

Elsevier Editorial System(tm) for Ecological Indicators
Manuscript Draft

Manuscript Number: ECOLIND-4097R1

Title: New frameworks for species surrogacy in monitoring highly variable coastal ecosystems: applying the BestAgg approach to Mediterranean coastal lagoons

Article Type: Research Paper

Keywords: Coastal lagoons, Indicators, Macrobenthic invertebrates, Null models, Species surrogacy, Taxonomic sufficiency

Corresponding Author: Dr. Stanislao Bevilacqua, PhD

Corresponding Author's Institution: University of Salento

First Author: Stanislao Bevilacqua, PhD

Order of Authors: Stanislao Bevilacqua, PhD; Antonio Terlizzi, Prof.; Michele Mistri, Prof.; Cristina Munari, Dr.

Abstract: Benthic invertebrates are good indicators of aquatic ecosystem health. Yet, environmental monitoring and assessment of community changes in relation to both natural and human sources of disturbance involves considerable efforts for sample processing and time-consuming identifications of organisms, which make challenging large-scale and continuous monitoring programs required under the current regulatory frameworks at European scale. The use of higher taxa (e.g. families) as surrogates for species is a mainstream approach to reduce cost and time associated to fine taxonomic resolution in environmental studies, especially concerning macro-invertebrate communities. However, this approach of 'Taxonomic Sufficiency' simply relies on the static grouping of organisms in taxa belonging to a single higher taxonomic level irrespective of their ecological relevance or difficulties in their taxonomic identification, leading to unnecessary losses of taxonomic detail or ecological information. A new approach, namely the Best Practicable Aggregation of Species (BestAgg), has been recently developed as an alternative procedure for species surrogacy. BestAgg is based on aggregating species in the minimum number of surrogates sufficient to depict species-level community patterns consistently, while capitalizing on ecological information. Although the approach has been successfully applied to marine and freshwater invertebrate assemblages, its effectiveness in transitional water systems, where the complex and highly variable environmental conditions may affect the performance of surrogacy methods, still remain untested. Here, we applied the BestAgg approach to quantifying spatio-temporal patterns of variability of macro-invertebrate assemblages from Mediterranean coastal lagoons (Northern Adriatic Sea). Surrogates were defined using species-level data from a representative lagoon system, which served as pilot study. Then, they were used to analyze macro-invertebrate assemblages in two independent lagoons in the same geographic area. Results showed that BestAgg surrogates were effective in depicting multivariate patterns of macro-invertebrate assemblages as at species level over and beyond potential variations among the investigated lagoons. The use of families, following a more classic approach based on Taxonomic Sufficiency, also provided results comparable to those obtained using species. However, with respect to families, BestAgg surrogates allowed an estimated timesaving of 40% higher while still retaining an equivalent amount of information on species-level patterns. More importantly, BestAgg allowed exploiting different criteria of species aggregation, leading to a set of surrogates more aligned with ecological/functional characteristics of organisms, suggesting that BestAgg approach may provide a fresh perspective for

optimizing trade-offs between pragmatism and the need for relevant ecological information in environmental assessment and monitoring.

Response to Reviewers: Response to Reviewer comment #1:

We quoted additional works in order to extend the range of references sustaining our arguments in the Discussion.

Response to Reviewer comment #2:

We take the points raised by the reviewer and recognize the need of additional discussion on these issues. We clearly stated in the manuscript that the application of any surrogacy method, BestAgg included, requires baseline knowledge of the system being investigated and at least a pilot assessment of surrogates for their definition and validation before they could be adopted as substitutes for species in similar study contexts and/or subsequent monitoring.

Approximation cannot be made a priori, as it is evident from the bulk of literature on surrogates for species, whether based on taxonomic sufficiency or other approaches. Several works pointed out that surrogates should not be considered as fixed operational units, requiring, especially on the long term, periodical refinements due to possible changes in assemblages structure.

It is worth noting that such limits do not apply exclusively to our approach, but are inherent to any attempt to define effective surrogates for species. Besides these intrinsic constraints, species surrogacy has a strong practical utility and offer undeniable benefits in routine monitoring, as we highlighted in the Introduction. With respect to mainstream surrogacy methods, however, the BestAgg approach provides additional advantages and relies on a more formal and transparent theoretical and statistical framework.

Moving again on the problem of variations in species identity and richness of the investigated system, it should be considered that (1) simulations underlying the identification of the minimum number of surrogates (i.e. G_{min}) imply a wide range of aggregation scenarios (with varying identity and number of species involved); (2) the choice of G_{min} is conservative because the true minimum number is comprised between the lowest G allowing >95% of analyses to be consistent with species-level analyses and the highest G at which analyses are consistent in less than 95% of cases, but the process is structured to select the former value as G_{min} ; (3) the number of surrogate should not necessarily equal G_{min} , just to be (and often the surrogate selection process may lead to be) $\geq G_{min}$. All these aspects likely confer robustness of BestAgg surrogates to variations in species identity and number over time.

Also, reduction in species richness will lead the fixed G_{min} to be even more conservative (as the ratio between the number of surrogates and the number of species actually increases). On the other hand, the disappearance of a surrogate will be compensated by a consequent reduction also in the number of species. Above all, the loss of surrogates, even when they comprise single species, will denote a substantial change in the system with respect to previous time of sampling, whereas, if new species appear that could not be included in any of the selected surrogates, such species could be simply treated following the criteria of relevance, easiness and resemblance to obtain new surrogates.

Therefore, variations in assemblage structure will likely exert limited effects on BestAgg surrogates, avoiding recurrent refinements at species level (which in turn will limit their practical utility). In addition, substantial changes in the system requiring a new calibration of the surrogate set will be evident, unless consisting in considerable increases in species richness.

We acknowledge that the above considerations have been in part neglected in the manuscript and we are grateful to the reviewer for raising the issue. A whole paragraph has been dedicated to these aspects in the revised version.

December 02, 2014

Dear Editor,

Please find attached the revised version (Bevilacqua et al_revised.doc) of the manuscript entitled '*New frameworks for species surrogacy in monitoring highly variable coastal ecosystems: applying the BestAgg approach to Mediterranean coastal lagoons*' by Bevilacqua S., Terlizzi, A., Mistri, M., and Munari A. (Ms. Ref. No.: ECOLIND-4097).

We considered all comments provided by the reviewers and modified the manuscript accordingly. We also carried out a careful check of the whole manuscript and supplementary material, providing some minor corrections. All changes to the original manuscript have been given in track-change mode in an additional copy of the revised version of the manuscript (Bevilacqua et al_track_changes.pdf). We are also providing a revised version of the supplementary Appendix A (Bevilacqua et al_Appendix A_revised.pdf).

Here enclosed please find our answers to the reviewers' comments.

We hope that the manuscript will be suitable for the final acceptance on *Ecological Indicators* in the present form.

Kind regards,

Stanislao Bevilacqua

Response to Reviewer #1:

We quoted additional works in order to extend the range of references sustaining our arguments in the Discussion.

Response to Reviewer #2:

We take the points raised by the reviewer and recognize the need of additional discussion on these issues. We clearly stated in the manuscript that the application of any surrogacy method, BestAgg included, requires baseline knowledge of the system being investigated and at least a pilot assessment of surrogates for their definition and validation before they could be adopted as substitutes for species in similar study contexts and/or subsequent monitoring.

Approximation cannot be made a priori, as it is evident from the bulk of literature on surrogates for species, whether based on taxonomic sufficiency or other approaches. Several works pointed out that surrogates should not be considered as fixed operational units, requiring, especially on the long term, periodical refinements due to possible changes in assemblages structure.

It is worth noting that such limits do not apply exclusively to our approach, but are inherent to any attempt to define effective surrogates for species. Besides these intrinsic constraints, species surrogacy has a strong practical utility and offer undeniable benefits in routine monitoring, as we highlighted in the Introduction. With respect to mainstream surrogacy methods, however, the BestAgg approach provides additional advantages and relies on a more formal and transparent theoretical and statistical framework.

Moving again on the problem of variations in species identity and richness of the investigated system, it should be considered that (1) simulations underlying the identification of the minimum number of surrogates (i.e. G_{min}) imply a wide range of aggregation scenarios (with varying identity and number of species involved); (2) the choice of G_{min} is conservative because the true minimum number is comprised between the lowest G allowing >95% of analyses to be consistent with species-level analyses and the highest G at which analyses are consistent in less than 95% of cases, but the process is structured to select the former value as G_{min} ; (3) the number of surrogate should not necessarily equal G_{min} , just to be (and often the surrogate selection process may lead to be) $\geq G_{min}$. All these aspects likely confer robustness of BestAgg surrogates to variations in species identity and number over time.

Also, reduction in species richness will lead the fixed G_{min} to be even more conservative (as the ratio between the number of surrogates and the number of species actually increases). On the other hand, the disappearance of a surrogate will be compensated by a consequent reduction also in the number of species. Above all, the loss of surrogates, even when they comprise single species, will denote a substantial change in the system with respect to previous time of sampling, whereas, if new species appear that could not be included in any of the selected surrogates, such species could be simply treated following the criteria of relevance, easiness and resemblance to obtain new surrogates.

Therefore, variations in assemblage structure will likely exert limited effects on BestAgg surrogates, avoiding recurrent refinements at species level (which in turn will limit their practical utility). In addition, substantial changes in the system requiring a new calibration of the surrogate set will be evident, unless consisting in considerable increases in species richness.

We acknowledge that the above considerations have been in part neglected in the manuscript and we are grateful to the reviewer for raising the issue. A whole paragraph has been dedicated to these aspects in the revised version.

1 **New frameworks for species surrogacy in monitoring highly variable coastal ecosystems:**
2 **applying the BestAgg approach to Mediterranean coastal lagoons**

3 Stanislaò Bevilacqua^{1*}, Antonio Terlizzi¹, Michele Mistri², Cristina Munari²

4
5 ¹*Laboratory of Zoology and Marine Biology, Department of Biological and Environmental Sciences*
6 *and Technologies, University of Salento, CoNISMa, I-73100 Lecce, Italy*

7
8
9 ²*Department of Life Science and Biotechnologies, University of Ferrara, I-44100 Ferrara, Italy*

10
11
12
13
14 ***Corresponding author:**

15 Dr Stanislaò Bevilacqua, Ph.D.

16
17 Department of Biological and Environmental Sciences and Technologies

18
19 University of Salento, Campus Ecotekne, I-73100 Lecce, Italy

20
21 Phone +39 0832 298803

22
23 e-mail: stanislao.bevilacqua@unisalento.it

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

14 **Abstract**

15 Benthic invertebrates are good indicators of aquatic ecosystem health. Yet, environmental
16 monitoring and assessment of community changes in relation to both natural and human sources of
17 disturbance involves considerable efforts for sample processing and time-consuming identifications
18 of organisms, which make challenging large-scale and continuous monitoring programs required
19 under the current regulatory frameworks at European scale. The use of higher taxa (e.g. families) as
20 surrogates for species is a mainstream approach to reduce cost and time associated to fine
21 taxonomic resolution in environmental studies, especially concerning macro-invertebrate
22 communities. However, this approach of 'Taxonomic Sufficiency' simply relies on the static
23 grouping of organisms in taxa belonging to a single higher taxonomic level irrespective of their
24 ecological relevance or difficulties in their taxonomic identification, leading to unnecessary losses
25 of taxonomic detail or ecological information. A new approach, namely the Best Practicable
26 Aggregation of Species (BestAgg), has been recently developed as an alternative procedure for
27 species surrogacy. BestAgg is based on aggregating species in the minimum number of surrogates
28 sufficient to depict species-level community patterns consistently, while capitalizing on ecological
29 information. Although the approach has been successfully applied to marine and freshwater
30 invertebrate assemblages, its effectiveness in transitional water systems, where the complex and
31 highly variable environmental conditions may affect the performance of surrogacy methods, still
32 remain untested. Here, we applied the BestAgg approach to quantifying spatio-temporal patterns of
33 variability of macro-invertebrate assemblages from Mediterranean coastal lagoons (Northern
34 Adriatic Sea). Surrogates were defined using species-level data from a representative lagoon
35 system, which served as pilot study. Then, they were used to analyze macro-invertebrate
36 assemblages in two independent lagoons in the same geographic area. Results showed that BestAgg
37 surrogates were effective in depicting multivariate patterns of macro-invertebrate assemblages as at
38 species level over and beyond potential variations among the investigated lagoons. The use of

Eliminato: ,

Eliminato: on

Eliminato: on

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

39 families, following a more classic approach based on Taxonomic Sufficiency, also provided results
40 comparable to those obtained using species. However, with respect to families, BestAgg surrogates
41 allowed an estimated timesaving of 40% higher while still retaining an equivalent amount of
42 information on species-level patterns. More importantly, BestAgg allowed exploiting different
43 criteria of species aggregation, leading to a set of surrogates more aligned with
44 ecological/functional characteristics of organisms, suggesting that BestAgg approach may provide a
45 fresh perspective for optimizing trade-offs between pragmatism and the need for relevant ecological
46 information in environmental assessment and monitoring.

47
48 **Keywords:** Coastal lagoons, Indicators, Macro-benthic invertebrates, Null models, Species
49 surrogacy, Taxonomic sufficiency

50 1. Introduction

51 Benthic macro-invertebrates have long been used as ecological indicators in monitoring the health
52 of marine, freshwater, and transitional water bodies because of their importance for the functioning
53 of aquatic ecosystems and their ability to reflect natural and anthropogenic pressures effectively
54 (Pearson and Rosenberg, 1978) through changes in abundance, taxonomic richness and
55 composition, biomass, body-size and biological traits (Bremner et al., 2006; Terlizzi et al., 2008;
56 Petchey and Belgrano, 2010; Basset et al., 2012; Munari and Mistri, 2012; Guarnieri et al., 2012;
57 Munari, 2013). Due to such desirable prerogatives, macro-invertebrates represent preferential
58 descriptors of benthic communities being considered in the current environmental policies for the
59 assessment of the Good Environmental Status (GES) of aquatic systems at European scale (Van
60 Hoey et al., 2010; Rice et al., 2012). Monitoring aquatic environments using macro-invertebrates
61 requires, however, considerable efforts in sample processing and taxonomic identifications of
62 organisms, which may be exacerbated by the limited availability of taxonomic expertise and
63 knowledge of invertebrate species (Terlizzi et al., 2003; Cardoso et al., 2011), resulting in
64 expensive and time-consuming procedures that make difficult coping with the continuous and large-
65 scale monitoring programs required under the current regulatory frameworks (e.g. Water
66 Framework Directive, WFD, EC 2000; Marine Strategy Framework Directive, MSFD, EC 2008).
67 Among the adopted approaches to reduce time and costs associated to sample processing and fine
68 taxonomic identifications when monitoring aquatic macro-invertebrate communities, the use of
69 higher taxa (e.g. families) as surrogates for species, has received particular attention (see Dauvin et
70 al., 2003; Terlizzi et al., 2003; Jones, 2008; Bevilacqua et al., 2012 for reviews). Such an approach
71 of *Taxonomic Sufficiency* (Ellis, 1985) relies on identifying the coarsest taxonomic level sufficient
72 to detect changes in assemblages exposed to environmental stress without a significant loss of
73 information with respect to species-level analysis. It has been successfully applied on benthic
74 communities from continental (e.g. Marshall et al., 2006; Heino and Soininen, 2007; Shokri and

Eliminato: Ecological

Eliminato: ,

75 Gladstone, 2009; Jiang et al., 2013) to marine water bodies (e.g. Vanderklift et al., 1996; Gomez-
76 Gesteira et al., 2003; Anderson et al., 2005; Włodarska-Kowalczyk and Kędra, 2007; Bertasi et al.,
1
77 2009; Terlizzi et al., 2009; Brind'Amour et al., 2014).

78 | The relevance of Taxonomic Sufficiency (TS) for environmental monitoring has been sustained
6
79 | invoking a hierarchical (from individuals to species and then higher taxonomic levels) biological
8
80 | response to environmental disturbance, and putative ecological similarities among closely related
9
81 | species (e.g. Ferraro and Cole, 1995; Warwick, 1993). The use of higher taxonomic levels would
11
82 | also reduce the potential effect of natural variability in obscuring changes related to environmental
13
83 | stress, as natural variations may affect community structure mainly by species replacement (Gray et
15
84 | al., 1988). However, also natural variations may affect community structure at higher taxonomic
17
85 | levels (Dethier and Schoch, 2006; Lovell et al., 2007; ~~Włodarska-Kowalczyk and Kędra 2007;~~
19
86 | Bevilacqua et al. 2009a; Bevilacqua et al., 2011), and ecological similarities among species within
20
87 | higher taxa may be largely idiosyncratic (Lenat and Resh, 2001; Heino and Soininen, 2007;
22
88 | Carranza et al., 2011), casting doubts on mechanisms proposed to explain the ability of higher
24
89 | taxonomic levels to reflect natural or human-driven community patterns at species level (Lasiak,
26
90 | 2003; Bertrand et al., 2006; Bevilacqua et al., 2009b; Heino, 2014).

Eliminato: Wodarska

28
29
30 | Indeed, TS does not rely on balancing costs and timesaving with the need for ecological
31
32 | information but, rather, it implies the static grouping of organisms in taxa belonging to a single
33
34 | higher taxonomic level (e.g. all organisms identified at the level of genus, family, etc.) irrespective
35
36 | of their ecological relevance or difficulties in their taxonomic identification. The approach,
37
38 | therefore, may lead to unnecessary loss of taxonomic details and, more importantly, may reduce the
39
40 | ability to infer about ecological processes underlying the observed patterns (Lenat and Resh, 2001;
41
42 | Lasiak, 2003; Terlizzi et al., 2003; Dethier and Schoch, 2006; Jones, 2008), which in turn are
43
44 | essential for defining effective environmental management strategies. Also, TS lacks any statistical

Eliminato: Taxonomic Sufficiency

Eliminato: Taxonomic Sufficiency

99 method to control for uncertainties when identifying a particular taxonomic level as sufficient to
100 detect community responses to environmental drivers as at species level (Mellin et al., 2011).

101 Recent attempts aimed at disentangling the mechanisms generating correlations between
102 community patterns at species and coarser taxonomic resolution reinforced the idea that higher taxa
103 of the Linnaean taxonomic hierarchy may be considered as arbitrary categories of species (e.g.
104 Bertrand et al., 2006) unlikely to convey consistent responses to natural or human-driven
105 environmental changes (Bevilacqua et al., 2012; Siqueira et al., 2012). Specifically, the use of null
106 models highlighted that information loss and the ensuing decrease in statistical power to detect
107 changes in assemblage structure at higher taxonomic levels depended on the degree of species
108 aggregation (exemplified by the ratio between the number of higher taxa and the number of
109 species), rather than on taxonomic relatedness of species *per se* (i.e. the relative closeness of species
110 in the Linnaean taxonomic hierarchy) (Bevilacqua et al., 2012). Such a dependence between
111 information loss and the fractal architecture of taxonomic systems has been found to be consistent
112 from terrestrial to aquatic organisms (see Bevilacqua et al., 2012 for a review) and validated by
113 several studies investigated factors affecting the performance of taxonomic surrogates (Giangrande
114 et al., 2005; Dethier and Schoch, 2006; Lovell et al., 2007; Siqueira et al., 2012; Neeson et al.,
115 2013).

Eliminato: higher

Eliminato: corroborated

116 In this new theoretical perspective, a novel approach to species surrogacy, the Best Practicable
117 Aggregation of Species (BestAgg), has been implemented (Bevilacqua et al., 2013). The approach
118 relies on the construction of null models based on real data in order to identify thresholds in species
119 aggregation (in terms of numerical ratios between the number of surrogates and the number of
120 species), and irrespective of taxonomic relatedness, below which the loss of ecological information
121 becomes relevant. In practice, BestAgg identifies the minimum number of surrogates, irrespective
122 of their type, allowing obtaining the same response as at species level. Surrogates are then defined

44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

123 based on their ecological importance (*relevance*), difficulty of identification (*easiness*), and shared
124 characteristics among organisms (*resemblance*).

125 As any other surrogacy method, BestAgg involves the selection of a set of surrogates based on a
126 pilot assessment assuming that they may be employed in similar environmental contexts or in
127 subsequent monitoring programs (Bevilacqua et al., 2013). The validity of this assumption,
128 however, may be limited when investigating peculiar and/or highly variable systems where habitat
129 heterogeneity shapes the interactions among organisms and modifies their response to perturbation,
130 potentially constraining the application of surrogates for species (White and Irvine, 2003; Munari
131 and Mistri, 2008). Transitional waters (e.g. estuaries, wetlands, and coastal lagoons, hereafter *Tws*)
132 exemplify the issue. *Tws* represent important and fragile ecosystems in the coastal landscape,
133 providing key ecosystem services such as water quality improvement, fisheries resources, habitat
134 and food for migratory and resident animals, and recreational areas for human populations. *Tws* are
135 characterized by highly variable physico-chemical and hydro-morphologic conditions. Progressive
136 changes in several environmental parameters (e.g. salinity, dissolved oxygen, temperature), often
137 mutually correlated, generates composite gradients, resulting in a mosaic of different habitats,
138 characterized by different marine water renewal, nutrient content, turbidity and sediment structure
139 (Tagliapietra et al., 2009). Moreover, *Tws* exhibit different characteristics depending on their
140 geographical, hydrodynamic and ecological features, and also display distinctive features in terms
141 of their extraordinary history of environmental management, the importance of their productivity
142 and associated economical value, which are reflected on the peculiarity of their fauna (Cognetti and
143 Maltagliati, 2008). The species living in such ecologically complex systems adapt to changes in
144 environmental conditions, to the presence of organic enrichment (Simonini et al., 2004), and metal
145 contamination (Munari and Mistri, 2007). Several studies have explored the effectiveness of **TS** on
146 benthic macro-invertebrates from *Tws* (Mistri and Rossi, 2001; De Biasi et al., 2003; Dethier and
147 Schoch, 2006; Munari et al., 2009; Shokri and Gladstone, 2009; Soares-Gomez et al., 2012),

Eliminato: Taxonomic Sufficiency

148 suggesting that the use of coarser level of taxonomic resolution, up to family level, might represent
149 a cost-effective solution for routine monitoring.

150 Here, we applied the BestAgg approach to Mediterranean transitional water systems, namely
151 coastal lagoons, in order to assess its performance in quantifying patterns of spatial and temporal
152 variability of macro_benthic assemblages. To check for the consistency of the approach when
153 applied to such peculiar and highly variable environments, we derived surrogates for species
154 following the BestAgg procedure based on a pilot lagoon system and then assessed the ability of
155 such surrogates to reflect species-level patterns in other lagoons in the same geographic area. A
156 classical approach for species surrogacy based on the concept of Taxonomic Sufficiency was also
157 applied to pilot data and results from the two approaches compared.

159 2. Methods

160 2.1. Study areas and datasets

161 The data sets consisted of macro_benthic species abundance gathered during several sampling
162 campaigns carried out between 2004 and 2008, from three coastal lagoons of the Po River Delta
163 (NE Italy), namely Marinetta-Vallona (MAR), Sacca di Scardovari (SCA), Sacca di Goro (GOR)
164 (Fig. 1). MAR is a micro_tidal, shallow (maximum depth 2.5 m) lagoon, extending for 11.5 km².
165 The lagoon is connected to the sea by a narrow mouth (about 100 m wide). A deltaic branch of the
166 Po River flows along the lagoon NW–SE axis. The westernmost part of the lagoon hosts clam-
167 farming activities. Macro_benthic samples were collected at four stations randomly distributed
168 within the lagoon in three random sampling times from 2007 to 2008. SCA is a micro_tidal, large
169 embayment (32 km²) located between two branches of the Po River Delta. The lagoon is connected
170 to the Adriatic Sea through a wide mouth that is partly obstructed by sand banks. It varies in depth

171 between 0.5 and 2.8 m. Its northern area receives nutrient-rich agricultural run-offs, whereas the
172 southern area hosts extensive cultures of clams and mussels. Samples were collected at five stations
173 in three sampling occasions. GOR is a wide (26 km²) micro-tidal lagoon whose maximum depth is
174 2.0 m. The lagoon receives nutrient-rich freshwater from deltaic branches of the Po River. Clams
175 are intensively cultured in its westernmost portion, representing the second largest clam-farming
176 ground in Europe, with an average annual yield of 8,000 tonnes (Solidoro et al., 2000). GOR is also
177 subjected to anthropogenic eutrophication, which causes extensive growth of opportunistic green
178 macro-algae leading to summer anoxia and dystrophy. Sampling was carried out at four stations in
179 four sampling times from 2004 to 2005. Sandy and/or silty clay sediments characterized soft
180 bottoms in all lagoons. For each of the three lagoons, three samples of soft sediments were collected
181 at each station using a Van Veen grab (0.027 m²). Samples were sieved (0.5 mm mesh size) and
182 macro-invertebrates were sorted, identified at the species level, and counted. Three data sets, one
183 for each lagoon (87 species × 36 samples for MAR, 90 species × 45 samples for SCA, and 38
184 species × 48 samples for GOR) were obtained.

Eliminato: , were obtained.
Eliminato: The data set consisted of

185 The three data sets were then used to test the performance of the BestAgg approach. **As for any**
186 **other** surrogacy **approaches, the identification of** surrogates **should be based on** a pilot assessment
187 **(Terlizzi et al., 2003; Defeo and Lercari, 2004; Dauvin et al., 2007; Musco et al., 2009).** We
188 selected one lagoon out of three, namely MAR, to serve as pilot system for the definition and
189 validation of BestAgg surrogates. We randomly selected the pilot lagoon between the two lagoons
190 accounting for the higher number of species (see above). This simulates the experimental condition
191 in which baseline information (e.g. a simple list of species or taxa) on systems being investigated is
192 available and could guide the choice of a representative pilot system (e.g. selection among more
193 speciose ones), in contrast to a completely random selection with the absence of any prior data.
194 Species-level data on community structure of the two remaining lagoons, namely SCA and GOR,

Eliminato: Irrespective of the
Eliminato: approach employed,
Eliminato: for species
Eliminato: not
Eliminato: defined *a priori* and should require at least
Eliminato: and validation before they could be assumed as suitable substitutes of species for subsequent monitoring or for very similar study contexts (Lenat and Resh, 2001;
Eliminato: Jones, 2008). In this respect BestAgg is no exception.

195 were used to show that BestAgg surrogates derived from the pilot lagoon could have been
196 successfully applied to investigate similar but independent ecosystems.

1
2
197
4
5
198
6
7

2.2. The Best Practicable Aggregation of Species (BestAgg)

199 Species surrogacy implies quantifying community structure through the use of operational units that
200 can substitute species successfully, in order to avoid difficulties related to fine identification of
201 organisms without ensuing significant loss of information. Such operational units, i.e. surrogates,
202 are indeed groups in which species have been aggregated following shared characteristics and their
203 abundance summed. The level of aggregation to which the original species-level information is
204 subjected when community structure is codified through the use of surrogates can be expressed as ϕ
205 = G/S , i.e. the ratio of the number of surrogate groups G to the number of species S (Bevilacqua et
206 al., 2012). The basic concept underlying the BestAgg approach is that the success of a given set of
207 surrogates to mirror species-level patterns depends mainly on the extent to which species-level
208 information is packed into surrogate groups, rather than on the actual, or putative, relationships
209 among species that such groups might embody (Bevilacqua et al., 2012). In this view, BestAgg
210 aims at identifying the minimum number of surrogates, irrespective of their type (whether
211 taxonomic, morphological, functional, etc.), sufficient to quantify species-level community patterns
212 consistently. The sufficient number of surrogates is determined through a null model assuming that
213 (1) surrogates can be considered as random groups of species from the original species pool, (2) the
214 ability of surrogates to reflect species-level community patterns depends on the correlation
215 (Spearman's ρ) between the species-level data matrix and the corresponding surrogate matrix, and
216 (3) correlation values (ρ) depends on ϕ . Such assumptions are largely sustained by empirical
217 evidence (for a review see Bevilacqua et al., 2012).

Eliminato: s

43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

218 BestAgg is based on constructing a null model from random aggregations of species-level data in
219 order to determine the value of ϕ , namely ϕ_{low} , sufficient to obtain results consistent with those
220 obtained analyzing species-level data. Once ϕ_{low} has been determined, surrogate groups in BestAgg
221 are identified on the basis of this threshold value and a set of selection criteria aiming to capitalize
222 on the ecological information in the surrogate data. Specifically, surrogates selection in BestAgg is
223 guided by the interplay among three macro-criteria: *relevance*, *easiness*, and *resemblance*.
224 *Relevance* concerns the importance of a given species from a general (e.g. key-stone species),
225 context-specific (e.g. indicators), or study-specific (e.g. species most contributing to the specific
226 pattern of interest) ecological perspective. *Easiness* regards the distinctiveness of a given species,
227 taxon, or group of organisms leading to be easily identified. *Resemblance*, instead, relates to
228 common characteristics among organisms allowing meaningful groupings (e.g. common ancestry,
229 functional similarity). In practice, BestAgg is structured to maximize the relevant ecological
230 information retained in the surrogate matrix, while reducing as much as possible the number of
231 surrogates and the difficulty in the identification of organisms and sample processing. A full
232 theoretical introduction to BestAgg, along with R codes for analyses, is provided in Bevilacqua et
233 al. (2013).

235 2.3. Construction of the null model and surrogate selection based on the pilot system

236 The first step in the application of the BestAgg approach concerned the construction of the null
237 model based on pilot data in order to identify the lowest ϕ value, i.e. ϕ_{low} , and consequently the
238 minimum number of surrogates G , i.e. G_{min} (where $G_{min} = \phi_{low} \times S$), sufficient to obtain results
239 consistent with those obtained using species with a probability of Type I error (i.e. the probability
240 of considering G_{min} as sufficient when actually it is not) of $P < 0.05$. In this framework, the S
241 species in the original species-level data matrix from the pilot lagoon (MAR) were randomly

242 aggregated in decreasing numbers of G groups in order to simulate decreasing values of ϕ . The
243 number of groups G was progressively decreased by step-wise reductions of a fixed decrement $d=5$,
244 starting from $G=S$ (see Bevilacqua et al., 2013). For each set of G groups from the step-wise
245 reduction procedure, 1,000 randomly aggregated matrices were obtained. A triangular matrix based
246 on Bray–Curtis dissimilarities was obtained from each randomly aggregated matrix and the
247 correlation value (ρ) with the species-level matrix was calculated. Then, a distance-based
248 permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) was performed
249 on each randomly aggregated matrix. The design for the analysis consisted of two crossed factors:
250 Time (Ti, 3 levels, random) and Station (St, 4 levels, random) with $n = 3$. Analyses were based on
251 Bray–Curtis dissimilarities of untransformed data. For the construction of the null model we focused
252 on the term $Ti \times St$ interaction as species-level analysis highlighted a significant interaction of
253 spatial and temporal variability in macro_benthic assemblages (see Results). The dependence
254 between ϕ and the effectiveness of the corresponding surrogate groups was checked by fitting a
255 semi-log model of ρ values against the corresponding ϕ values and calculating, out of $n = 1,000$ and
256 for each level of aggregation ϕ , the percentage of significant PERMANOVA tests for the term of
257 interest, consistent with results obtained at species level. The lowest practicable aggregation level
258 ϕ_{low} was determined by identifying the lowest ϕ value allowing the 95% of PERMANOVAs on
259 randomly aggregated matrices to give results consistent with those obtained at species level (for
260 further details on the whole procedure see Bevilacqua et al., 2013).

261 The next step focused on selecting the set of surrogates, given G_{min} . For each species found in the
262 pilot lagoon MAR, we compiled information related to the three macro-criteria of *relevance*,
263 *easiness*, and *resemblance* (see previous section and Table A1 in Appendix A). Relevant species
264 were identified based on available knowledge from scientific literature on their importance for the
265 structure and functioning of coastal lagoon ecosystems. Species were classified by their feeding
266 behaviour to one of the following guilds: predators, scavengers (carrion feeders),

267 herbivores/grazers, suspension/filter feeders, detritus feeders (including surface deposit feeders,
268 subsurface deposit feeders/limnivores, that feed on sedimentary detritus and bacteria), species that
269 are both suspension feeders and surface deposit feeders, and omnivores. Ecological information
270 (e.g. trophic guild, tolerance, indicator species) was mined from the literature (see Table A1) and
271 existing databases (ERMS, [ww.marbef.org](http://www.marbef.org); MarLIN, Marine life information network; WoRMS,
272 www.marinespecies.org). Moreover, study-specific relevant species most contributing to the
273 observed community patterns were identified through Similarity Percentage analysis (SIMPER) by
274 considering only species whose contribution to dissimilarities was $\geq 3\%$. Easiness of species was
275 defined based on expert opinion according to the difficulty and intensity of analyses required to
276 identify organisms, whereas species groupings (*resemblance*) were decided combining information
277 on *relevance* and *easiness* with available taxonomic and ecological knowledge of groups. As the
278 identification of G_{\min} goes beyond any surrogacy scheme, we exploited simultaneously the potential
279 of different surrogate types potentially leading to retain ecological information and/or to reduce
280 efforts for the identification of organisms and sample processing (for a detailed conceptual
281 framework on surrogate selection in BestAgg see also Bevilacqua et al., 2013).

282 283 2.4. Testing BestAgg surrogates

284 As BestAgg assumes that surrogates are random subsets of the original pool of species, the
285 information retained using the selected BestAgg surrogates (expressed as ρ) should fall at least
286 within random expectations from the null model. If selected surrogates allow retaining significantly
287 greater information, the selection procedure has led to a set of surrogates able to represent species-
288 level community patterns better than what expected by chance. In contrast, ρ values below
289 expectations would indicate a set of surrogates that is likely unsuitable. The BestAgg approach

290 allows checking for this through a randomization test built following the same simulation procedure
291 underlying the construction of the general null model.

292 The original S species from MAR were randomly aggregated in G_{BestAgg} groups, where G_{BestAgg} is
293 the number of BestAgg surrogates obtained from MAR. Random aggregations were repeated 1,000
294 times. Correlation values (ρ) between the original species-level matrix and each randomly
295 aggregated matrix were then calculated obtaining a frequency distribution against which testing
296 ρ_{BestAgg} , i.e. the correlation between the original species-level matrix and the matrix aggregated
297 using the BestAgg surrogates. Also, PERMANOVA was performed on randomly aggregated
298 matrices (design as above), and the percentage of tests for the term of interest showing consistent
299 results with those obtained at species level was calculated, representing the P -value (i.e. the
300 probability of Type I error) specific for G_{BestAgg} . Finally, the ability of BestAgg surrogates to
301 represent species-level community patterns was checked against results from analyses based on
302 species with those obtained using the surrogates. BestAgg surrogates might substitute effectively
303 species-level information in subsequent sampling programs or very similar case studies if (i) results
304 using surrogates were consistent with those obtained using species, (ii) ρ_{BestAgg} fell within or above
305 random expectations, and (iii) the associated probability of Type I error was < 0.05 (Bevilacqua et
306 al., 2013).

308 2.5. Taxonomic Sufficiency and BestAgg compared

309 The performance of BestAgg surrogates was compared with results obtained applying a more
310 classical approach for species surrogacy based on the concept of Taxonomic Sufficiency (see for
311 reviews Dauvin et al., 2003; Terlizzi et al., 2003; Jones, 2008). Previous studies on macro_benthic
312 assemblages from the investigated lagoons indicated the level of family as effective in describing
313 species-level community patterns (Munari et al., 2009). Species-level data from MAR were

314 aggregated at family level following the Linnaean taxonomy and a PERMANOVA based on Bray-
315 Curtis dissimilarity on untransformed data (design as above) was performed. Data ordinations
316 (nMDS) were plotted to depict community patterns when using species, families, and BestAgg
317 surrogates. Also, the original species matrix was aggregated at Genus, Family, Order, Class, and
318 Phylum level, and a second-stage nMDS ordination was plotted to visualize patterns of correlation
319 among species, BestAgg, and higher-taxon matrices. Finally, we estimated savings derived from the
320 use of BestAgg surrogates and families following the approach proposed by Ferraro and Cole
321 (1995). In this approach, savings are estimated by taking the ratio of the number of surrogates to the
322 number of species, assuming that the time spent to identify organisms and sample processing is
323 proportional to the number of categories (e.g. species, higher taxa, groups) to which they must be
324 assigned (see also Thompson et al., 2003).

Eliminato: genus, family, order, class

Eliminato: phylum

326 2.6. Application of BestAgg surrogates

327 BestAgg surrogates derived from the pilot lagoon were employed to analyze macro_benthic
328 assemblages from the other lagoons, in order to ascertain whether the BestAgg approach is
329 conducive to a set of surrogates that might be successfully applied to comparable study contexts.
330 Species-level data from SCA and GOR were therefore aggregated using the set of BestAgg
331 surrogates from MAR (the pilot lagoon). PERMANOVAs based on Bray-Curtis dissimilarity
332 matrices (untransformed data) and nMDS ordinations were performed for SCA and GOR using
333 BestAgg surrogates and results compared with those obtained from analyses at species level. In
334 both cases, designs for analyses consisted of two random crossed factors, Time (Ti) and Station
335 (St), although differing in the number of levels of each factor (see section Study area), with $n = 3$.
336 SIMPER analyses were done in order to identify species most contributing to the observed patterns
337 by considering only species whose contribution to dissimilarities was $\geq 3\%$.

338 Analyses were performed using R (R Development Core Team, 2013) and the PRIMER v 6
339 software package (Clarke and Gorley, 2006), including the PERMANOVA+ add-on package
340 (Anderson and Gorley, 2008).

342 3. Results

343 The linear regression of ρ values against $\ln(\phi)$ from the pilot lagoon (MAR) was significant ($P <$
344 0.001) with a correlation value of $R^2 = 0.69$, indicating that the loss information associated to a
345 given set of surrogates strongly depended on the level of aggregation following a semi-log model
346 (Fig. 2).

347 The null model based on MAR predicted that the lowest ϕ value allowing 95% of PERMANOVAs
348 on aggregated data to give consistent results with those obtained at species level was $\phi_{\text{low}} = 0.29$,
349 corresponding to $G_{\text{min}} = 25$ (Table 1, Fig. 2), under the null hypothesis that surrogates are random
350 groups of species from the original species pool. This meant that 25 surrogates could substitute the
351 original 87 species found in the lagoon while still allowing quantifying spatio-temporal community
352 patterns as well as at species level.

353 BestAgg surrogate selection based on MAR led to a set of $G_{\text{BestAgg}} = G_{\text{min}} = 25$ surrogates (see
354 Table A2 in Appendix A, and Table B1 in Appendix B). Most surrogates were formally taxonomic,
355 although in some cases they corresponded to trophic or indicator groups (Table A2), with 11
356 species (*Anadara transversa*, *Crassostrea gigas*, *Arcuatula senhousia*, *Ruditapes philippinarum*,
357 *Armandia cirrhosa*, *Alitta succinea*, *Cyathura carinata*, *Carcinus aestuarii*, *Dyspanopeus sayi*,
358 *Upogebia pusilla*, *Chironomus salinarius*), four families (Tellinidae, Semelidae, Capitellidae,
359 Spionidae), three orders (Phyllodocida, Eunicida, Amphipoda), six classes (Anthozoans, Other

360 Bivalves, Gastropods, Oligochaetes, Other Polychaetes, Other Crustaceans), and one morphological
361 group (Flat/ribbon worms).

362 The test based on randomizations for MAR showed that the probability of G_{BestAgg} to fail in
363 depicting species-level community patterns was $P = 0.036$, and correlation value ρ_{BestAgg} fell within
364 the frequency distribution specific for G_{BestAgg} (Fig. 3), indicating that the selected BestAgg
365 surrogates satisfied the assumptions underlying the null model.

366 For the pilot lagoon MAR, PERMANOVA carried out at species level showed a significant $Ti \times St$
367 interaction (Table 2), indicating significant variation among stations that varied in time. Such
368 patterns were clearly depicted in the nMDS ordination of $Ti \times St$ centroids (Fig. 4). Analyses based
369 on data aggregated using the set of BestAgg surrogates allowed obtaining consistent results (Table
370 2) and an almost interchangeable ordination of $Ti \times St$ centroids (Fig. 4). The same results were
371 obtained with family-level data (Table 2, Fig. 4), indicating that both families (64 taxa) and
372 BestAgg surrogates (25 taxa/groups) were effective in representing species-level community
373 patterns. With respect to species-level analyses, the expected timesaving during sample processing
374 and identifications was higher when using the BestAgg surrogates (>70%) than for family-level
375 analyses (~26%).

376 The second stage nMDS ordination showed a decreasing correlation with species-level information
377 at decreasing taxonomic resolution. The genus level had the strongest relationships with species,
378 followed by the family level, with a drop in correlations beyond the order level (Fig. 5). BestAgg
379 surrogates clustered with species, genus, and family levels (Fig. 5) indicating that, despite the
380 limited number of surrogates compared to such taxonomic levels, the information retained in the
381 BestAgg data matrix was strongly correlated with species-level information as well.

382 PERMANOVA on macro_benthic assemblages from SCA and GOR identified using the BestAgg
383 surrogates derived from MAR (pilot lagoon) showed results consistent with those that would have
384 been obtained if species were analyzed, i.e. a significant $T_i \times St$ interaction (Table 2).

385 For the Scardovari lagoon (SCA), the nMDS ordination plot based on species-level data clearly
386 depicted temporal changes in patterns of variability among stations (Fig. 6a). A clear gradient from
387 more confined stations towards stations at the mouth of the **embayment** (see Fig. 1) was visible in
388 time 1 but not in times 2 and 3 (Fig. 6a). The same patterns were noticed when using BestAgg
389 surrogates (Fig. 6a). The SIMPER analysis on BestAgg surrogates most contributing to such
390 temporal changes showed a general increase, from time 1 to times 2 and 3, of opportunistic/tolerant
391 organisms (e.g. Spionidae, Capitellidae, *C. salinarius*, *A. senhousia*) associated to a reduction in
392 sensitive groups (e.g. Amphipoda), and Oligochaeta (Table 4).

Eliminato: embayment

393 For the Goro lagoon (GOR), the nMDS ordination at species level showed a clear separation
394 between station centroids from the first two times and the remaining times of sampling, which
395 clustered **together** respectively with exception of centroids of station 3 (see Fig. 1) that lay apart
396 from centroids of the remaining stations in all times (Fig. 6b). Such patterns were clearly depicted
397 when using BestAgg surrogates, except station 3 in time 1 (Fig. 6b). The SIMPER analysis on
398 BestAgg aggregated data showed a lower abundance of Spionidae and Oligochaeta in station 3 with
399 respect to the **remaining** stations, associated to a higher abundance of Capitellidae and *R.*
400 *philippinarum* (Table 5).

Eliminato: togheter

Eliminato: remaing

Eliminato: V.

402 4. Discussion

403 The effectiveness of the BestAgg approach in providing surrogates able to depict community
404 patterns as at species level has been demonstrated in soft bottom and rocky reef macro_benthic

405 communities, whether composed by individual or colonial organisms, subjected to both natural and
406 anthropogenic drivers of change (Bevilacqua et al., 2013). Also, the approach has been successfully
407 applied to freshwater macro-invertebrate in exploring environmental gradients in continental river
408 basins (Milošević et al., 2014). It could be argued, however, that the approach might exhibit optimal
409 responses when analysing sharp environmental variations that cause gross changes in community
410 structure, which are likely detectable although using operational units other than species.

411 Consequently, the approach might perform less effectively in approximating complex species-level
412 patterns driven by multiple and interacting environmental factors, as could occur in invertebrate
413 communities inhabiting transitional ecosystems such as estuaries and coastal lagoons. Indeed, the
414 strong relationship between information loss and the aggregation level (ϕ), which underlies the
415 BestAgg approach, has been confirmed analysing published case studies on the application of
416 surrogates in Tws (e.g. estuaries, coastal lagoons, salt marshes) worldwide (Bevilacqua et al.,
417 2012). In this respect, our findings provide first direct evidence of the robustness of BestAgg when
418 applied to Tws, demonstrating the ability of the approach to quantifying consistently complex
419 patterns of spatio-temporal variability of assemblage structure in northern Adriatic lagoons, where
420 the effects of natural environmental gradients and human disturbance largely overlap.

421 Northern Adriatic lagoons share features that are typical of Tws, namely low number of species and
422 taxonomic diversity with respect to marine ecosystems, dominance of few species adapted to the
423 particular environmental conditions and generally tolerant to sharp regimes of natural disturbance
424 (Elliott and Quintino, 2007; Pérez-Ruzafa et al., 2010; Borja et al., 2011). However, beyond such
425 general commonalities, each lagoon holds its own distinctiveness. The pilot lagoon (MAR) and the
426 SCA lagoon were characterized by a relatively high number of species, whereas in GOR the
427 community appeared quite poor in species. The three lagoons were affected by different levels of
428 anthropogenic disturbance related to clam harvesting, and differed also in the amount of nutrient-
429 rich freshwater inputs, with GOR being the one with the greater influx of freshwater through two

Eliminato: underlying

430 deltaic branches, and many drains. Such differences are likely to concur in shaping assemblage
431 structure and composition within the three lagoons, and determining dissimilarities occurring
432 among them (Munari and Mistri, 2008; Pérez-Ruzafa et al., 2013; Sfriso et al., 2014).
433 One of the required prerogatives of any surrogacy approach is to provide surrogates for species that
434 perform consistently in comparable experimental contexts or in subsequent reiterations of the same
435 study (e.g. in monitoring programs) (Olsgard and Somerfield, 2000). Our results showed that the
436 BestAgg approach led to a reliable set of surrogates that are effective in quantifying the overall
437 community patterns as at species level across distinct lagoons from the same geographic area,
438 working over and beyond the environmental heterogeneity and peculiarity characterizing such
439 highly variable systems. This ability likely depends on the fact that surrogate selection in BestAgg
440 tends to maximize ecological information and, therefore, is more prone to capture salient ecological
441 and functional properties of systems under study. In the present study, 5 out of 11 surrogates at
442 species level were alien species, thus reflecting the vulnerability to invasion of northern Adriatic
443 lagoons (Occhipinti-Ambrogi, 2000). The remaining species included sensitive organisms (e.g.
444 *Armandia cirrhosa*), indicators of environmental stress (e.g. *Chironomus salinarius*), or commercial
445 targets (e.g. *Carcinus aestuarii*), providing direct information on the ecological status of the
446 investigated lagoons. Also, surrogate selection in BestAgg prioritizes ecological information even
447 when species are grouped, so that surrogate groups (taxa, morphological groups, etc.) are still
448 meaningful from an ecological perspective. In our case study, for instance, several surrogate groups
449 may be considered as indicator taxa (e.g. Capitellidae, Spionidae) (Simboura and Zenetos, 2002), or
450 trophic guilds (e.g. Eunicida, which were all predators) (Table A2). Only species with low
451 relevance and/or of difficult identification are grouped based exclusively on shared features (e.g.
452 taxonomic relatedness, morphology) (Bevilacqua et al., 2013). In this view, surrogate selection in
453 BestAgg may help the interpretation of patterns, which could result less intuitive whether based on
454 a single taxonomic rank. In the SCA lagoon, for example, the disruption of the spatial pattern of

Eliminato: meaningful

Eliminato: interpretation

Eliminato:

455 confinement from time 1 (May 2007) to times 2 and 3 (November 2007 and 2008) could be a
456 consequence of increased agricultural run-offs, which are likely to follow increased rainfall regimes
457 in autumn, especially in time 3 (Barbi et al., 2012), resulting in an increased dominance of
458 opportunistic/tolerant polychaetes (e.g. Capitellidae, Spionidae) associated to a decreased
459 abundance of amphipods (Dauvin and Ruellet, 2007). For the Goro lagoon, instead, the peculiarity
460 of station 3, beyond seasonal variations between times 1-2 (autumn-winter 2004) and times 3-4
461 (spring-summer 2005), could depend on the fact that the remaining stations were subjected to
462 nutrient inputs from the Po river branches whereas station 3 is mostly marine and within an area
463 affected by clam harvesting. Such factors probably led in this last station to the decreased
464 abundance of groups tolerant to organic enrichment (e.g. Spionidae, Oligochaeta), and to the
465 increased abundance of *R. philippinarum* that found in this area of the lagoon the optimal
466 environmental conditions for growth (Vincenzi et al., 2006).

467 Indeed, also a more classic approach to species surrogacy based on Taxonomic Sufficiency, and
468 specifically the use of family-level identification of organisms, was effective in elucidating species-
469 level patterns for all lagoons, confirming the results obtained in other transitional water ecosystems
470 (Mistri and Rossi, 2001; De Biasi et al., 2003; Chainho et al., 2007; Munari et al., 2009; Shokri and
471 Gladstone, 2009). Such findings, however, may be simply due to the effect of a close numerical
472 overlapping between species and family richness, which could generate interchangeable community
473 patterns when using either species or family-level taxonomic resolution, rather than to the similarity
474 in ecological response of species within families (Bevilacqua et al., 2012). For instance, in the three
475 lagoons about 80% of families was monotypic, the 11-14% had two species, and only the 9-6%
476 included three or more species. This probably reflects the outcomes of ecological and evolutionary
477 mechanisms driving community composition, which have led only very few species able to tolerate
478 the stressful condition prevailing in such environments to evolve for each family (Mistri and Rossi,
479 2001).

Eliminato: remaing

Eliminato: V.

480 Irrespective of the surrogacy approach employed, surrogates for species should not be defined *a*
481 *priori* and should require at least a pilot assessment and validation before they could be assumed as
482 suitable substitutes of species for subsequent monitoring or for very similar study contexts (Lenat
483 and Resh, 2001; Terlizzi et al., 2003; Defeo and Lercari, 2004; Dauvin et al., 2007; Jones, 2008).
484 Also, once fixed, surrogates for species should not be considered as unchangeable operational units,
485 and periodical adjustments involving species-level analyses should alternate routine monitoring
486 based on surrogates, due to possible changes in assemblage structure over time (Tataranni et al.,
487 2009; Musco et al., 2011). In this respect BestAgg is no exception. It should be noted, however, that
488 random simulations underlying the identification of ϕ_{low} provide a wide range of potential scenarios
489 of aggregated assemblage structure, which likely increases the robustness of the approach to
490 variations in assemblage composition. In addition, variations in the number of species might have
491 little outcomes on the effectiveness of the fixed ϕ_{low} , unless consisting in considerable increases in
492 species richness. Our findings showed the robustness of the approach to capture temporal variations
493 in community structure confirming the results of other case studies (Bevilacqua et al., 2013),
494 although further investigations focusing on long-term data are needed to ascertain whether and,
495 eventually, how often, a refinement of BestAgg surrogates should be recommended.
496 Species surrogacy should be a more formal and transparent process than is often practised, and
497 should conjugate solid theory and quantitative assessments of surrogates (Jones, 2008). The
498 BestAgg approach relies on a rigorous theoretical and statistical framework aimed at quantifying
499 the sufficient number of surrogates objectively and identifying the type of effective surrogates
500 based on objective macro-criteria (Bevilacqua et al., 2012; Bevilacqua et al., 2013). Above all, the
501 BestAgg approach demonstrated several advantages with respect to the classic approach based on
502 Taxonomic Sufficiency. Our results on northern Adriatic lagoons showed that BestAgg surrogates
503 might lead saving an additional 40% of time during sample processing and identifications of
504 organisms if compared to family-level analyses, while still retaining a greater correspondence with

505 species-level information. The approach, therefore, may enhance cost-effectiveness of species
506 surrogacy, also providing a statistical framework controlling for uncertainty in the application of the
507 selected set of surrogates (Bevilacqua et al., 2013). More importantly, surrogate selection in
508 BestAgg takes into account the ecological meaning of surrogates, and unleashes the investigator to
509 exploit different surrogate types, from morphological groups to functional groups, or mixing
510 different surrogate types, such as taxa from different taxonomic ranks. This last prerogative might
511 improve the effectiveness of the selected set of surrogates, beyond the static approach of focusing
512 exclusively on taxa of a single taxonomic rank higher than species (Groc et al., 2010).

513 The WFD and the MSFD directives are leading to an increasing demand for monitoring programs
514 and, consequently, for increasing research efforts and expertise in invertebrate taxonomy in Europe.
515 Sooner or later this will impose dilemmas regarding the spatial coverage, the sampling intensity and
516 the sampling frequency in order to obtain reliable results, particularly because the information
517 obtained is expected to guide management actions (Dauvin et al., 2012). This may be particularly
518 relevant for transitional water ecosystems, for which there is a urgent need for understanding the
519 links between the effects of human activities and changes in communities (Munari and Mistri,
520 2008), as natural and anthropogenic environmental drivers have created a mosaic of conditions
521 making difficult separating natural variations from human-induced alterations (Estuarine Quality
522 Paradox: Elliott and Quintino, 2007; De Paz et al., 2008). Further research aiming at defining sets
523 of operational units for rapid assessments of community changes from local to regional scale, able
524 to depict species-level patterns and to inform about underlying ecological processes while minimize
525 efforts during sample processing and taxonomic identifications, are strongly needed ([Olsgard and](#)
526 [Sommerfield, 2000](#); [Dauvin et al. 2007](#); [De Paz et al., 2008](#); [Borja et al. 2011](#)). In this direction, the
527 BestAgg approach provides a new perspective for optimizing trade-offs between practical
528 constraints and the need for essential ecological information in environmental monitoring.

530 **Acknowledgements**

531 The research leading to these results was instrumental for the preparation of the EU Community's
2
532 Seventh Framework (FP7/2007–2013) Programs COCONET, Grant Agreement No. 287844) and
4
533 PERSEUS (Grant Agreement No. 287600).
6

534
7
8
9

10 **References**

11
12
13 Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance.
14
15 Aust. Ecol. 26, 32–46.
16
17 Anderson, M.J., Connell, S.D., Gillanders, B.M., Diebel, C.E., Blom, W.M., Saunders, J.E.,
18
19 Landers, T.J., 2005. Relationships between taxonomic resolution and spatial scales of
20
21 multivariate variation. J. Anim. Ecol. 74, 636–646.
22
23 Anderson, M.J., Gorley, R.N., 2008. PERMANOVA+ for PRIMER: Guide to Software and
24
25 Statistical Methods. PRIMER-E, Plymouth, UK.
26
27
28 Barbi, A., Monai, M., Racca, R., Rossa, A. M., 2012. Recurring features of extreme autumnal
29
30 rainfall events on the Veneto coastal area. Nat. Hazards Earth Syst. Sci. 12, 2463–2477.
31
32 Basset, A., Barbone, E., Borja, A., Brucet, S., Pinna, M., Quintana, X.D., Reizopoulou, S.,
33
34 Rosati, I., Simboura, N., 2012. A benthic macroinvertebrate size spectra index for implementing
35
36 the Water Framework Directive in coastal lagoons in Mediterranean and Black Sea ecoregions.
37
38 Ecol. Indic. 12, 72–83.
39
40 Bertasi, F., Colangelo, M. A., Colosio, F., Gregorio, G., Abbiati, M., Ceccherelli, V.U., 2009.
41
42 Comparing efficacy of different taxonomic resolutions and surrogates in detecting changes in
43
44 soft bottom assemblages due to coastal defence structures. Mar. Pollut. Bull. 58, 686–694.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

552 Bertrand, Y., Pteijel, F., Rouse, G.W., 2006. Taxonomic surrogacy in biodiversity assessments,
553 and the meaning of Linnaean ranks. *Syst. Biodivers.* 4, 149–159.
1
2
534 Bevilacqua, S., Fraschetti, S., Terlizzi, A., Boero, F., 2009a. The use of taxonomic distinctness
4 indices in assessing patterns of biodiversity in modular organisms. *Mar. Ecol.* 30, 151–163.
555
6
556 Bevilacqua, S., Fraschetti, S., Musco, L., Terlizzi, A., 2009b. Taxonomic sufficiency in the
8 detection of natural and human-induced changes in marine assemblages: A comparison of
557 habitats and taxonomic groups. *Mar. Pollut. Bull.* 58, 1850–1859.
10
558
12
13
559 Bevilacqua, S., Fraschetti, S., Musco, L., Guarnieri, G., Terlizzi, A., 2011. Low sensitiveness of
15 taxonomic distinctness indices to human impacts: Evidences across marine benthic organisms
560 and habitat types. *Ecol. Indic.* 11, 448–455.
17
561
18
19
562 Bevilacqua, S., Terlizzi, A., Fraschetti, S., Claudet, J., Boero, F., 2012. Taxonomic relatedness
21 does not matter for species surrogacy in the assessment of community responses to
563 environmental drivers. *J. Appl. Ecol.* 49, 357–366.
23
564
25
565 Bevilacqua, S., Claudet, J., Terlizzi, A., 2013. Best Practicable Aggregation of Species: a step
27 forward for species surrogacy in environmental assessment and monitoring. *Ecol. Evol.* 3,
566 3780–3793.
28
29
30
31
32
568 [Borja, A., Barbone, E., Basset, A., Borgersen, G., Brkljacic, M., Elliott, M., Garmendia, J.M.,
34 Marques, J.C., Mazik, K., Muxika, I., Neto, J.M., Norling, K., Rodríguez, J.G., Rosati, I., Rygg,
569 B., Teixeira, H., Trayanova, A., 2011. Response of single benthic metrics and multi-metric
36 methods to anthropogenic pressure gradients, in five distinct European coastal and transitional
37 ecosystems. *Mar. Poll. Bullet.* 62, 499–513.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65](#)

573 Bremner, J., Rogers, S.I., Frid, C.L.J., 2006. Methods for describing ecological functioning of
574 marine benthic assemblages using biological traits analysis (BTA). *Ecol. Indic.* 6, 609–622.
1
2
575 Brind'Amour, A., Laffargue, P., Morin, J., Vaz, S., Foveau, A., Le Bris, H., 2014.
4
576 Morphospecies and taxonomic sufficiency of benthic megafauna in scientific bottom trawl
6
577 surveys. *Cont. Shelf Res.* 72, 1–9.
8
9
578 Cardoso, P., Erwin, T.L., Borges, P.A.V., New, T.R., 2011. The seven impediments in
10
579 invertebrate conservation and how to overcome them. *Biol. Conserv.* 144, 2647–2655.
12
13
580 Carranza, A., Defeo, O., Arim, M., 2011. Taxonomic relatedness and spatial structure of a shelf
14
581 benthic gastropod assemblage. *Divers. Distribut.* 17, 25–34.
16
17
582 [Chainho, P., Lane, M.F., Chaves, M.L., Costa, J.L., Costa, M.J., Dauer, D.M., 2007. Taxonomic](#)
19
583 [sufficiency as a useful tool for typology in a poikilohaline estuary. *Hydrobiologia* 587, 63–78.](#)
21
22
584 Clarke, K.R., Gorley, R.N., 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.
23
24
585 Cognetti, G., Maltagliati, F., 2008. Perspectives on the ecological assessment of transitional
26
586 waters. *Mar. Pollut. Bull.* 56, 607-608.
28
29
587 Dauvin, J.-C., Gomez Gesteira, J.L., Salvande Fraga, M., 2003. Taxonomic sufficiency: an
30
588 overview of its use in the monitoring of sublittoral benthic communities after oil spills. *Mar.*
31
589 *Pollut. Bull.* 46, 552–555.
34
35
590 Dauvin, J.-C., Bellan, G., Bellan-Santini, D., 2007. Benthic indicators: From subjectivity to
37
591 objectivity – Where is the line? *Mar. Pollut. Bull.* 60, 947–953.
39
40
592 Dauvin, J.-C., Ruellet, 2007. Polychaete/amphipod ratio revisited. *Mar. Pollut. Bull.* 55, 215–
41
42
593 224.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

594 Dauvin, J.-C., Alizier, S., Rolet, C., Bakalem, A., Bellan, G., Gomez Gesteira, J.L., Grimes, S.,
595 de-la-Ossa-Carretero, J.A., Del-Pilar-Ruso, Y., 2012. Response of different benthic indices to
1
596 diverse human pressures. *Ecol. Indic.* 12, 143–153.
3
4
597 De Biasi, A.M., Bianchi, C.N., Morri, C., 2003. Analysis of macrobenthic communities at
6
598 different taxonomic levels: an example from an estuarine environment in the Ligurian Sea (NW
8
599 Mediterranean). *Est. Coast. Shelf Sci.* 58, 99–106.
9
10
600 Defeo, O., Lercari, D., 2004. Testing taxonomic resolution levels for ecological monitoring in
12
601 sandy beach macrobenthic communities. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 14, 65–74.
14
15
602 De Paz L., Patricio J., Marques J.C., Borja A., Laborda A.J., 2008. Ecological status assessment
17
603 in the lower Ebso estuary (Spain). The challenge of habitat heterogeneity integration: a benthic
18
19
604 perspective. *Mar. Pollut. Bull.* 56, 1275–1283.
20
21
605 Dethier, M.N., Schoch, G.C., 2006. Taxonomic sufficiency in distinguishing natural spatial
23
606 patterns on an estuarine shoreline. *Mar. Ecol. Progr. Ser.* 306, 41–49.
25
26
607 EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council establishing a
27
28
608 framework for community action in the field of water policy.
29
30
609 EC, 2008. Directive 2008/56/EC of the European Parliament and of the Council establishing a
32
610 framework for community action in the field of marine environmental policy.
34
35
611 Elliott, M., Quintino, V., 2007. The estuarine quality paradox: environmental homeostasis and
36
37
612 the difficulty of detecting anthropogenic stress in naturally stressed areas. *Mar. Pollut. Bull.* 54,
38
613 640–645.
40
41
614 Ellis, D., 1985. Taxonomic sufficiency in pollution assessment. *Mar. Pollut. Bull.* 16, 459.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

615 Ferraro, S.P., Cole, F.A., 1995. Taxonomic level sufficient for assessing pollution impacts on
616 the southern Californian Bight macrobenthos – revisited. *Environ. Toxicol. Chem.* 14, 1031–
617 1040.
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645

Giangrande, A., Licciano, M., Musco, L., 2005. Polychaetes as environmental indicators
revisited. *Mar. Pollut. Bull.* 50, 1153–1162.

Gomez Gesteira, J.L., Dauvin, J.C., Fraga, M.S., 2003. Taxonomic level for assessing oil spill
effects on soft-bottom sublittoral benthic communities. *Mar. Pollut. Bull.* 46, 562–572.

Gray, J.S., Aschan, M., Carr, M.R., Clarke, K.R., Green, R.H., Pearson, T.H., Rosenberg, R.,
Warwick, R.M., 1988. Analysis of community attributes of the benthic macrofauna of
Frierfjord/Langesundfjord and in a mesocosm experiment. *Mar. Ecol. Prog. Ser.* 46, 151-165.

Groc, S., Delabie, J.H.C., Longino, J.T., Orivel, J., Majer, J.D., Vasconcelos, H.L., Dejean, A.,
2010. A new method based on taxonomic sufficiency to simplify studies on Neotropical ant
assemblages. *Biol. Conserv.* 143, 2832–2839.

Guarnieri, G., Terlizzi, A., Bevilacqua, S., Fraschetti, S., 2012. Increasing heterogeneity of
sensitive assemblages as a consequence of human impact in submarine caves. *Mar. Biol.* 159,
1155–1164.

Heino, J., 2014. Taxonomic surrogacy, numerical resolution and responses of stream
macroinvertebrate communities to ecological gradients: Are the inferences transferable among
regions? *Ecol. Indic.* 36, 186–194.

Heino, J., Soininen, J., 2007. Are higher taxa adequate surrogates for species-level assemblage
patterns and species richness in stream organisms? *Biol. Conserv.* 137, 78–89.

636 Jiang, X., Xiong, J., Song, Z., Morse, J.C., Jones, F.C., Xie, Z., 2013. Is coarse taxonomy
637 sufficient for detecting macroinvertebrate patterns in floodplain lakes? *Ecol. Indic.* 27, 48–55.
638
639 Jones, F.C., 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater
640 bioassessments using benthic macroinvertebrates. *Environ. Rev.* 16, 45–69.
641
642 Lasiak, T., 2003. Influence of taxonomic resolution, biological attributes and data
643 transformations on multivariate comparisons of rocky macrofaunal assemblages. *Mar. Ecol.*
644 *Progr. Ser.* 250, 29–34.
645
646 Lenat, D.R., Resh, V.H., 2001. Taxonomy and stream ecology – the benefits of genus- and
647 species-level identifications. *J. North Am. Benthol. Soc.* 20, 287–298.
648
649 Lovell, S., Hamer, M., Slotow, R., Herbert, D., 2007. Assessment of congruency across
650 invertebrate taxa and taxonomic levels to identify potential surrogates. *Biol. Conserv.* 139, 113–
651 125.
652
653 Marshall, J.C., Steward, A.L., Harch, B.D., 2006. Taxonomic resolution and quantification of
654 freshwater macroinvertebrate samples from an Australian dryland river: The benefits and costs
655 of using species abundance data. *Hydrobiologia* 572, 171–194.
656
657 Mellin, C., Delean, S., Caley, J., Edgar, G., Meekan, M., Pitcher, R., Przeslawski, R., Williams,
658 A., Bradshaw, C., 2011. Effectiveness of biological surrogates for predicting patterns of marine
659 biodiversity: a global meta-analysis. *PLoS One* 6, e20141.
660
661 Milošević D., Stojković, M., Čerba, D., Petrović, A., Paunović, M., Simić, V., 2014. Different
662 aggregation approaches in the chironomid community and the threshold of acceptable
663 information loss. *Hydrobiologia* 727, 35–50.

657 Mistri, M., Rossi, R., 2001. Taxonomic sufficiency in lagoonal ecosystems. J. Mar. Biol. Ass.
658 U.K. 81, 339–340.
1
2
659 Munari, C., 2013. Benthic community and biological trait composition in respect to artificial
4 coastal defence structures: A study case in the northern Adriatic Sea. Mar. Environ. Res. 90,
660 47–54.
6
661
8
662 Munari, C. Mistri, M., 2007. Effect of copper on the scope for growth of clams (*Tapes*
10 *philippinarum*) from a farming area in the Northern Adriatic Sea. Mar. Environ. Res. 64, 347–
163 357.
12
164
14
15
665 Munari, C. Mistri, M., 2008. Biodiversity of soft-sediment benthic communities from Italian
17 transitional waters. J. Biogeog. 35, 1622–1637.
666
18
19
667 Munari, C., Mistri, M., 2012. Ecological status assessment and response of benthic communities
21 to environmental variability: The Valli di Comacchio (Italy) as a study case. Mar. Environ. Res.
668 81, 53–61.
23
669
25
26
670 Munari, C., Warwick, R.M., Mistri, M., 2009. Monitoring with benthic fauna in Italian coastal
27 lagoons: new tools for new prospects. Aquat. Conserv. Mar. Freshw. Ecosyst. 19, 575–587.
28
671
29
30
672 [Musco, L., Mikac, B., Tataranni, M., Giangrande, A., Terlizzi, A., 2011. The use of coarser](#)
32 [taxonomy in the detection of long-term changes in polychaete assemblages. Mar. Environ. Res.](#)
673 [71, 131–138.](#)
34
674
36
37
675 [Musco, L., Terlizzi, A., Licciano, M., Giangrande, A., 2009. Taxonomic structure and the](#)
38 [effectiveness of surrogates in environmental monitoring: a lesson from polychaetes. Mar. Ecol.](#)
676 [Progr. Ser. 383, 199–210.](#)
40
677
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

678 Neeson, T., van Rijn, I., Mandelik, Y., 2013. How taxonomic diversity, community structure
679 and sample size determine the reliability of higher taxon surrogates. *Ecol. Appl.* 23, 1216–1225.
1
2
680 [Olsgard, F., Somerfield, P.J., 2000. Surrogates in marine benthic investigations - which](#)
4 [taxonomic unit to target? *J. Aquat. Ecosyst. Stress Rec.* 7, 25–42.](#)
6
682 Occhipinti-Ambrogi, A., 2000. Biotic invasions in a Mediterranean lagoon. *Biol. Invasions* 2:
8
683 165-176.
10
11
684 Pearson, T.H., Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment
13 and pollution of the marine environment. *Oceanogr. Mar. Biol. Annu. Rev.* 16, 229–311.
14
15
686 [Pérez-Ruzafa, A., Marcos, C., Pérez-Ruzafa, I., Pérez-Marcos, M., 2010. Coastal lagoons:](#)
17 [“transitional ecosystems” between transitional and coastal waters. *J. Coast. Conserv.* 36, 160–](#)
18 [166.](#)
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

699 | [Sfriso, A., Facca, C., Bon, D., Giovannone, F., Buosi, A., 2014. Using phytoplankton and](#)
700 | [macrophytes to assess the trophic and ecological status of some Italian transitional systems.](#)
701 | [Cont. Shelf Sci. 81, 88–98.](#)
702 | Shokri, M.R., Gladstone, W., 2009. Higher taxa are effective surrogates for species in the
703 | selection of conservation reserves in estuaries. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 19, 626–
704 | 636.
705 | [Simboura, N., Zenetos, A., 2002. Benthic indicators to use ecological quality classification of](#)
706 | [Mediterranean soft bottom marine ecosystems, including a new biotic index. *Mediterr. Mar. Sci.*](#)
707 | [3, 77–111.](#)
708 | Simonini, R., Ansaloni, I., Bonvicini Pagliai, A.M., Prevedelli, D., 2004. Organic enrichment
709 | and structure of the macrozoobenthic community in the northern Adriatic Sea in an area facing
710 | Adige and Po mouths. *ICES J. Mar. Sci.* 61, 871–881.
711 | Siqueira, T., Bini, L.M., Roque, F.O., Cottenie, K., 2012. A metacommunity framework for
712 | enhancing the effectiveness of biological monitoring strategies. *PLoS One* 7, e43626.
713 | Soares-Gomes, A., Mendes, C.L.T., Tavares, M., Santi, L., 2012. Taxonomic sufficiency of
714 | polychaete taxocenes for estuary monitoring. *Ecol. Indic.* 15, 149–156.
715 | Solidoro, C., Pastres, S., Melaku Canu, D., Pellizzato, M., Rossi, R., 2000. Modelling the
716 | growth of *Tapes philippinarum* in Northern Adriatic lagoons. *Mar. Ecol. Progr. Ser.* 199, 137–
717 | 148.
718 | Tagliapietra, D., Sigovini, M., Volpi Ghirardini, A., 2009. A review of terms and definitions to
719 | categorise estuaries, lagoons and associated environments. *Mar. Freshw. Res.* 60, 497–509.

- 720 Terlizzi, A., Bevilacqua, S., Fraschetti, S., Boero, F., 2003. Taxonomic sufficiency and the
721 increasing insufficiency of taxonomic expertise. *Mar. Pollut. Bull.* 46, 544–560.
1
2
- 722 [Tataranni, M., Maltagliati, F., Floris, A., Castelli, A., Lardicci, C., 2009. Variance estimate and](#)
4 [taxonomic resolution: an analysis of macrobenthic spatial patterns at different scales in a](#)
723 [Western Mediterranean coastal lagoon. *Mar. Environ. Res.* 67, 219–229.](#)
6
724
- 725 Terlizzi, A., Bevilacqua, S., Scuderi, D., Fiorentino, D., Guarnieri, G., Giangrande, A.,
10
726 Licciano, M., Felling, S., Fraschetti, S., 2008. Effects of offshore platforms on soft bottom
12 macrobenthic assemblages: a case study in a Mediterranean gas field. *Mar. Pollut. Bull.* 56,
727 1303–1309.
14
728
- 16 Terlizzi, A., Anderson, M.J., Bevilacqua, S., Fraschetti, S., Włodarska-Kowalczyk, M.,
729 Ellingsen, K.E., 2009. Beta diversity and taxonomic sufficiency: Do higher-level taxa reflect
18 heterogeneity in species composition? *Divers. Distribut.* 15, 450–458.
730
20
21
- 22 Thompson, B.W., Riddle, M.J., Stark, J.S., 2003. Cost-efficient methods for marine pollution
232 monitoring at Casey Station, East Antarctica: the choice of sieve mesh-size and taxonomic
24 resolution. *Mar. Pollut. Bull.* 46, 232–243.
253
26
274
- 28 Van Hoey, G., Borja, A., Birchenough, S., Buhl-Mortensen, L., Degraer, S., Fleischer, D.,
735 Kerckhof, F., Magni, P., Muxika, I., Reiss, H., Schröder, A., Zettler, M.L., 2010. The use of
31 benthic indicators in Europe: From the Water Framework Directive to the Marine Strategy
736 Framework Directive. *Mar. Pollut. Bull.* 60, 2187–2196.
32
337
34
- 358 Vanderklift, M.A., Ward, T.J., Jacoby, C.A., 1996. Effect of reducing taxonomic resolution on
36 ordinations to detect pollution-induced gradients in macrobenthic infaunal assemblages. *Mar.*
739 *Ecol. Progr. Ser.* 136, 137–145.
38
390
40
411
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

742 Vincenzi, S., Caramori, G., Rossi, R., De Leo, G. A., 2006. Estimating clam yield potential in
743 the Sacca di Goro lagoon (Italy) by using a two-part conditional model. *Aquaculture* 261, 1281–
1
744 1291.
3
4
745 Warwick, R.M., 1993. Environmental impact studies on marine communities - pragmatical
6
746 considerations. *Aust. J. Ecol.* 18, 63–80.
8
747 White, J., Irvine, K., 2003. The use of littoral mesohabitats and their macroinvertebrate
10
748 assemblages in the ecological assessment of lakes. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 13,
12
749 331–351. [Mar. Envir. Res.](#)
14
15
750 Włodarska-Kowalczyk, M., Kędra, M., 2007. Surrogacy in natural patterns of benthic
17
751 distribution and diversity: selected taxa versus lower taxonomic resolution. *Mar. Ecol. Progr.*
18
752 Ser. 351, 53–63.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

753 **Figure captions**

754 **Figure 1.** Map of investigated lagoon systems. The white circle in the down left panel indicates the
2 study area where all the three studied lagoons were located. The remaining panels show, clock-
755 4 wise, the three lagoons: Marinetta (MAR), Goro (GOR), and Scardovari (SCA). Numbers in panels
756 6 indicate sampling stations in each lagoon.
757 8

Eliminato: are located
Eliminato: .
Eliminato: remaing

758 **Figure 2.** Semilog plot of ρ values (Spearman's correlation) between the species-level matrix and
10 each randomly aggregated matrix against the corresponding ϕ values for MAR, which served as
759 12 pilot lagoon system for BestAgg surrogate selection (see Method section for further details). Fading
760 14 gray zones indicate the range of ϕ values for which results were consistent with those obtained
761 16 analyzing species-level data. Dotted lines indicate ϕ_{low} (i.e. the lowest practicable aggregation
762 18 level), sufficient to reflect species-level patterns consistently. Results of regression analysis are also
763 20 reported.
764 22

Eliminato: Result
Eliminato: were

765 **Figure 3.** Frequency distribution ($n = 1,000$) of ρ values between the species-level matrix and
24 matrices in which species were randomly aggregated in $G_{BestAgg}$ groups (see text), based on the pilot
766 26 lagoon (MAR). Dotted lines indicate $\rho_{BestAgg}$, that is, the correlation value between the species-level
767 28 matrix and the matrix aggregated using BestAgg surrogates, which fall within random expectations.
768 30

769 **Figure 4.** Non-metric multidimensional scaling ordinations (nMDS) of $T_i \times S_t$ centroids for MAR
33 based on Bray-Curtis dissimilarity for species-level data, and data aggregated using family and
770 34 BestAgg surrogates. Symbols indicate different times of sampling, whereas numbers indicate the
35 36 sampling stations (as in Figure 1).
771 37
772 38
39

Eliminato: -
Eliminato: time

403 **Figure 5.** Second-stage nMDS ordination plot of inter-matrix rank correlations for MAR (pilot
41 lagoon) based based on Bray-Curtis dissimilarity matrices at the species (S), genus (G), family (F),
42 43
44

Eliminato: -

775 order (O), class level and using the BestAgg surrogates (BA). Numbers in brackets indicate the
776 number of taxa or surrogates for each aggregation.

777
778 **Figure 6.** Non-metric multidimensional scaling ordinations (nMDS) of $T_i \times S_t$ centroids for SCA
779 (a) and GOR (b) based on Bray–Curtis dissimilarity for species-level data and data aggregated
780 using BestAgg surrogates. Numbers indicate sampling stations as reported in Figure 1. For SCA
781 lagoon (a), lines connect stations' centroids in the three times of sampling (Time 1 = black line;
782 Time 2 = dark grey line; Time 3 = light grey line). For GOR lagoon, circles indicate the four times
783 of sampling (Time 1 = black circles; Time 2 = dark grey circles; Time 3 = light grey circles; Time 4
784 = white circles) whereas clusters group centroids with average similarity $\geq 50\%$.

- Monitoring benthic assemblages at species level requires considerable efforts
- Taxonomic Sufficiency represented the mainstream approach to reduce costs and time
- The BestAgg approach has been proposed as a new method for species surrogacy
- Here, BestAgg was assessed in transitional water systems for the first time
- BestAgg surrogates performed better than families, providing additional advantages

1 **New frameworks for species surrogacy in monitoring highly variable coastal ecosystems:**
2 **applying the BestAgg approach to Mediterranean coastal lagoons**

3
4
5 3 Stanislao Bevilacqua^{1*}, Antonio Terlizzi¹, Michele Mistri², Cristina Munari²

6
7
8
9 4 ¹*Laboratory of Zoology and Marine Biology, Department of Biological and Environmental Sciences*
10
11 5 *and Technologies, University of Salento, CoNISMa, I-73100 Lecce, Italy*

12
13
14 6 ²*Department of Life Science and Biotechnologies, University of Ferrara, I-44100 Ferrara, Italy*
15

16 7

17
18
19
20 8 *Corresponding author:

21
22 9 Dr Stanislao Bevilacqua, Ph.D.

23
24 10 Department of Biological and Environmental Sciences and Technologies

25
26
27 11 University of Salento, Campus Ecotekne, I-73100 Lecce, Italy

28
29 12 Phone +39 0832 298803

30
31 13 e-mail: stanislao.bevilacqua@unisalento.it
32

33
34 14

35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

15 **Abstract**

1
2
3
4 16 Benthic invertebrates are good indicators of aquatic ecosystem health. Yet, environmental
5
6 17 monitoring and assessment of community changes in relation to both natural and human sources of
7
8 18 disturbance involves considerable efforts for sample processing and time-consuming identifications
9
10
11 19 of organisms, which make challenging large-scale and continuous monitoring programs required
12
13 20 under the current regulatory frameworks at European scale. The use of higher taxa (e.g. families) as
14
15 21 surrogates for species is a mainstream approach to reduce cost and time associated to fine
16
17
18 22 taxonomic resolution in environmental studies, especially concerning macro-invertebrate
19
20 23 communities. However, this approach of 'Taxonomic Sufficiency' simply relies on the static
21
22
23 24 grouping of organisms in taxa belonging to a single higher taxonomic level irrespective of their
24
25 25 ecological relevance or difficulties in their taxonomic identification, leading to unnecessary losses
26
27
28 26 of taxonomic detail or ecological information. A new approach, namely the Best Practicable
29
30 27 Aggregation of Species (BestAgg), has been recently developed as an alternative procedure for
31
32 28 species surrogacy. BestAgg is based on aggregating species in the minimum number of surrogates
33
34
35 29 sufficient to depict species-level community patterns consistently, while capitalizing on ecological
36
37 30 information. Although the approach has been successfully applied to marine and freshwater
38
39
40 31 invertebrate assemblages, its effectiveness in transitional water systems, where the complex and
41
42 32 highly variable environmental conditions may affect the performance of surrogacy methods, still
43
44
45 33 remain untested. Here, we applied the BestAgg approach to quantifying spatio-temporal patterns of
46
47 34 variability of macro-invertebrate assemblages from Mediterranean coastal lagoons (Northern
48
49 35 Adriatic Sea). Surrogates were defined using species-level data from a representative lagoon
50
51
52 36 system, which served as pilot study. Then, they were used to analyze macro-invertebrate
53
54 37 assemblages in two independent lagoons in the same geographic area. Results showed that BestAgg
55
56
57 38 surrogates were effective in depicting multivariate patterns of macro-invertebrate assemblages as at
58
59 39 species level over and beyond potential variations among the investigated lagoons. The use of
60
61
62
63
64
65

40 families, following a more classic approach based on Taxonomic Sufficiency, also provided results
1
21 comparable to those obtained using species. However, with respect to families, BestAgg surrogates
3
42 allowed an estimated timesaving of 40% higher while still retaining an equivalent amount of
6
73 information on species-level patterns. More importantly, BestAgg allowed exploiting different
8
94 criteria of species aggregation, leading to a set of surrogates more aligned with
10
11 ecological/functional characteristics of organisms, suggesting that BestAgg approach may provide a
1245
13 fresh perspective for optimizing trade-offs between pragmatism and the need for relevant ecological
1446
15 information in environmental assessment and monitoring.
16
1747
18
19
2048
21
22
23

2449 **Keywords:** Coastal lagoons, Indicators, Macro-benthic invertebrates, Null models, Species
25
2650 surrogacy, Taxonomic sufficiency
27
28
29
3051
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1. Introduction

Benthic macro-invertebrates have long been used as ecological indicators in monitoring the health of marine, freshwater, and transitional water bodies because of their importance for the functioning of aquatic ecosystems and their ability to reflect natural and anthropogenic pressures effectively (Pearson and Rosenberg, 1978) through changes in abundance, taxonomic richness and composition, biomass, body-size and biological traits (Bremner et al., 2006; Terlizzi et al., 2008; Petchey and Belgrano, 2010; Basset et al., 2012; Munari and Mistri, 2012; Guarnieri et al., 2012; Munari, 2013). Due to such desirable prerogatives, macro-invertebrates represent preferential descriptors of benthic communities being considered in the current environmental policies for the assessment of the Good Environmental Status (GES) of aquatic systems at European scale (Van Hoey et al., 2010; Rice et al., 2012). Monitoring aquatic environments using macro-invertebrates requires, however, considerable efforts in sample processing and taxonomic identifications of organisms, which may be exacerbated by the limited availability of taxonomic expertise and knowledge of invertebrate species (Terlizzi et al., 2003; Cardoso et al., 2011), resulting in expensive and time-consuming procedures that make difficult coping with the continuous and large-scale monitoring programs required under the current regulatory frameworks (e.g. Water Framework Directive, WFD, EC 2000; Marine Strategy Framework Directive, MSFD, EC 2008). Among the adopted approaches to reduce time and costs associated to sample processing and fine taxonomic identifications when monitoring aquatic macro-invertebrate communities, the use of higher taxa (e.g. families) as surrogates for species has received particular attention (see Dauvin et al., 2003; Terlizzi et al., 2003; Jones, 2008; Bevilacqua et al., 2012 for reviews). Such an approach of *Taxonomic Sufficiency* (Ellis, 1985) relies on identifying the coarsest taxonomic level sufficient to detect changes in assemblages exposed to environmental stress without a significant loss of information with respect to species-level analysis. It has been successfully applied on benthic communities from continental (e.g. Marshall et al., 2006; Heino and Soininen, 2007; Shokri and

77 Gladstone, 2009; Jiang et al., 2013) to marine water bodies (e.g. Vanderklift et al., 1996; Gomez-
1
278 Gesteira et al., 2003; Anderson et al., 2005; Włodarska-Kowalczyk and Kędra, 2007; Bertasi et al.,
3
479 2009; Terlizzi et al., 2009; Brind'Amour et al., 2014).
6
7
80 The relevance of Taxonomic Sufficiency (TS) for environmental monitoring has been sustained
9
10
1181 invoking a hierarchical (from individuals to species and then higher taxonomic levels) biological
12
1382 response to environmental disturbance, and putative ecological similarities among closely related
14
1583 species (e.g. Ferraro and Cole, 1995; Warwick, 1993). The use of higher taxonomic levels would
16
17
1884 also reduce the potential effect of natural variability in obscuring changes related to environmental
19
2085 stress, as natural variations may affect community structure mainly by species replacement (Gray et
21
22
2386 al., 1988). However, also natural variations may affect community structure at higher taxonomic
24
2587 levels (Dethier and Schoch, 2006; Lovell et al., 2007; Włodarska-Kowalczyk and Kędra 2007;
26
27
2888 Bevilacqua et al. 2009a; Bevilacqua et al., 2011), and ecological similarities among species within
29
3089 higher taxa may be largely idiosyncratic (Lenat and Resh, 2001; Heino and Soininen, 2007;
31
32
3390 Carranza et al., 2011), casting doubts on mechanisms proposed to explain the ability of higher
34
3591 taxonomic levels to reflect natural or human-driven community patterns at species level (Lasiak,
36
3792 2003; Bertrand et al., 2006; Bevilacqua et al., 2009b; Heino, 2014).
38
39
40
4193 Indeed, TS does not rely on balancing costs and timesaving with the need for ecological
42
4394 information but, rather, it implies the static grouping of organisms in taxa belonging to a single
44
45
4695 higher taxonomic level (e.g. all organisms identified at the level of genus, family, etc.) irrespective
47
4896 of their ecological relevance or difficulties in their taxonomic identification. The approach,
49
50
5197 therefore, may lead to unnecessary loss of taxonomic details and, more importantly, may reduce the
52
5398 ability to infer about ecological processes underlying the observed patterns (Lenat and Resh, 2001;
54
5599 Lasiak, 2003; Terlizzi et al., 2003; Dethier and Schoch, 2006; Jones, 2008), which in turn are
56
57
5800 essential for defining effective environmental management strategies. Also, TS lacks any statistical
59
60
61
62
63
64
65

101 method to control for uncertainties when identifying a particular taxonomic level as sufficient to
1
102 detect community responses to environmental drivers as at species level (Mellin et al., 2011).
3
4
5
103 Recent attempts aimed at disentangling the mechanisms generating correlations between
7
104 community patterns at species and coarser taxonomic resolution reinforced the idea that higher taxa
9
105 of the Linnaean taxonomic hierarchy may be considered as arbitrary categories of species (e.g.
11
12
106 Bertrand et al., 2006) unlikely to convey consistent responses to natural or human-driven
14
15
107 environmental changes (Bevilacqua et al., 2012; Siqueira et al., 2012). Specifically, the use of null
16
17
108 models highlighted that information loss and the ensuing decrease in statistical power to detect
19
20
109 changes in assemblage structure at higher taxonomic levels depended on the degree of species
21
22
110 aggregation (exemplified by the ratio between the number of higher taxa and the number of
24
25
111 species), rather than on taxonomic relatedness of species *per se* (i.e. the relative closeness of species
26
27
112 in the Linnaean taxonomic hierarchy) (Bevilacqua et al., 2012). Such a dependence between
28
29
113 information loss and the fractal architecture of taxonomic systems has been found to be consistent
31
32
114 from terrestrial to aquatic organisms (see Bevilacqua et al., 2012 for a review) and validated by
33
34
115 several studies investigated factors affecting the performance of taxonomic surrogates (Giangrande
36
37
116 et al., 2005; Dethier and Schoch, 2006; Lovell et al., 2007; Siqueira et al., 2012; Neeson et al.,
38
39
117 2013).
41
42
118 In this new theoretical perspective, a novel approach to species surrogacy, the Best Practicable
44
45
119 Aggregation of Species (BestAgg), has been implemented (Bevilacqua et al., 2013). The approach
47
48
120 relies on the construction of null models based on real data in order to identify thresholds in species
49
50
121 aggregation (in terms of numerical ratios between the number of surrogates and the number of
52
53
122 species), and irrespective of taxonomic relatedness, below which the loss of ecological information
54
55
123 becomes relevant. In practice, BestAgg identifies the minimum number of surrogates, irrespective
56
57
124 of their type, allowing obtaining the same response as at species level. Surrogates are then defined
59
60
61
62
63
64
65

125 based on their ecological importance (*relevance*), difficulty of identification (*easiness*), and shared
1
126 characteristics among organisms (*resemblance*).

127 As any other surrogacy method, BestAgg involves the selection of a set of surrogates based on a
7
128 pilot assessment assuming that they may be employed in similar environmental contexts or in
9
129 subsequent monitoring programs (Bevilacqua et al., 2013). The validity of this assumption,
11
130 however, may be limited when investigating peculiar and/or highly variable systems where habitat
14
131 heterogeneity shapes the interactions among organisms and modifies their response to perturbation,
16
132 potentially constraining the application of surrogates for species (White and Irvine, 2003; Munari
19
133 and Mistri, 2008). Transitional waters (e.g. estuaries, wetlands, and coastal lagoons, hereafter *Tws*)
21
134 exemplify the issue. *Tws* represent important and fragile ecosystems in the coastal landscape,
24
135 providing key ecosystem services such as water quality improvement, fisheries resources, habitat
26
136 and food for migratory and resident animals, and recreational areas for human populations. *Tws* are
28
137 characterized by highly variable physico-chemical and hydro-morphologic conditions. Progressive
31
138 changes in several environmental parameters (e.g. salinity, dissolved oxygen, temperature), often
33
139 mutually correlated, generates composite gradients, resulting in a mosaic of different habitats,
36
140 characterized by different marine water renewal, nutrient content, turbidity and sediment structure
38
141 (Tagliapietra et al., 2009). Moreover, *Tws* exhibit different characteristics depending on their
41
142 geographical, hydrodynamic and ecological features, and also display distinctive features in terms
43
143 of their extraordinary history of environmental management, the importance of their productivity
46
144 and associated economical value, which are reflected on the peculiarity of their fauna (Cognetti and
48
145 Maltagliati, 2008). The species living in such ecologically complex systems adapt to changes in
50
146 environmental conditions, to the presence of organic enrichment (Simonini et al., 2004), and metal
53
147 contamination (Munari and Mistri, 2007). Several studies have explored the effectiveness of TS on
55
148 benthic macro-invertebrates from *Tws* (Mistri and Rossi, 2001; De Biasi et al., 2003; Dethier and
58
149 Schoch, 2006; Munari et al., 2009; Shokri and Gladstone, 2009; Soares-Gomez et al., 2012),
60

150 suggesting that the use of coarser level of taxonomic resolution, up to family level, might represent
1
151 a cost-effective solution for routine monitoring.

152 Here, we applied the BestAgg approach to Mediterranean transitional water systems, namely
153 coastal lagoons, in order to assess its performance in quantifying patterns of spatial and temporal
154 variability of macro-benthic assemblages. To check for the consistency of the approach when
155 applied to such peculiar and highly variable environments, we derived surrogates for species
156 following the BestAgg procedure based on a pilot lagoon system and then assessed the ability of
157 such surrogates to reflect species-level patterns in other lagoons in the same geographic area. A
158 classical approach for species surrogacy based on the concept of Taxonomic Sufficiency was also
159 applied to pilot data and results from the two approaches compared.

161 **2. Methods**

162 *2.1. Study areas and datasets*

163 The data sets consisted of macro-benthic species abundance gathered during several sampling
164 campaigns carried out between 2004 and 2008, from three coastal lagoons of the Po River Delta
165 (NE Italy), namely Marinetta-Vallona (MAR), Sacca di Scardovari (SCA), Sacca di Goro (GOR)
166 (Fig. 1). MAR is a micro-tidal, shallow (maximum depth 2.5 m) lagoon, extending for 11.5 km².
167 The lagoon is connected to the sea by a narrow mouth (about 100 m wide). A deltaic branch of the
168 Po River flows along the lagoon NW–SE axis. The westernmost part of the lagoon hosts clam-
169 farming activities. Macro-benthic samples were collected at four stations randomly distributed
170 within the lagoon in three random sampling times from 2007 to 2008. SCA is a micro-tidal, large
171 embayment (32 km²) located between two branches of the Po River Delta. The lagoon is connected
172 to the Adriatic Sea through a wide mouth that is partly obstructed by sand banks. It varies in depth

173 between 0.5 and 2.8 m. Its northern area receives nutrient-rich agricultural run-offs, whereas the
1
174 southern area hosts extensive cultures of clams and mussels. Samples were collected at five stations
3
175 in three sampling occasions. GOR is a wide (26 km²) micro-tidal lagoon whose maximum depth is
4
6
176 2.0 m. The lagoon receives nutrient-rich freshwater from deltaic branches of the Po River. Clams
8
177 are intensively cultured in its westernmost portion, representing the second largest clam-farming
9
10
11
1278 ground in Europe, with an average annual yield of 8,000 tonnes (Solidoro et al., 2000). GOR is also
13
14
179 subjected to anthropogenic eutrophication, which causes extensive growth of opportunistic green
15
16
180 macro-algae leading to summer anoxia and dystrophy. Sampling was carried out at four stations in
18
19
181 four sampling times from 2004 to 2005. Sandy and/or silty clay sediments characterized soft
20
21
22 bottoms in all lagoons. For each of the three lagoons, three samples of soft sediments were collected
23
24
2483 at each station using a Van Veen grab (0.027 m²). Samples were sieved (0.5 mm mesh size) and
25
26
284 macro-invertebrates were sorted, identified at the species level, and counted. Three data sets, one
27
28
285 for each lagoon (87 species × 36 samples for MAR, 90 species × 45 samples for SCA, and 38
30
31
326 species × 48 samples for GOR) were obtained.

33
34
357 The three data sets were then used to test the performance of the BestAgg approach. As for any
36
37
388 other surrogacy approaches, the identification of surrogates should be based on a pilot assessment
39
4089 (Terlizzi et al., 2003; Defeo and Lercari, 2004; Dauvin et al., 2007; Musco et al., 2009). We
41
42
4390 selected one lagoon out of three, namely MAR, to serve as pilot system for the definition and
44
45
4591 validation of BestAgg surrogates. We randomly selected the pilot lagoon between the two lagoons
46
47
492 accounting for the higher number of species (see above). This simulates the experimental condition
48
49
503 in which baseline information (e.g. a simple list of species or taxa) on systems being investigated is
51
52
5394 available and could guide the choice of a representative pilot system (e.g. selection among more
53
54
5595 speciose ones), in contrast to a completely random selection with the absence of any prior data.
56
57
5796 Species-level data on community structure of the two remaining lagoons, namely SCA and GOR,
58
59
60
61
62
63
64
65

197 were used to show that BestAgg surrogates derived from the pilot lagoon could have been
1
198 successfully applied to investigate similar but independent ecosystems.

3
4
5
199
6
7
8

200 *2.2. The Best Practicable Aggregation of Species (BestAgg)*

10
11
12
201 Species surrogacy implies quantifying community structure through the use of operational units that
13
14
202 can substitute species successfully, in order to avoid difficulties related to fine identification of
15
16
17
203 organisms without ensuing significant loss of information. Such operational units, i.e. surrogates,
18
19
204 are indeed groups in which species have been aggregated following shared characteristics and their
20
21
22
205 abundance summed. The level of aggregation to which the original species-level information is
23
24
206 subjected when community structure is codified through the use of surrogates can be expressed as ϕ
25
26
207 = G/S , i.e. the ratio of the number of surrogate groups G to the number of species S (Bevilacqua et
27
28
29
308 al., 2012). The basic concept underlying the BestAgg approach is that the success of a given set of
30
31
32
309 surrogates to mirror species-level patterns depends mainly on the extent to which species-level
33
34
35
310 information is packed into surrogate groups, rather than on the actual, or putative, relationships
36
37
311 among species that such groups might embody (Bevilacqua et al., 2012). In this view, BestAgg
38
39
40
41
42
412 aims at identifying the minimum number of surrogates, irrespective of their type (whether
43
44
413 taxonomic, morphological, functional, etc.), sufficient to quantify species-level community patterns
45
46
47
414 consistently. The sufficient number of surrogates is determined through a null model assuming that
48
49
415 (1) surrogates can be considered as random groups of species from the original species pool, (2) the
50
51
52
416 ability of surrogates to reflect species-level community patterns depends on the correlation
53
54
517 (Spearman's ρ) between the species-level data matrix and the corresponding surrogate matrix, and
55
56
518 (3) correlation values (ρ) depends on ϕ . Such assumptions are largely sustained by empirical
57
58
519 evidence (for a review see Bevilacqua et al., 2012).

59
60
61
62
63
64
65

220 BestAgg is based on constructing a null model from random aggregations of species-level data in
1
221 order to determine the value of ϕ , namely ϕ_{low} , sufficient to obtain results consistent with those
2
3
4
222 obtained analyzing species-level data. Once ϕ_{low} has been determined, surrogate groups in BestAgg
5
6
223 are identified on the basis of this threshold value and a set of selection criteria aiming to capitalize
7
8
9
224 on the ecological information in the surrogate data. Specifically, surrogates selection in BestAgg is
10
11
225 guided by the interplay among three macro-criteria: *relevance*, *easiness*, and *resemblance*.
12
13
226 *Relevance* concerns the importance of a given species from a general (e.g. key-stone species),
14
15
227 context-specific (e.g. indicators), or study-specific (e.g. species most contributing to the specific
16
17
228 pattern of interest) ecological perspective. *Easiness* regards the distinctiveness of a given species,
18
19
20
21
229 taxon, or group of organisms leading to be easily identified. *Resemblance*, instead, relates to
22
23
24
250 common characteristics among organisms allowing meaningful groupings (e.g. common ancestry,
24
25
26
271 functional similarity). In practice, BestAgg is structured to maximize the relevant ecological
27
28
29
30
232 information retained in the surrogate matrix, while reducing as much as possible the number of
29
30
31
233 surrogates and the difficulty in the identification of organisms and sample processing. A full
31
32
33
234 theoretical introduction to BestAgg, along with R codes for analyses, is provided in Bevilacqua et
34
35
36
235 al. (2013).
36
37
38
39

4036 41 42 43 237 2.3. Construction of the null model and surrogate selection based on the pilot system 44

45
46
4738 The first step in the application of the BestAgg approach concerned the construction of the null
47
48
49
239 model based on pilot data in order to identify the lowest ϕ value, i.e. ϕ_{low} , and consequently the
49
50
51
240 minimum number of surrogates G , i.e. G_{min} (where $G_{min} = \phi_{low} \times S$), sufficient to obtain results
52
53
54
241 consistent with those obtained using species with a probability of Type I error (i.e. the probability
54
55
56
242 of considering G_{min} as sufficient when actually it is not) of $P < 0.05$. In this framework, the S
57
58
59
243 species in the original species-level data matrix from the pilot lagoon (MAR) were randomly
60
61
62
63
64
65

244 aggregated in decreasing numbers of G groups in order to simulate decreasing values of ϕ . The
1
245 number of groups G was progressively decreased by step-wise reductions of a fixed decrement $d=5$,
2
3
4
246 starting from $G=S$ (see Bevilacqua et al., 2013). For each set of G groups from the step-wise
5
6
247 reduction procedure, 1,000 randomly aggregated matrices were obtained. A triangular matrix based
7
8
9
248 on Bray–Curtis dissimilarities was obtained from each randomly aggregated matrix and the
10
11
249 correlation value (ρ) with the species-level matrix was calculated. Then, a distance-based
12
13
250 permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) was performed
14
15
16
251 on each randomly aggregated matrix. The design for the analysis consisted of two crossed factors:
17
18
19
252 Time (Ti, 3 levels, random) and Station (St, 4 levels, random) with $n = 3$. Analyses were based on
20
21
253 Bray-Curtis dissimilarities of untransformed data. For the construction of the null model we focused
22
23
24
254 on the term $Ti \times St$ interaction as species-level analysis highlighted a significant interaction of
25
26
255 spatial and temporal variability in macro-benthic assemblages (see Results). The dependence
27
28
29
256 between ϕ and the effectiveness of the corresponding surrogate groups was checked by fitting a
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

The next step focused on selecting the set of surrogates, given G_{\min} . For each species found in the pilot lagoon MAR, we compiled information related to the three macro-criteria of *relevance*, *easiness*, and *resemblance* (see previous section and Table A1 in Appendix A). Relevant species were identified based on available knowledge from scientific literature on their importance for the structure and functioning of coastal lagoon ecosystems. Species were classified by their feeding behaviour to one of the following guilds: predators, scavengers (carrion feeders),

269 herbivores/grazers, suspension/filter feeders, detritus feeders (including surface deposit feeders,
1 subsurface deposit feeders/limnivores, that feed on sedimentary detritus and bacteria), species that
270 3 are both suspension feeders and surface deposit feeders, and omnivores. Ecological information
4 271 5 are both suspension feeders and surface deposit feeders, and omnivores. Ecological information
6 (e.g. trophic guild, tolerance, indicator species) was mined from the literature (see Table A1) and
272 8 existing databases (ERMS, www.marbef.org; MarLIN, Marine life information network; WoRMS,
9 273 10 www.marinespecies.org). Moreover, study-specific relevant species most contributing to the
11 274 12 observed community patterns were identified through Similarity Percentage analysis (SIMPER) by
14 275 15 considering only species whose contribution to dissimilarities was $\geq 3\%$. Easiness of species was
16 276 17 defined based on expert opinion according to the difficulty and intensity of analyses required to
19 277 20 identify organisms, whereas species groupings (*resemblance*) were decided combining information
21 278 22 on *relevance* and *easiness* with available taxonomic and ecological knowledge of groups. As the
24 279 25 identification of G_{\min} goes beyond any surrogacy scheme, we exploited simultaneously the potential
26 280 27 of different surrogate types potentially leading to retain ecological information and/or to reduce
28 281 29 efforts for the identification of organisms and sample processing (for a detailed conceptual
31 282 32 framework on surrogate selection in BestAgg see also Bevilacqua et al., 2013).
33 283 34

35 284 36 37 285 38 *2.4. Testing BestAgg surrogates* 39 40 41 285 42

43 286 44 As BestAgg assumes that surrogates are random subsets of the original pool of species, the
45 287 46 information retained using the selected BestAgg surrogates (expressed as ρ) should fall at least
48 288 49 within random expectations from the null model. If selected surrogates allow retaining significantly
51 289 52 greater information, the selection procedure has led to a set of surrogates able to represent species-
53 290 54 level community patterns better than what expected by chance. In contrast, ρ values below
55 291 56 expectations would indicate a set of surrogates that is likely unsuitable. The BestAgg approach
58 291 59
60
61
62
63
64
65

292 allows checking for this through a randomization test built following the same simulation procedure
1
293 underlying the construction of the general null model.
2
3
4

5
294 The original S species from MAR were randomly aggregated in G_{BestAgg} groups, where G_{BestAgg} is
6
7
295 the number of BestAgg surrogates obtained from MAR. Random aggregations were repeated 1,000
8
9
10
296 times. Correlation values (ρ) between the original species-level matrix and each randomly
11
12
297 aggregated matrix were then calculated obtaining a frequency distribution against which testing
13
14
15
298 ρ_{BestAgg} , i.e. the correlation between the original species-level matrix and the matrix aggregated
16
17
299 using the BestAgg surrogates. Also, PERMANOVA was performed on randomly aggregated
18
19
20
300 matrices (design as above), and the percentage of tests for the term of interest showing consistent
21
22
301 results with those obtained at species level was calculated, representing the P -value (i.e. the
23
24
25
302 probability of Type I error) specific for G_{BestAgg} . Finally, the ability of BestAgg surrogates to
26
27
28
303 represent species-level community patterns was checked against results from analyses based on
29
30
304 species with those obtained using the surrogates. BestAgg surrogates might substitute effectively
31
32
305 species-level information in subsequent sampling programs or very similar case studies if (i) results
33
34
306 using surrogates were consistent with those obtained using species, (ii) ρ_{BestAgg} fell within or above
35
36
37
307 random expectations, and (iii) the associated probability of Type I error was < 0.05 (Bevilacqua et
38
39
40
308 al., 2013).
41
42
43

44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65

2.5. *Taxonomic Sufficiency and BestAgg compared*

66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

The performance of BestAgg surrogates was compared with results obtained applying a more classical approach for species surrogacy based on the concept of Taxonomic Sufficiency (see for reviews Dauvin et al., 2003; Terlizzi et al., 2003; Jones, 2008). Previous studies on macro-benthic assemblages from the investigated lagoons indicated the level of family as effective in describing

315 species-level community patterns (Munari et al., 2009). Species-level data from MAR were
1
316 aggregated at family level following the Linnaean taxonomy and a PERMANOVA based on Bray-
3
317 Curtis dissimilarity on untransformed data (design as above) was performed. Data ordinations
4
6
318 (nMDS) were plotted to depict community patterns when using species, families, and BestAgg
8
319 surrogates. Also, the original species matrix was aggregated at Genus, Family, Order, Class, and
10
11
320 Phylum level, and a second-stage nMDS ordination was plotted to visualize patterns of correlation
13
321 among species, BestAgg, and higher-taxon matrices. Finally, we estimated savings derived from the
14
15
322 use of BestAgg surrogates and families following the approach proposed by Ferraro and Cole
16
18
323 (1995). In this approach, savings are estimated by taking the ratio of the number of surrogates to the
20
21
324 number of species, assuming that the time spent to identify organisms and sample processing is
22
23
325 proportional to the number of categories (e.g. species, higher taxa, groups) to which they must be
25
26
326 assigned (see also Thompson et al., 2003).
27
28
29
30
31
32

33 2.6. Application of BestAgg surrogates 34

35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

BestAgg surrogates derived from the pilot lagoon were employed to analyze macro-benthic assemblages from the other lagoons, in order to ascertain whether the BestAgg approach is conducive to a set of surrogates that might be successfully applied to comparable study contexts. Species-level data from SCA and GOR were therefore aggregated using the set of BestAgg surrogates from MAR (the pilot lagoon). PERMANOVAs based on Bray-Curtis dissimilarity matrices (untransformed data) and nMDS ordinations were performed for SCA and GOR using BestAgg surrogates and results compared with those obtained from analyses at species level. In both cases, designs for analyses consisted of two random crossed factors, Time (Ti) and Station (St), although differing in the number of levels of each factor (see section Study area), with $n = 3$.

338 SIMPER analyses were done in order to identify species most contributing to the observed patterns
1
339 by considering only species whose contribution to dissimilarities was $\geq 3\%$.

340 Analyses were performed using R (R Development Core Team, 2013) and the PRIMER v 6
341 software package (Clarke and Gorley, 2006), including the PERMANOVA+ add-on package
342 (Anderson and Gorley, 2008).

344 3. Results

345 The linear regression of ρ values against $\ln(\phi)$ from the pilot lagoon (MAR) was significant ($P <$
346 0.001) with a correlation value of $R^2 = 0.69$, indicating that the loss information associated to a
347 given set of surrogates strongly depended on the level of aggregation following a semi-log model
348 (Fig. 2).

349 The null model based on MAR predicted that the lowest ϕ value allowing 95% of PERMANOVAs
350 on aggregated data to give consistent results with those obtained at species level was $\phi_{low} = 0.29$,
351 corresponding to $G_{min} = 25$ (Table 1, Fig. 2), under the null hypothesis that surrogates are random
352 groups of species from the original species pool. This meant that 25 surrogates could substitute the
353 original 87 species found in the lagoon while still allowing quantifying spatio-temporal community
354 patterns as well as at species level.

355 BestAgg surrogate selection based on MAR led to a set of $G_{BestAgg} = G_{min} = 25$ surrogates (see
356 Table A2 in Appendix A, and Table B1 in Appendix B). Most surrogates were formally taxonomic,
357 although in some cases they corresponded to trophic or indicator groups (Table A2), with 11
358 species (*Anadara transversa*, *Crassostrea gigas*, *Arcuatula senhousia*, *Ruditapes philippinarum*,
359 *Armandia cirrhosa*, *Alitta succinea*, *Cyathura carinata*, *Carcinus aestuarii*, *Dyspanopeus sayi*,
360 *Upogebia pusilla*, *Chironomus salinarius*), four families (Tellinidae, Semelidae, Capitellidae,

361 Spionidae), three orders (Phyllodocida, Eunicida, Amphipoda), six classes (Anthozoans, Other
1
362 Bivalves, Gastropods, Oligochaetes, Other Polychaetes, Other Crustaceans), and one morphological
3
4
363 group (Flat/ribbon worms).
5
6
7

364 The test based on randomizations for MAR showed that the probability of G_{BestAgg} to fail in
9
10
365 depicting species-level community patterns was $P = 0.036$, and correlation value ρ_{BestAgg} fell within
11
12
366 the frequency distribution specific for G_{BestAgg} (Fig. 3), indicating that the selected BestAgg
14
15
367 surrogates satisfied the assumptions underlying the null model.
16
17
18

368 For the pilot lagoon MAR, PERMANOVA carried out at species level showed a significant $Ti \times St$
20
21
369 interaction (Table 2), indicating significant variation among stations that varied in time. Such
22
23
370 patterns were clearly depicted in the nMDS ordination of $Ti \times St$ centroids (Fig. 4). Analyses based
25
26
371 on data aggregated using the set of BestAgg surrogates allowed obtaining consistent results (Table
28
29
372 2) and an almost interchangeable ordination of $Ti \times St$ centroids (Fig. 4). The same results were
30
31
373 obtained with family-level data (Table 2, Fig. 4), indicating that both families (64 taxa) and
33
34
374 BestAgg surrogates (25 taxa/groups) were effective in representing species-level community
35
36
375 patterns. With respect to species-level analyses, the expected timesaving during sample processing
38
39
376 and identifications was higher when using the BestAgg surrogates (>70%) than for family-level
40
41
377 analyses (~26%).
42
43
44

378 The second stage nMDS ordination showed a decreasing correlation with species-level information
46
47
379 at decreasing taxonomic resolution. The genus level had the strongest relationships with species,
48
49
380 followed by the family level, with a drop in correlations beyond the order level (Fig. 5). BestAgg
51
52
381 surrogates clustered with species, genus, and family levels (Fig. 5) indicating that, despite the
53
54
382 limited number of surrogates compared to such taxonomic levels, the information retained in the
56
57
383 BestAgg data matrix was strongly correlated with species-level information as well.
58
59
60
61
62
63
64
65

384 PERMANOVA on macro-benthic assemblages from SCA and GOR identified using the BestAgg
1
385 surrogates derived from MAR (pilot lagoon) showed results consistent with those that would have
2
386 been obtained if species were analyzed, i.e. a significant $Ti \times St$ interaction (Table 2).
3
4
5
6
7
8
387 For the Scardovari lagoon (SCA), the nMDS ordination plot based on species-level data clearly
9
10
388 depicted temporal changes in patterns of variability among stations (Fig. 6a). A clear gradient from
11
12
389 more confined stations towards stations at the mouth of the embayment (see Fig. 1) was visible in
13
14
390 time 1 but not in times 2 and 3 (Fig. 6a). The same patterns were noticed when using BestAgg
15
16
391 surrogates (Fig. 6a). The SIMPER analysis on BestAgg surrogates most contributing to such
17
18
392 temporal changes showed a general increase, from time 1 to times 2 and 3, of opportunistic/tolerant
19
20
21
22
23
24
25
26
27
28
29
295 For the Goro lagoon (GOR), the nMDS ordination at species level showed a clear separation
30
31
316 between station centroids from the first two times and the remaining times of sampling, which
32
33
317 clustered together respectively with exception of centroids of station 3 (see Fig. 1) that lay apart
34
35
36
37
38
39
398 from centroids of the remaining stations in all times (Fig. 6b). Such patterns were clearly depicted
39
40
400 when using BestAgg surrogates, except station 3 in time 1 (Fig. 6b). The SIMPER analysis on
41
42
401 BestAgg aggregated data showed a lower abundance of Spionidae and Oligochaeta in station 3 with
43
44
45
46
402 respect to the remaining stations, associated to a higher abundance of Capitellidae and *R.*
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

404 **4. Discussion**

405 The effectiveness of the BestAgg approach in providing surrogates able to depict community
56
57
58
59
406 patterns as at species level has been demonstrated in soft bottom and rocky reef macro-benthic

407 communities, whether composed by individual or colonial organisms, subjected to both natural and
1 anthropogenic drivers of change (Bevilacqua et al., 2013). Also, the approach has been successfully
408 applied to freshwater macro-invertebrate in exploring environmental gradients in continental river
3
4
409 basins (Milošević et al., 2014). It could be argued, however, that the approach might exhibit optimal
5
6
410 responses when analysing sharp environmental variations that cause gross changes in community
7
8
411 structure, which are likely detectable although using operational units other than species.
9
10
11
12
13
14
413 Consequently, the approach might perform less effectively in approximating complex species-level
15
16
414 patterns driven by multiple and interacting environmental factors, as could occur in invertebrate
17
18
415 communities inhabiting transitional ecosystems such as estuaries and coastal lagoons. Indeed, the
19
20
21
416 strong relationship between information loss and the aggregation level (ϕ), which underlies the
22
23
24
417 BestAgg approach, has been confirmed analysing published case studies on the application of
25
26
418 surrogates in Tws (e.g. estuaries, coastal lagoons, salt marshes) worldwide (Bevilacqua et al.,
27
28
419 2012). In this respect, our findings provide first direct evidence of the robustness of BestAgg when
29
30
31
420 applied to Tws, demonstrating the ability of the approach to quantifying consistently complex
32
33
34
421 patterns of spatio-temporal variability of assemblage structure in northern Adriatic lagoons, where
35
36
422 the effects of natural environmental gradients and human disturbance largely overlap.
37
38
39
423 Northern Adriatic lagoons share features that are typical of Tws, namely low number of species and
40
41
42
424 taxonomic diversity with respect to marine ecosystems, dominance of few species adapted to the
43
44
425 particular environmental conditions and generally tolerant to sharp regimes of natural disturbance
45
46
426 (Elliott and Quintino, 2007; Pérez-Ruzafa et al., 2010; Borja et al., 2011). However, beyond such
47
48
49
427 general commonalities, each lagoon holds its own distinctiveness. The pilot lagoon (MAR) and the
50
51
52
428 SCA lagoon were characterized by a relatively high number of species, whereas in GOR the
53
54
55
429 community appeared quite poor in species. The three lagoons were affected by different levels of
56
57
58
430 anthropogenic disturbance related to clam harvesting, and differed also in the amount of nutrient-
59
60
431 rich freshwater inputs, with GOR being the one with the greater influx of freshwater through two

432 deltaic branches, and many drains. Such differences are likely to concur in shaping assemblage
1
433 structure and composition within the three lagoons, and determining dissimilarities occurring
2
3
4
434 among them (Munari and Mistri, 2008; Pérez-Ruzafa et al., 2013; Sfriso et al., 2014).
5
6
435 One of the required prerogatives of any surrogacy approach is to provide surrogates for species that
7
8
436 perform consistently in comparable experimental contexts or in subsequent reiterations of the same
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

deltaic branches, and many drains. Such differences are likely to concur in shaping assemblage structure and composition within the three lagoons, and determining dissimilarities occurring among them (Munari and Mistri, 2008; Pérez-Ruzafa et al., 2013; Sfriso et al., 2014). One of the required prerogatives of any surrogacy approach is to provide surrogates for species that perform consistently in comparable experimental contexts or in subsequent reiterations of the same study (e.g. in monitoring programs) (Olsgard and Somerfield, 2000). Our results showed that the BestAgg approach led to a reliable set of surrogates that are effective in quantifying the overall community patterns as at species level across distinct lagoons from the same geographic area, working over and beyond the environmental heterogeneity and peculiarity characterizing such highly variable systems. This ability likely depends on the fact that surrogate selection in BestAgg tends to maximize ecological information and, therefore, is more prone to capture salient ecological and functional properties of systems under study. In the present study, 5 out of 11 surrogates at species level were alien species, thus reflecting the vulnerability to invasion of northern Adriatic lagoons (Occhipinti-Ambrogi, 2000). The remaining species included sensitive organisms (e.g. *Armandia cirrhosa*), indicators of environmental stress (e.g. *Chironomus salinarius*), or commercial targets (e.g. *Carcinus aestuarii*), providing direct information on the ecological status of the investigated lagoons. Also, surrogate selection in BestAgg prioritizes ecological information even when species are grouped, so that surrogate groups (taxa, morphological groups, etc.) are still meaningful from an ecological perspective. In our case study, for instance, several surrogate groups may be considered as indicator taxa (e.g. Capitellidae, Spionidae) (Simboura and Zenetos, 2002), or trophic guilds (e.g. Eunicida, which were all predators) (Table A2). Only species with low relevance and/or of difficult identification are grouped based exclusively on shared features (e.g. taxonomic relatedness, morphology) (Bevilacqua et al., 2013). In this view, surrogate selection in BestAgg may help the interpretation of patterns, which could result less intuitive whether based on a single taxonomic rank. In the SCA lagoon, for example, the disruption of the spatial pattern of

457 confinement from time 1 (May 2007) to times 2 and 3 (November 2007 and 2008) could be a
1
458 consequence of increased agricultural run-offs, which are likely to follow increased rainfall regimes
3
459 in autumn, especially in time 3 (Barbi et al., 2012), resulting in an increased dominance of
4
5
6
460 opportunistic/tolerant polychaetes (e.g. Capitellidae, Spionidae) associated to a decreased
7
8
9
461 abundance of amphipods (Dauvin and Ruellet, 2007). For the Goro lagoon, instead, the peculiarity
10
11
462 of station 3, beyond seasonal variations between times 1-2 (autumn-winter 2004) and times 3-4
12
13
463 (spring-summer 2005), could depend on the fact that the remaining stations were subjected to
14
15
16
464 nutrient inputs from the Po river branches whereas station 3 is mostly marine and within an area
17
18
465 affected by clam harvesting. Such factors probably led in this last station to the decreased
19
20
21
466 abundance of groups tolerant to organic enrichment (e.g. Spionidae, Oligochaeta), and to the
22
23
467 increased abundance of *R. philippinarum* that found in this area of the lagoon the optimal
24
25
26
468 environmental conditions for growth (Vincenzi et al., 2006).
27
28
469 Indeed, also a more classic approach to species surrogacy based on Taxonomic Sufficiency, and
29
30
31
470 specifically the use of family-level identification of organisms, was effective in elucidating species-
32
33
471 level patterns for all lagoons, confirming the results obtained in other transitional water ecosystems
34
35
472 (Mistri and Rossi, 2001; De Biasi et al., 2003; Chainho et al., 2007; Munari et al., 2009; Shokri and
36
37
38
473 Gladstone, 2009). Such findings, however, may be simply due to the effect of a close numerical
39
40
474 overlapping between species and family richness, which could generate interchangeable community
41
42
475 patterns when using either species or family-level taxonomic resolution, rather than to the similarity
43
44
45
476 in ecological response of species within families (Bevilacqua et al., 2012). For instance, in the three
46
47
48
477 lagoons about 80% of families was monotypic, the 11-14% had two species, and only the 9-6%
49
50
51
478 included three or more species. This probably reflects the outcomes of ecological and evolutionary
52
53
479 mechanisms driving community composition, which have led only very few species able to tolerate
54
55
480 the stressful condition prevailing in such environments to evolve for each family (Mistri and Rossi,
56
57
58
481 2001).
59
60
61
62
63
64
65

482 Irrespective of the surrogacy approach employed, surrogates for species should not be defined *a*
1
483 *priori* and should require at least a pilot assessment and validation before they could be assumed as
3
484 suitable substitutes of species for subsequent monitoring or for very similar study contexts (Lenat
4
5
6
485 and Resh, 2001; Terlizzi et al., 2003; Defeo and Lercari, 2004; Dauvin et al., 2007; Jones, 2008).
8
9
486 Also, once fixed, surrogates for species should not be considered as unchangeable operational units,
10
11
487 and periodical adjustments involving species-level analyses should alternate routine monitoring
12
13
488 based on surrogates, due to possible changes in assemblage structure over time (Tataranni et al.,
14
15
489 2009; Musco et al., 2011). In this respect BestAgg is no exception. It should be noted, however, that
16
17
490 random simulations underlying the identification of ϕ_{low} provide a wide range of potential scenarios
18
19
20
491 of aggregated assemblage structure, which likely increases the robustness of the approach to
21
22
23
492 variations in assemblage composition. In addition, variations in the number of species might have
24
25
26
493 little outcomes on the effectiveness of the fixed ϕ_{low} , unless consisting in considerable increases in
27
28
29
494 species richness. Our findings showed the robustness of the approach to capture temporal variations
30
31
32
495 in community structure confirming the results of other case studies (Bevilacqua et al., 2013),
33
34
496 although further investigations focusing on long-term data are needed to ascertain whether and,
35
36
37
497 eventually, how often, a refinement of BestAgg surrogates should be recommended.
38
39
498 Species surrogacy should be a more formal and transparent process than is often practised, and
40
41
499 should conjugate solid theory and quantitative assessments of surrogates (Jones, 2008). The
42
43
500 BestAgg approach relies on a rigorous theoretical and statistical framework aimed at quantifying
44
45
501 the sufficient number of surrogates objectively and identifying the type of effective surrogates
46
47
48
502 based on objective macro-criteria (Bevilacqua et al., 2012; Bevilacqua et al., 2013). Above all, the
49
50
503 BestAgg approach demonstrated several advantages with respect to the classic approach based on
51
52
504 Taxonomic Sufficiency. Our results on northern Adriatic lagoons showed that BestAgg surrogates
53
54
55
505 might lead saving an additional 40% of time during sample processing and identifications of
56
57
506 organisms if compared to family-level analyses, while still retaining a greater correspondence with
58
59
60
61
62
63
64
65

507 species-level information. The approach, therefore, may enhance cost-effectiveness of species
1
508 surrogacy, also providing a statistical framework controlling for uncertainty in the application of the
3
509 selected set of surrogates (Bevilacqua et al., 2013). More importantly, surrogate selection in
4
6
510 BestAgg takes into account the ecological meaning of surrogates, and unleashes the investigator to
8
511 exploit different surrogate types, from morphological groups to functional groups, or mixing
9
10
11
512 different surrogate types, such as taxa from different taxonomic ranks. This last prerogative might
13
14
513 improve the effectiveness of the selected set of surrogates, beyond the static approach of focusing
15
16
514 exclusively on taxa of a single taxonomic rank higher than species (Groc et al., 2010).
18
19
515 The WFD and the MSFD directives are leading to an increasing demand for monitoring programs
20
21
516 and, consequently, for increasing research efforts and expertise in invertebrate taxonomy in Europe.
23
24
517 Sooner or later this will impose dilemmas regarding the spatial coverage, the sampling intensity and
25
26
518 the sampling frequency in order to obtain reliable results, particularly because the information
27
28
519 obtained is expected to guide management actions (Dauvin et al., 2012). This may be particularly
30
31
520 relevant for transitional water ecosystems, for which there is a urgent need for understanding the
32
33
521 links between the effects of human activities and changes in communities (Munari and Mistri,
35
36
522 2008), as natural and anthropogenic environmental drivers have created a mosaic of conditions
37
38
523 making difficult separating natural variations from human-induced alterations (Estuarine Quality
40
41
524 Paradox: Elliott and Quintino, 2007; De Paz et al., 2008). Further research aiming at defining sets
42
43
525 of operational units for rapid assessments of community changes from local to regional scale, able
44
45
526 to depict species-level patterns and to inform about underlying ecological processes while minimize
47
48
527 efforts during sample processing and taxonomic identifications, are strongly needed (Olsgard and
49
50
528 Somerfield, 2000; Dauvin et al. 2007; De Paz et al., 2008; Borja et al. 2011). In this direction, the
52
53
529 BestAgg approach provides a new perspective for optimizing trade-offs between practical
54
55
530 constraints and the need for essential ecological information in environmental monitoring.
57

531
59
60
61
62
63
64
65

532 **Acknowledgements**

1
2
33 The research leading to these results was instrumental for the preparation of the EU Community's
4
5
34 Seventh Framework (FP7/2007–2013) Programs COCONET, Grant Agreement No. 287844) and
6
7
35 PERSEUS (Grant Agreement No. 287600).
8
9

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

37 **References**

38 Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance.
39 Aust. Ecol. 26, 32–46.

40 Anderson, M.J., Connell, S.D., Gillanders, B.M., Diebel, C.E., Blom, W.M., Saunders, J.E.,
41 Landers, T.J., 2005. Relationships between taxonomic resolution and spatial scales of
42 multivariate variation. J. Anim. Ecol. 74, 636-646.

43 Anderson, M.J., Gorley, R.N., 2008. PERMANOVA+ for PRIMER: Guide to Software and
44 Statistical Methods. PRIMER-E, Plymouth, UK.

45 Barbi, A., Monai, M., Racca, R., Rossa, A. M., 2012. Recurring features of extreme autumnal
46 rainfall events on the Veneto coastal area. Nat. Hazards Earth Syst. Sci. 12, 2463–2477.

47 Basset, A., Barbone, E., Borja, A., Brucet, S., Pinna, M., Quintana, X.D., Reizopoulou, S.,
48 Rosati, I., Simboura, N., 2012. A benthic macroinvertebrate size spectra index for implementing
49 the Water Framework Directive in coastal lagoons in Mediterranean and Black Sea ecoregions.
50 Ecol. Indic. 12, 72–83.

51 Bertasi, F., Colangelo, M. A., Colosio, F., Gregorio, G., Abbiati, M., Ceccherelli, V.U., 2009.
52 Comparing efficacy of different taxonomic resolutions and surrogates in detecting changes in
53 soft bottom assemblages due to coastal defence structures. Mar. Pollut. Bull. 58, 686–694.

554 Bertrand, Y., Pteijel, F., Rouse, G.W., 2006. Taxonomic surrogacy in biodiversity assessments,
1 and the meaning of Linnaean ranks. *Syst. Biodivers.* 4, 149–159.

555
2
3
4
5
556 Bevilacqua, S., Fraschetti, S., Terlizzi, A., Boero, F., 2009a. The use of taxonomic distinctness
6 indices in assessing patterns of biodiversity in modular organisms. *Mar. Ecol.* 30, 151–163.

557
7
8
9
10
11
558 Bevilacqua, S., Fraschetti, S., Musco, L., Terlizzi, A., 2009b. Taxonomic sufficiency in the
12 detection of natural and human-induced changes in marine assemblages: A comparison of
13 habitats and taxonomic groups. *Mar. Pollut. Bull.* 58, 1850–1859.

559
14
15
16
560
17
18
19
561 Bevilacqua, S., Fraschetti, S., Musco, L., Guarnieri, G., Terlizzi, A., 2011. Low sensitiveness of
20 taxonomic distinctness indices to human impacts: Evidences across marine benthic organisms
21 and habitat types. *Ecol. Indic.* 11, 448–455.

562
22
23
24
563
25
26
27
564 Bevilacqua, S., Terlizzi, A., Fraschetti, S., Claudet, J., Boero, F., 2012. Taxonomic relatedness
28 does not matter for species surrogacy in the assessment of community responses to
29 environmental drivers. *J. Appl. Ecol.* 49, 357–366.

565
30
31
32
566
33
34
35
567 Bevilacqua, S., Claudet, J., Terlizzi, A., 2013. Best Practicable Aggregation of Species: a step
36 forward for species surrogacy in environmental assessment and monitoring. *Ecol. Evol.* 3,
37 3780–3793.

568
38
39
40
41
569
42
43
44
570 Borja, A., Barbone, E., Basset, A., Borgersen, G., Brkljacic, M., Elliott, M., Garmendia, J.M.,
45 Marques, J.C., Mazik, K., Muxika, I., Neto, J.M., Norling, K., Rodríguez, J.G., Rosati, I., Rygg,
46 B., Teixeira, H., Trayanova, A., 2011. Response of single benthic metrics and multi-metric
47 methods to anthropogenic pressure gradients, in five distinct European coastal and transitional
48 ecosystems. *Mar. Poll. Bulletin.* 62, 499–513.

571
49
50
51
52
572
53
54
55
56
57
58
59
60
61
62
63
64
65

575 Bremner, J., Rogers, S.I., Frid, C.L.J., 2006. Methods for describing ecological functioning of
1 marine benthic assemblages using biological traits analysis (BTA). *Ecol. Indic.* 6, 609–622.
576
3
4
5
577 Brind'Amour, A., Laffargue, P., Morin, J., Vaz, S., Foveau, A., Le Bris, H., 2014.
7
578 Morphospecies and taxonomic sufficiency of benthic megafauna in scientific bottom trawl
9 surveys. *Cont. Shelf Res.* 72, 1–9.
10
579
11
12
13
580 Cardoso, P., Erwin, T.L., Borges, P.A.V., New, T.R., 2011. The seven impediments in
15 invertebrate conservation and how to overcome them. *Biol. Conserv.* 144, 2647–2655.
16
581
17
18
19
582 Carranza, A., Defeo, O., Arim, M., 2011. Taxonomic relatedness and spatial structure of a shelf
20 benthic gastropod assemblage. *Divers. Distribut.* 17, 25–34.
21
583
23
24
25
584 Chainho, P., Lane, M.F., Chaves, M.L., Costa, J.L., Costa, M.J., Dauer, D.M., 2007. Taxonomic
26 sufficiency as a useful tool for typology in a poikilohaline estuary. *Hydrobiologia* 587, 63–78.
27
585
29
30
31
586 Clarke, K.R., Gorley, R.N., 2006. *PRIMER v6: User Manual/Tutorial*. PRIMER-E, Plymouth.
32
33
34
587 Cognetti, G., Maltagliati, F., 2008. Perspectives on the ecological assessment of transitional
36 waters. *Mar. Pollut. Bull.* 56, 607–608.
37
588
38
39
40
589 Dauvin, J.-C., Gomez Gesteira, J.L., Salvande Fraga, M., 2003. Taxonomic sufficiency: an
42 overview of its use in the monitoring of sublittoral benthic communities after oil spills. *Mar.*
590
44
45
591
46
47
48
592 Dauvin, J.-C., Bellan, G., Bellan-Santini, D., 2007. Benthic indicators: From subjectivity to
50 objectivity – Where is the line? *Mar. Pollut. Bull.* 60, 947–953.
51
593
52
53
54
594 Dauvin, J.-C., Ruellet, 2007. Polychaete/amphipod ratio revisited. *Mar. Pollut. Bull.* 55, 215–
56
57
595
58
59
60
61
62
63
64
65

596 Dauvin, J.-C., Alizier, S., Rolet, C., Bakalem, A., Bellan, G., Gomez Gesteira, J.L., Grimes, S.,
1
597 de-la-Ossa-Carretero, J.A., Del-Pilar-Ruso, Y., 2012. Response of different benthic indices to
3
4
598 diverse human pressures. *Ecol. Indic.* 12, 143–153.
5
6
7
599 De Biasi, A.M., Bianchi, C.N., Morri, C., 2003. Analysis of macrobenthic communities at
9
10
600 different taxonomic levels: an example from an estuarine environment in the Ligurian Sea (NW
11
12
601 Mediterranean). *Est. Coast. Shelf Sci.* 58, 99–106.
14
15
16
602 Defeo, O., Lercari, D., 2004. Testing taxonomic resolution levels for ecological monitoring in
17
18
603 sandy beach macrobenthic communities. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 14, 65–74.
20
21
22
604 De Paz L., Patricio J., Marques J.C., Borja A., Laborda A.J., 2008. Ecological status assessment
23
24
605 in the lower Ebso estuary (Spain). The challenge of habitat heterogeneity integration: a benthic
26
27
606 perspective. *Mar. Pollut. Bull.* 56, 1275–1283.
28
29
30
607 Dethier, M.N., Schoch, G.C., 2006. Taxonomic sufficiency in distinguishing natural spatial
32
33
608 patterns on an estuarine shoreline. *Mar. Ecol. Progr. Ser.* 306, 41–49.
34
35
36
609 EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council establishing a
38
610 framework for community action in the field of water policy.
40
41
42
611 EC, 2008. Directive 2008/56/EC of the European Parliament and of the Council establishing a
43
44
612 framework for community action in the field of marine environmental policy.
46
47
48
613 Elliott, M., Quintino, V., 2007. The estuarine quality paradox: environmental homeostasis and
49
50
614 the difficulty of detecting anthropogenic stress in naturally stressed areas. *Mar. Pollut. Bull.* 54,
52
53
615 640–645.
54
55
56
616 Ellis, D., 1985. Taxonomic sufficiency in pollution assessment. *Mar. Pollut. Bull.* 16, 459.
58
59
60
61
62
63
64
65

617 Ferraro, S.P., Cole, F.A., 1995. Taxonomic level sufficient for assessing pollution impacts on
1
618 the southern Californian Bight macrobenthos – revisited. *Environ. Toxicol. Chem.* 14, 1031–
3
4
619 1040.
5
6
7
620 Giangrande, A., Licciano, M., Musco, L., 2005. Polychaetes as environmental indicators
9
621 revisited. *Mar. Pollut. Bull.* 50, 1153–1162.
10
11
12
13
622 Gomez Gesteira, J.L., Dauvin, J.C., Fraga, M.S., 2003. Taxonomic level for assessing oil spill
15
623 effects on soft-bottom sublittoral benthic communities. *Mar. Pollut. Bull.* 46, 562–572.
16
17
18
19
624 Gray, J.S., Aschan, M., Carr, M.R., Clarke, K.R., Green, R.H., Pearson, T.H., Rosenberg, R.,
21
625 Warwick, R.M., 1988. Analysis of community attributes of the benthic macrofauna of
22
23
24
626 Frierfjord/Langesundfjord and in a mesocosm experiment. *Mar. Ecol. Prog. Ser.* 46, 151-165.
26
27
28
627 Groc, S., Delabie, J.H.C., Longino, J.T., Orivel, J., Majer, J.D., Vasconcelos, H.L., Dejean, A.,
29
30
628 2010. A new method based on taxonomic sufficiency to simplify studies on Neotropical ant
32
629 assemblages. *Biol. Conserv.* 143, 2832–2839.
34
35
630 Guarnieri, G., Terlizzi, A., Bevilacqua, S., Fraschetti, S., 2012. Increasing heterogeneity of
37
38
631 sensitive assemblages as a consequence of human impact in submarine caves. *Mar. Biol.* 159,
39
40
632 1155–1164.
42
43
633 Heino, J., 2014. Taxonomic surrogacy, numerical resolution and responses of stream
45
46
634 macroinvertebrate communities to ecological gradients: Are the inferences transferable among
47
48
635 regions? *Ecol. Indic.* 36, 186–194.
50
51
52
636 Heino, J., Soininen, J., 2007. Are higher taxa adequate surrogates for species-level assemblage
53
54
637 patterns and species richness in stream organisms? *Biol. Conserv.* 137, 78–89.
56
57
58
59
60
61
62
63
64
65

638 Jiang, X., Xiong, J., Song, Z., Morse, J.C., Jones, F.C., Xie, Z., 2013. Is coarse taxonomy
639 sufficient for detecting macroinvertebrate patterns in floodplain lakes? *Ecol. Indic.* 27, 48–55.
640
641 Jones, F.C., 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater
642 bioassessments using benthic macroinvertebrates. *Environ. Rev.* 16, 45–69.
643
644 Lasiak, T., 2003. Influence of taxonomic resolution, biological attributes and data
645 transformations on multivariate comparisons of rocky macrofaunal assemblages. *Mar. Ecol.*
646 *Progr. Ser.* 250, 29–34.
647
648 Lenat, D.R., Resh, V.H., 2001. Taxonomy and stream ecology – the benefits of genus- and
649 species-level identifications. *J. North Am. Benthol. Soc.* 20, 287–298.
650
651 Lovell, S., Hamer, M., Slotow, R., Herbert, D., 2007. Assessment of congruency across
652 invertebrate taxa and taxonomic levels to identify potential surrogates. *Biol. Conserv.* 139, 113–
653 125.
654
655 Marshall, J.C., Steward, A.L., Harch, B.D., 2006. Taxonomic resolution and quantification of
656 freshwater macroinvertebrate samples from an Australian dryland river: The benefits and costs
657 of using species abundance data. *Hydrobiologia* 572, 171–194.
658
659 Mellin, C., Delean, S., Caley, J., Edgar, G., Meekan, M., Pitcher, R., Przeslawski, R., Williams,
660 A., Bradshaw, C., 2011. Effectiveness of biological surrogates for predicting patterns of marine
661 biodiversity: a global meta-analysis. *PLoS One* 6, e20141.
662
663 Milošević D., Stojković, M., Čerba, D., Petrović, A., Paunović, M., Simić, V., 2014. Different
664 aggregation approaches in the chironomid community and the threshold of acceptable
665 information loss. *Hydrobiologia* 727, 35–50.

659 Mistri, M., Rossi, R., 2001. Taxonomic sufficiency in lagoonal ecosystems. J. Mar. Biol. Ass.
1
660 U.K. 81, 339–340.
2
3
4
5
661 Munari, C., 2013. Benthic community and biological trait composition in respect to artificial
7
662 coastal defence structures: A study case in the northern Adriatic Sea. Mar. Environ. Res. 90,
9
10
663 47–54.
11
12
13
664 Munari, C. Mistri, M., 2007. Effect of copper on the scope for growth of clams (*Tapes*
15
665 *philippinarum*) from a farming area in the Northern Adriatic Sea. Mar. Environ. Res. 64, 347–
17
18
666 357.
20
21
22
667 Munari, C. Mistri, M., 2008. Biodiversity of soft-sediment benthic communities from Italian
23
24
668 transitional waters. J. Biogeog. 35, 1622–1637.
26
27
28
669 Munari, C., Mistri, M., 2012. Ecological status assessment and response of benthic communities
29
670 to environmental variability: The Valli di Comacchio (Italy) as a study case. Mar. Environ. Res.
30
671 81, 53–61.
32
33
34
35
36
672 Munari, C., Warwick, R.M., Mistri, M., 2009. Monitoring with benthic fauna in Italian coastal
38
673 lagoons: new tools for new prospects. Aquat. Conserv. Mar. Freshw. Ecosyst. 19, 575–587.
40
41
42
674 Musco, L., Mikac, B., Tataranni, M., Giangrande, A., Terlizzi, A., 2011. The use of coarser
44
675 taxonomy in the detection of long-term changes in polychaete assemblages. Mar. Environ. Res.
46
47
676 71, 131–138.
48
49
50
677 Musco, L., Terlizzi, A., Licciano, M., Giangrande, A., 2009. Taxonomic structure and the
52
53
678 effectiveness of surrogates in environmental monitoring: a lesson from polychaetes. Mar. Ecol.
54
55
679 Progr. Ser. 383, 199–210.
57
58
59
60
61
62
63
64
65

680 Neeson, T., van Rijn, I., Mandelik, Y., 2013. How taxonomic diversity, community structure
1 and sample size determine the reliability of higher taxon surrogates. *Ecol. Appl.* 23, 1216–1225.
2
3
4
5
682 Olsgard, F., Somerfield, P.J., 2000. Surrogates in marine benthic investigations - which
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

681 and sample size determine the reliability of higher taxon surrogates. *Ecol. Appl.* 23, 1216–1225.

682 Olsgard, F., Somerfield, P.J., 2000. Surrogates in marine benthic investigations - which
683 taxonomic unit to target? *J. Aquat. Ecosyst. Stress Rec.* 7, 25–42.

684 Occhipinti-Ambrogi, A., 2000. Biotic invasions in a Mediterranean lagoon. *Biol. Invasions* 2:
685 165-176.

686 Pearson, T.H., Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment
687 and pollution of the marine environment. *Oceanogr. Mar. Biol. Annu. Rev.* 16, 229–311.

688 Pérez-Ruzafa, A., Marcos, C., Pérez-Ruzafa, I., Pérez-Marcos, M., 2010. Coastal lagoons:
689 “transitional ecosystems” between transitional and coastal waters. *J. Coast. Conserv.* 36, 160–
690 166.

691 Pérez-Ruzafa, A., Marcos, C., Pérez-Ruzafa, I., Pérez-Marcos, M., 2013. Are coastal lagoons
692 physically or biologically controlled ecosystems? Revisiting r vs. K strategies in coastal lagoons
693 and estuaries. *Est. Coast. Shelf Sci.* 132, 17–33.

694 Petchey, O.L., Belgrano, A., 2010. Body-size distributions and size-spectra: universal indicators
695 of ecological status? *Biol. Lett.* 6, 434–437.

696 R Development Core Team, 2013. R: a language and environment for statistical computing. R
697 Foundation for statistical computing, Vienna. Available at <http://www.R-project.org>

698 Rice, J., Arvanitidis, C., Borja, A., Frid, C., Hiddink, J.C., Krause, J., Lorange, P., Ragnarsson,
699 S.A., Sköld, M., Trabucco, B., Enserink, L., Norkko, A., 2012. Indicators for Sea-floor Integrity
700 under the European Marine Strategy Framework Directive. *Ecol. Indic.* 12, 174–184.

701 Sfriso, A., Facca, C., Bon, D., Giovannone, F., Buosi, A., 2014. Using phytoplankton and
1
702 macrophytes to assess the trophic and ecological status of some Italian transitional systems.
3
703 Cont. Shelf Sci. 81, 88–98.
4
6
7
704 Shokri, M.R., Gladstone, W., 2009. Higher taxa are effective surrogates for species in the
9
10 selection of conservation reserves in estuaries. Aquat. Conserv. Mar. Freshw. Ecosyst. 19, 626–
11
12 636.
13
14
15
16
707 Simboura, N., Zenetos, A., 2002. Benthic indicators to use ecological quality classification of
17
18 Mediterranean soft bottom marine ecosystems, including a new biotic index. Mediterr. Mar. Sci.
19
20 3, 77–111.
21
22
23
24
2510 Simonini, R., Ansaloni, I., Bonvicini Pagliai, A.M., Prevedelli, D., 2004. Organic enrichment
26
2711 and structure of the macrozoobenthic community in the northern Adriatic Sea in an area facing
28
29 Adige and Po mouths. ICES J. Mar. Sci. 61, 871–881.
30
31
32
3313 Siqueira, T., Bini, L.M., Roque, F.O., Cottenie, K., 2012. A metacommunity framework for
34
35 enhancing the effectiveness of biological monitoring strategies. PLoS One 7, e43626.
36
37
38
3915 Soares-Gomes, A., Mendes, C.L.T., Tavares, M., Santi, L., 2012. Taxonomic sufficiency of
40
41 polychaete taxocenes for estuary monitoring. Ecol. Indic. 15, 149–156.
42
43
44
4517 Solidoro, C., Pastres, S., Melaku Canu, D., Pellizzato, M., Rossi, R., 2000. Modelling the
46
47 growth of *Tapes philippinarum* in Northern Adriatic lagoons. Mar. Ecol. Progr. Ser. 199, 137–
48
49 148.
50
51
52
5320 Tagliapietra, D., Sigovini, M., Volpi Ghirardini, A., 2009. A review of terms and definitions to
54
55 categorise estuaries, lagoons and associated environments. Mar. Freshw. Res. 60, 497–509.
56
57
58
59
60
61
62
63
64
65

722 Terlizzi, A., Bevilacqua, S., Fraschetti, S., Boero, F., 2003. Taxonomic sufficiency and the
1 increasing insufficiency of taxonomic expertise. *Mar. Pollut. Bull.* 46, 544–560.
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Tataranni, M., Maltagliati, F., Floris, A., Castelli, A., Lardicci, C., 2009. Variance estimate and taxonomic resolution: an analysis of macrobenthic spatial patterns at different scales in a Western Mediterranean coastal lagoon. *Mar. Environ. Res.* 67, 219–229.

Terlizzi, A., Bevilacqua, S., Scuderi, D., Fiorentino, D., Guarnieri, G., Giangrande, A., Licciano, M., Fellingine, S., Fraschetti, S., 2008. Effects of offshore platforms on soft bottom macrobenthic assemblages: a case study in a Mediterranean gas field. *Mar. Pollut. Bull.* 56, 1303–1309.

Terlizzi, A., Anderson, M.J., Bevilacqua, S., Fraschetti, S., Włodarska-Kowalczyk, M., Ellingsen, K.E., 2009. Beta diversity and taxonomic sufficiency: Do higher-level taxa reflect heterogeneity in species composition? *Divers. Distribut.* 15, 450–458.

Thompson, B.W., Riddle, M.J., Stark, J.S., 2003. Cost-efficient methods for marine pollution monitoring at Casey Station, East Antarctica: the choice of sieve mesh-size and taxonomic resolution. *Mar. Pollut. Bull.* 46, 232–243.

Van Hoey, G., Borja, A., Birchenough, S., Buhl-Mortensen, L., Degraer, S., Fleischer, D., Kerckhof, F., Magni, P., Muxika, I., Reiss, H., Schröder, A., Zettler, M.L., 2010. The use of benthic indicators in Europe: From the Water Framework Directive to the Marine Strategy Framework Directive. *Mar. Pollut. Bull.* 60, 2187–2196.

Vanderklift, M.A., Ward, T.J., Jacoby, C.A., 1996. Effect of reducing taxonomic resolution on ordinations to detect pollution-induced gradients in macrobenthic infaunal assemblages. *Mar. Ecol. Progr. Ser.* 136, 137–145.

744 Vincenzi, S., Caramori, G., Rossi, R., De Leo, G. A., 2006. Estimating clam yield potential in
1 the Sacca di Goro lagoon (Italy) by using a two-part conditional model. *Aquaculture* 261, 1281–
2 745 1291.
3
4
5 746
6
7
8 747 Warwick, R.M., 1993. Environmental impact studies on marine communities - pragmatical
9 considerations. *Aust. J. Ecol.* 18, 63–80.
10
11 748
12
13
14 749 White, J., Irvine, K., 2003. The use of littoral mesohabitats and their macroinvertebrate
15 assemblages in the ecological assessment of lakes. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 13,
16 750 331–351. *Mar. Envir. Res.*
17
18
19 751
20
21
22 752 Włodarska-Kowalczyk, M., Kędra, M., 2007. Surrogacy in natural patterns of benthic
23 distribution and diversity: selected taxa versus lower taxonomic resolution. *Mar. Ecol. Progr.*
24 Ser. 351, 53–63.
25 753
26
27 754
28
29
30
31 755
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

756 **Figure captions**

1
2

757 **Figure 1.** Map of investigated lagoon systems. The white circle in the down left panel indicates the
4
5
758 study area where all the three studied lagoons were located. The remaining panels show, clock-
7
759 wise, the three lagoons: Marinetta (MAR), Goro (GOR), and Scardovari (SCA). Numbers in panels
9
10
160 indicate sampling stations in each lagoon.

12
13

1761 **Figure 2.** Semilog plot of ρ values (Spearman's correlation) between the species-level matrix and
15
16
1762 each randomly aggregated matrix against the corresponding ϕ values for MAR, which served as
18
19
1763 pilot lagoon system for BestAgg surrogate selection (see Method section for further details). Fading
20
21
1764 gray zones indicate the range of ϕ values for which results were consistent with those obtained
23
24
1765 analyzing species-level data. Dotted lines indicate ϕ_{low} (i.e. the lowest practicable aggregation
25
26
1766 level), sufficient to reflect species-level patterns consistently. Results of regression analysis are also
28
29
1767 reported.

30
31
32

3768 **Figure 3.** Frequency distribution ($n = 1,000$) of ρ values between the species-level matrix and
34
35
3769 matrices in which species were randomly aggregated in $G_{BestAgg}$ groups (see text), based on the pilot
36
37
3870 lagoon (MAR). Dotted lines indicate $\rho_{BestAgg}$, that is, the correlation value between the species-level
39
40
1771 matrix and the matrix aggregated using BestAgg surrogates, which fall within random expectations.

42
43

4472 **Figure 4.** Non-metric multidimensional scaling ordinations (nMDS) of $T_i \times S_t$ centroids for MAR
45
46
4773 based on Bray-Curtis dissimilarity for species-level data, and data aggregated using family and
47
48
4974 BestAgg surrogates. Symbols indicate different times of sampling, whereas numbers indicate the
50
51
1775 sampling stations (as in Figure 1).

52
53
54

5576 **Figure 5.** Second-stage nMDS ordination plot of inter-matrix rank correlations for MAR (pilot
56
57
1777 lagoon) based based on Bray-Curtis dissimilarity matrices at the species (S), genus (G), family (F),
58

59
60
61
62
63
64
65

778 order (O), class level and using the BestAgg surrogates (BA). Numbers in brackets indicate the
1
779 number of taxa or surrogates for each aggregation.

780 **Figure 6.** Non-metric multidimensional scaling ordinations (nMDS) of $T_i \times St$ centroids for SCA
7
781 (a) and GOR (b) based on Bray–Curtis dissimilarity for species-level data and data aggregated
9
10 using BestAgg surrogates. Numbers indicate sampling stations as reported in Figure 1. For SCA
11782 lagoon (a), lines connect stations' centroids in the three times of sampling (Time 1 = black line;
12
13783 Time 2 = dark grey line; Time 3 = light grey line). For GOR lagoon, circles indicate the four times
14
15784 of sampling (Time 1 = black circles; Time 2 = dark grey circles; Time 3 = light grey circles; Time 4
16
17785 = white circles) whereas clusters group centroids with average similarity $\geq 50\%$.
18
19
20
21786

Figure

[Click here to download high resolution image](#)

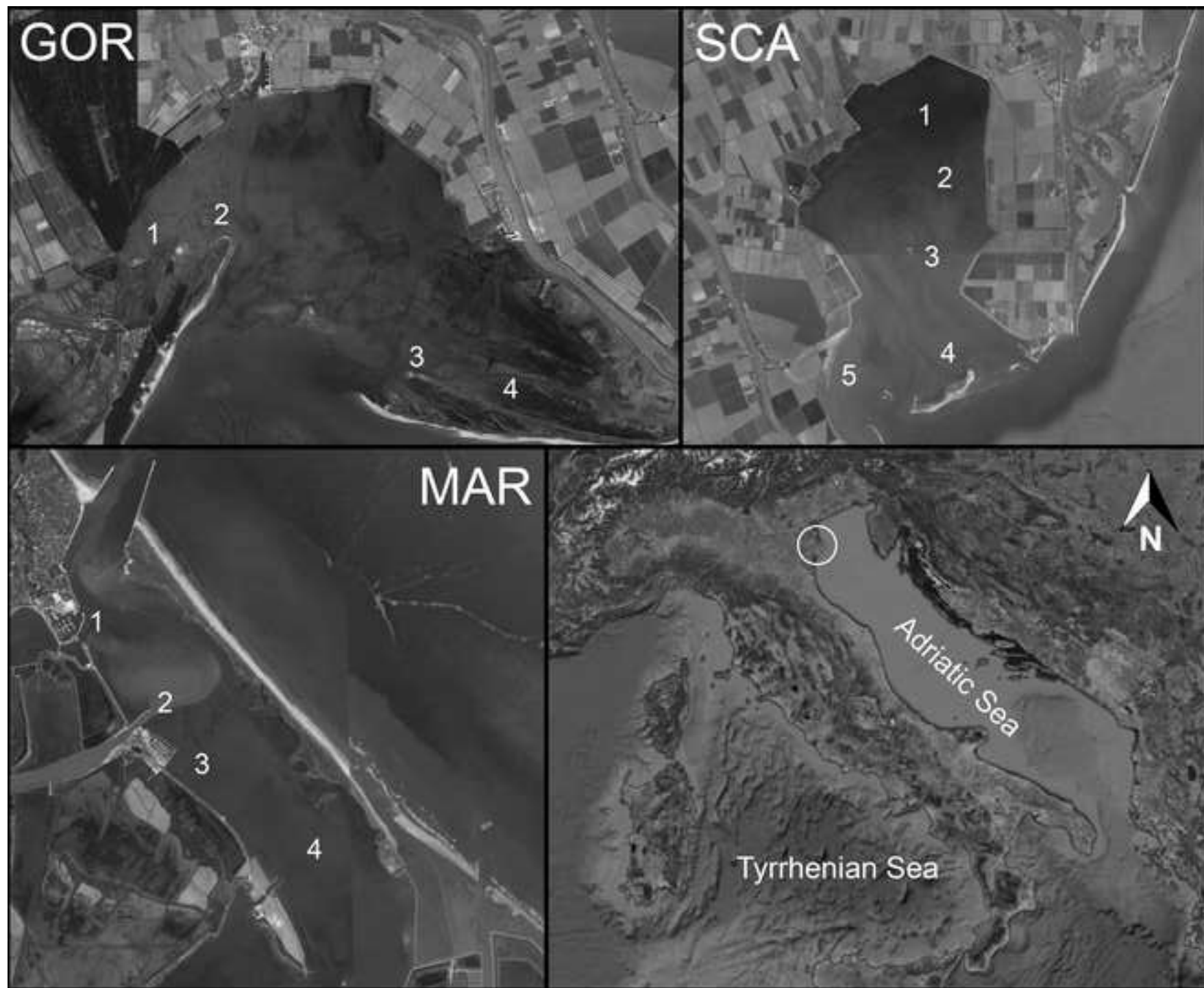
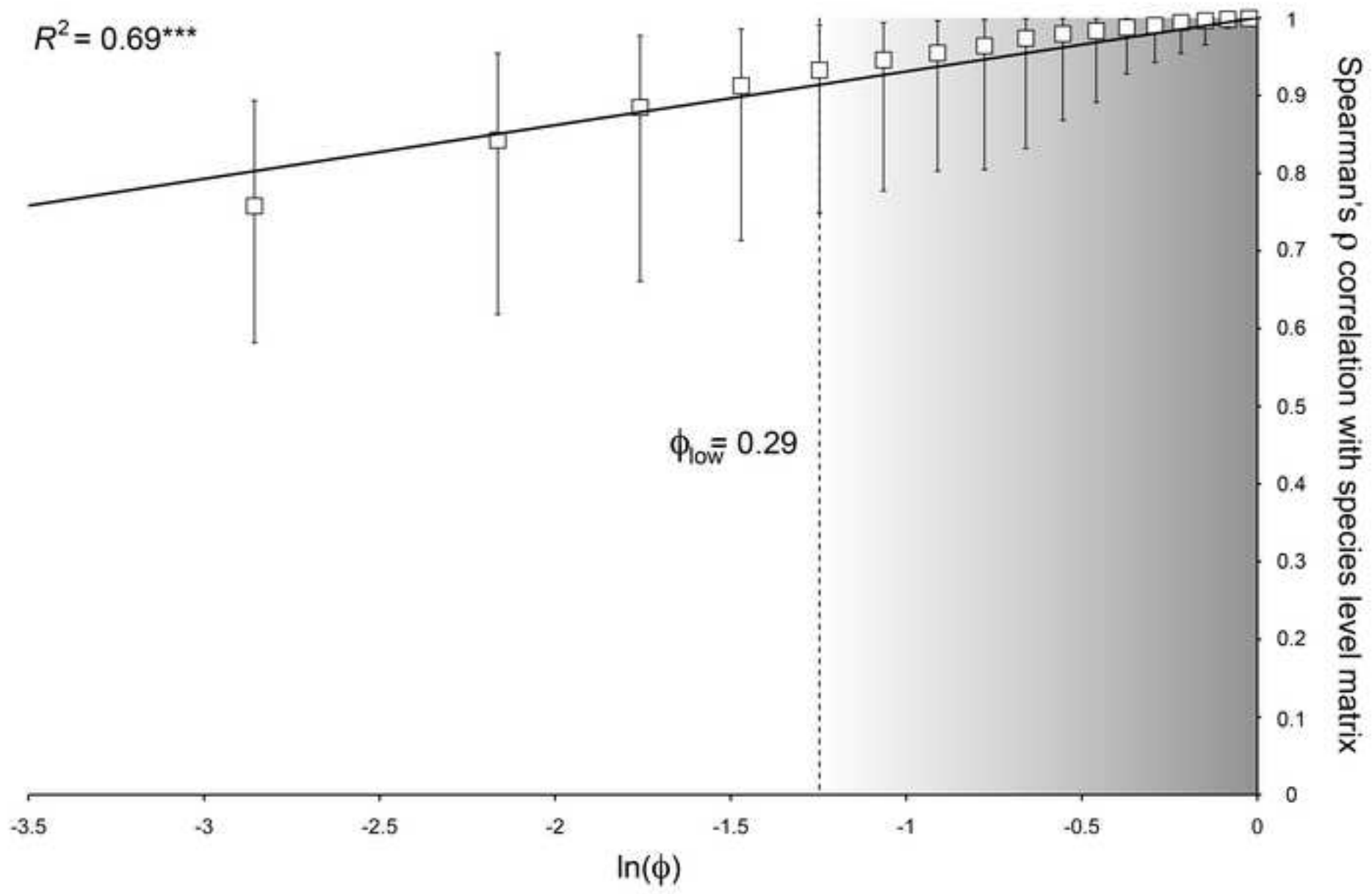
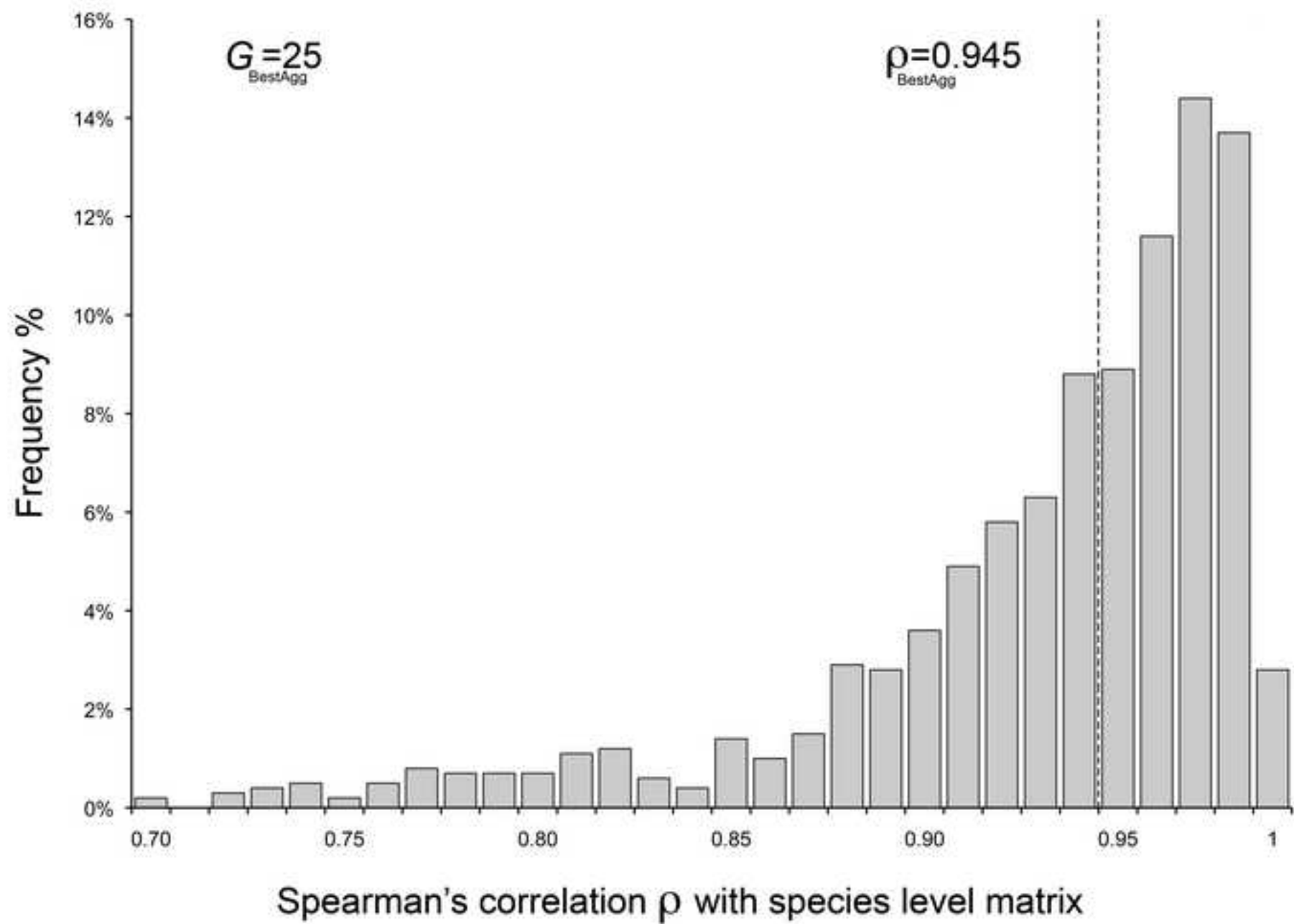


Figure
[Click here to download high resolution image](#)



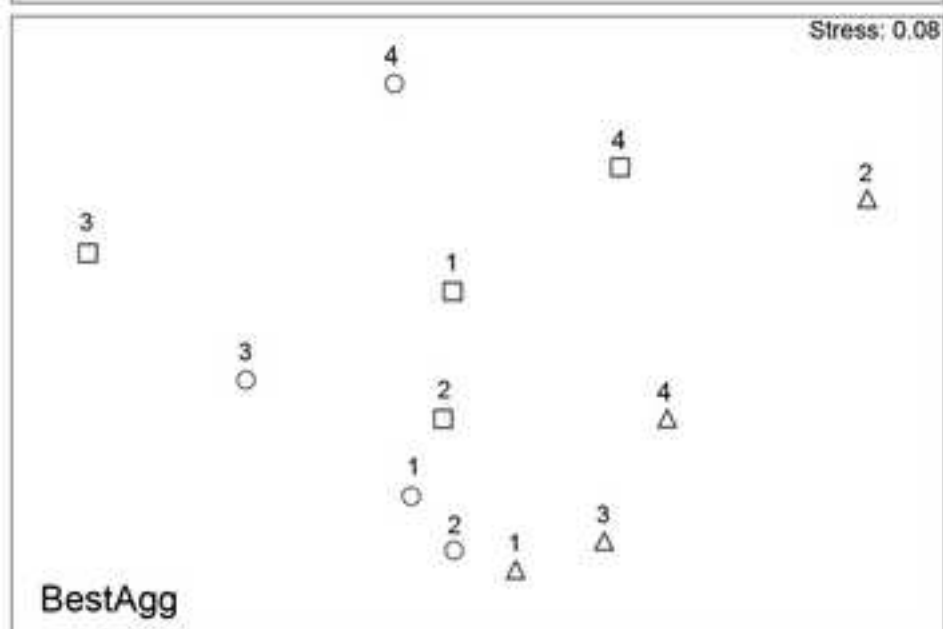
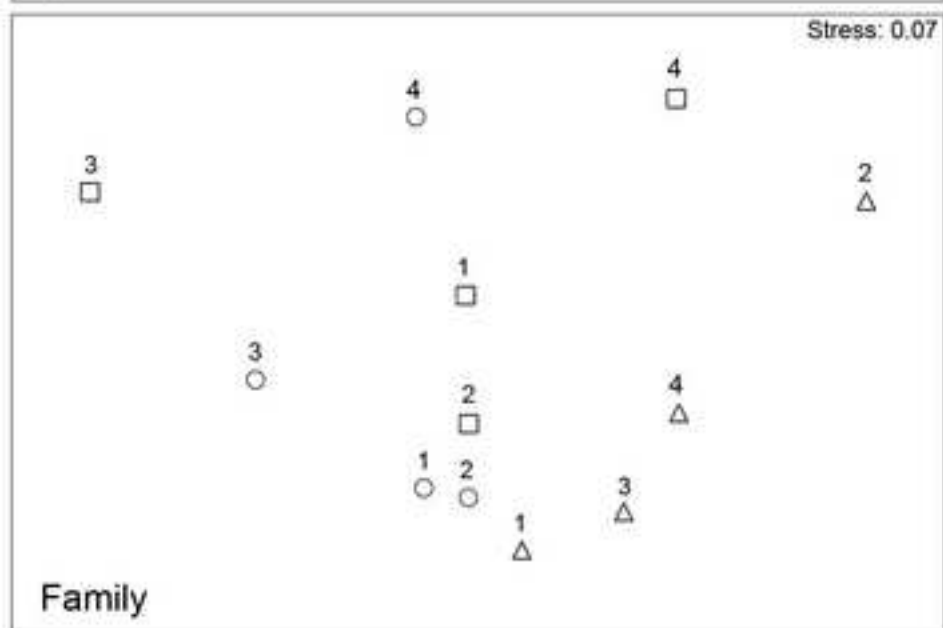
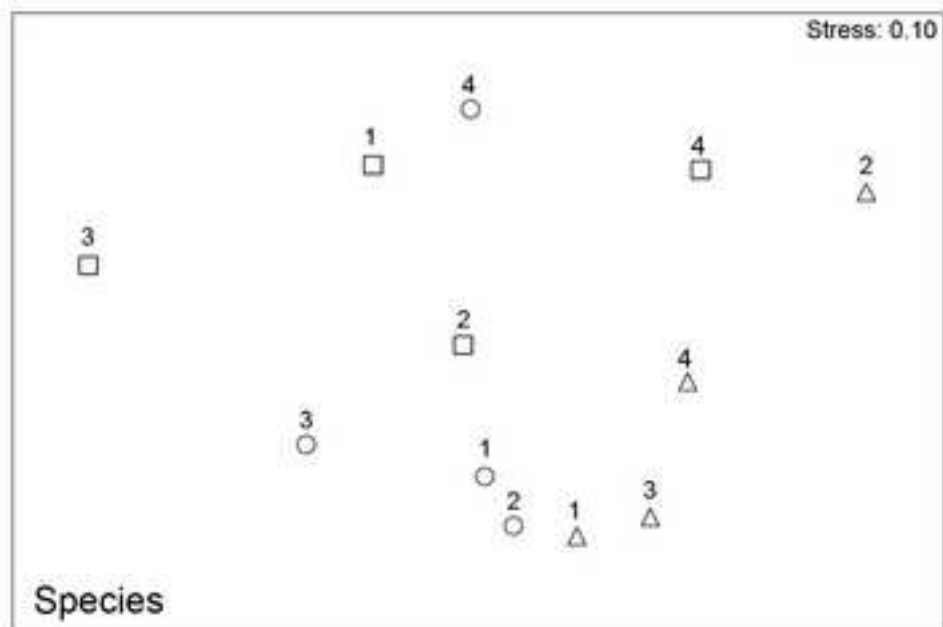
Figure

[Click here to download high resolution image](#)



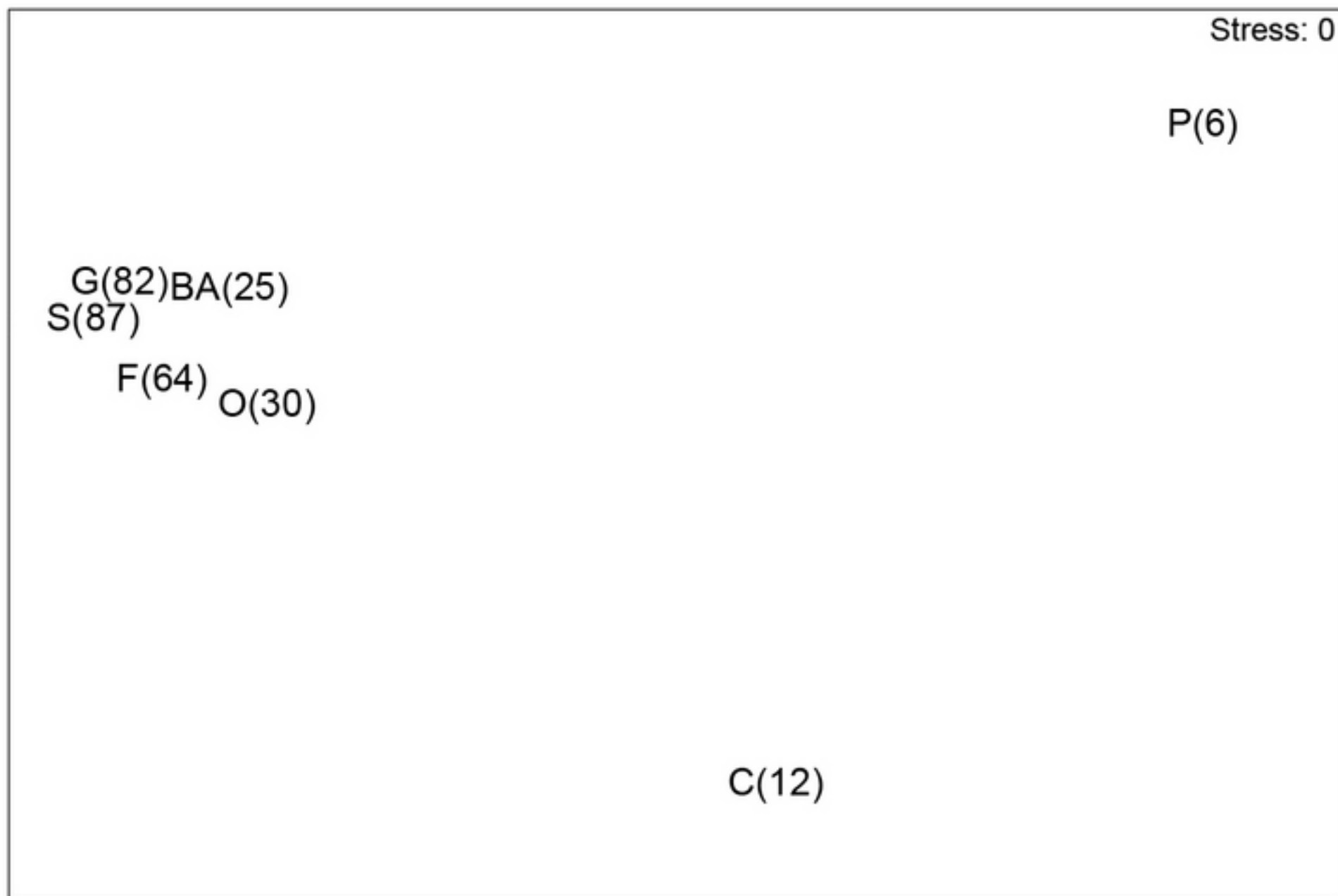
Figure

[Click here to download high resolution image](#)



Figure

[Click here to download high resolution image](#)



Figure

[Click here to download high resolution image](#)

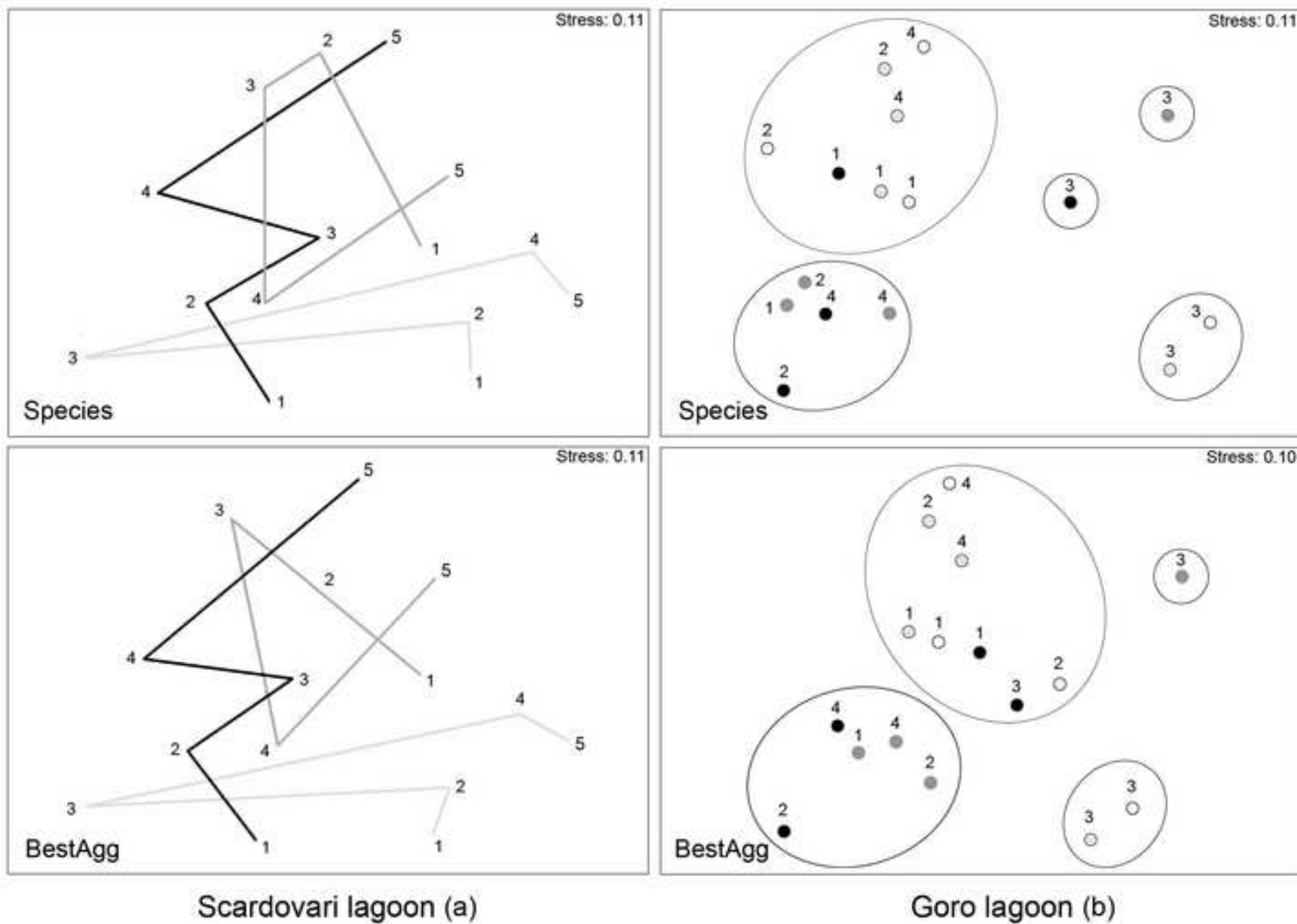


Table 1. Percentage of tests from PERMANOVA on randomly aggregated matrices consistent with those from species-level analyses (pilot lagoon MAR), at decreasing levels of aggregation (ϕ). The corresponding number of surrogates (G) is also provided. The lowest practicable aggregation ϕ_{low} and the corresponding minimum number of surrogates G_{min} are given in bold.

Number of surrogates (G)	Aggregation ratio (ϕ)	% of significant tests consistent with species level analyses
85	0.98	100%
80	0.92	100%
75	0.86	100%
70	0.80	100%
65	0.75	100%
60	0.69	100%
55	0.63	99.8%
50	0.57	99.8%
45	0.52	99.8%
40	0.46	99.2%
35	0.40	98.7%
30	0.34	98.1%
25	0.29	96.4%
20	0.23	93.7%
15	0.17	84.6%
10	0.11	71.4%
5	0.06	28.5%

Table 2. Results of PERMANOVA based on data from the pilot lagoon (MAR) at species level, and aggregated using families or BestAgg surrogates (see Table A2 in Appendix A). Significant results for the term of interest were given in bold.

Source of variation	d.f.	Species			BestAgg			Family		
		MS	<i>F</i>	<i>P</i>	MS	<i>F</i>	<i>P</i>	MS	<i>F</i>	<i>P</i>
Time=T	2	8884.7	2.39	0.031	7229.4	2.23	0.042	7275.3	2.17	0.041
Station=St	3	5337.6	1.43	0.155	4484.6	1.39	0.201	4772.1	1.42	0.177
T×St	6	3723.4	3.61	0.000	3230.2	3.31	0.000	3348.5	3.32	0.000
Residual	24	1032.8			977.0			1009.3		

Table 3. Results of PERMANOVA on the two lagoons in the same geographic area of the pilot lagoon (MAR). Analyses were done at species level and using the BestAgg surrogates obtained from the study on the pilot lagoon. Significant results for the term of interest were given in bold.

Scardovari lagoon (SCA)							
Source of variation	d.f.	Species			BestAgg		
		MS	<i>F</i>	<i>P</i>	MS	<i>F</i>	<i>P</i>
Time=T	2	11553.0	2.36	0.015	10343.0	2.49	0.013
Station=St	4	6622.1	1.35	0.132	5802.0	1.39	0.136
T×St	8	4910.3	3.45	0.000	4158.8	3.31	0.000
Residual	30	1421.7			1255.0		

Goro lagoon (GOR)							
Source of variation	d.f.	Species			BestAgg		
		MS	<i>F</i>	<i>P</i>	MS	<i>F</i>	<i>P</i>
Time=T	3	6494.8	1.30	0.227	5920.3	1.23	0.278
Station=St	3	14348.0	2.88	0.003	7508.8	1.57	0.116
T×St	9	4981.5	5.03	0.000	4796.8	5.27	0.000
Residual	32	990.2			911.0		

Table 4. Results of SIMPER on invertebrate assemblages from SCA. Species most contributing to dissimilarities among times across stations were identified. Only species whose contribution exceeded the 3% were considered. Pair-wise average dissimilarity between times are also reported.

Species	Time 1	Time 2	Time 3	Avg Contrib. (%)
Spionidae	2784.9	6909.1	14114.3	29.8
Oligochaeta	2348.3	5680.7	81.4	19.7
Capitellidae	715.3	1919.1	3495.3	12.8
Polychaeta	1122.3	1186.5	8835.6	9.2
Crustacea	5355.1	86.3	54.3	6.2
<i>Chironomus salinarius</i>	0.0	226.9	2298.9	5.8
Amphipoda	1396.1	1036.0	900.3	5.3
<i>Arcuatula senhousia</i>	64.1	1344.3	367.5	4.2
Semelidae	37.0	59.2	604.3	3.2

Average dissimilarity among times

	Time 1	Time 2	Time 3
Time 1	-		
Time 2	63%	-	
Time 3	81%	69%	-

Table 5. Results of SIMPER on invertebrate assemblages from GOR. Species most contributing to dissimilarities among stations across times were identified. Only species whose contribution exceeded the 3% were considered. Pair-wise average dissimilarity between stations are also reported.

Species	Station 1	Station 2	Station 3	Station 4	Avg Contrib. (%)
Spionidae	6000.2	9354.8	1942.5	5466.8	50.5
Oligochaeta	1427.6	1245.7	126.4	5099.8	25.2
Capitellidae	30.8	70.9	1350.5	508.8	11.5
Amphipoda	450.2	58.6	107.9	1.7	5.7
<i>Ruditapes philippinarum</i>	3.1	9.3	385.4	40.1	4.1

Average dissimilarity among stations

	Station 1	Station 2	Station 3	Station 4
Station 1	-			
Station 2	58%	-		
Station 3	69%	77%	-	
Station 4	55%	54%	82%	-

**New frameworks for species surrogacy in monitoring highly variable coastal ecosystems:
applying the BestAgg approach to Mediterranean coastal lagoons**

Stanislao Bevilacqua, Antonio Terlizzi, Michele Mistri, Cristina Munari

Appendix A

Table A1. List of species found in the pilot lagoon

Table A2. Selected BestAgg surrogates

Table A1. List and taxonomic tree of species found in the pilot lagoon MAR (see Method section). Ecological notes and functional roles were mined from the literature and existing databases (see text for further details). Difficulty of taxonomic identifications were based on expert opinion.

Species	Genus	Family	Order	Class	Phylum	Ecological notes	Identification notes	Functional role
<i>Actinia</i> sp.	<i>Actinia</i>	Actiniidae	Actiniaria	Anthozoa	Cnidaria	-	easy at class level	Predator
<i>Lineus</i> sp.	<i>Lineus</i>	Lineidae	Heteronemertea	Anopla	Nemertea	-	easy at phylum level	Predator
<i>Valencinia</i> sp.	<i>Valencinia</i>	Valenciiniidae	Heteronemertea	Anopla	Nemertea	-	easy at phylum level	Predator
<i>Tetrastemma helvolum</i>	<i>Tetrastemma</i>	Tetrastemmatidae	Monostilifera	Enopla	Nemertea	-	easy at phylum level	Predator
<i>Tubulanus polymorphus</i>	<i>Tubulanus</i>	Tubulanidae	Tubulanidae	Palaeonemertea	Nemertea	-	easy at phylum level	Predator
Leptoplanidae sp.	Leptoplanidae	Leptoplanidae	Polycladida	Rhabditophora	Platyhelminthes	-	easy at phylum level	Predator
<i>Cestoplanea</i> sp.	<i>Cestoplanea</i>	Cestoplanidae	Polycladida	Rhabditophora	Platyhelminthes	-	easy at phylum level	Predator
<i>Moerella distorta</i>	<i>Moerella</i>	Tellinidae	Veneroida	Bivalvia	Mollusca	-	easy at family level	Detritus feeder
<i>Angulus tenuis</i>	<i>Angulus</i>	Tellinidae	Veneroida	Bivalvia	Mollusca	-	easy at family level	Detritus feeder, suspension feeder
<i>Abra alba</i>	<i>Abra</i>	Semelidae	Veneroida	Bivalvia	Mollusca	-	easy at family level	Detritus feeder, suspension feeder
<i>Abra nitida</i>	<i>Abra</i>	Semelidae	Veneroida	Bivalvia	Mollusca	-	easy at family level	Detritus feeder
<i>Cerastoderma glaucum</i>	<i>Cerastoderma</i>	Cardiidae	Veneroida	Bivalvia	Mollusca	-	easy	Suspension/filter feeder
<i>Crassostrea gigas</i>	<i>Crassostrea</i>	Ostreidae	Ostreoida	Bivalvia	Mollusca	Invasive, tolerant	easy	Suspension/filter feeder
<i>Donax semistriatus</i>	<i>Donax</i>	Donacidae	Veneroida	Bivalvia	Mollusca	-	easy at family level	Suspension/filter feeder
<i>Arcuatula senhousia</i>	<i>Arcuatula</i>	Mytilidae	Mytiloida	Bivalvia	Mollusca	Invasive, tolerant	easy	Suspension/filter feeder
<i>Mytilus galloprovincialis</i>	<i>Mytilus</i>	Mytilidae	Mytiloida	Bivalvia	Mollusca	-	easy	Suspension/filter feeder

<i>Ruditapes philippinarum</i>	<i>Ruditapes</i>	Veneridae	Veneroidea	Bivalvia	Mollusca	Invasive, commercial	easy	Suspension/filter feeder
<i>Polititapes aureus</i>	<i>Polititapes</i>	Veneridae	Veneroidea	Bivalvia	Mollusca	-	easy at family level	Suspension/filter feeder
<i>Anadara transversa</i>	<i>Anadara</i>	Arcidae	Arcoidea	Bivalvia	Mollusca	Alien	easy at family level	Suspension/filter feeder
<i>Bittium reticulatum</i>	<i>Bittium</i>	Cerithiidae	unassigned Caenogastropoda	Gastropoda	Mollusca	-	easy	Detritus feeder, herbivorous
<i>Parthenina indistincta</i>	<i>Parthenina</i>	Pyramidellidae	unassigned Heterobranchia	Gastropoda	Mollusca	-	easy at family level	Predator (ectoparasites)
<i>Nassarius nitidus</i>	<i>Nassarius</i>	Nassariidae	Neogastropoda	Gastropoda	Mollusca	-	easy	Scavenger, omnivores
<i>Hydrobia acuta</i>	<i>Hydrobia</i>	Hydrobiidae	Littorinimorpha	Gastropoda	Mollusca	Indicator, tolerant	easy at family level	Detritus feeder, herbivorous
<i>Pusillina sarsii</i>	<i>Pusillina</i>	Rissoidae	Littorinimorpha	Gastropoda	Mollusca	-	easy at family level	Herbivorous
<i>Haminoea navicula</i>	<i>Haminoea</i>	Haminoeidae	Cephalaspidea	Gastropoda	Mollusca	-	easy at order level	Predator
<i>Philine aperta</i>	<i>Philine</i>	Philinidae	Cephalaspidea	Gastropoda	Mollusca	-	easy at order level	Predator
<i>Runcina</i> sp.	<i>Runcina</i>	Runcinidae	Runcinacea	Gastropoda	Mollusca	-	easy at class level	Herbivorous
<i>Berthella</i> sp.	<i>Berthella</i>	Pleurobranchidae	Pleurobranchomorpha	Gastropoda	Mollusca	-	easy at class level	Herbivorous, predator
<i>Tubificoides</i> sp.	<i>Tubificoides</i>	Tubificidae	Haplotaxida	Oligochaeta	Annelida	Opportunistic	easy at class level	Detritus feeder/Limivores
<i>Amphicorina armandi</i>	<i>Amphicorina</i>	Sabellidae	Sabellida	Polychaeta	Annelida	-	easy at family level	Suspension/filter feeder
<i>Desdemona ornata</i>	<i>Desdemona</i>	Sabellidae	Sabellida	Polychaeta	Annelida	Alien	easy at family level	Suspension/filter feeder
<i>Sabellaria alcocki</i>	<i>Sabellaria</i>	Sabellariidae	Sabellida	Polychaeta	Annelida	-	easy at family level	Suspension/filter feeder
<i>Spirobranchus triqueter</i>	<i>Spirobranchus</i>	Serpulidae	Sabellida	Polychaeta	Annelida	-	easy at family level	Suspension/filter feeder
<i>Spirobranchus polytrema</i>	<i>Spirobranchus</i>	Serpulidae	Sabellida	Polychaeta	Annelida	-	easy at family level	Suspension/filter feeder
<i>Hydroides dianthus</i>	<i>Hydroides</i>	Serpulidae	Sabellida	Polychaeta	Annelida	Alien	easy at family level	Suspension/filter feeder

<i>Hydroides elegans</i>	<i>Hydroides</i>	Serpulidae	Sabellida	Polychaeta	Annelida	Alien	easy at family level	Suspension/filter feeder
<i>Neodexiospira pseudocorrugata</i>	<i>Neodexiospira</i>	Serpulidae	Sabellida	Polychaeta	Annelida	-	easy at family level	Suspension/filter feeder
<i>Armandia cirrhosa</i>	<i>Armandia</i>	Opheliidae	Scolecida	Polychaeta	Annelida	Sensitive	easy	Detritus feeder
<i>Heteromastus filiformis</i>	<i>Heteromastus</i>	Capitellidae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Detritus feeder/Limivores
<i>Capitella capitata</i>	<i>Capitella</i>	Capitellidae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Detritus feeder/Limivores
<i>Mediomastus fragilis</i>	<i>Mediomastus</i>	Capitellidae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Detritus feeder/Limivores
<i>Pseudocapitella fauveli</i>	<i>Pseudocapitella</i>	Capitellidae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Detritus feeder/Limivores
<i>Cirrophorus furcatus</i>	<i>Cirrophorus</i>	Paraonidae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Detritus feeder/Limivores
<i>Paradoneis lyra</i>	<i>Paradoneis</i>	Paraonidae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Detritus feeder/Limivores
<i>Cossura soyeri</i>	<i>Cossura</i>	Cossuridae	Scolecida	Polychaeta	Annelida	Indicator, tolerant	easy	Detritus feeder/Limivores
<i>Mysta picta</i>	<i>Mysta</i>	Phyllodocidae	Phyllodocida	Polychaeta	Annelida	-	easy at family level	Predator
<i>Gattyana cirrhosa</i>	<i>Gattyana</i>	Polynoidae	Phyllodocida	Polychaeta	Annelida	-	easy at family level	Predator
<i>Glycera tridactyla</i>	<i>Glycera</i>	Glyceridae	Phyllodocida	Polychaeta	Annelida	-	easy	Predator
<i>Websterinereis glauca</i>	<i>Websterinereis</i>	Nereididae	Phyllodocida	Polychaeta	Annelida	-	easy at family level	Detritus feeder, herbivorous
<i>Platynereis dumerilii</i>	<i>Platynereis</i>	Nereididae	Phyllodocida	Polychaeta	Annelida	Indicator, tolerant	easy at family level	Herbivorous
<i>Alitta succinea</i>	<i>Alitta</i>	Nereididae	Phyllodocida	Polychaeta	Annelida	Indicator, tolerant	easy	Detritus feeder, omnivorous
<i>Nephtys hombergii</i>	<i>Nephtys</i>	Nephtyidae	Phyllodocida	Polychaeta	Annelida	Sensitive	easy at family level	Predator
<i>Micronephthys maryae</i>	<i>Micronephthys</i>	Nephtyidae	Phyllodocida	Polychaeta	Annelida	-	easy at family level	Predator
<i>Syllis gracilis</i>	<i>Syllis</i>	Syllidae	Phyllodocida	Polychaeta	Annelida	-	easy at family level	Detritus feeder, predator, omnivorous

<i>Podarkeopsis capensis</i>	<i>Podarkeopsis</i>	Hesionidae	Phyllodocida	Polychaeta	Annelida	Alien	easy at family level	Predator
<i>Schistomeringos rudolphii</i>	<i>Schistomeringos</i>	Dorvilleidae	Eunicida	Polychaeta	Annelida	-	easy at family level	Predator
<i>Protodorvillea kefersteini</i>	<i>Protodorvillea</i>	Dorvilleidae	Eunicida	Polychaeta	Annelida	-	easy at family level	Predator
<i>Lumbrineris latreilli</i>	<i>Lumbrineris</i>	Lumbrineridae	Eunicida	Polychaeta	Annelida	-	easy at family level	Detritus feeder, omnivorous
<i>Malacoceros fuliginosus</i>	<i>Malacoceros</i>	Spionidae	Spionida	Polychaeta	Annelida	Tolerant	easy at family level	Detritus feeder
<i>Prionospio cirrifera</i>	<i>Prionospio</i>	Spionidae	Spionida	Polychaeta	Annelida	Tolerant	easy at family level	Detritus feeder
<i>Prionospio multibranchiata</i>	<i>Prionospio</i>	Spionidae	Spionida	Polychaeta	Annelida	Tolerant	easy at family level	Detritus feeder
<i>Polydora ciliata</i>	<i>Polydora</i>	Spionidae	Spionida	Polychaeta	Annelida	Tolerant	easy at family level	Detritus feeder
<i>Spio decoratus</i>	<i>Spio</i>	Spionidae	Spionida	Polychaeta	Annelida	Tolerant	easy at family level	Detritus feeder
<i>Streblospio shrubsolii</i>	<i>Streblospio</i>	Spionidae	Spionida	Polychaeta	Annelida	Tolerant	easy at family level	Detritus feeder
<i>Polycirrus aurantiacus</i>	<i>Polycirrus</i>	Terebellidae	Terebellida	Polychaeta	Annelida	-	easy at family level	Detritus feeder
<i>Ampithoe ferox</i>	<i>Ampithoe</i>	Ampithoidae	Amphipoda	Malacostraca	Arthropoda	-	easy at family level	Herbivorous
<i>Ampithoe valida</i>	<i>Ampithoe</i>	Ampithoidae	Amphipoda	Malacostraca	Arthropoda	Alien	easy at family level	Herbivorous
<i>Ampelisca sarsi</i>	<i>Ampelisca</i>	Ampeliscidae	Amphipoda	Malacostraca	Arthropoda	-	easy	Detritus feeder, suspension feeder
<i>Caprella scaura</i>	<i>Caprella</i>	Caprellidae	Amphipoda	Malacostraca	Arthropoda	Alien	easy at family level	Omnivorous, predator, scavenger
<i>Monocorophium insidiosum</i>	<i>Monocorophium</i>	Corophiidae	Amphipoda	Malacostraca	Arthropoda	Biodiffuser	easy	Detritus feeder
<i>Perioculodes aequimanus</i>	<i>Perioculodes</i>	Oedicerotidae	Amphipoda	Malacostraca	Arthropoda	-	easy	Omnivorous, predator, detritus feeder
<i>Erichthonius brasiliensis</i>	<i>Erichthonius</i>	Ischyroceridae	Amphipoda	Malacostraca	Arthropoda	-	easy	Detritus feeder
<i>Gammarus aequicauda</i>	<i>Gammarus</i>	Gammaridae	Amphipoda	Malacostraca	Arthropoda	-	easy	Herbivorous

<i>Dexamine spinosa</i>	<i>Dexamine</i>	Dexaminidae	Amphipoda	Malacostraca	Arthropoda	-	easy	Detritus feeder, herbivorous
<i>Melita palmata</i>	<i>Melita</i>	Melitidae	Amphipoda	Malacostraca	Arthropoda	Indicator, tolerant	easy	Herbivorous, detritus feeder, suspension feeder
<i>Microdeutopus gryllotalpa</i>	<i>Microdeutopus</i>	Aoridae	Amphipoda	Malacostraca	Arthropoda	Biodiffuser	easy	Herbivorous, suspension feeder,
<i>Stenothoe monoculoides</i>	<i>Stenothoe</i>	Stenothoidae	Amphipoda	Malacostraca	Arthropoda	-	easy at family level	Suspension feeder, predator
<i>Uromunna petiti</i>	<i>Uromunna</i>	Munnidae	Isopoda	Malacostraca	Arthropoda	-	easy at family level	Detritus feeder
<i>Cyathura carinata</i>	<i>Cyathura</i>	Anthuridae	Isopoda	Malacostraca	Arthropoda	-	easy	Detritus feeder
<i>Idotea balthica</i>	<i>Idotea</i>	Idoteidae	Isopoda	Malacostraca	Arthropoda	-	easy	Herbivorous
<i>Iphinoe serrata</i>	<i>Iphinoe</i>	Bodotriidae	Cumacea	Malacostraca	Arthropoda	-	easy at order level	Detritus feeder
<i>Carcinus aestuarii</i>	<i>Carcinus</i>	Portunidae	Decapoda	Malacostraca	Arthropoda	Commercial target	easy	Predator
<i>Dyspanopeus sayi</i>	<i>Dyspanopeus</i>	Panopeidae	Decapoda	Malacostraca	Arthropoda	Alien	easy	Predator
<i>Palaemon adspersus</i>	<i>Palaemon</i>	Palaemonidae	Decapoda	Malacostraca	Arthropoda	Commercial target	easy at family level	Predator
<i>Upogebia pusilla</i>	<i>Upogebia</i>	Upogebiidae	Decapoda	Malacostraca	Arthropoda	Commercial target, bioturbation	easy	Suspension feeder, detritus feeder
<i>Amphibalanus eburneus</i>	<i>Amphibalanus</i>	Balanidae	Sessilia	Maxillopoda	Arthropoda	-	easy at family level	Suspension/filter feeder
<i>Chironomus salinarius</i>	<i>Chironomus</i>	Chironomidae	Diptera	Insecta	Arthropoda	Indicator	easy	Detritus feeder, omnivorous

References

Patton, W.K., 1968. Feeding habitats, behaviour and host specificity of *Caprella grahmi*, an amphipod commensal with the starfish *Asterias forbesi*. Biol. Bull. 134, 148–153.

Inyangn, M., 1978. Notes on the food of the Baltic Palaemonid shrimp *Palaemon adspersus* var. *fabricii* (Rathke). Meeresforschung Rep. Mar. Res. 26, 42–46.

Pearson, T.H., Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanogr. Mar. Biol. Annu. Rev.* 16, 229–311.

Fauchald, K., Jumars, P.A., 1979. The diet of worms: a study of polychaete feeding guilds. *Oceanogr. Mar. Biol. Annu. Rev.* 17, 193–284.

Levinton, J.S., 1982. *Marine ecology*. Prentice Hall, New Jersey, pp. 526.

Dauvin, J.-C., Ibanez, F. 1986. Variations a long-terms (1977–1985) du peuplement des sables fins de la Pierre Noire (baie de Morlaix, Manche occidentale): analyse statistique de l'évolution structurale. *Hydrobiologia* 142, 171–186.

Guerao, G., 1994. Feeding habits of the prawns *Processa edulis* and *Palaemon adspersus* (Crustacea, Decapoda, Caridea) in the Alfacs Bay, Ebro Delta (NW Mediterranean). *Misc. Zool.* 17, 115–122.

Giangrande, A., Licciano, M., Pagliara, P., 2000. The diversity of diets in Syllidae Annelida: Polychaeta. *Cah. Biol. Mar.* 41, 55–65.

Simboura, N., Zenetos, A., 2002. Benthic indicators to use ecological quality classification of Mediterranean soft bottom marine ecosystems, including a new biotic index. *Mediterr. Mar. Sci.* 3, 77–111.

Scipione M.B., Zupo V., 2010. Crustacean amphipods from the seagrasses *Zostera marina*, *Cymodocea nodosa* and *Posidonia oceanica* in the Adriatic Sea (Italy): a first comparison. *Zoologica baetica* 21, 15–32.

de-la-Ossa-Carretero, J.A., Del-Pilar-Ruso, Y., Giménez-Casalduero, F., Sánchez-Lizaso, J.L., 2011. Assessing reliable indicators to sewage pollution in coastal soft-bottom communities. *Environ. Monit. Assess.* 184, 2133–2149.

Navarro-Barranco, C., Tierno de Figueroa, J.M., Guerra-García, J.M., Sánchez Tocino, L., García-Gómez, J.C., 2013. Feeding habits of amphipods (Crustacea: Malacostraca) from shallow soft bottom communities; comparison between marine caves and open habitats. *J. Sea Res.* 78, 1–7.

Table A2. Selected surrogates for BestAgg based on the pilot lagoon MAR (see Method section). *Relevance* is reported according to evidence from literature and expert opinion (see Table A1) and SIMPER analysis (see Method section, see also TableB1 in Appendix B). For *Easiness*: y = easy identification, n = difficult identification. In the *Resemblance* column were reported aggregation criteria if applied (NA = not applied). Priority has been assigned following the procedure described in Bevilacqua et al. (2013). Numbers in brackets are the number of species included in taxa or surrogates.

Phylum	Species	Relevance	Easiness	Resemblance	BestAgg	Priority
Cnidaria (1)	<i>Actinia</i> sp.	Not relevant	n	Could be difficult at species and/or genus level, aggregated in Anthozoans	Anthozoans (1)	Low
Nemertea (4)	<i>Lineus</i> sp.	Not relevant	n	Predators, difficult to identify at species level, aggregated in Flat/ribbon worms	Flat/ribbon worms (6)	Low
	<i>Valencinia</i> sp.					
	<i>Tetrastemma helvolum</i>					
Platyhelminthes (2)	<i>Tubulanus polymorphus</i>	Not relevant	n			
	Leptoplanidae sp.					
Mollusca (21)	<i>Cestoplana</i> sp.					
	<i>Anadara transversa</i>	Alien/invasive	y	NA	<i>Anadara transversa</i> (1)	High
	<i>Crassostrea gigas</i>	Invasive	y	NA	<i>Crassostrea gigas</i> (1)	High
	<i>Arcuatula senhousia</i>	Simper, invasive	y	NA	<i>Arcuatula senhousia</i> (1)	High
	<i>Ruditapes philippinarum</i>	Invasive	y	NA	<i>Ruditapes philippinarum</i> (1)	High
	<i>Moerella distorta</i>	Relevant as detritus feeder bivalves	y	Easier at Family level, aggregated in Tellinidae	Tellinidae (2)	Medium
	<i>Angulus tenuis</i>					
	<i>Abra alba</i>	Relevant as detritus feeder bivalves	n	Difficult to identify at species level, aggregated in Semelidae	Semelidae (2)	Medium
	<i>Abra nitida</i>					
	<i>Cerastoderma glaucum</i>					
<i>Donax semistriatus</i>	Relevant as suspension filter feeder bivalves	n	Bivalves, all suspension filter feeders, aggregated in Other Bivalves	Other Bivalves (4)	Medium	
<i>Mytilus galloprovincialis</i>						
<i>Polititapes aureus</i>						

Annelida (37)	<i>Bittium reticulatum</i>			Mostly predator/herbivorous, general sensitiveness to pollution, difficult to identify at species level, aggregated in Gastropods	Gastropods (9)	Medium
	<i>Parthenina indistincta</i>					
	<i>Nassarius nitidus</i>					
	<i>Hydrobia acuta</i>					
	<i>Pusillina sarsii</i>	Not relevant	n			
	<i>Haminoea navicula</i>					
	<i>Philine aperta</i>					
	<i>Runcina sp.</i>					
	<i>Berthella sp.</i>					
	<i>Tubificoides sp.</i>	Simper, opportunistic	n			
<i>Armandia cirrhosa</i>	Sensitive, typical of lagoon systems	y	NA	<i>Armandia cirrhosa</i> (1)	High	
<i>Alitta succinea</i>	Simper	y	NA	<i>Alitta succinea</i> (1)	High	
<i>Heteromastus filiformis</i>	Not relevant		Difficult to identify at species level, aggregated in Capitellidae (Indicator)	Capitellidae (4)	Medium	
<i>Capitella capitata</i>	Simper, indicator	n				
<i>Mediomastus fragilis</i>	Not relevant					
<i>Pseudocapitella fauveli</i>	Indicator					
<i>Mysta picta</i>	Relevant as predator polychaetes		Generally difficult to identify at species level, mostly predators	Phyllodocida (9)	Medium	
<i>Gattyana cirrhosa</i>	Relevant as predator polychaetes					
<i>Glycera tridactyla</i>	Relevant as predator polychaetes					
<i>Websterinereis glauca</i>	Not relevant					
<i>Platynereis dumerilii</i>	Not relevant	n				
<i>Nephtys hombergii</i>	Relevant as predator polychaetes					
<i>Micronephthys maryae</i>	Relevant as predator polychaetes					
<i>Syllis gracilis</i>	Relevant as predator polychaetes					
<i>Podarkeopsis capensis</i>	Alien					

	<i>Schistomeringos rudolphii</i>	Relevant as predator polychaetes					
	<i>Protodorvillea kefersteini</i>	Relevant as predator polychaetes	n	Generally difficult to identify at species level, mostly predators		Eunicida (3)	Medium
	<i>Lumbrineris latreilli</i>	Not relevant					
	<i>Malacoceros fuliginosus</i>	Not relevant					
	<i>Prionospio cirrifera</i>	Tolerant/Indicator					
	<i>Prionospio multibranchiata</i>	Not relevant	n	Easier at Family level, aggregated in Spionidae (Tolerant/indicator)		Spionidae (6)	Medium
	<i>Polydora ciliata</i>	Simper					
	<i>Spio decoratus</i>	Not relevant					
	<i>Streblospio shrubsolii</i>	Simper					
	<i>Amphicorina armandi</i>	Not relevant					
	<i>Desdemona ornata</i>	Simper, Alien					
	<i>Sabellaria alcocki</i>	Not relevant					
	<i>Spirobranchus triqueter</i>	Not relevant					
	<i>Spirobranchus polytrema</i>	Not relevant					
	<i>Hydroides dianthus</i>	Not relevant	n	Mostly not relevant, difficult to identify at species level, aggregated in Other Polychaetes		Other Polychaetes (10)	Low
	<i>Hydroides elegans</i>	Not relevant					
	<i>Neodexiospira pseudocorrugata</i>	Not relevant					
	<i>Cirrophorus furcatus</i>	Not relevant					
	<i>Paradoneis lyra</i>	Not relevant					
	<i>Cossura soyeri</i>	Not relevant					
	<i>Polycirrus aurantiacus</i>	Not relevant					
Arthropoda (22)	<i>Cyathura carinata</i>	Simper	y		NA	<i>Cyathura carinata</i> (1)	High
	<i>Carcinus aestuarii</i>	Relevant for fisheries and as predator	y		NA	<i>Carcinus aestuarii</i> (1)	High
	<i>Dyspanopeus sayi</i>	Alien	y		NA	<i>Dyspanopeus sayi</i> (1)	High
	<i>Upogebia pusilla</i>	Relevant as bait for fisheries, bioturbation	y		NA	<i>Upogebia pusilla</i> (1)	High
	<i>Chironomus salinarius</i>	Simper, indicator	y		NA	<i>Chironomus salinarius</i> (1)	High
	<i>Ampithoe ferox</i>	Not relevant	n	Generally difficult to identify at species level,		Amphipoda (12)	Medium

<i>Ampithoe valida</i>	Alien		aggregated in Amphipoda		
<i>Ampelisca sarsi</i>	Not relevant				
<i>Caprella scaura</i>	Alien				
<i>Monocorophium insidiosum</i>	Simper, biodiffuser				
<i>Perioculodes aequimanus</i>	Simper				
<i>Erichthonius brasiliensis</i>	Not relevant				
<i>Gammarus aequicauda</i>	Not relevant				
<i>Dexamine spinosa</i>	Simper, biodiffuser				
<i>Melita palmata</i>	Not relevant				
<i>Microdeutopus gryllotalpa</i>	Biodiffuser				
<i>Stenothoe monoculoides</i>	Not relevant				
<i>Uromunna petiti</i>					
<i>Idotea balthica</i>					
<i>Iphinoe serrata</i>	Not relevant	n	Generally difficult at species level, aggregated in Other Crustacea	Other Crustaceans (5)	Low
<i>Palaemon adspersus</i>					
<i>Amphibalanus eburneus</i>					

**New frameworks for species surrogacy in monitoring highly variable coastal ecosystems:
applying the BestAgg approach to Mediterranean coastal lagoons**

Stanislao Bevilacqua, Antonio Terlizzi, Michele Mistri, Cristina Munari

Appendix B

Table B1. Result of SIMPER on the pilot lagoon

APPENDIX B

Table B1. Results of SIMPER on data from the pilot lagoon (MAR). Species most contributing to dissimilarities among times and among stations were identified. Only species whose contribution exceeded the 3% were considered.

SIMPER among times (across stations)

Time 1 vs. Time 2 (Average dissimilarity = 58.11)

Species	Average abundance Time 1	Average abundance Time 2	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	8534.67	3228.25	15.3	26.32	26.32
<i>Tubificoides</i> sp.	3351.58	3348.5	7.29	12.55	38.87
<i>Capitella capitata</i>	1538.58	897.25	4.15	7.14	46.01
<i>Desdemona ornata</i>	1177.83	255.92	3.56	6.13	52.13
<i>Polydora ciliata</i>	40.08	1291.92	3.48	6	58.13
<i>Dexamine spinosa</i>	2130.58	1510.83	3.45	5.95	64.08
<i>Monocorophium insidiosum</i>	3.08	647.5	2.58	4.43	68.51
<i>Alitta succinea</i>	428.58	9.25	2.48	4.26	72.77
<i>Chironomus salinarius</i>	6.17	425.5	2.32	3.99	76.76

Time 1 vs. Time 3 (Average dissimilarity = 66.70)

Species	Average abundance Time 1	Average abundance Time 3	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	8534.67	9135.92	29.67	44.48	44.48
<i>Tubificoides</i> sp.	3351.58	249.75	12.37	18.55	63.03
<i>Dexamine spinosa</i>	2130.58	6.17	5.49	8.23	71.26
<i>Capitella capitata</i>	1538.58	1063.75	4.12	6.17	77.43
<i>Desdemona ornata</i>	1177.83	0	2.97	4.45	81.88
<i>Alitta succinea</i>	428.58	95.58	2.52	3.79	85.67

Time 2 vs. Time 3 (Average dissimilarity = 75.49)

Species	Average abundance Time 2	Average abundance Time 3	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	3228.25	9135.92	29.53	39.12	39.12
<i>Tubificoides</i> sp.	3348.5	249.75	11.22	14.87	53.99
<i>Polydora ciliata</i>	1291.92	15.42	5.38	7.12	61.11
<i>Dexamine spinosa</i>	1510.83	6.17	4.48	5.94	67.04
<i>Capitella capitata</i>	897.25	1063.75	3.16	4.19	71.23
<i>Monocorophium insidiosum</i>	647.5	12.33	2.87	3.8	75.03
<i>Chironomus salinarius</i>	425.5	9.25	2.54	3.37	78.4
<i>Cyathura carinata</i>	629	46.25	2.42	3.21	81.6

SIMPER among stations (across times)

Station 1 vs. Station 2 (Average dissimilarity = 52.82)

Species	Average abundance Station 1	Average abundance Station 2	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	8152.33	7687.78	28.84	51.67	51.67
<i>Capitella capitata</i>	3280.67	439.89	9.14	16.37	68.04
<i>Desdemona ornata</i>	1660.89	209.67	3.94	7.05	75.09
<i>Tubificoides</i> sp.	1336.11	1628	3.56	6.37	81.46
<i>Dexamine spinosa</i>	715.33	131.56	1.88	3.36	84.82

Station 1 vs. Station 3 (Average dissimilarity = 59.32)

Species	Average abundance Station 1	Average abundance Station 3	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	8152.33	6376.33	14.52	24.48	24.48
<i>Tubificoides</i> sp.	1336.11	4604.44	9.22	15.55	40.03
<i>Capitella capitata</i>	3280.67	291.89	7.71	13	53.03
<i>Dexamine spinosa</i>	715.33	3909.67	7.17	12.09	65.12
<i>Desdemona ornata</i>	1660.89	12.33	3.37	5.68	70.8
<i>Perioculodes aequimanus</i>	12.33	1212.78	2.83	4.77	75.57
<i>Arcuatula senhousia</i>	90.44	883.89	2.09	3.53	79.1

Station 1 vs. Station 4 (Average dissimilarity = 64.85)

Species	Average abundance Station 1	Average abundance Station 4	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	8152.33	5648.67	20.02	30.87	30.87
<i>Capitella capitata</i>	3280.67	653.67	9.61	14.82	45.69
<i>Tubificoides</i> sp.	1336.11	1697.89	6.29	9.7	55.39
<i>Desdemona ornata</i>	1660.89	28.78	5.19	8	63.39
<i>Polydora ciliata</i>	830.44	28.78	3.76	5.8	69.19
<i>Monocorophium insidiosum</i>	16.44	661.89	2.83	4.37	73.56
<i>Chironomus salinarius</i>	16.44	555	2.82	4.35	77.92
<i>Alitta succinea</i>	115.11	587.89	2.22	3.42	81.34
<i>Dexamine spinosa</i>	715.33	106.89	2.14	3.3	84.63

Station 2 vs. Station 3 (Average dissimilarity = 64.84)

Species	Average abundance Station 2	Average abundance Station 3	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	7687.78	6376.33	29.75	45.89	45.89
<i>Tubificoides</i> sp.	1628	4604.44	8.14	12.56	58.45
<i>Dexamine spinosa</i>	131.56	3909.67	8.06	12.43	70.88
<i>Perioculodes aequimanus</i>	61.67	1212.78	2.63	4.05	74.93
<i>Arcuatula senhousia</i>	32.89	883.89	2.11	3.26	78.19

Station 2 vs. Station 4 (Average dissimilarity = 61.68)

Species	Average abundance Station 2	Average abundance Station 4	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	7687.78	5648.67	34.17	55.4	55.4
<i>Tubificoides</i> sp.	1628	1697.89	4.89	7.93	63.33
<i>Polydora ciliata</i>	686.56	28.78	2.86	4.63	67.96
<i>Monocorophium insidiosum</i>	8.22	661.89	2.54	4.13	72.09
<i>Capitella capitata</i>	439.89	653.67	2.54	4.13	76.21
<i>Chironomus salinarius</i>	0	555	2.44	3.95	80.17
<i>Alitta succinea</i>	8.22	587.89	2.08	3.38	83.55
<i>Websterinereis glauca</i>	382.33	0	1.36	2.21	85.76
<i>Cyathura carinata</i>	296	8.22	1.25	2.02	87.78
<i>Microdeutopus gryllotalpa</i>	28.78	69.89	0.82	1.32	89.1
<i>Dexamine spinosa</i>	131.56	106.89	0.82	1.32	90.43

Station 3 vs. Station 4 (Average dissimilarity = 68.69)

Species	Average abundance Station 3	Average abundance Station 4	Average Dissimilarity	Contribution%	Cumulative%
<i>Streblospio shrubsolii</i>	6376.33	5648.67	25.08	36.52	36.52
<i>Dexamine spinosa</i>	3909.67	106.89	9.91	14.43	50.95
<i>Tubificoides</i> sp.	4604.44	1697.89	9.18	13.36	64.31
<i>Perioculodes aequimanus</i>	1212.78	8.22	3.32	4.83	69.14
<i>Arcuatula senhousia</i>	883.89	4.11	2.57	3.74	72.88