0.565**	0.921**
II — All larvae, distinct taxonomic levels			
None	0.532**	0.417**	0.802**
Square root	0.633**	0.479**	0.808**
Fourth root	0.659**	0.479**	0.792**
Log(x + 1)	0.576**	0.448**	0.809**
Presence/absence	0.632**	0.434**	0.767**

tively large disparity among genus/species and family estimates of density and richness (Fig. 2b,e) highlights the number of families in the Gulf of Mexico for which species-level descriptions are sorely lacking or absent (e.g. Serranidae, Triglidae, Haemulidae, Sparidae, Lutjanidae). Although the estimates of total density and richness differed in magnitude among the data sets and levels of taxonomic inclusion, the overall trends were very similar. The results for exp H' differed, however, in that family-level estimates were the least sensitive to detecting seasonal changes. This was evident by the significant decrease in genus- and species-level diversity estimates observed in October that was not detected by the family-level estimate (Fig. 2g,h). Unlike the genus:species ratio, the family:genus and family:species ratios for the Gulf of Mexico (0.27 and 0.14, respectively) are relatively poor, so this result is not surprising. Many of the dominant taxa collected in our study were representative of families with relatively large numbers of genera, e.g. Clupeidae (*Brevoortia*, *Etrumeus*, *Harengula*, *Opisthonema*, *Sardinella*), Carangidae (*Caranx*, *Chloroscombrus*, *Decapterus*, *Elagatis*, *Oligoplites*, *Selar*, *Selene*), and Sciaenidae (*Bairdiella*, *Cynoscion*, *Larimus*, *Leiostomus*, *Menticirrus*, *Micropogonias*, *Sciaenops*, *Stel-*

lifer) (Hernandez et al. 2010a,b). Therefore, family-level analysis of ichthyoplankton for this region has a homogenizing effect on diversity indices such as exp *H'*.

The MDS plots obtained when using family-, genus- and species-level data were similar in that all of the graphical representations revealed statistically significant (and similar) patterns of seasonal changes in larval densities (Figs. 3 & 4). The strong statistical correlations between genus- and species-level analyses (Table 3), as well as the relatively low (and similar) stress values from the MDS plots, suggest that genus-level analyses provided the same level of information as the species-level analyses, and therefore may provide a valuable proxy for research programs faced with managing the time constraints involved with ichthyoplankton identifications. The MDS plots from the family-level analyses, in comparison, had relatively poor stress values (>0.2), with the exception of the Stage I analysis with a presence/absence transformation (0.16), yet the graphical resolution of monthly patterns were largely similar to those of the genus- and species-level analyses. Therefore, though not as statistically rigorous, a family-level analysis may be informative as a first-order assessment of ichthyoplankton assemblage structure related to seasonality or other environmental parameters (e.g. Carassou et al. 2012). In contrast to the single taxonomic level approaches, the inclusion of multiple taxonomic levels (Stage III) yielded MDS plots with little resolution in monthly patterns (Fig. 5), and poor statistical performance (stress values ≥0.22 for all transformations). These results are significant, considering the number of ichthyoplankton studies with analyses using mixed taxonomic levels, both in our region and elsewhere.

Overall, our results suggest that for the purpose of descriptions of ichthyoplankton structure, or for 'rapid' assessment of changes in these assemblages in response to pollution or other anthropogenic impacts, family-level estimates should suffice. In our study region, for example, a tremendous amount of resources and time have gone into biological and oceanographic data collection to assess damages to the marine environment and its resources during and after the 2010 Deepwater Horizon oil spill (Lubchenco et al. 2012, NOAA 2012). In order to enact restoration and mitigation measures, studies on the impacts of the Deepwater Horizon oil spill on the structure of marine communities (including ichthyoplankton) need to be addressed in a relatively short time frame, while at the same time covering many aspects of marine community structure and processes. For estimates of impacts on biodiversity, our results have shown that family-level data are not ideal in detecting potential changes in ichthyoplankton diversity, since family-level data may have a homogenizing effect on diversity indices, as discussed above. However, genus-level data are sufficient, since results obtained with genus- and species-level data provided similar statistical performances in every aspect of our analyses. Conversely, for the detection of potential changes in processes affecting the dynamics of ichthyoplankton assemblages, such as interannual and seasonal patterns in abundance and relative composition tested herein, or changes in community structure in response to an impact such as the Deepwater Horizon event, family-level data may be adequate, at least for a first-order analysis. This result is consistent with studies conducted on invertebrate assemblages, which have shown that coarse taxonomic levels are sufficient for detecting the effect of pollution on benthic community structure (Warwick 1988, Ferraro & Cole 1990, Dauvin et al. 2003).

Ultimately, however, analyses conducted at taxonomic levels lower than families may remain a necessary step for the assessment of pollution impacts of moderate or unknown severity. Similar to the detection of fishing effects on invertebrate assemblages, pollution may affect differently the taxa combined within a given family (Jimenez et al. 2010). Also, the detection of the impact of pollution on communities may depend on the severity of this impact, as pollution will affect individual organisms first, and then increasing taxonomic levels as the severity of the pollution increases (Dauvin et al. 2003). Furthermore, family-level taxonomic information will not adequately address impacts to economically-important fishes (e.g. snappers), which have species-specific management plans. Accordingly, a pragmatic approach for addressing the impacts of the Deepwater Horizon oil spill on ichthyoplankton from the northern Gulf of Mexico could include (1) a 'rapid' assessment of variations in overall assemblage structure, based on family-level data, which will help in identifying the groups of larvae responsible for the observed changes, and (2) genus-level analyses (at minimum) for these particular groups, in order to further understand their particular responses to the pollution, the severity of which remains largely unknown so far.

While our results suggest the TS approach has some analytical utility for studies in the Gulf of Mexico, the efficacy of the approach will vary from region to region based on taxonomic richness, and to some

degree, the ratios of higher taxonomic levels to species level. The genus:species ratios for the Mediterranean Sea and North Sea, for example, are approximately 0.61 and 0.81, respectively, which suggest the TS method may be useful for these regions (ratios determined by analyses of marine fish species reported in www.fishbase.org for each region [Froese & Pauly 2013]). In contrast, the genus:species ratio for the species-rich marine waters of Australia is relatively low, approximately 0.33 (Froese & Pauly 2013). The number of described larval forms in this region is also very low, approximately 10% for the tropical Indo-Pacific and 18% for the temperate waters of Australia (Fahay 2007). So although the TS method is not a rigorous approach for ichthyoplankton studies in Australian waters based on the family:species and genus:species ratios, researchers are left with few alternatives but to analyze ichthyoplankton assemblages at higher taxonomic levels (often family level).

In summary, this study provides the first critical examination of the taxonomic sufficiency approach as applied to ichthyoplankton assemblages. Our results suggest that the common approach of analyzing ichthyoplankton assemblages at mixed taxonomic levels (family, genus or species levels combined) is not as statistically rigorous as single taxonomic-level analyses, and should be used with caution, particularly when discerning temporal patterns in density and richness, and to a lesser extent, taxonomic diversity. Results from seasonal (monthly) patterns examined using multidimensional methods were similar among family-, genus- and species-level analyses. Genus- and species-level similarity matrices were highly correlated, which suggests analyses at the genus level could serve as a good proxy for species when examining assemblage diversity, among other community attributes. Overall, our results provide validation of the TS approach for ichthyoplankton in the northern Gulf of Mexico, however, the efficacy of this approach for other regions with different taxa may vary. This study outlines an analytical framework for validating the use of TS approaches for ichthyoplankton in other marine ecosystems, and provides a means of weighing tradeoffs between the level of taxonomic detail needed to address different scientific questions and the time constraints and need for efficient or rapid response to management needs.

Acknowledgements. This study was conducted through the Fisheries Oceanography of Coastal Alabama (FOCAL) program at the Dauphin Island Sea Lab (DISL), which was supported by the Alabama Department of Conservation and Natural Resources (ADCNR). We thank the field and laboratory personnel of the FOCAL program at DISL, as well as all the graduate students who participated on our research cruises. We also especially thank the captains and crew of the RV 'E. O. Wilson' and the DISL technical support team. We are very grateful to M. Konieczna and the scientific staff at the Plankton Sorting and Identification Center for larval fish identifications. H. Fletcher and L. Hu (DISL) provided database management support.

LITERATURE CITED

Bates CR, Scott G, Tobin M, Thompson R (2007) Weighing the costs and benefits of reduced sampling resolution in biomonitoring studies: perspectives from the temperate rocky intertidal. Biol Conserv 137:617–625

Bertasi F, Colangelo MA, Colosio F, Gregorio G, Abbiata M, Ceccherelli VU (2009) Comparing efficacy of different taxonomic resolutions and surrogates in detecting changes in soft bottom assemblages due to coastal defence structures. Mar Pollut Bull 58:686–694

Bevilacqua S, Terlizzi A, Claudet J, Fraschetti S, Boero F (2012) Taxonomic relatedness does not matter for species surrogacy in the assessment of community responses to environmental drivers. J Appl Ecol 49:357–366

Boeing WJ, Duffy-Anderson JT (2008) Ichthyoplankton dynamics and biodiversity in the Gulf of Alaska: responses to environmental change. Ecol Indic 8:292–302

Brodeur RD, Peterson WT, Auth TD, Soulen HL, Parnel MM, Emerson AA (2008) Abundance and diversity of coastal fish larvae as indicators of recent changes in ocean and climate conditions in the Oregon upwelling zone. Mar Ecol Prog Ser 366:187–202

Carassou L, Ponton D (2007) Spatio-temporal structure of pelagic larval and juvenile fish assemblages in coastal areas of New Caledonia, southwest Pacific. Mar Biol 150: 697–711

Carassou L, Hernandez F, Powers S, Graham W (2012) Cross-shore, seasonal and depth-related structure of ichthyoplankton assemblages in coastal Alabama. Trans Am Fish Soc 141:1137–1150

Cardoso P, Silva I, Oliveira NG, Serrano ARM (2004) Higher taxa surrogates of spider (Araneae) diversity and their efficiency in conservation. Biol Conserv 117:453–459

Carneiro FM, Nini LM, Rodrigues LC (2010) Influence of taxonomic and numerical resolution on the analysis of temporal changes in phytoplankton communities. Ecol Indic 10:249–255

Clarke KR (1993) Non parametric multivariate analysis of changes in community structure. Aust J Ecol 18:117–143

Clarke KR, Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. Mar Ecol Prog Ser 92:205–219

D'Alessandro EK, Sponagle S, Serafy JE (2010) Larval ecology of a suite of snappers (family: Lutjanidae) in the Straits of Florida, western Atlantic Ocean. Mar Ecol Prog Ser 410:159–175

Dauvin JC, Gomez Gesteira JL, Salvande Fraga M (2003) Taxonomic sufficiency: an overview of its use in the monitoring of sublittoral benthic communities after oil spills. Mar Pollut Bull 46:552–555

Drass DM, Bootes KL, Lyczkowski-Shultz J, Comyns BH, Holt GJ, Riley CM, Phelps RP (2000) Larval development of red snapper, *Lutjanus campechanus*, and comparisons with co-occurring snapper species. Fish Bull 98:507–527

Duffy-Anderson JT, Busby MS, Mier KL, Deliyanides CM, Stabeno PJ (2006) Spatial and temporal patterns in summer ichthyoplankton assemblages on the eastern Bering Sea shelf 1996–2000. Fish Oceanogr 15:80–94

Ellis D (1985) Taxonomic sufficiency in pollution assessment. Mar Pollut Bull 16:459

Espinosa-Fuentes ML, Flores-Coto C (2004) Cross-shelf and vertical structure of ichthyoplankton assemblages in continental shelf waters of the Southern Gulf of Mexico. Estuar Coast Shelf Sci 59:333–352

Fahay MP (2007) Early stages of fishes in the western North Atlantic Ocean, Vol I. Northwest Atlantic Fisheries Organization, Dartmouth, NS

Faria A, Morais P, Chícharo MA (2006) Ichthyoplankton dynamics in the Guadiana estuary and adjacent coastal area, south-east Portugal. Estuar Coast Shelf Sci 70: 85–97

Ferraro SP, Cole FA (1990) Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California Bight macrobenthos. Mar Ecol Prog Ser 67: 251–262

Froese R, Pauly D (2013) FishBase. www.fishbase.org (accessed 1 Jul 2013)

Gomez Gesteira JL, Dauvin JC, Salvande Fraga M (2003) Taxonomic level for assessing oil spill effects on soft-bottom sublittoral benthic communities. Mar Pollut Bull 46:562–572

Grelle CEV (2002) Is higher-taxon analysis a useful surrogate of species richness in studies of neotropical mammal diversity? Biol Conserv 108:101–106

Hernandez FJ, Powers SP, Graham WM (2010a) Detailed examination of ichthyoplankton seasonality from a high-resolution time series in the northern Gulf of Mexico during 2004–2006. Trans Am Fish Soc 139:1511–1525

Hernandez FJ, Powers SP, Graham WM (2010b) Seasonal variability in ichthyoplankton abundance and assemblage composition in the northern Gulf of Mexico off Alabama. Fish Bull 108:193–207

Jimenez H, Dumas P, Bigot L, Amouroux JM, Ferraris J (2010) Taxonomic resolution needed to describe invertebrate assemblages and to detect harvesting effects on coral reef ecosystems. Mar Ecol Prog Ser 406:211–222

Jost L (2006) Entropy and diversity. Oikos 113:363–375

Karakassis I, Hatziyanni E (2000) Benthic disturbance due to fish farming analyzed under different levels of taxonomic resolution. Mar Ecol Prog Ser 203:247–253

Lindeman KC, Richards WJ, Lyczkowski-Shultz J, Drass DM and others (2005) Lutjanidae: snappers. In: Richards WJ (ed) Early stages of Atlantic fishes: an identification guide for the western Central North Atlantic, Vol II. CRC Press, Boca Raton, FL, p 1549–1586

Lubchenco J, McNutt MK, Dreyfus G, Murawski SA and others (2012) Science in support of the Deepwater Horizon response. Proc Natl Acad Sci USA 109: 20212–20221

Luthy SA, Cowen RK, Serafy JE, McDowell JR (2005) Toward identification of larval sailfish (*Istiophorus platypterus*), white marlin (*Tetrapturus albidus*), and blue marlin (*Makaira nigricans*) in the western North Atlantic Ocean. Fish Bull 103:588–600

Lyczkowski-Shultz J, Bond P (2005) Paralichthyidae: sand flounders. In: Richards WJ (ed) Early stages of Atlantic fishes: an identification guide for the western Central North Atlantic, Vol II. CRC Press, Boca Raton, FL, p 2291–2326

Marancik KE, Clough LM, Hare JA (2005) Cross-shelf and seasonal variation in larval fish assemblages on the southeast United States continental shelf off the coast of Georgia. Fish Bull 103:108–129

Marancik KE, Richardson DE, Lyczkowski-Shultz J, Konieczna M, Cowen RK (2010) Evaluation of morphological characters to identify grouper (Serranidae: Epinephelini) larvae in the Gulf of Mexico using genetically identified specimens. Bull Mar Sci 86:1–54

McEachran JD (2009) Fishes (Vertebrata: Pisces) of the Gulf of Mexico. In: Felder DL, Camp DK (eds) Gulf of Mexico origin, waters, and biota: Vol 1, Biodiversity. Texas A&M University Press, College Station, TX, p 1223–1316

Mendes CLT, Tavares M, Soares-Gomes A (2007) Taxonomic sufficiency for soft-bottom sublittoral mollusks assemblages in a tropical estuary, Guanabara Bay, Southeast Brazil. Mar Pollut Bull 54:377–384

Muhling BA, Beckley LE, Koslow JA, Pearce AF (2008) Larval fish assemblages and water mass structure off the oligotrophic south-western Australian coast. Fish Oceanogr 17:16–31

Muhling BA, Lamkin JT, Richards WJ (2012) Decadal-scale responses of larval fish assemblages to multiple ecosystem processes in the northern Gulf of Mexico. Mar Ecol Prog Ser 450:37–53

NOAA (2012) Natural resource damage assessment April 2012: Status update for the Deepwater Horizon oil spill. www.gulfspillrestoration.noaa.gov/wp-content/uploads/FINAL_NRDA_StatusUpdate_April2012.pdf

Olsgard F, Somerfield PJ, Carr MR (1998) Relationships between taxonomic resolution, macrobenthic community patterns and disturbance. Mar Ecol Prog Ser 172:25–36

Pik AJ, Oliver I, Beattie AJ (1999) Taxonomic sufficiency in ecological studies of terrestrial invertebrates. Aust J Ecol 24:555–562

Powell AB, Greene MD (2005) Sparidae: porgies. In: Richards WJ (ed) Early stages of Atlantic fishes: an identification guide for the western Central North Atlantic, Vol II. CRC Press, Boca Raton, FL, p 1651–1668

Puente A, Juanes JA (2008) Testing taxonomic resolution, data transformation and selection of species for monitoring macroalgae communities. Estuar Coast Shelf Sci 78: 327–340

Quattrini AM, Lindquist DG, Bingham FM, Lankford TE, Govoni JJ (2005) Distribution of larval fishes among water masses in Onslow Bay, North Carolina: implications for cross-shelf exchange. Fish Oceanogr 14:413–431

Rakocinski CF, Lyczkowski-Shultz J, Richardson SL (1996) Ichthyoplankton assemblage structure in Mississippi Sound as revealed by canonical correspondence analysis. Estuar Coast Shelf Sci 43:237–257

Richards WJ (2005) Introduction. In: Richards WJ (ed) Early stages of Atlantic fishes: an identification guide for the western Central North Atlantic, Vol I. CRC Press, Boca Raton, FL, p 1–70

Richards WJ, Baldwin CC, Röpke A (2005) Serranidae: sea basses. In: Richards WJ (ed) Early stages of Atlantic fishes: an identification guide for the western Central North Atlantic, Vol I. CRC Press, Boca Raton, FL, p 1225–1331

Richardson DE, Vanwye JD, Exum AM, Cowen RK, Crawford DL (2007) High-throughput species identification: from DNA isolation to bioinformatics. Mol Ecol Notes 7: 199–207

Smale DA, Kendrick GA, Wernberg T (2010) Assemblage

turnover and taxonomic sufficiency of subtidal macroalgae at multiple spatial scales. J Exp Mar Biol Ecol 384: 76–86

Suthers IM, Taggart CT, Rissik D, Baird ME (2006) Day and night ichthyoplankton assemblages and zooplankton biomass size spectrum in a deep ocean island wake. Mar Ecol Prog Ser 322:225–238

Syahailatua A, Roughan M, Suthers IM (2011) Characteristic ichthyoplankton taxa in the separation zone of the East Australian Current: larval assemblages as tracers of coastal mixing. Deep-Sea Res II 58:678–690

Tolan JM, Holt SA, Onuf CP (1997) Distribution and community structure of ichthyoplankton in Laguna Madre seagrass meadows: potential impact of seagrass species change. Estuaries 20:450–464

Vandersea MW, Lem S, Litaker RW, West MA and others (2008) Identification of larval sea basses (*Centropristis* spp.) using ribosomal DNA-specific molecular assays. Fish Bull 106:183–193

Vásquez-Yeomans L (2000) Seasonal variation of ichthyoplankton in a western Caribbean bay system. Environ Biol Fishes 58:379–392

Warwick RM (1988) The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar Pollut Bull 19:259–268

Xu H, Jiang Y, Zhang W, Zhu M, Al-Rasheid KAS (2011) An approach to determining potential surrogates for analyzing ecological patterns of planktonic ciliate communities in marine ecosystems. Environ Sci Pollut Res 18: 1433–1441

Editorial responsibility: Stylianos Somarakis, Heraklion, Greece

Submitted: March 28, 2013; Accepted: July 9, 2013
Proofs received from author(s): September 16, 2013