Proceedings of the Royal Society B, 2015

Electronic Supplementary Material

Differential responses of marine communities to natural and anthropogenic changes

Michał Kowalewski[1], Jacalyn M. Wittmer[2], Troy A. Dexter[1], Alessandro Amorosi[3] and Daniele Scarponi[3]

[1]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA e-mail: kowalewski@ufl.edu

[2]Department of Geology, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA

[3]Dipartimento di Scienze Biologiche, Geologiche e Ambientali, University of Bologna, Bologna, 40126, Italy

ESM-Data S1

Study Area

Due to high rates of tectonic subsidence, the Quaternary succession of the Po Basin (Northern Italy) represents a spectacular example of an expanded stratigraphic package. At proximal locations (Po Plain), the upper Quaternary succession of the Po Basin is made up entirely of alluvial facies that grades distally (Po coastal plain and Po delta) into characteristic alternations of shallow-marine and alluvial deposits. This vertically cyclic pattern of facies, which has been documented to fall in the Milankovitch (100 ka) band [1], has been investigated in detail in the last 15 years via integrated studies of several cores up to 200 m in length. Consequently, a considerable amount of data is now available for the upper Pleistocene to Recent succession. Detailed geologic information include facies analysis [1–3], sequence stratigraphy [3–5], benthic mollusks [6–8], micropaleontology [9–10], pollen profiles [1–3], geochemistry [11–12], paleosol stratigraphy [13], and geochronology [4]. The general chronologic framework relies upon a variety of data, including pollen [1], radiocarbon and amino acid racemization dating [4], and electron spin resonance [14].

The allostratigraphic approach to the stratigraphy of the Po Basin has had a profound impact on the official geological cartography of Italy by motivating the subdivision of the Quaternary succession into vertically stacked 'synthems', bounded by stratigraphic unconformities of tectonic origin (i.e., third-order depositional sequences). These, in turn, include 'sub-synthems', corresponding to glacio-eustatically controlled, fourth-order depositional cycles. The internal stratigraphic architecture of the individual sub-synthems can be tracked from the fluvial to the marine realm due to identification and lateral tracking of prominent bounding surfaces that represent transgressive surfaces.

Two prominent stratigraphic markers are recorded in the subsurface of the Po coastal plain. These wedge-shaped coastal sedimentary bodies are located at approximately 0-30 m and 95-120 m core depth, respectively (figure 1). They show comparable transgressive-regressive stacking patterns and were deposited during the two major transgressive pulsations and subsequent sea-level highstands of the last 150 ky.

ESM-Table S1. Summary of the datasets. The final dataset excludes unidentified species and non-marine samples. Small samples and rare taxa are retained. Apparent extirpations represent the number of Late Pleistocene species not reported in the Holocene. Apparent originations (new species arrivals) represent the number of Holocene species not found in the Late Pleistocene samples. Singletons are species found in one of the two time intervals only. Two-timers are species found in both time intervals.

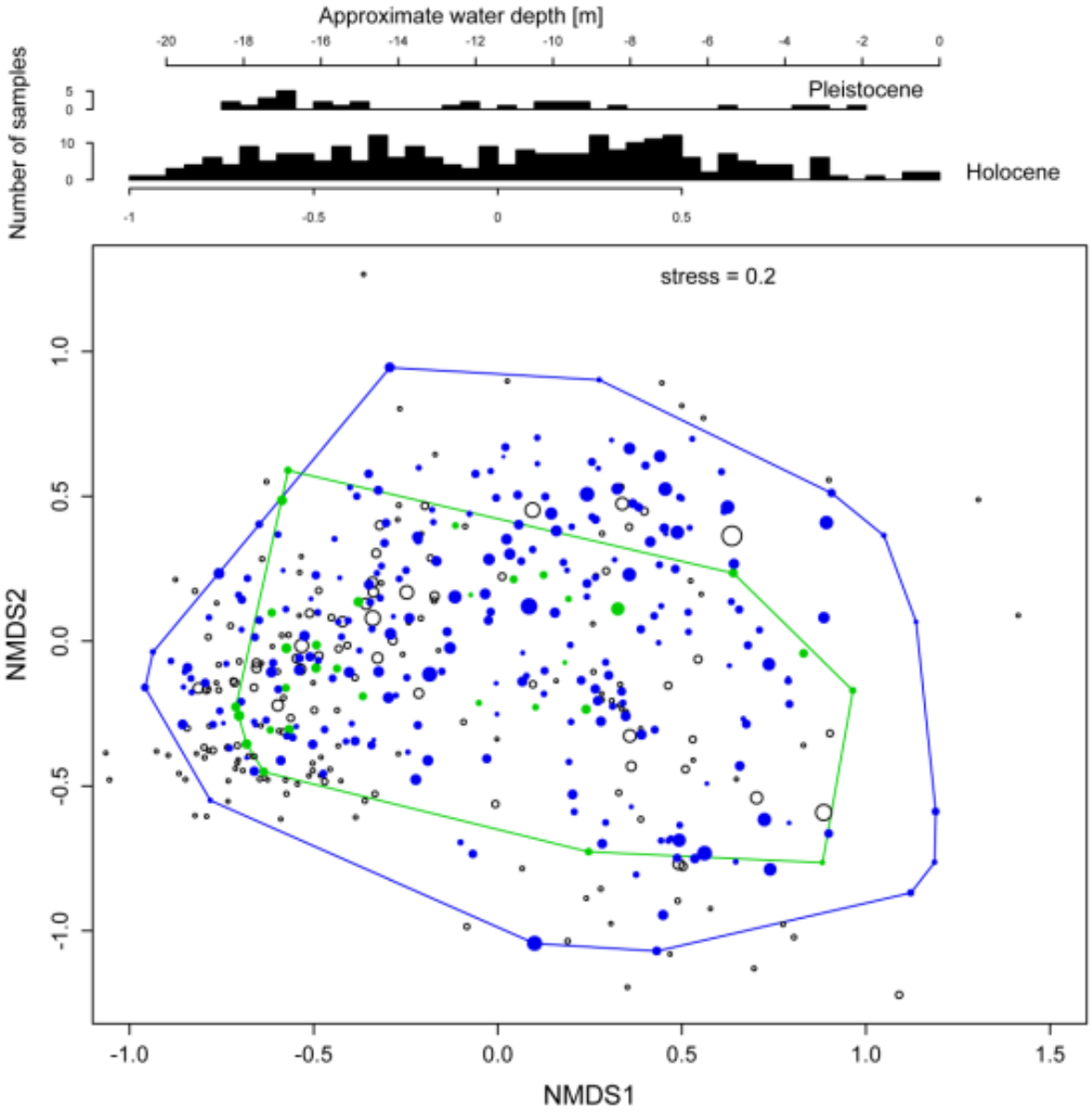
Variable	Total dataset	Final Dataset (excluding non-marine samples)	Final Dataset (Holocene)	Final Dataset (Holocene-HST)	Final Dataset (Late Pleistocene)	Modern Dataset
Number of samples	611	453	414	263	39	78
Number of species	333	221	210	129	97	91
Number of specimens	131780	125558	110387	72911	15171	91552
Mean sample size	215.7	277.2	266.6	277.2	389.0	1173.7
Number of apparent extirpations	-----	-----	11	-----	-----	-----
Number of apparent originations	-----	-----	124	-----	-----	-----
Number of singletons	-----	135	-----	-----	-----	-----
Number of two-timers	-----	86	-----	-----	-----	-----

These two sand bodies, assigned respectively to the Marine Isotope Stage (MIS) 5e and the Holocene, are separated by a thick package of alluvial sediments. The stratigraphic architecture of post-MIS 5e deposits shows consistent patterns of coastal development with changing sea-level position. Lowering of sea level between ca. 116 and 70 ky (MIS 5d to 4) resulted in extensive and repeated basinward shifts of facies, which are recorded by closely-spaced unconformity surfaces developed distally (MIS 5) followed by alluvial sedimentation (MIS 4). This prolonged phase of sea-level fall was punctuated by short transgressive pulses, which led to widespread deposition of thin packages of organic-rich, lagoonal (MIS 5c) and swamp (MIS 5a and subordinately MIS 3) deposits. Uppermost Pleistocene (MIS 2) deposits are either lacking or represented by a laterally extensive fluvial channel belt. The Pleistocene/Holocene boundary is marked by a prominent hiatal surface, which is overlain by the onlapping Holocene transgressive deposits.

The two transgressive surfaces bounding the upper Pleistocene succession are paralleled invariably by distinctive changes in pollen spectra [1]. A tight relationship is observed between initial transgression and the development of mixed, broad-leafed vegetation, indicating the onset of warm (interglacial) periods. Similarly, good correlation exists between middle and upper parts of T-R sequences and the development of cold-climate vegetation. Due to their combined depositional facies and pollen signature, the bounding surfaces of the two stratigraphic markers (see above) in the Po Basin can be easily identified from core data. These surfaces are interpreted to represent the MIS 6/5e and MIS 2/1 transitions, respectively.

Core Sample Processing

A total of 611 comparable bulk samples were acquired from a network of 16 cores drilled on the Po Plain (figure 1). Each sample represents a 5 cm core interval (~375 cm^3 of sediment). Samples were soaked in ~4% H_2O_2 (≤ 4 hours, depending on lithology) and dried for 24 hours at 45°C. For some samples (massive-clay), this process was repeated at least two times. The resulting processed sediment was wet sieved using 1 mm screens. For each sample all mollusk specimens were identified to the species level (when possible) and counted. Only complete fossils or fragments which could be identified as a unique individual (e.g., apex for gastropods or umbo for bivalves) were counted. In the case of bivalves, each valve or unique fragment was counted as a 0.5 specimen. Other macrobenthic fossils (e.g., serpulids, crustaceans) were very rare. The raw dataset produced from the cores included 131,780 specimens overwhelmingly dominated by bivalves and gastropods (table S1).



ESM-Fig. S1. - Indirect ordination (nMDS, Bray-Curtis distance, k=2 dimensions) derived using "metaMDS" function in "Vegan" package [28] in *R* [27]. Although the stress value is high, nMDS ordinations derived for three dimensions yielded comparable results. The results are also consistent with previous indirect ordination analyses based on Correspondence Analysis and Detrended Correspondence Analysis, [8]. Black open circles represent species and solid circles represent samples (blue – Holocene, green – Late Pleistocene). The size of each point depicts sample size (number of specimens representing a given species or a given sample). Bar charts summarize the bathymetric distribution of samples for the Late Pleistocene and Holocene datasets. Approximate water depth (top axis) based on previous ordination analyses of the Po Plain data calibrated *a posteriori* against water depth estimates obtained from modern populations of mollusk species that were most common in the cores [7–8]. Small samples and rare taxa were excluded from this analysis.

Data Filtering

Core samples included numerous samples representing proximal settings dominated by non-marine mollusks. These samples were removed prior to analyses to restrict the study to the marine part of the ecosystem. Also, taxa that could not be unambiguously identified to a unique genus and species ("indet.", "spp.") were removed from the dataset. These taxa were typically represented by one or a few specimens and their removal did not reduce the dataset noticeably. In addition, multiple species included specimens with uncertain species identification (cf.). This creates additional species and may also result in apparent extirpations and originations. For final analyses these species were treated as correctly identified to prevent an inflation of diversity and turnover rates across the compared datasets. However, when the "cf." designation was retained, analytical outcomes did not differ in any substantial way from those reported here. All analyses reported here are based on the restricted, species level dataset with "cf." designations ignored.

In addition, for analyses in which individual sample sizes and species rarity can generate substantial volatility and extreme outliers (multivariate ordinations, gradient beta diversity, etc.), the data were further filtered by removing small samples (n<20) and taxa that occurred in only one sample. However, all marine samples and all rare taxa were included in analyses for which rare taxa are meaningful and size of individual samples does not directly impact the analysis (e.g., rank abundance distribution [RAD] of species for pooled data). The exclusion or inclusion of small samples and rare taxa is indicated in table and figure captions.

Modern Dataset

To evaluate the present-day communities, a numerical dataset was downloaded from the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) database: www.santateresa.enea.it. The database contains ~20,000 records for 901 species sampled at 663 localities [15]. These data are based on quantitative bulk samples obtained in surveys (years: 1958-1998) of marine seafloors in the region. To make these data maximally comparable with the core data, the samples were restricted to a water depth ≤20 m (the bathymetric range represented by the restricted core samples used in the final analyses). Also, dead and live specimens obtained at each site were both included in specimen counts. The inclusion of dead specimens makes those samples more comparable to time-averaged core samples. The inclusion of dead specimens makes the dataset also more likely to show ecological congruence with Holocene core data by potential encompassing older specimens that predate extensive industrialization of the region starting in the early 20th century. This is a conservative approach given the outcome of the study points to substantial shift despite this potential partial overlap in the age range of the compared datasets. In addition, taxonomic names were synonymized to make them consistent with the nomenclature used in the core dataset. Modern dataset was used in key comparisons while employing the same methodology as outlined above for the core data. The only analytical exception includes depth estimates: in the case of Modern dataset, the water depth is known directly (as recorded during the surveys). All the above criteria aim at maximizing similarities between the Holocene and Modern datasets, which makes the notable differences reported here even more noteworthy. The present-day dataset is referred to in text and figures as "Modern".

Multivariate Ordinations and Bathymetric Estimates

The Po Plain dataset represents a multivariate data matrix that can be explored using dimension-reducing indirect ordination techniques. Here, Nonmetric Multidimensional Scaling [NMDS] has been used. Data were standardized by double relativization by standardizing sequentially for both species and sample abundance [16]. Bray-Curtis similarity, a measure of pairwise distance often recommended for ecological ordinations [17], was used to develop NMDS ordinations. The NMDS ordinations fitted into k=2 and k=3 dimensions yielded comparable outcomes. Only the results for k=2 are reported here (figure S1). Other indirect ordination techniques, including Correspondence Analysis, and Detrended Correspondence Analysis yielded ordination outcomes comparable to the NMDS results [8].

The previous indirect ordination analyses of the Po core data [7–8] revealed a pronounced bathymetric gradient, with species ordinated along the first ordination axes according to their preferred bathymetry. The species preferred bathymetry was obtained from the ENEA database (see above). The recently published analyses [8], based on Weighted Averaging [16] and Detrended Correspondence [18] approaches, demonstrated *a*

posteriori that ordination scores are a robust predictor of water depth for both samples and species. The sample water depth estimates reported previously [8] have been used here to place all core samples along the bathymetric gradient (figures 3, S1). Note that this analysis was not applied to the Modern samples, because water depth was measured directly for samples collected from present-day seafloors.

Diversity Analyses

Diversity analyses have been conducted at multiple scales. For sample-level (alpha) analyses, species richness was estimated using sampling standardization with all samples subsampled without replacement down to 20 specimens. Dominance was estimated using Berger-Parker index (i.e., relative abundance of the most common taxon). Because standardized species richness at small sample size tends to be driven primarily by evenness, the two above metrics are inversely correlated and somewhat redundant. In general, for mollusk-dominated fossil samples, standard metrics of diversity, evenness and dominance tend to be strongly correlated [19], so the choice of a specific diversity/evenness measure used for data interpretations tends to be inconsequential.

To explore beta diversity within each of the two time intervals, the turnover beta diversity approach was used here [20], where pairwise distances between samples (measured here as Bray-Curtis similarity) were evaluated as a function of "environmental distance" between those sample pairs [21]. Difference in the estimated sample water depth was used as a measure of "environmental distance" (figure 2). In addition, distance-decay exponential models were fit to the data to evaluate turnover rate. The parameter λ was used to compare turnover rates expressed as δY/δX: a ratio between change in environmental distance and change in pairwise similarity [20].

Overall (gamma) diversity for each time interval was estimated for pooled data, with rare species and small samples included. For Late Pleistocene-Holocene comparisons, the total species richness of the time interval and observed extirpations (here synonymous with the number of Late Pleistocene singletons) and origination rates (here synonymous with the number of Holocene singletons) are affected by unbalanced sampling, with Holocene samples being much more numerous but smaller in terms of mean sample size. Resampling models (discussed below) were applied to evaluate the effect of the unbalanced sampling. Rank abundance distribution [RAD] analyses were employed to compare the Late Pleistocene, Holocene, and Modern RADs. Maximum likelihood was used to determine most suitable RAD model for each time interval: both Akaike [AIC] and Bayesian [BIC] information criteria yielded consistent outcomes for selecting the best model. The Late Pleistocene, Holocene, and Modern RADs were compared using Kolmogorov-Smirnov D and the correlation between rank abundances of the two time intervals was measured using Spearman Rank Correlation calculated with singleton species either included or excluded.

Multivariate Tests

A series of permutation-based multivariate methods were employed to evaluate two comparisons: (1) the Late Pleistocene versus Holocene dataset and (2) the Holocene-HST versus Modern dataset. The three conceptually related permutation-based methods, which included Mantel test [22], ANOSIM [23], and PERMANOVA [24], were carried out here to evaluate multivariate differences between the compared pairs of datasets (table S5). Mantel tests were based on distances matrices of pairwise species similarities across samples and used to measure the strength of relationship between the species with sample sets grouped by database. The higher value of the Mantel statistic r_M is potentially indicative of stronger relationships between data matrices. ANOSIM was based on ranks of pairwise distances between samples. The difference between the mean rank for between-dataset comparisons and the mean rank for within-dataset comparisons, scaled by number of pairwise comparisons, was used as a measure of multivariate difference between the two datasets. The higher value of ANOSIM R_A suggests greater multivariate differences between the two datasets. PERMANOVA, similarly based on within-dataset versus between-dataset sample distances, was used to compute sums of squares of distances and compute pseudo F ratio. The higher value of F suggests greater multivariate differences between the two datasets. This measure tends to be less influenced by differences in dispersion relative to ANOSIM R_A [25]. All three measured statistics (table S5) consistently suggest that that the multivariate differences between the Holocene-HST and Modern dataset are notably greater than those observed between the Pleistocene and Holocene dataset. All three methods are used here only

ESM-Table S2 (Part 1). – Comparison of Late Pleistocene and Holocene datasets (summary of simulation outcomes, 10000 iterations each). For the Randomization Model (RDM) simulations are based on random reassignment of sample membership without replacement (RDM model). For the Subsampling Model simulations are based on random subsampling without replacement of Holocene samples to mimic the number of Late Pleistocene samples. The Adjusted Randomization Model (ARDM) was derived from RDM simulations to adjust for differences in mean sample size. Small samples and rare taxa were retained in these analyses. See text above for details.

Parameter	Observed value	Number of iterations	Simulated mean	*p* value	Min	Q25	Median	Q75	Max
Randomization Model [RDM]									
Total number of specimens (Holocene)	110387.00	10000	114753.88	0.183	80556.00	112302.75	116389.00	119169.00	123789.00
Total number of species (Holocene)	210.00	10000	214.93	0.103	197.00	213.00	215.00	217.00	221.00
Total number of specimens (Pleistocene)	15171.00	10000	10367.87	0.166	1632.00	6010.00	8638.50	12677.25	40958.00
Total number of species (Pleistocene)	97.00	10000	85.15	0.167	40.00	77.00	85.00	93.00	131.00
Number of apparent species extirpations	11.00	10000	6.07	0.103	0.00	4.00	6.00	8.00	24.00
Number of apparent species originations	124.00	10000	135.85	0.167	90.00	128.00	136.00	144.00	181.00
Number of singleton	135.00	10000	136.37	0.467	90.00	128.00	136.00	144.00	181.00
Number of two-timers	86.00	10000	84.63	0.467	40.00	77.00	85.00	93.00	131.00
Most abundant apparently extirpated species	4.00	10000	3.80	0.415	0.00	2.00	3.00	4.00	95.00
Mean specimen number (extirpated species)	1.46	10000	1.76	0.427	0.00	1.25	1.50	2.00	48.00
Mean sample size (Holocene)	266.64	10000	277.18	0.183	194.58	271.26	281.13	287.85	299.01
Mean sample size (Pleistocene)	389.00	10000	265.84	0.166	41.85	154.10	221.50	325.06	1050.21
Zipf (param. 1) (Pleistocene)	0.58	10000	0.59	0.430	0.26	0.51	0.60	0.67	0.92
Zipf (gamma) (Pleistocene)	-1.90	10000	-1.99	0.432	-3.98	-2.19	-1.95	-1.71	-1.17
AIC (Pleistocene)	2042.58	10000	1707.21	0.245	218.92	691.76	1087.59	2010.33	9565.49
Kolmogorov-Smirnov D	0.08	10000	0.16	0.013	0.05	0.14	0.16	0.19	0.31
Subsampling Model (SSM)									
Total number of specimens (Holocene)	110387.00	-----	-----	-----	-----	-----	-----	-----	-----
Total number of species (Holocene)	210.00	-----	-----	-----	-----	-----	-----	-----	-----
Total number of specimens (Pleistocene)	15171.00	10000	10388.11	0.170	1557.00	5933.75	8611.50	12755.50	43947.00
Total number of species (Pleistocene)	97.00	10000	85.18	0.170	46.00	77.00	85.00	93.00	129.00
Number of apparent species extirpations	11.00	-----	-----	-----	-----	-----	-----	-----	-----
Number of apparent species originations	124.00	10000	135.82	0.170	92.00	128.00	136.00	144.00	175.00
Number of singleton	135.00	10000	135.82	0.486	92.00	128.00	136.00	144.00	175.00
Number of two-timers	86.00	10000	85.18	0.486	46.00	77.00	85.00	93.00	129.00
Most abundant apparently extirpated species	4.00	-----	-----	-----	-----	-----	-----	-----	-----
Mean specimen number (extirpated species)	1.46	-----	-----	-----	-----	-----	-----	-----	-----
Mean sample size (Holocene)	266.64	-----	-----	-----	-----	-----	-----	-----	-----
Mean sample size (Pleistocene)	389.00	10000	260.86	0.170	32.46	146.35	216.60	322.02	1165.31
Zipf (param. 1) (Pleistocene)	0.58	10000	0.61	0.413	0.26	0.53	0.62	0.69	0.92
Zipf (gamma) (Pleistocene)	-1.90	10000	-2.05	0.415	-3.96	-2.28	-2.01	-1.76	-1.13
AIC (Pleistocene)	2042.58	10000	1642.57	0.261	253.82	636.14	1014.57	1929.59	9917.94
Kolmogorov-Smirnov D	0.08	10000	0.17	0.011	0.05	0.15	0.17	0.19	0.33
Adjusted Randomization Model [ARDM]									
Total number of specimens (Holocene)	110387.00	4728	114792.58	0.182	84981.00	112373.00	116438.50	119180.75	123668.00
Total number of species (Holocene)	210.00	4728	214.91	0.103	197.00	213.00	215.00	217.00	221.00
Total number of specimens (Pleistocene)	15171.00	4728	15173.10	0.349	9009.00	10707.75	13004.50	18299.50	40958.00
Total number of species (Pleistocene)	97.00	4728	86.88	0.210	40.00	79.00	87.00	95.00	127.00
Number of apparent species extirpations	11.00	4728	6.09	0.103	0.00	4.00	6.00	8.00	24.00
Number of apparent species originations	124.00	4728	134.12	0.210	94.00	126.00	134.00	142.00	181.00
Number of singleton	135.00	4728	134.67	0.509	94.00	127.00	135.00	142.00	181.00
Number of two-timers	86.00	4728	86.33	0.509	40.00	79.00	86.00	94.00	127.00
Most abundant apparently extirpated species	4.00	4728	3.83	0.420	0.00	2.00	3.00	4.00	95.00
Mean specimen number (extirpated species)	1.46	4728	1.76	0.428	0.00	1.25	1.50	2.00	48.00
Mean sample size (Holocene)	266.64	4728	277.28	0.182	205.27	271.43	281.25	287.88	298.72
Mean sample size (Pleistocene)	389.00	4728	389.05	0.349	231.00	274.56	333.45	469.22	1050.21
Zipf (param. 1) (Pleistocene)	0.58	4728	0.68	0.111	0.43	0.61	0.66	0.74	0.92
Zipf (gamma) (Pleistocene)	-1.90	4728	-2.27	0.111	-3.98	-2.47	-2.17	-1.99	-1.54
AIC (Pleistocene)	2042.58	4728	2689.71	0.487	343.63	1354.28	2094.60	3741.44	9565.49
Kolmogorov-Smirnov D	0.08	4728	0.16	0.015	0.05	0.14	0.16	0.18	0.31

ESM-Table S2 (Part 2). – Comparison of the Holocene-HST and Modern dataset using the Randomization Model (RDM) (10000 iterations each). See table caption for part 1 of this table above and text for details.

Parameter	Observed value	Number of iterations	Simulated mean	*p* value	Min	Q25	Median	Q75	Max
Randomization Model [Modern Data]									
Total number of specimens (Holocene HST)	72911.00	10000	126897.92	<0.0001	79092	119843.8	128009	134871.5	154636
Total number of species (Holocene HST)	129.00	10000	156.05	<0.0001	139	154	156	159	167
Total number of specimens (Modern)	91552.00	10000	37565.08	<0.0001	9827	29591.5	36454	44619.25	85371
Total number of species (Modern)	91.00	10000	102.58	0.0760	73	97	103	108	128
Number of apparent species extirpations	76.00	10000	64.42	0.0760	39	59	64	70	94
Number of apparent species originations	38.00	10000	10.95	<0.0001	0	8	11	13	28
Number of singleton	114.00	10000	75.37	<0.0001	55	71	75	79	100
Number of two-timers	53.00	10000	91.63	<0.0001	67	88	92	96	112
Most abundant apparently extirpated species	451.00	10000	47.42	0.0002	4	18	24	95	1143
Mean specimen number (extirpated species)	15.07	10000	2.53	0.0001	0.88	1.82	2.25	3.27	18.13
Mean sample size (Holocene HST)	277.23	10000	482.5	<0.0001	300.73	455.68	486.73	512.82	587.97
Mean sample size (Modern)	1173.74	10000	481.6	<0.0001	125.99	379.38	467.36	572.04	1094.50
Zipf (param. 1) (Modern)	0.43	10000	0.54	0.0327	0.36	0.49	0.54	0.59	0.80
Zipf (gamma) (Modern)	-1.53	10000	-1.82	0.0303	-2.76	-1.94	-1.79	-1.67	-1.39
AIC (Modern)	20528.08	10000	6240.7	0.0100	493.63	2858.48	4944.55	8425.85	30353.64
Kolmogorov-Smirnov D	0.45	10000	0.16	<0.0001	0.03	0.13	0.16	0.18	0.29

for the purpose of the above relative comparison, and not to evaluate statistical hypotheses. Reported *p* values are suspect for various reasons. First, the use of the same variables in compared datasets is inappropriate for Mantel test [17] and all tests are differentially sensitive to differences in dispersion, unbalanced sampling, and other spurious effects [25]. The Gower distance, a measure deemed effective in community studies [26], was used in all tests. However, the reported outcome (table S5) remained qualitatively consistent when other common distance metrics and various data standardization procedures were applied.

Resampling Models

To evaluate effects of difference in sampling intensity, four models were used to develop estimates predicted under the null hypothesis that compared datasets came from a single underlying ecosystem.

1. Randomization Model [RDM] was based on random reassignment of sample membership. We will use here the Holocene-Pleistocene comparison as an example of how this model was implemented. The Holocene and Late Pleistocene samples were pooled together and then reassigned randomly and without replacement to the two analyzed time intervals. Subsampling with replacement (bootstrapping) is not appropriate here because such an approach would underestimate the total diversity of the system (i.e., removal of samples can only result in loss of species, but duplication cannot increase diversity). In every iteration 39 samples were assigned to the Late Pleistocene and 414 to the Holocene to correctly reflect the sample structure of the actual data. All relevant parameters (e.g., total diversity, RAD coefficients, number of singletons) were computed at each iteration. Sampling distributions for all parameters of interests (table S2) were constructed based on 10,000 iterations (preliminary analyses indicate that standard errors and *p* values stabilized well below 10,000 iterations). *P* values were computed using the percentile approach. Because all tests evaluate the same dataset and are thus highly dependent, Bonferroni correction was not applied. However, most of the statistical decisions would not have changed even if the stringent Bonferroni correction had been applied. The medians of the sampling distributions were used to estimate the parameter value expected under the null hypotheses that the Holocene and Late Pleistocene samples represent the same ecosystem. RDM approach corrects for difference in the number of samples, but not for difference in number of specimens. Because most samples are Holocene, randomized sets of samples mimicking the Late Pleistocene dataset have an average sample size that approximates the mean sample size of the actual Holocene dataset. The actual Late Pleistocene samples have a notably higher average sample size (table S2). Consequently, due to sample-level undersampling, the RDM estimates may underestimate extirpation rates, overestimate origination rates, and underestimate the gamma diversity of the Pleistocene samples (this problem is addressed by the third and fourth model explained below). The same RDM protocol was applied to compare Holocene-HST and Pleistocene samples.

2. For Holocene-Late Pleistocene comparison, Subsampling Model [SSM] was based on random subsetting of the larger Holocene samples. A subset of Holocene samples was reassigned randomly and without replacement to the Late Pleistocene dataset. In each iterative run, 39 samples were assigned to Late Pleistocene and compared against the actual 414 Holocene samples to correctly reflect the sample structure of the actual data. All subsequent steps followed the protocol described above for RDM. The SSM simulates a model under the null hypothesis that Late Pleistocene samples came from the Holocene ecosystem and is thus subtly different from the RDM model. The SSM resampling model is also more conservative than the RDM model. First, it excludes Late Pleistocene singletons. However, given that there are only 11 Late Pleistocene singletons, which represent a total of 16 specimens, the effect of the exclusion of Late Pleistocene singletons is expected to be minimal. Second, it excludes all Late Pleistocene samples from simulations, and consequently, replicate "Pleistocene" datasets cannot incorporate any of the actual Late Pleistocene samples. As was the case for RDM, the SSM model corrects only for difference in number of samples and not average number of specimens. The SSM and RDM models are remarkably consistent further pointing to extreme homogeneity of the two datasets. Because SSM model subsets Holocene samples, some parameters that can be computed for RDM cannot be estimated under this model (table S2). For example, the number of extirpations has to be 0 because simulated Late Pleistocene datasets are subsets of the Holocene dataset and cannot include any species absent in the Holocene.

3. Adjusted Randomization Model [ARDM] was implemented to correct for both uneven number of samples and uneven mean sample size. ARDM was derived by modifying the output of the RDM simulation described above. The 10000 replicate samples generated in RDM were filtered out by removing iterations for which the average sample size of 39 "Late Pleistocene" samples was dramatically smaller than that observed for the actual Late Pleistocene data. The process was continued iteratively by increasing the threshold sample size until the mean sample size pooled across all remaining iterations approximated the observed mean. The resulting subset retained 4814 of the original 10000 replicate samples. Note that this protocol does not remove small samples from the simulated datasets, but rather removes entire replicate datasets with small mean sample size. Consequently, the retained "Late Pleistocene" datasets have sample size variance comparable to sample size variance observed in the actual dataset. The results of these simulations are consistent with those for RDM and SSM, although the offset between the model predictions and the observed values tend to be smaller. For example, ARDM model predicts that under the null model 92.8 species should be recovered, on average, from the Late Pleistocene samples. In contrast, RDM's estimate is 90.6 species and SSM's estimate is 90.5 species, which represents a more notable (if still small) offset from the observed value of 98 species. These differences are relatively minor and qualitatively trivial. Consequently, the three models are reported for the Holocene-Late Pleistocene comparisons, only RDM estimates are reported here for the comparison of the Holocene-HST and Modern datasets (table S2).

4. Double Subsampling Standardization Model [DSM] – In this approach, which differs substantially from the three approaches above, a two-step protocol was applied to simultaneously compare all four datasets (Late Pleistocene, Holocene, Holocene-HST, and Modern). The oldest and also smallest dataset (Late Pleistocene) was used as a reference baseline. Each iteration of the model involved two sequential steps. First, a given dataset was subsampled down to 39 samples (the number of samples present in the smallest, Late Pleistocene dataset) to generate a replicate dataset. In the second step all large samples ($n>k$) included in the replicate dataset were subsampled to ensure comparable mean sample size across the standardized datasets. The value of k was established by trial and error and varied across datasets from k=70 to k=185. All replicate datasets were compared to the actual Late Pleistocene dataset to estimate number of extirpations and originations. For each dataset, 1000 iterations were used. The replicates derived from the Late Pleistocene dataset were used as a null model (apparent origination and extirpation rates for subsampled Late Pleistocene datasets compared to the actual Late Pleistocene dataset).

Computations

All analyses were performed in *R* [27]. The "vegan" [28] and "ade" [29] packages were used to implement some of the analyses performed in this study, including Mantel test, NMDS, ANOSIM, PERMANOVA, and evaluation of RAD models. Resampling models were written using standard *R* functions.

ESM-Table S3. – Late Pleistocene species apparently extirpated (not found in the Holocene samples).

Genus	Species	Authorship	Number of Late Pleistocene specimens	Number of Late Pleistocene samples	Present-day status in the Adriatic Sea
Ensis	*ensis*	(Linné, 1758)	2	1	Extant
Heteranomia	*squamula*	(Linné, 1758)	1	1	Extant
Myrtea	*spinifera*	(Montagu, 1803)	2	2	Extant
Rocellaria	*dubia*	(Pennant, 1777)	4	1	Extant
Bela	*menkhorsti*	van Aartsen, 1988	1	1	Extant
Parthenina[1]	*decussata*	(Montagu, 1803)	1	1	Extant
Monophorus	*perversus*	(Linné, 1758)	1	1	Extant
Raphitoma	*concinna*	(Scacchi, 1836)	1	1	Extant
Roxania	*utriculus*	(Brocchi, 1814)	1	1	Extant
Valvata[2]	*cristata*	Müller, 1774	1	1	N/A
Weinkauffia	*turgidula*	(Forbes, 1844)	1	1	Extant

[1] This species is referred to as *Chrysallida decussata* in the actual dataset. However, *Chrysallida decussata* is a junior synonym of *Parthenina decussata*.

[2] This is a freshwater species represented by one specimen in the dataset. Most likely, the specimen was transported downstream from more proximal (non-marine) habitats. This species is also extant.

ESM-Table S4. – Summary of RAD model parameters for Holocene, Holocene-HST, Late Pleistocene, and Modern data. Except for Modern data, the Zipf model provided the best fit, based on both AIC and BIC criteria. Analyses performed using "radfit" function, "Vegan" package [28] in *R* [27]. Small samples and rare taxa were retained in this analysis. For the Zipf Model the parameters are: p_1 – fitted proportion of the most abundant taxon, p_2 – decay coefficient (λ). For the Lognormal Model the parameters are: p_1 – mean, p_2 – standard deviation.

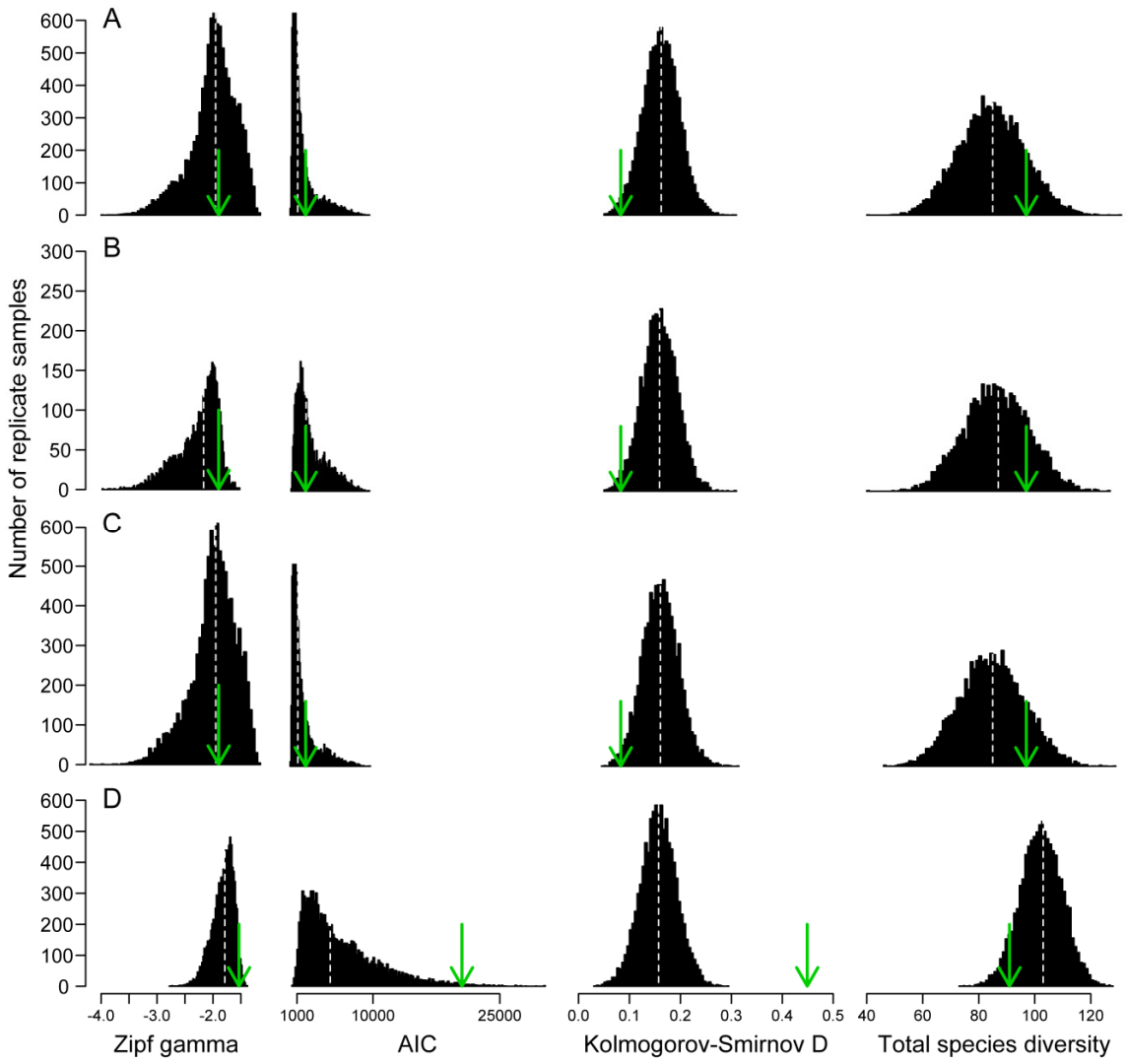
	Best model	p_1	p_2
Holocene	Zipf	0.610	-2.000
Holocene-HST	Zipf	0.765	-2.592
Pleistocene	Zipf	0.576	-1.897
Modern	Lognormal	4.145	2.507

ESM-Table S5 - Results of permutation-based multivariate tests for the two comparisons: Late Pleistocene vs. Holocene and Holocene-HST vs. Modern. See text above for caveats and further discussion.

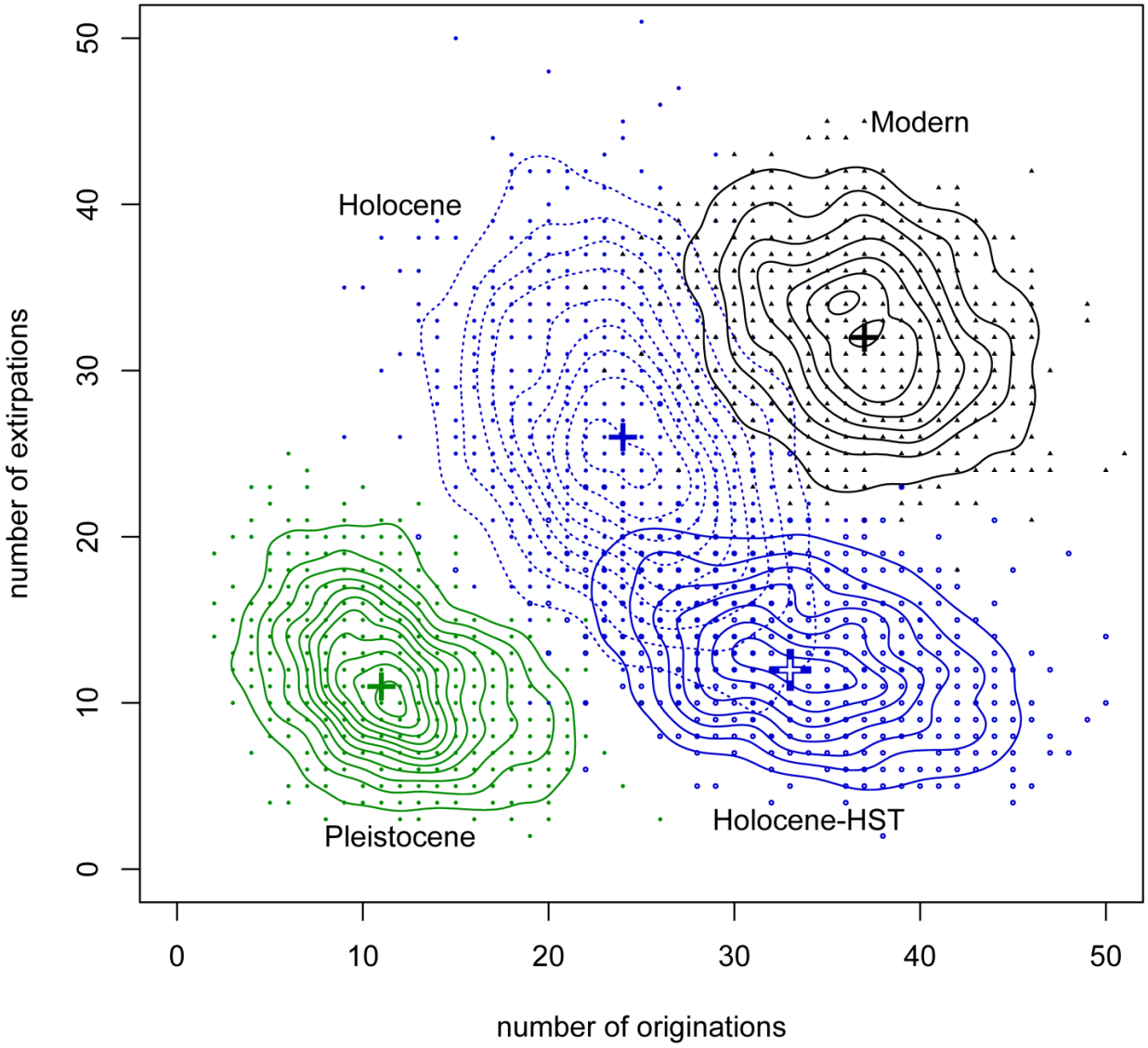
Method	Implementation	Pleistocene versus Holocene	Holocene-HST versus Modern	Interpretation
Mantel Test (comparison of species distances across samples)	Based on comparison of dissimilarity matrices (Gower distances between species)	MANTEL r_M=0.89 p=0.0001 9999 iterations	r_M=0.40 p=0.0106 9999 iterations	The strength of multivariate relationship among species is notably weaker for the Holocene-HST - Modern comparison
ANOSIM (rank-based analysis of differences between samples)	Based on difference between among-group and within-group mean rank distances (derived from Gower distances among samples)	ANOSIM R_A=-0.05 p=0.771 999 iterations	R_A=0.347 p=0.002 999 iterations	The multivariate difference between the two groups of samples is greater for the Holocene-HST - Modern comparison ($R_A \sim 0$ reported for the Pleistocene-Holocene comparison is indicative of complete overlap of the two groups). Compare also fig S1 and fig. S4.
PERMANOVA (permutation-based MANOVA)	Based on ratio of among-group and within-group sums of squares of Gower distances among samples	pseudo F=3.52 p=0.0017	pseudo F=7.01 p=0.0001	The multivariate difference between the two groups of samples is greater for the Holocene-HST - Modern comparison

ESM-Table S6. -Estimates of the temporal structure of samples and specimens from the Holocene-HST dataset (i.e., the highstand systems tract samples from uppermost parts of the cores). Samples and specimens grouped by time bins based on estimates of net accumulation rates (NAR) in the study area. NAR was computed as thickness (meters) divided by time duration (years). The time duration was estimated by two geochronologically constrained horizons (i.e., core top-most layer and a dated horizon). NAR estimates used for grouping samples from each targeted cores (i.e., localities 1, 2, 4-6) based on previously reported estimates [14 Table DR2]).

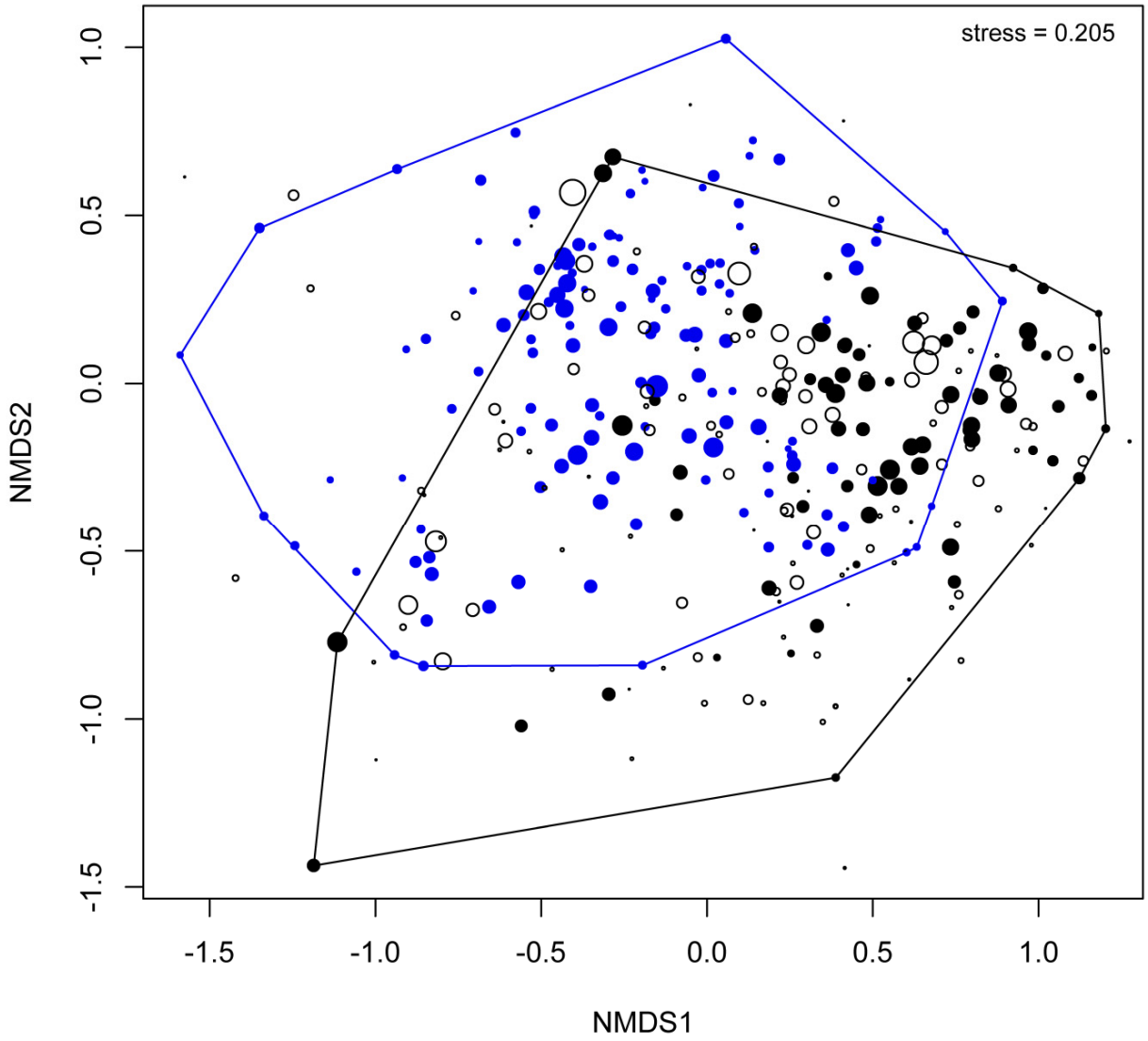
Age range (yrs BP)	Number of samples	Percent of samples	Number of specimens	Percent of specimens	Cunmulative percent of samples	Cumulative percent of specimens
500 – 0	35	29	17892	74	29	74
1000 – 500	25	21	3835	16	50	90
2000 – 1000	36	30	1357	6	80	96
Base of HST – 2000	24	20	940	4	100	100
Total	120	100	24024	100	100	100



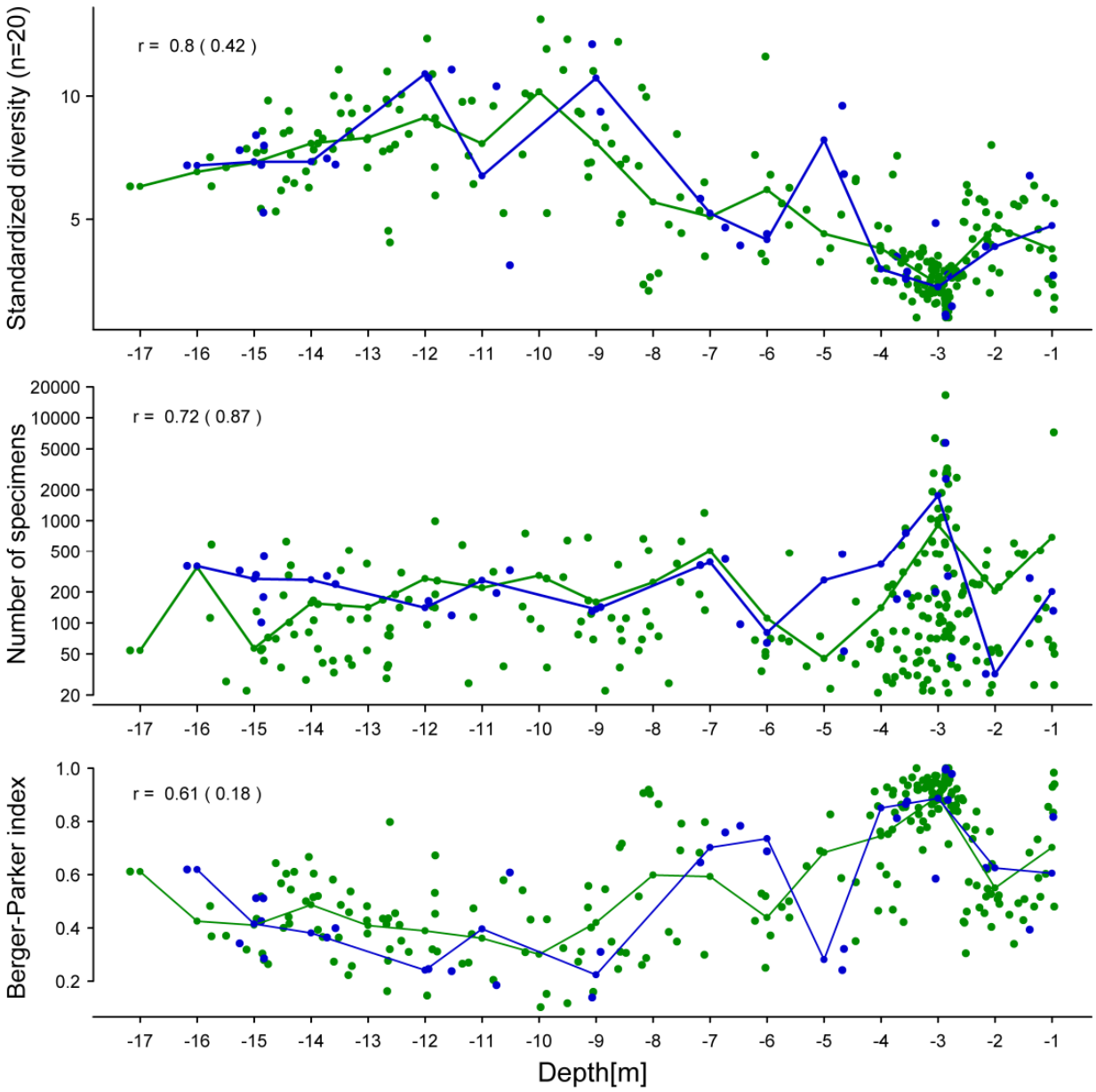
ESM-Fig. S2. – Sampling distributions for "Zipf gamma", "AIC" (Akiaki Information Criterion), Kolmogorov-Smirnov *D*, and "Total species diversity" generated by resampling models. Each distribution is based on 10,000 replicate samples. For each iteration, "Zipf gamma", "AIC" (Akiaki Information Criterion), and "Total species diversity" values were computed for the randomized datasets (table S2). The sampling distributions are for the Pleistocene (**A–C**) and Modern (**D**) datasets, respectively. Kolmogorov-Smirnov *D* values are based on pairwise comparisons of the pairs of the randomized datasets obtained at each iteration: the Holocene-Pleistocene comparisons (**A–C**) and Modern-Holocene HST comparisons (**D**). Dashed lines mark the location of each distribution's median. Arrows include the values observed in actual data. Zipf gamma and AIC computed using "radfit" function in "Vegan" package [28] in *R* [27]. Small samples and rare taxa were retained in this analysis. **A.** Randomization [RDM] model for the Holocene versus Late Pleistocene comparison based on random reassignment of sample membership without replacement (10000 iterations). **B.** Adjusted Randomization Model [ARDM] foe the Holocene versus Late Pleistocene comparison derived from the RDM simulation output (4814 iterations). **C.** Subsampling [SSM] model for the Holocene versus Late Pleistocene comparison based on subsetting without replacement sets of Holocene samples to mimic the Late Pleistocene data dimensions (10,000 iterations). **D.** Randomization [RDM] model for the Modern versus Holocene-HST data (10000 iterations). See text above for details.



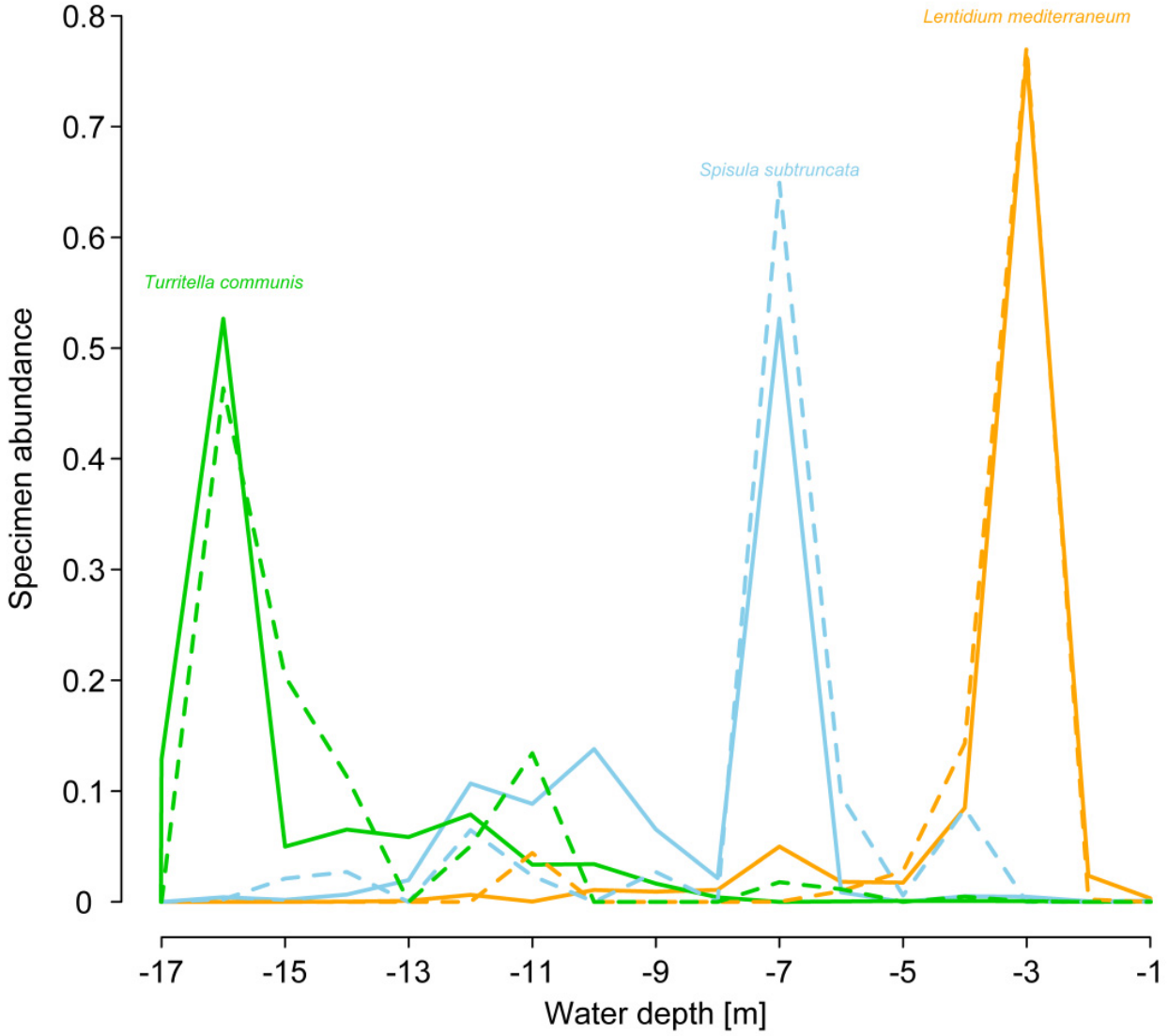
ESM-Fig. S3. – Double subsampling standardization [DSM] model (1000 iterations) used to simultaneously compare Modern (black lines and symbols), Holocene (dashed blue lines and symbols), Holocene-HST (blue lines and symbols), and Late Pleistocene (green lines and symbols). Late Pleistocene replicate samples represent estimates of expected originations and extirpations expected under the null model assuming that all samples came from the same ecosystems. Large crosses mark median values for each distribution. Lines represent two-dimensional kernel density estimates. See text above for detailed description of the DSM subsampling protocol. The modern dataset displays strongest departures from the null model, in terms of both the number of originations and the number of extirpations.



ESM-Fig. S4. - Indirect ordination (nMDS, Bray-Curtis distance, k=2 dimensions) derived using "metaMDS" function in "Vegan" package [28] in *R* [27]. Although the stress value is high, nMDS ordination derived for three dimensions yielded comparable results (partial overlap). Black open circles represent species and solid circles represent samples (blue – Holocene-HST, black – Modern). The size of each point depicts sample size (number of specimens representing a given species or a given sample). Small samples and rare taxa were excluded from this analysis.



ESM-Fig. S5. – Ecological characteristics of samples plotted along the bathymetric gradient. Sample water depth estimated based on previous ordination analyses of the Po Plain data calibrated *a posteriori* against water depth estimates obtained from modern populations of mollusk species that were most common in the cores [7–8]. Solid circles represent individual samples: blue – Holocene, green – Late Pleistocene. Small samples and rare taxa were excluded from this analysis. Symbols: *r* – Spearman rank correlation. Numbers in parenthesis represent rho values for detrended data (first differences).



ESM-Fig. S6. – Bathymetric distribution of the three species most abundant in the Late Pleistocene samples. Solid lines represent the Holocene samples and dashed lines represent the Late Pleistocene samples. Sample species abundance (y axis) estimated as relative specimen abundance. Sample water depth estimated based on previous analyses of the Po Plain data (axis 1 of Detrended Correspondence Analysis calibrated against water depth estimates obtained from modern populations of genera most common in the cores; [7-8]). Small samples (n<20) were excluded from this analysis.

ESM-References

1. Amorosi A, Colalongo ML, Fusco F, Pasini G, Fiorini F. 1999 Glacio-eustatic control of continental–shallow marine cyclicity from Late Quaternary deposits of the southeastern Po plain, Northern Italy. *Quat. Res.* **52**, 1–13.
2. Amorosi A, Dinelli E, Rossi V, Vaiani SC, Sacchetto M. 2008 Late Quaternary palaeoenvironmental evolution of the Adriatic coastal plain and the onset of Po River Delta. *Palaeogeogr. Palaeoclim. Palaeoecol.* **268**, 80–90.
3. Amorosi A, Colalongo ML. 2005 The linkage between alluvial and coeval nearshore marine successions: evidence from the Late Quaternary record of the Po River Plain, Italy *Fluvial Sedimentology VII,* eds Blum MD, Marriott SB Leclair SF (Blackwell Publishing Ltd., Oxofrd), pp. 257–275.
4. Scarponi D, Kaufman D, Amorosi A, Kowalewski M. 2013 Sequence stratigraphy and the resolution of the fossil record. *Geology* **41**, 239–242.
5. Scarponi D, Angeletti L. 2008 Integration of palaeontological patterns in the sequence stratigraphy paradigm: a case study from Holocene deposits of the Po Plain (Italy). *GeoActa* **7**, 1–13.
6. Scarponi D, Kowalewski M. 2007 Sequence stratigraphic anatomy of diversity patterns: Late Quaternary benthic mollusks of the Po Plain, Italy. *Palaios* **22**, 296–305.
7. Scarponi D, Kowalewski M. 2004 Stratigraphic paleoecology: bathymetric signatures and sequence overprint of mollusk associations from upper Quaternary sequences of the Po Plain, Italy. *Geology* **32**, 989–992.
8. Wittmer J, Dexter TA, Scarponi D, Amorosi A, Kowalewski M. 2014 Quantitative bathymetric models for Late Quaternary transgressive-regressive cycles of the Po Plain, Italy. *J. Geol.* **122***,* 649–670.
9. Rossi V, Vaiani SC. 2008 Benthic foraminiferal evidence of sediment supply changes and fluvial drainage reorganization in Holocene deposits of the Po Delta, Italy *Mar. Micropaleont.* **69**, 106–118.
10. Rossi V, Horton BP. 2009 The application of a subtidal foraminifera-based transfer function to reconstruct Holocene paleobathymetry of the Po Delta, Northern Adriatic Sea. *J. Foram. Res.* **39**, 180–190.
11. Amorosi A, Colalongo ML, Dinelli E, Lucchini F, Vaiani SC. 2007 Cyclic variations in sediment provenance from late Pleistocene deposits of the eastern Po Plain, Italy. *GSA Special Paper* **420**, 13–24.
12. Amorosi A. 2012 Chromium and nickel as indicators of source-to-sink sediment transfer in a Holocene alluvial and coastal system (Po Plain, Italy). *Sediment. Geol.* **280**, 260–269.
13. Amorosi A, Bruno L, Rossi V, Severi P, Hajdas I. 2014 Paleosol architecture of a late Quaternary basin-margin sequence and its implications for high-resolution, non-marine sequence stratigraphy. *Global Planet. Change* **112**, 12–25.
14. Ferranti L, *et al.* 2006 Markers of the last interglacial sea-level high stand along the coast of Italy: Tectonic implications. *Quater. Int.* **145-146**, 30–54.
15. Bedulli D, Bassignani F, Bruschi A. 2002 Use of biodiversity hotspots for conservation of Marine Molluscs: a regional approach. *Mediterr. Mar. Sci.* **3**, 113–121.
16. McCune B, Grace JB. 2002 *Analysis of Ecological Communities* (MjM Software Design, Glenden Beach), 284 pp.
17. Legendre P, Legendre L. 2012 *Numerical Ecology* (Elsevier, Amsterdam), 969 pp.

18. Hill MO, Gauch HG. 1980 Detrended correspondence analysis: an improved ordination technique. *Vegetatio* **42**, 47–58.
19. Kowalewski M, *et al.* 2006 Ecological, taxonomic, and taphonomic components of the post-Paleozoic increase in sample-level species diversity of marine benthos. *Paleobiology* **32**, 533–561.
20. Anderson MJ, *et al.* 2011 Navigating the multiple meanings of beta diversity: a roadmap for the practicing ecologist. *Ecol Lett* **14**, 19–28.
21. Condit R, *et al*. 2002 Beta-diversity in tropical forest trees. *Science* **295**, 666–669.
22. Mantel N. 1967 The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220.
23. Clarke KR. 1993 Non-parametric multivariate analyses of changes in community structure. *Austr. J. Ecol.* **18**, 117–143.
24. Anderson MJ. 2001 A new method for non-parametric multivariate analysis of variance. *J. Austr. Ecol.* **26**, 32–46.
25. Anderson MJ, Walsh DC. 2013 PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? *Ecol. Monogr.* **83**, 557–574.
26. Faith DP, Minchin PR, Belbin L. 1987 Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* **69**, 57–68.
27. R Core Team. 2013 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
28. Oksanen J, *et al.* 2013 *Vegan: Community Ecology Package. R package version 2.0-7,* http://CRAN.R-project.org/package=vegan.
29. Dray S, Dufour AB. 2007 The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Software* **22**, 1–20.