Deep-sea benthic megafaunal habitat suitability modelling: A global-scale maximum entropy model for xenophyophores

Oliver S. Ashford*[a], Andrew J. Davies[b], Daniel O. B. Jones[a]

a. National Oceanography Centre, European Way, Southampton, SO14 3ZH, UK

b. School of Ocean Sciences, Bangor University, Menai Bridge, LL59 5AB, UK.

* Corresponding author:

Present address: Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford, OX1 3PS, UK.

Email: oliver.ashford@zoo.ox.ac.uk

Telephone: (+44) 7763 018136

Abstract

The Xenophyophorea is a group of exclusively deep-sea agglutinating rhizarian protozoans, at least some of which are foraminifera. They are an important constituent of the deep-sea megafauna that are sometimes found in sufficient abundance to act as a significant source of habitat structure for meiofaunal and macrofaunal organisms. This study utilised maximum entropy modelling (Maxent) and a high-resolution environmental database to explore the environmental factors controlling the presence of Xenophyophorea and two frequently sampled xenophyophore species that are taxonomically stable: *Syringammina fragilissima* and *Stannophyllum zonarium*. These factors were

also used to predict the global distribution of each taxon. Areas of high habitat suitability for xenophyophores were highlighted throughout the world's oceans, including in a large number of areas yet to be sampled, but the Northeast and Southeast Atlantic Ocean, Gulf of Mexico and Caribbean Sea, the Red Sea and deep-water regions of the Malay Archipelago represented particular hotspots. The two species investigated showed more specific habitat requirements when compared to the model encompassing all xenophyophore records, perhaps in part due to the smaller number and relatively more clustered nature of the presence records available for modelling at present. The environmental variables depth, oxygen parameters, nitrate concentration, carbon-chemistry parameters and temperature were of greatest importance in determining xenophyophore distributions, but, somewhat surprisingly, hydrodynamic parameters were consistently shown to have low importance, possibly due to the paucity of well-resolved global hydrodynamic datasets. The results of this study (and others of a similar type) have the potential to guide further sample collection, environmental policy, and spatial planning of marine protected areas and industrial activities that impact the seafloor, particularly those that overlap with these conspicuously large single-celled eukaryotes.

Keywords: Maxent; species distribution modelling; Xenophyophorea; *Syringammina fragilissima*; *Stannophyllum zonarium*.

1. Introduction

Xenophyophores (Schulze, 1904) represent some of the most remarkable megafauna in the deep-sea. These giant rhizarian protozoans build agglutinated tests that, in some cases, reach diameters of over 20 cm (Tendal, 1972; Levin and Thomas, 1988; Gooday *et al.* 2011), and were first described in the late 19th century. Initially they were interpreted as a type of primitive foraminifera (Brady, 1883) or alternatively as a group of horny sponges living in symbiosis with hydroids (Haeckel, 1889). It was not until the early 20th century that xenophyophores were recognised and named as a well defined group at a high taxonomic level within rhizopod protozoans (Schulze, 1904, 1907; Tendal, 1972; Pawlowski *et al.* 2003). Even with this taxonomic recognition, xenophyophores remained in relative obscurity for much of the 20th century, and only after the publication of a landmark monograph in 1972 (Tendal, 1972) did the group became widely known

amongst marine biologists in general (Gooday *et al.* 1993; Riemann *et al.* 1993; Pawlowski *et al.* 2003; Hughes and Gooday, 2004; Gooday *et al.* 2011).

Xenophyophores are a large, conspicuous component of the benthic megafauna found in all major ocean basins (Tendal, 1972; Levin, 1991; Levin and Gooday, 1992; Tendal, 1996) and can be enumerated in deep-water photographs (Kamenskaya *et al.* 2013) due to their often visually distinctive agglutinated tests. These tests enclose a branching system of organic tubes containing the cell body (the granellare system) together with often voluminous masses and strings of waste material (stercomata) enclosed within an organic membrane (Tendal, 1972; Gooday *et al.* 2011). However, observations of living specimens are limited, and so many aspects of xenophyophore biology, reproduction and life cycle remains obscure (Pawlowski *et al.* 2003).

Two major xenophyophore lineages are recognised based on morphological criteria: the Psamminida (4 families, 14 genera and over 50 described species), most of which have rigid tests, and the Stannomida (1 family, 2 genera and ~17 described species), which have more flaccid tests, in general, ramified by proteinaceous fibres (linellae) (Tendal, 1972; Tendal, 1996; Gooday and Tendal, 2002; Bisby *et al.* 2010). Opinions about the phylogenetic position of xenophyophores have developed over time. Following initial attempts at classifying xenophyophores (Brady, 1883; Haeckel, 1889) (see above), Schultze (1907) concluded that they represent a distinct group of rhizopod protozoans, an opinion followed by many later workers (Tendal, 1972; Gooday and Tendal, 2002). Recently, however, phylogenetic analysis of small sub-unit ribosomal RNA sequences from *Syringammina corbicula* (Richardson, 2001) (Pawlowski *et al.* 2003), *Aschemonella ramuliformis* (Brady, 1884), *Shinkaiya lindsayi* (Lecroq *et al.* 2009) and *Reticulammina cerebreformis* (Gooday *et al.* 2011) support Brady's (1883) conclusions that xenophyophores are foraminiferans. However, no sequence data yet exists for stannomids, and so it remains to be proven that all xenophyophores are foraminiferans.

Confined to depths greater than about 500 meters, xenophyophores reach peak densities where particle flux to the seafloor is enhanced, such as beneath productive surface waters, in canyons, on areas of raised topography (seamounts or ridges, for instance), or on continental slopes (Tendal, 1972; Tendal and Gooday, 1981; Levin *et al.* 1986; Levin and Thomas, 1988; Levin, 1994; Buhl-Mortensen *et al.* 2010; Gooday *et al.* 2011). Xenophyophores may live infaunally in soft mud, but

most are epifaunal and live on soft sediment or attached to hard substrates (e.g. Tendal and Gooday, 1981; Gooday *et al.* 2011; Kamenskaya *et al.* 2013). They are likely to feed on a diet comprised mainly of detrital particles that are obtained via suspension feeding, surface-deposit feeding or by being trapped within the complex morphology of the test (Tendal, 1972; Lemche *et al.* 1976; Levin and Thomas, 1988; Gooday *et al.* 1993). It has further been suggested that xenophyophores are able to prey on small metazoans (Levin and Gooday, 1992; Smith and Demopoulos, 2003), and may 'farm' microbes as secondary food sources (Tendal, 1979; Laureillard *et al.* 2004; Hori *et al.* 2013), although there is no direct evidence for either of these feeding modes (A. Gooday, personal communication).

Xenophyophores sometimes play a significant role in biological processes that occur at the sediment-water interface (Tendal, 1972; Levin and Thomas, 1988; Levin and Gooday, 1992) and large morphologically complex species of genera such as *Reticulammina* (Tendal, 1972) and *Syringammina* (Brady, 1883) can be considered as autogenic ecosystem engineers. For example, xenophyophore tests provide a focus for organic carbon deposition, serving as traps of organic-rich sedimenting particles and add physical heterogeneity to seafloor mineralisation processes (Levin and Thomas, 1988; Levin and Gooday, 1992). Xenophyophore tests represent important habitat-forming structures on the seafloor, contributing significantly to deep-sea biological heterogeneity (Levin and Thomas, 1988; Levin, 1991; Levin and Gooday, 1992; Smith and Demopoulos, 2003; Buhl-Mortensen *et al.* 2010; Hori *et al.* 2013). As a result, large complex tests appear to constitute faunal hotspots in the deep sea (Levin, 1991; Levin, 1994; Hughes and Gooday, 2004), with enhanced faunal densities and species richness (particularly of crustaceans, molluscs, echinoderms, foraminifera and bacteria) in their close vicinity (Levin *et al.* 1986; Hori *et al.* 2013).

As such, some xenophyophore species may represent an effective umbrella taxon, and knowledge of their distributions has the potential to be used as a guide, in addition to further information on the distribution of vulnerable marine ecosystems, as to which regions could be designated as marine protected areas (MPAs). Certain areas with abundant xenophyophores may be important for deep-sea biodiversity and so should be considered for protection from human practices that disturb the seafloor, like deep-sea trawling, oil and gas extraction and mining.

The great utility of global habitat suitability modelling is in the determination of the potential distributions of taxa that cannot be easily mapped using traditional methods. This is commonly the case in deep-water marine environments, which are difficult to sample due to barriers of cost and isolation. Species distribution models are informative in the context of general scientific investigations (e.g. in targeting regions for further research), but additionally, in cases where the investigated group is of conservation concern (being vulnerable to anthropogenic disturbance and/or important in the functioning of ecosystems) these models can be instructive in directing the designation of protected areas (Davies and Guinotte, 2011; Yesson *et al*. 2012). Maximum entropy modelling (Maxent) (Phillips *et al.* 2006) is a machine-learning habitat suitability modelling method that produces a niche model by minimising the relative entropy between two probability densities; one estimated from the input presence data and the other from the environmental parameters of the landscape in question (Tittensor *et al.* 2010; Elith *et al.* 2011). Maxent has been shown to be one of the highest performing (i.e. most accurate) habitat suitability modelling techniques available (Elith *et al.* 2006; Ortega-Huerta and Peterson, 2008; Wisz, *et al.* 2008). However, its accuracy can drop substantially if a suitable number and variety of presence records are not available to guide the model, and/or if the environmental data available are not reliable or do not fully encompass the range of environmental factors experienced by the focus taxon. There are known issues associated with species distribution models based on small numbers of presence records (Feely and Silman, 2011), including over-prediction, resulting in false positives (Anderson and Gonzalez, 2011), and false negatives. In terms of environmental variables, obtaining a high-resolution global dataset currently requires up-scaling from lower-resolution data, and this inevitably introduces some error, which grows as the difference between native and required resolution increases (Davies and Guinotte, 2011). On balance however, distribution modelling at a global scale remains an instructive and informative technique.

This manuscript uses predictive habitat modelling to assess the global probability of occurrence of xenophyophores as a whole taxon and to explore the potential global distribution of the two most commonly recorded xenophyophore species that are also considered taxonomically stable: *Syringammina fragilissima* (Brady, 1883) and *Stannophyllum zonarium* (Haeckel, 1889). Xenophyophore distributions are modelled using a 30 arc-second environmental database, which allows for global modelling at relatively fine spatial scales.

2. Methods

2.1 Xenophyophore presence data

A total of 837 independent presence records representing 68 xenophyophore species were obtained from peer-reviewed journals, cruise reports, the Global Biodiversity Information Facility (GBIF) and Ocean Biogeographic Information System (OBIS) (Table 1; Supplementary Table 1). Prior to analysis, this dataset was revised so that only a single record was retained within each 30-arc second cell since multiple presence localities within a single 30-arc second cell can cause habitat suitability values to be weighted in favour of the environmental conditions that exist in that cell (Davies and Guinotte, 2011). As a result, 569 presence records were retained for the distribution modelling of Xenophyophorea, 40 for *Syringammina fragilissima,* and 31 for *Stannophyllum zonarium* (Table 1).

Producing a robust species-level distribution model requires a sufficient number of presence records to represent the full observed niche of the species in question as well as confidence in the taxonomic stability and consistency of identification of the chosen species (i.e. that all available records represent a single species, rather than a suite of morphologically similar species). *Aschemonella ramuliformis*, *Aschemonella scabra* (Brady, 1879) and *Reticulammina labyrinthica* (Tendal, 1972) fit this first requirement, with a large number of geo-referenced samples available relative to other xenophyophore species (Table 1). However, there is considerable doubt concerning the taxonomic status of these *Aschemonella* (Brady, 1879) species, and they are probably morphotypes that encompass several similar species rather than representing discrete and consistently identified species (A. Gooday, personal communication). This concern also extends to *Reticulammina labyrinthica*, since it is unlikely that this name has been applied consistently in the literature (A. Gooday, personal communication). As a result, only two species remained for which numerically sufficient taxonomically reliable geo-referenced records were available: *Syringammina fragilissima* and *Stannophyllum zonarium.* Hence, only these species were subjected to distribution modelling.

2.2 Environmental data

In total, 25 environmental layers were produced for use in the Maxent (maximum entropy; Phillips *et al.* 2006) models (Table 2, and see Figure S4 for a correlation matrix of these layers). These were chosen based both on their ecological relevance and their availability at a global scale, and can be split into six broad categories: bathymetric variables (layers derived from a bathymetric grid), carbonate chemistry variables (measures of calcite saturation state), chemical variables (general chemical parameters including salinity, alkalinity and dissolved inorganic carbon amongst others), hydrodynamic variables (current flow), oxygen variables (combination of variables relating to oxygen availability and utilisation), and temperature variables (after Yesson *et al.* 2012). Productivity variables were not available for use owing to a rapid decline in data quality at latitudes greater than ~70°N and S. Since all environmental grids must be of the same latitudinal extent for use in Maxent, and about 22% of xenophyophore presence records were from over 70°N, the decision was made to abandon the use of productivity variables in order to maximise the number of presence records used in the models. However, apparent oxygen utilisation (AOU – the difference in dissolved oxygen concentration of a body of water and its equilibrium oxygen saturation concentration under the same physical and chemical parameters – relating to the use of oxygen due to organismal respiration (Garcia *et al.* 2006a)) was available as a variable, and can be considered a proxy for respiration, which in turn correlates with rates of particulate organic carbon (POC) reaching the benthos (Pfannkuche, 1993).

Terrain attributes were extracted from bathymetric data (SRTM30 – a topographical layer produced from a combination of data from the U.S. 'Shuttle Radar Topography Mission' and the U.S. Geological Survey's 'Global 30 Arc-Second Elevation Data Set') following techniques and algorithms described in Wilson *et al.* (2007). Individual approaches are detailed in footnotes within Table 2; a brief description of each variable is given here. Topographic position index (TPI) is an approach to determine topographical features based on their relative position within a neighbourhood, and can be calculated over fine or broad scales to capture smaller or larger terrain features respectively. This calculation has been developed into a GDAL tool (Geospatial Data Abstraction Library) and the approach is described in Wilson *et al.* (2007). Slope was calculated using DEM Tools for ArcGIS developed by Jenness (2012), in particular the 4-cell method of calculating slope, which is accepted as the most accurate approach (Jones, 1998). Here, slope is defined as the gradient in the direction of the maximum slope. Curvature attempts to describe general terrain features and may

provide an indication of how water interacts with the terrain. Plan and tangential curvature describe how water converges or diverges as it flows over relief, whilst profile curvature describes how water accelerates or decelerates as it flows over relief (Jenness, 2012). Aspect is defined as the direction of maximum slope and was converted to continuous radians following Wilson *et al.* (2007). Rugosity, terrain ruggedness index and roughness all describe the variability of the relief of the seafloor (Wilson *et al.* 2007). Rugosity is defined as the ratio of the surface area to the planar area across a neighbourhood of a central pixel (Jenness, 2012) while terrain ruggedness index is defined as the mean difference between a central pixel and its surrounding cells and roughness as the largest inter-cell difference of a central pixel and its surrounding cell (Wilson *et al.* 2007). Roughness is calculated as the difference in value between the minimum and maximum bathymetry within a neighborhood (Wilson *et al.* 2007).

All other variables were created using the up-scaling approach presented within Davies and Guinotte (2011). All data were available in a gridded form partitioned into standardised depth bins ('z-layers') with a depth range of ~0-5500 m. These z-layers facilitated the determination of approximate benthic habitat conditions (of greatest interest owing to the benthic nature of xenophyophores) on a global scale. This was achieved by the projection of each z-layer to its corresponding area of seafloor using the up-scaling approach (Davies and Guinotte, 2011). This process involved three steps: 1) each z-layer was initially interpolated using inverse-distance weighting to a slightly higher spatial resolution (usually 0.1°) in order to minimise potential gaps that could appear between adjacent z-layers due to non-overlap following projection on to the bathymetric layer. 2) These layers were re-sampled to match SRTM30 (Becker *et al.* 2009) resolution (the highest resolution global bathymetric dataset available) and so preserve as high a spatial resolution as possible. 3) Each re-sampled z-layer was draped over the SRTM30 bathymetric layer to provide an indication of conditions near the seabed. Due to the limitations of the global datasets currently available, it had to be assumed that conditions below the deepest z-layer available were stable to the seabed. However, this approach has been demonstrated to work well over global and regional scales (Davies and Guinotte, 2011; Guinotte and Davies, 2012).

Since the incorporation of too many variables into a habitat suitability model can cause over-fitting of the model (Beaumont *et al.* 2005), a small subset of the available environmental layers were

selected for use in the final analyses. Owing to the co-variant nature of many of the layers in each of the six variable categories (Figure S4), a single variable from each category was selected to represent the influence of that category (following the method of Yesson *et al.* 2012). Variable selection for each category was based on the predictive power of models based on single environmental layers. This was measured using the Test AUC statistic (area under the receiver operating characteristic – ROC – curve; Fielding and Bell, 1997). The AUC statistic can be defined as the probability that a presence site is ranked above a random background site. Values vary from 0 (model performance worse than random) to 0.5 (model performance indistinguishable from random) to 1 (model is maximally predictive) (Fielding and Bell, 1997). Test AUC is quoted since it is more reliable than training AUC scores (Warren and Seifert, 2011). Thus the single variable that produced the greatest test AUC value in isolation for each biological category was selected to represent that category in the final analyses, and hence the final analyses utilised six environmental layers (Table 3 and Table 4).

2.3 Maximum entropy predictions

Maxent (Phillips *et al.* 2006) version 3.3.3k was used to perform the global distribution prediction analyses. Default model parameters were used (convergence threshold of 10^{-5}, regularisation parameter of 1 and a maximum iterations value of 500, with 10000 points randomly selected as background data to construct the model and each model run setting aside 30% of presence records for model evaluation) since these settings have been shown to produce reliable results (Phillips and Dudik, 2008; Davies and Guinotte, 2011; Yesson *et al.* 2012). Higher regularisation parameter values were trialled for the more taxonomically inclusive model of Xenophyophorea to produce smoother response curves (Figure 2). However, this resulted in the production of over-generalised models, lower test AUC, increased differences between training and test AUC, and a less discriminatory model output.

Model performance was evaluated by considering entropy, test AUC, test gain, and test omission scores (see Phillips *et al.* 2006). The importance of each environmental variable was assessed using a jack-knifing procedure (by comparing the gain achieved by variables in isolation – jack-knife of regularised training gain). Response curves were produced to visualise how xenophyophore habitat suitability varied with each environmental factor analysed.

253

254 # 3. Results

255 ## 3.1 Sample locations

256 Xenophyophore sampling to date is patchily distributed throughout the world's oceans (Figure 1).
257 Areas where the highest numbers of xenophyophores have been collected include the North Atlantic
258 (the Porcupine Abyssal Plain, Rockall Bank, Monaco Basin, Cape Verde Plateau and along the Mid-
259 Atlantic Ridge), the Gulf of Mexico, the South Atlantic (especially around the Rio Grande Rise and
260 Mid-Atlantic, Atlantic-Indian and Walvis Ridges) and Atlantic portion of the Southern Ocean, the
261 Arctic Ocean (Baffin Basin, Barents Sea, Nansen Basin, Amundsen Basin and Makarov Basin in
262 particular), patchily in the Indian Ocean (Somali Basin and off the coast of South Africa in
263 particular) the South China Sea, the Northwest Pacific Basin, the Peru Basin, and around New
264 Zealand. In contrast, xenophyophores have been only sparsely collected from the majority of the
265 Indian and Southern Oceans, the western Arctic Ocean, and the South Pacific Ocean.

266 The global distribution of xenophyophore samples (Figure 1) cannot be directly interpreted in terms
267 of overall sampling effort, but the patterns described above suggest that deep-water investigations
268 are concentrated close to nations with more established sampling programmes, as well as hinting of
269 potential bias against more remote locations (the Southern Ocean, for instance).

270 ## 3.2 Variable selection

271 Test AUC scores for the models based on a single variable varied greatly – from a minimum value of
272 0.393 (topographic position index, *Syringammina fragilissima*) to a maximum of 0.987 (depth,
273 *Syringammina fragilissima*) (Table 3). Considering the six variable groupings (Table 2), depth
274 performed best of all bathymetric variables across the taxa, whilst the curvature variables and
275 topographic position index consistently produced some of the lowest AUC scores. Calcite saturation
276 state returned high AUC scores for all taxa analysed, as did all chemical variables analysed (with
277 nitrate, phosphate and silicate in particular performing consistently well), oxygen variables (with
278 apparent oxygen utilisation, in particular, scoring highly) and temperature. AUC values for
279 hydrodynamic variables were generally low, but regional flow rate consistently outperformed

vertical flow rate (Table 3). The highest scoring variables in each variable group for each taxon were chosen for use in the multivariate Maxent models (see Table 3 and Table 4).

3.3 Multivariate model evaluation

Test AUC scores for the multivariate Maxent models were high (Table 4), ranging from 0.836 (Xenophyophorea) to 0.997 (*Syringammina fragilissima*). Test gain values ranged from 0.841 (Xenophyophorea) to 4.580 (*Syringammina fragilissima*), while entropy values ranged from 8.632 (Xenophyophorea) to 4.512 (*Syringammina fragilissima*). Test omission scores were low, ranging from 0.200 (Xenophyophorea) to 0.000 (*Syringammina fragilissima* and *Stannophyllum zonarium*) (based on the maximum sensitivity plus specificity of the test dataset). These low omission scores indicate that few known presences were wrongly classified as absences by the models, and that the predicted presences were significantly more probable than that of random background pixels (Table 4).

3.4 Taxa niches

Jack-knife assessment of model regularised training gain was used to determine which three variables were most important in the production of each of the multivariate Maxent models (Table 4). Combining this with information present in Figure 2, the main environmental conditions for peak habitat suitability (defined as a logistic habitat suitability of ≥0.5) - i.e. the niche - for each taxon were estimated. For *Syringammina fragilissima* these were a depth of between ~835 and 1180 m, a calcite saturation state of between ~2.56 and 3.36, and a temperature of between ~5.3 and 7.7 °C. For *Stannophyllum zonarium*, these were a nitrate concentration greater than 37.5 μmol l^{-1}, apparent oxygen utilisation values of between 4.50 and 6.32 mol O_2 m^{-3}, and temperatures of between 1.6 and 4.7 °C.

It is more complex to estimate conditions of peak habitat suitability for Xenophyophorea since the model encompasses the varied habitat requirements of multiple species (including those described above) and hence produced variable responses that had multiple peaks (Figure 2). Considering this, high habitat suitability for the taxon occurred at nitrate concentrations of ~12.5 to 29.2 μmol l^{-1} and above 38.0 μmol l^{-1}, oxygen saturations of between 6.6 and 42.6 $\%O_2^S$, between 69.2 and 74.3

307 $\%O_2^S$ and between 82.1 and 89.7 $\%O_2^S$, and temperatures ranging from ~-0.8 to -0.6 °C and ~2.4 to
308 8.7 °C.

309 For Xenophyophorea and *Syringammina fragilissima*, temperature was the variable that both
310 reduced the training gain by the greatest amount when omitted from the multivariate Maxent
311 model and produced the highest gain when used in isolation. Hence, this variable contained the
312 most useful information that was not present in the other variables used to construct the models
313 and the most useful information when used in isolation (Table 4). For *Stannophyllum zonarium*,
314 apparent oxygen utilisation contained the most information that was not present in the other
315 variables and the most useful information when used in isolation (Table 4).

316 At the other end of the spectrum, regional flow rate consistently contributed very little to the
317 Maxent multivariate models, whilst depth and calcite saturation state contributed relatively little
318 to the niche model of *Stannophyllum zonarium* (Table 4).

319 **3.5 Areas of maximal habitat suitability**

320 For Xenophyophorea (see Figure 3 and Figure S1), areas of peak habitat suitability were centred on
321 a range of bathymetric features, including continental slopes, subduction trenches, semi-enclosed
322 seas, ridges, seamounts and plateaus. In the Atlantic, xenophyophore habitat suitability was high
323 along all continental slopes, around the Rio Grande Rise, along the Walvis, Reykjanes and Mid-
324 Atlantic Ridges, in the Gulf of Guinea, on the Cape Verde Plateau and plain, in the Angola,
325 Porcupine and Biscay abyssal plains, the most westerly extent of the Mediterranean Sea, around
326 Rockall Bank and the Icelandic Plateau, along the Davies Strait and in Baffin Bay, around the
327 Flemish Cap, in the Gulf of Mexico, and in deep water areas off Florida and in the Caribbean Sea.
328 Habitat suitability was essentially zero on all continental shelves, in all but the very western extent
329 of the Mediterranean and Sargasso seas, on the Sohm and Hatteras Plains, and in the, Sierra Leone,
330 Guinea, Brazil, Argentine and Cape Verde Basins.

331 Habitat suitability was moderate in the Arctic Ocean (between 0.3 – 0.8 logistic suitability), and
332 hotspots were centred upon the continental slopes, Voring Plateau, Greenland Sea, Denmark Strait,
333 Baffin Bay, and Lomonosov Ridge (Figure 3).

The Southern and Indian oceans exhibited only isolated areas of high habitat suitability for xenophyophores, relative to the Atlantic Ocean. These included points along continental slopes (save for the Antarctic continental slope), the South Tasman Rise and the Exmouth Plateau, along Broken Ridge, scattered points along Ninetyeast Ridge, parts of the Agulhas, Madagascar and Mozambique plateaus, regions of the Carlsberg ridge, along the Chagos-Laccadive ridge, the Mascarene Plateau, and regions of high suitability in the north of the Bay of Bengal, the Lakshadweep Sea, Gulf of Aden and deepest parts of the Red Sea. Further south, the South Sandwich Trench is also notable for relatively high habitat suitability (Figure 3).

The Malay Archipelago exhibited very high habitat suitability for xenophyophores in general. Particularly suitable areas included the Andaman Sea (particularly Dreadnought Bank) and the South China Sea, Sulu Sea, Celebes Sea and Banda Sea. In the southwest Pacific, the Bismarck Sea, Ontong Java Rise and regions of the Coral Sea showed areas of relatively high habitat suitability. Regions of suitable habitat were also found around New Zealand - particularly on the Challenger Plateau and Chatham Rise, and along the Kermadec and Tonga trenches and associated ridges (Figure 3).

In the Pacific proper, high habitat suitability was generally centred along subduction trenches, continental slopes and numerous seamounts. For example, along the Mariana, Ryukyu, Izu-Ogasawara, Japan and Kuril-Kamchatka trenches and associated ridges to the west, the Mid-Pacific Mountains, Emperor Seamount chain, and around the Hawaiian ridge and Islands. High habitat suitability was also highlighted along the Cocos and Carnegie Ridges, along the length of the Peru-Chile Trench, in the deep water off the Californian coast, the northern-most extent of the Bering Sea, in the deeper regions of the Sea of Okhotsk and to the west of the Ryukyu Islands (Figure 3).

The model for *Syringammina fragilissima* produced the smallest area of suitable habitat of the taxa investigated (Table 4, Figure 4, Figure S2), being restricted to around Rockall Bank, the Hebrides Terrace and Anton Dohrn Seamounts, Rosemary Bank, along the Wyville Thomson Ridge, points on the continental slope along the west of the United Kingdom, the Iceland-Faeroe Rise, the continental slope around Iceland and the Reykjanes Ridge, along the Mid-Atlantic Ridge close to the Azores, around the northernmost extent of the Labrador Sea, north of the Bahamas, along the shallowest regions of the Madagascar Plateau, and points around New Zealand (particularly in areas

of the Campbell Plateau). However, it must be stressed that, due to the limited number of presence records available for this species, and their relatively clustered nature (and so the requirement of extrapolation over relatively large areas in the model – such as over the South Atlantic, North and East Pacific, Arctic Ocean and Indian Ocean), the above does not represent a definitive map of distribution for the species. The addition of further presence records (particularly in regions yet to be sampled) may alter the area of apparent high habitat suitability for *Syringammina fragilissima*.

The model for *Stannophyllum zonarium* (Figure 5 and Figure S3) highlights a broader distribution than for *Syringammina fragilissima*, with areas of maximal habitat suitability centred on the Pacific Ocean rather than the Atlantic Ocean. Areas of high habitat suitability include much of the East and Northeast Pacific (Guatemala Basin and Albatross Plateau in particular), along the northern slope of the Aleutian Trench, around the Hawaiian Islands and Ridge, the Mid-Pacific Mountains, along Sculpin Ridge, in the Aleutian Basin (particularly the northernmost extent), along the Emperor Seamount Chain, on the Hess and Shatsky rises, in the Kuril Basin, along the Mariana, Ryukyu, Izu-Ogasawara, Japan and Kuril-Kamchatka trenches and associated ridges to the west, to the South of Japan, the Ontong Java Rise, Caroline Seamounts, and isolated areas in the Coral Sea. In the Malay Archipelago, areas of high habitat suitability include deep areas of the Andaman, Sulu and South China seas, the Celebes Sea, the Makassar Strait and North Banda Basin, and areas of the Molucca and Flores Sea. In the Indian Ocean, the northernmost extent of the Bay of Bengal, regions of the Arabian Sea and Gulf of Aden, points along the Mascarene and Chagos-Laccadive plateaus and areas of continental slope along the northern shores of the ocean show high habitat suitability for this species (Figure 5). However, as for *Syringammina fragilissima* above, this distribution should not be interpreted as definitive due to the relatively small number of presence records available (and the requirement of extrapolation in the model such as over the Indian, Atlantic and polar oceans). The addition of further presence records may alter the area of apparent high habitat suitability for *Stannophyllum zonarium*.

4. Discussion

4.1 Habitat predictions and applications

391 This exploratory study enhances both our knowledge of xenophyophore distributions and illuminates
392 the controlling physical factors of these distributions. It is generally accepted that xenophyophores
393 reach highest densities in regions of high surface productivity (Tendal, 1972), and in areas where
394 the flux of organic particles is enhanced by topography (Levin and Thomas, 1988; Levin, 1994;
395 Gooday *et al.* 2011), including seamounts, mid-ocean ridges, canyons, subduction trenches,
396 plateaus and continental slopes (Lemche *et al.* 1976; Tendal and Lewis, 1978; Levin and Thomas,
397 1988; Levin, 1994; Gooday *et al.* 2011). These topographic features are associated with localised
398 currents (e.g. Roden, 1987), and thus it is hypothesised that organic particles are concentrated in
399 their vicinity, increasing food availability for xenophyophores. As an alternative, Levin and Thomas
400 (1988) suggest that the localised current regimes around these topographic features result in an
401 increased flux and/or deposition of xenophyophore propogules. Whilst productivity or localised
402 current flow data were not available for use in this analysis (although apparent oxygen utilisation
403 can be thought of as a proxy for productivity, and terrain variables can capture topographical
404 driven flow information), predicted xenophyophore distributions (Figures 3, 4 and 5) were broadly
405 concordant with the accepted views above. High habitat suitability scores were commonly obtained
406 for mid-ocean ridges, continental slopes, plateaus, seamounts and the slopes of subduction
407 trenches. This suggests that these topographic features may be associated with additional
408 environmental characteristics positive to xenophyophore growth. Interestingly, in addition to the
409 topographic features outlined above, this analysis suggests that deep semi-enclosed seas and bays
410 may also be favourable to xenophyophore growth. For example, Baffin Bay, the Gulf of Mexico, the
411 Caribbean Sea, the South China Sea, Andaman Sea, Sulu Sea, Celebes Sea and Banda Sea all exhibit
412 high habitat suitability (Figure 3), although the Mediterranean Sea is an exception to this.

413 Comparison of the global sampling distribution (Figure 1) and the habitat suitability map for
414 xenophyophores (Figure 3) reveals that some xenophyophores have been sampled from areas with
415 relatively low predicted habitat suitability. These include the Mediterranean Sea, the Southern
416 Ocean and along the coast of Antarctica, around the southernmost extent of the Brazil Basin and
417 northernmost extent of the Argentine Basin, to the southeast of Sri Lanka, to the north of
418 Madagascar and parts of the Northwest Pacific and Arctic Ocean. Incorrect identification or spatial
419 referencing errors may explain some of these records. Alternatively, the xenophyophore
420 distribution model may not fully reflect the potential distribution of the group, or these samples

421 may represent collection of xenophyophores in fringe habitats where they naturally occur at low
422 densities.

423 When xenophyophore sampling locations (Figure 1) are compared with the Maxent habitat
424 suitability maps (Figures 3, 4 and 5) it is demonstrated that a significant number of locations with
425 high predicted xenophyophore habitat suitability are yet to be sampled. For Xenophyophorea, these
426 include much of the western Arctic Ocean, the Icelandic Plateau and Reykjanes Ridge, most of
427 Baffin Bay and the Labrador Sea, around the Flemish Cap, much of the Caribbean Sea, the Angola
428 Basin, many locations along the continental slopes of the East and West Atlantic, most of the Mid-
429 Atlantic Ridge, the Madagascar and Mascarene plateaus, the Gulf of Aden, the north of the Bay of
430 Bengal and much of the Indian Ocean continental slopes of Australia, the Andaman Sea and seas
431 around Sulawesi in the Malay Archipelago, around the Ryukyu Islands, the Ontong Java Rise, the
432 northernmost extent of the Bearing Sea, much of the Sea of Okhotsk, and numerous seamounts in
433 the Pacific Ocean which are yet to be investigated. For *Syringammina fragilissima*, such areas are
434 less numerous, but include the Reykjanes Ridge and the Mid-Atlantic Ridge around the Azores, the
435 northernmost extent of the Labrador Sea, and potentially to the north of the Bahamas and on
436 Walters Shoal of the Madagascar Plateau. For *Stannophyllum zonarium,* such areas are numerous
437 and include much of the East and Northeast Pacific, along the Hawaiian and Boudeuse ridges, the
438 deepest parts of the Bering Sea (particularly the northern slopes of the Aleutian Basin), along the
439 Emperor Seamount Chain, along the major trenches of the West Pacific, on the Ontong Java Rise,
440 the Andaman, South China, Sulu, Celebes and Banda Seas, the north of the Bay of Bengal, and the
441 Arabian Sea. These locations represent key targets for future sampling.

442 In regions where there is a perceived threat to benthic environments, the presence of a high
443 density of large, morphologically complex xenophyophore species, in conjunction with other
444 vulnerable marine ecosystems, could be considered as further motivation for the instalment of
445 deep-water MPAs, since many xenophyophore species are important but fragile autogenic
446 ecosystem engineers, playing a significant role in biological processes that occur at the sediment-
447 water interface (Tendal, 1972; Levin and Thomas, 1988; Levin, 1991; Levin and Gooday, 1992;
448 Smith and Demopoulos, 2003; Hughes and Gooday, 2004; Hori *et al.* 2013). For instance, this
449 analysis provides possible direction for the instalment of MPAs in areas of high predicted

xenophyophore habitat suitability that are currently subjected to deep-sea bottom trawling. Such areas include a large number of seamounts, banks, ridges and plateaus across the world's oceans (Figure 3) (e.g. see Koslow *et al.* 2000 and Thrush and Dayton, 2002), and, indeed, some such areas have already been protected based on evidence of the presence of vulnerable marine ecosystems (including xenophyophore aggregations) (e.g. the Darwin Mounds off the NW coast of Scotland (De Santo and Jones, 2007)). This analysis also potentially provides spatial guidance for the protection of areas vulnerable to local negative ecosystem impacts associated with deep-sea oil and gas drilling, such as around the 'Atlantic Frontier' drilling sites near the Faroe Islands and sites in the Gulf of Mexico (Glover and Smith, 2003) (Figure 3). Similar knowledge and data could be applied to help propose MPAs in areas earmarked for deep-sea mining operations, such as areas of the Manus Basin off New Guinea and the Havre Trough off New Zealand (Glover and Smith, 2003), and the Clarion-Clipperton Fracture zone in the Eastern Pacific (where xenophyophores are known to reach quite high abundances) (Kamenskaya *et al.* 2013) (Figures 3 and 5). However it should be stressed that, although the resolution of this analysis is very high for its global scale, it is not adequate for probing the fine-scale distributions of xenophyophores within areas of high apparent habitat suitability, and as such, targeted surveys and distribution modelling of potential MPA locations at local or regional scales should be undertaken before designation to ensure they are based on the highest quality observational data available (Rengstorf *et al.* 2012; Ross and Howell, 2012; Guinotte and Davies, 2012; Rengstorf *et al.* 2013).

4.2 Taxa niches

Depth was one of the most important variables defining habitat suitability for the taxa analysed (Table 4). Moving from sea-level to greater depths, habitat suitability increased to values over 0.5 only in depths greater than about 500 m (Figure 2). This agrees well with the accepted observation that xenophyophores are found in water depths greater than ~500 m (Tendal, 1972; Levin, 1994; Buhl-Mortensen *et al.* 2010). The importance of depth was not unexpected since multiple factors of biological importance also change with depth, including light intensity, pressure, temperature, productivity, salinity, calcium carbonate saturation states, and many more chemical variables. The trough in xenophyophore habitat suitability between about 4800 and 6350m depth was unexpected, however (Figure 2). It is possible that this is caused by unfavourable environmental conditions for

xenophyophores at these depths, such as nutrient-depletion. Alternatively this trough could reflect the lack of environmental data available at depths of >5500 m from many global data products (i.e. World Ocean Atlas). What is most likely, however, is that this habitat suitability trough represents an artefact of poor sampling effort at these depths.

Nitrate concentration was found to be an important environmental parameter for Xenophyophorea as a whole and for *Stannophyllum zonarium*. Peak xenophyophore habitat suitability was found to occur in waters with relatively high nitrate concentrations, and at particularly high nitrate concentrations for *Stannophyllum zonarium* in particular (maximal habitat suitability at concentrations greater than 37.5 μmol l^{-1}) (Figure 2). This finding agrees well with the notion of xenophyophores being most common in relatively nutrient-enriched waters (Tendal, 1972; Levin and Thomas, 1988; Levin, 1994; Gooday *et al.* 2011).

The importance of calcite saturation state as a habitat characteristic relevant to xenophyophore distributions (Table 4) has, to the authors' knowledge, never been explicitly stated, although it has been observed that xenophyophores exhibit a 'preference' for sand-sized particles in test construction (Levin and Thomas, 1988; Levin, 1994), and that foraminiferal tests are a common sand-sized component in many xenophyophore tests (A. Gooday, personal communication). It seems that for multiple (but certainly not all) xenophyophore species, a calcite saturation state >1 is associated with test production from recycled calcareous foraminifera. This would appear to be the case for *Syringammina fragilissima*, for instance, with this commonly foraminiferal-tested (Tendal, 1972) species occurring well above the carbonate compensation depth and experiencing peak habitat suitability in waters with a calcite saturation state of between ~2.56 and 3.36 (Figure 2).

Oxygen variables were important in model construction for Xenophyophorea and for *Stannophyllum zonarium* in this study. Per cent oxygen saturation was an important variable for Xenophyophorea (Table 4), with the taxon exhibiting peaks of habitat suitability at a range of saturations – from 6.6 up to 89.7 $\%O_2^S$ (Figure 2). Such a broad range of suitable oxygen saturations demonstrates a high level of variability in oxygen requirements and tolerance amongst species in this taxon, although most xenophyophores have been sampled from relatively well-oxygenated regions (A. Gooday, personal communication). Apparent oxygen utilisation was the most important variable in the construction of the Maxent model for *Stannophyllum zonarium*. Interestingly, this species reaches

peak densities at apparent oxygen utilisation values of between 4.50 and 6.32 mol O_2m^{-3} (Figure 2). Such high values link well with the high nitrate concentration preferences of *Stannophyllum zonarium*, suggesting that this species is often sampled from productive nutrient enriched regions with particularly high associated biological activity.

Temperature was of consistent importance to all taxa investigated, and was the single most important variable in the construction of the Maxent models of Xenophyophorea and *Syringammina fragilissima* (Table 4). The two species investigated in detail exhibited discrete temperature windows of peak habitat suitability (between ~5.3 and 7.7 °C for *Syringammina fragilissima*, compared to between ~1.6 and 4.7 °C for *Stannophyllum zonarium*) (Figure 2). The relationship between xenophyophore occurrence and temperature has not been discussed in detail, although there is some mention by Tendal (1972). Here Tendal argues that, as xenophyophores are to be considered members of a distinct cold-water fauna, it is to be expected that they are restricted in their upper vertical distributions by increasing temperature with decreasing depth, although he does not expand as to why this should be. More generally, such relationships are well documented for marine invertebrates (e.g. see Orton, 1920 and Barras *et al.* 2009), with temperature being an important factor controlling growth rate and various aspects of reproductive physiology, mediated by its influence on biochemical reactions (Brown *et al.* 2004).

Finally, it is interesting to note the low AUC values (Table 3) and low jack-knife training gains (Table 4) that were obtained for the hydrodynamic variables for all taxa investigated (regional flow for *Syringammina fragilissima* being a potential exception). This was surprising since the importance of water flow for xenophyophores as suspension feeders has been stressed by many authors (Tendal, 1972; Tendal and Lewis, 1978; Levin and Thomas, 1988; Levin, 1994). However, hydrodynamic variables also performed badly in a recent study of cold-water coral distributions (known suspension feeders) (Yesson *et al.* 2012) which used a similar environmental dataset. Thus this poor performance is likely to represent a scale issue - with these global scale layers not accurately portraying local scale variations in current velocity associated with small topographic features (Yesson *et al.* 2012). Higher resolution data is required to shed further light on the importance of current flow for the distribution of megafaunal suspension-feeders (Mohn *et al.* 2014).

537 **4.3 Model evaluation and limitations**

538 Model performance was good for all taxa, with high test AUC and gain scores, and low test omission
539 values. Test AUC and gain values were higher for the species investigated compared to
540 xenophyophores as a whole, and were higher for *Syringammina fragilissima* than for *Stannophyllum*
541 *zonarium*. This was probably a result of the greater level of clustering of *Syringammina fragilissima*
542 sample locations relative to *Stannophyllum zonarium.* The models for *Syringammina fragilissima*
543 and *Stannophyllum zonarium* had a less variable dataset to fit than for Xenophyophorea, with
544 smaller total variance in the environmental parameters at their sampling localities (smaller entropy
545 values – see Table 4) as a result of the smaller number of presence records used in the models.
546 Thus the Maxent model could be fitted more tightly around the presence data.

547 There are known issues associated with species distribution models produced using small numbers
548 of presence records (Feely and Silman, 2011), these chiefly being over-prediction, resulting in false
549 positives (Anderson and Gonzalez, 2011), and false negatives. The use of presence records that are
550 distributed across a large longitudinal and latitudinal range (as for the model for Xenophyophorea)
551 should lower the risk of over-prediction, and in general, the models appear to have performed well.
552 However, there is some evidence of small areas of false positives. For example, in Figure 3
553 (Xenophyophorea), relatively shallow areas (<500 m depth) of the Norwegian trough are highlighted
554 as potentially suitable habitat (0.4-0.6 logistic habitat suitability), while in Figure 5 (*Stannophyllum*
555 *zonarium*), small areas of the Shelikof Strait are highlighted as suitable habitat (0.7-0.9 logistic
556 habitat suitability) in water depths of around 200m. Considering our current knowledge of
557 xenophyophore bathymetric distributions (see above and Figure 2), these predictions almost
558 certainly represent false positives, although only ground-truthing can confirm this. False negatives
559 are harder to pinpoint in the Maxent predictions, but almost certainly occur to some extent in the
560 models for *S. fragilissima* and *S. zonarium* presented here (considering the level of extrapolation
561 across ocean basins from a relatively small number of presence records). The addition of further
562 presence records (particularly in regions yet to be sampled) will help to better define the
563 distributions of these two species and highlight any false negatives present in the current models.

564 Comparison of the habitat suitability predictions obtained in the present study with those of the
565 only other Maxent model yet produced for a xenophyophore species (Ross and Howell, 2012)

represents a further way in which model performance can be evaluated. In general, the two models show a high level of similarity. High habitat suitability (>0.6) for *Syringammina fragilissima* in the NE Atlantic is demonstrated in both models along the continental slope off Ireland and the United Kingdom, around the Hebrides Terrace and Anton Dohrn seamounts, around Rosemary Bank, and along the slopes of Rockall and Hatton banks. There are some areas where the two models disagree, however. For instance, the model of Ross and Howell predicts higher *S. fragilissima* habitat suitability around the slopes of Edoras and Fangorn banks and along the slopes of the Porcupine Seabight and Goban Spur relative to the model presented in this paper. In addition, the model presented in this paper predicts larger areas of high habitat suitability, relative to Ross and Howell (2012), for *S. fragilissima* to the south of the Wyville-Thomson Ridge, between Bill Bailey's Bank and Rosemary Bank, and in the Hatton-Rockall Basin. The overall similarity of the two models, however, gives further confidence to their predictions, especially so considering that they are produced from different data sets – Ross and Howell choosing only to use topographic data.

Choice of modelling resolution is an important factor when producing predictive species distribution models (Guisan *et al.* 2007). While higher resolution outputs are preferable when we need to capture environmental variability at small spatial scales (like the rapid changes in temperature that occur with distance across the Faroe-Shetland Channel (Oey, 1997)) and for visualising predictions, they do carry certain associated error and limitations (Davies *et al.* 2008; Davies and Guinotte, 2011). Since, other than depth, global environmental layers are not available at 30 arc-second resolution, variables have to be up-scaled from their native resolution to that required (30 arc-seconds in this analysis – see Methods). Up-scaling inevitably introduces some error, with this error growing as the difference between native and required resolution increases (Davies and Guinotte, 2011). Examples of the manifestation of this error include the generalisation and smoothing of variables and the failure to incorporate small scale variability in the up-scaled layers that is not present in the lower resolution source data (Davies and Guinotte, 2011; Rengstorf *et al.* 2012). Further, the majority of global layers currently available that can be up-scaled represent annual means of values (in order to ensure a high number of samples to maximise certainty in the variables (Davies and Guinotte, 2011)). As a result, these layers do not capture any component of annual variability, a particular drawback when modelling highly seasonal high latitude regions. However, comparison of up-scaled data with GLODAP (Global Ocean Data Analysis Project) test bottle water

data by Davies and Guinotte (2011) found the two datasets to be highly correlated, and hence the authors concluded that any issues associated with the up-scaling method are outweighed by its benefits.

Whilst the dataset utilised in this study comprised a high number and diversity of variables, there were still variables which may have been informative but were not available for use. Chief amongst these were productivity variables, such as measures of particulate organic carbon reaching the seafloor, and surface water chlorophyll a concentrations. These variables were not available for use owing to a rapid decline in data quality at latitudes greater than ~70° (see section 2.2). Substratum type is a further variable that would have been interesting to incorporate into this analysis as there is evidence for sediment-type preference in xenophyophores (Levin and Thomas, 1988). Unfortunately, a global environmental layer containing details of sediment type is not yet available, but progress is being made towards this goal (e.g. Shumchenia and King, 2010). Furthermore, the hydrodynamic variables used in this study under-performed and were not of sufficient sensitivity to capture local scale variation in flow rates associated with isolated topographic features such as seamounts. Thus, considering the current uncertainty surrounding xenophyophore feeding methods, it would be particularly interesting to incorporate a high resolution local current flow into future analyses. Such a layer is currently unavailable at a global scale, although advances are being made at the regional scale (Mohn *et al.* 2014).

Potential evidence for xenophyophore sampling bias has been mentioned in section 3.1. Firm evidence of sampling bias would imply that the current distribution of presence localities used in this study is not adequate to represent all potential environments from which xenophyophores can be sampled. This would potentially lead to false negatives in the Maxent outputs. Whether this is the case will only become apparent following further sampling and analyses.

The most conclusive way to validate or refute the predictions of this analysis (Figures 3-5) would be to directly test them in the field via 'ground-truthing' (Guinotte and Davies, 2012). Do we find xenophyophores in areas of predicted high habitat suitability that have not yet been sampled, like the Andaman Sea, or do these predictions represent false positives? There are some issues with this method. Assuming that a cruise to undertake this task could be funded, it would be a huge undertaking to systematically search an entire 30 arc-second cell of high predicted habitat

suitability using ROVs or camera equipment, and subsampling may miss specimens as xenophyophore distributions may be patchy within this cell. However, it should be noted that xenophyophores have been recorded at very high densities in areas of suitable habitat (Tendal and Gooday, 1981), increasing the likelihood of discovery.

Finally, although species distribution modelling has progressed rapidly in recent years in terms of resolution (compare the results of Davies *et al.* (2008) with those of Davies and Guinotte (2011) and Yesson *et al.* (2012)) and model performance criteria (e.g. Warren and Seifert, 2011), clearly the availability of additional relevant environmental variables with global coverage at high resolution, and a growing number of reliable presence localities will continue to lead to ever an increasing accuracy of models suitable for a number of research and industrial applications.

4.4 Conclusive remarks

This study represents the first of its kind for xenophyophores at a global scale and serves to improve knowledge of their distributions and further illuminate details of their ecology. Additionally, this analysis draws attention to the possible use of these fragile and remarkable deep-sea megafaunal ecosystem engineers in enhancing MPA planning and designation. However, this work represents but a first step and aims to motivate continued research into these intriguing and important organisms; testing model predictions with further sampling, performing local-scale high-resolution analyses and addressing some of the still unanswered questions concerning xenophyophore ecology and physiology.

5. Acknowledgements

The authors would like to thank Prof. Andrew Gooday for his insightful comments and suggestions which have much improved this manuscript, Dr. Ole Tendal for advice concerning the availability of records of xenophyophore sampling locations, especially for *Stannophyllum zonarium*, and three anonymous reviewers for their detailed comments. Dr. Daniel Jones was funded for this work by the

651 UK National Environment Research Council as part of the Marine Environmental Mapping Programme
652 (MAREMAP).

653 6. References

654 Anderson, R.P., and Gonzalez, I. Jr. 2011. Species-specific tuning increases robustness to sampling
655 bias in models of species distributions: An implementation with Maxent. *Ecological Modelling* 222:
656 2796-2811.

657 Barras, C., Geslin, E., Duplessy, J-C., and Jorissen, F.J. 2009. Reproduction and growth of the
658 deep-sea benthic foraminifer *Bulimina marginata* under different laboratory conditions. *Journal of*
659 *Foraminiferal Research* 39: 155-165.

660 Beaumont, L.J., Hughes, L., and Poulsen, M. 2005. Predicting species distributions: use of climatic
661 parameters in BIOCLIM and its impact on predictions of species' current and future distributions.
662 *Ecological Modelling* 186: 251-270.

663 Becker, J.J., Sandwell, D.T., Smith, W.H.F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor,
664 J., Ingalls, S., Kim, S-H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., Von
665 Rosenberg, J., Wallace, G., and Weatherall, P. 2009. Global bathymetry and elevation data at 30
666 arc seconds resolution: SRTM30_PLUS. *Marine Geodesy* 32: 355–371.

667 Bisby, F.A., Roskov, Y.R., Orrell, T.M., Nicolson, D., Paglinawan, L.E., Bailly, N., Kirk, P.M.,
668 Bourgoin, T., and Baillargeon, G. (Eds.). 2010. Species 2000 & ITIS Catalogue of Life: 2010 Annual
669 Checklist Taxonomic Classification. DVD; Species 2000: Reading, UK.

670 Biodiversity occurrence data published by: PANGAEA - Publishing Network for Geoscientific and
671 Environmental Data; Danish Biodiversity Information Facility; Museum of Comparative Zoology,
672 Harvard University; Yale University Peabody Museum; (Accessed through GBIF Data Portal,
673 data.gbif.org, 2013-01-11).

674 Boyer, T.P., Levitus, S., Garcia, H.E., Locamini, R.A., Stephens, C., and Antonov, J.I. 2005.
675 Objective analyses of annual, seasonal, and monthly temperature and salinity for the World Ocean
676 on a 0.25° grid. *International Journal of Climatology* 25:931-945.

677 Brady, H.B. 1879. Notes on some of the Reticularian Rhizopoda of the "Challenger" Expedition. Part
678 1. On new or little known arenaceous types. *Quarterly Journal of Microscopical Science* 19: 20-63.

679 Brady, H.B. 1883. *Syringammina*, a new type of arenaceous Rhizopoda. *Proceedings of the Royal*
680 *Society* 35: 55–161.

681 Brady, H.B. 1884. Report on the Foraminifera. *Report on the scientific results of the voyage of H.*
682 *M. S. Challenger* 9: 1-814.

683 Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M., and West, G.B. 2004. Toward a metabolic
684 theory of ecology. *Ecology* 85(7): 1771-1789.

685 Buhl-Mortensen, L., Vanreusel, A., Gooday, A.J., Levin, L.A., Priede, I.G., Buhl-Mortensen, P.,
686 Gheerardyn, H., King, N.J., and Raes, M. 2010. Biological structures as a source of habitat
687 heterogeneity and biodiversity on the deep ocean margins. *Marine Ecology* 31: 21-50.

688 Carton, J.A., Giese, B.S., and Grodsky, S.A. 2005. Sea level rise and the warming of the oceans in
689 the SODA ocean reanalysis. *Journal of Geophysical Research* 110: C09006.

690 Davies, A.J., and Guinotte, J.M. 2011. Global habitat suitability for framework-forming cold-water
691 corals. *PLoS ONE* 6: e18483.

692 Davies, A.J., Wisshak, M., Orr, J.C., and Roberts, J.M. 2008. Predicting suitable habitat for the
693 cold-water reef framework-forming coral *Lophelia pertusa* (Scleractinia). *Deep Sea Research Part I:*
694 *Oceanographic Research Papers* 55: 1048–1062.

695 De Santo, E.M., and Jones, P.J.S. 2007. The Darwin Mounds: from undiscovered coral to the
696 development of an offshore marine protected area regime. *Bulletin of Marine Science* 81: 147-156.

697 Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J.,
698 Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G.,
699 Moritz, C., Nakamura, M., Nakazawa, Y., McC. Overton, J., Peterson, A.T., Phillips, S.J.,
700 Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., and
701 Zimmermann, N.E. 2006. Novel methods improve prediction of species' distributions from
702 occurrence data. *Ecography* 29: 129-151.

703 Elith, J., Phillips, S.J., Hastie, T., Dudik, M., En Chee, Y., and Yates, C.J. 2011. A statistical
704 explanation of MaxEnt for ecologists. *Diversity and Distributions* 17: 43-57.

705 Feely, K.J., and Silman, M.R. 2011. Keep collecting: accurate species distribution modelling
706 requires more collections than previously thought. *Diversity and Distributions* 17: 1132-1140.

707 Fielding, A.H., and Bell, J.F. 1997. A review of methods for the assessment of prediction errors in
708 conservation presence/absence models. *Environmental Conservation* 24: 38-49.

709 Garcia, H.E., Locarnini, R.A., Boyer, T.P., and Antonov, J.I. 2006a. World Ocean Atlas 2005,
710 Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. Levitus, S. (Ed.).
711 NOAA Atlas NESDIS 63, U.S. Government Printing Office, Washington, D.C., 342 pp.

712 Garcia, H.E., Locarnini, R.A., Boyer, T.P., and Antonov, J.I. 2006b. World Ocean Atlas 2005,
713 Volume 4: Nutrients (phosphate, nitrate, silicate). Levitus, S. (Ed.). NOAA Atlas NESDIS 64, U.S.
714 Government Printing Office, Washington, D.C., 396 pp.

715 Glover, A.G., and Smith, C.R. 2003. The deep-sea floor ecosystem: current status and prospects of
716 anthropogenic change by the year 2025. *Environmental Conservation* 30: 219-241.

717 Gooday, A.J., Aranda da Silva, A., and Pawlowski, J. 2011. Xenophyophores (Rhizaria, Foraminifera)
718 from the Nazare Canyon (Portuguese margin, NE Atlantic). *Deep-Sea Research II* 58: 2401-2419.

719 Gooday, A.J., Bett, B.J., and Pratt, D.N. 1993. Direct observation of episodic growth in an abyssal
720 xenophyophore (Protista). *Deep-Sea Research I* 40: 2131-2143.

721 Gooday, A.J., and Tendal, O.S. 2002. Class Xenophyophorea. In: Lee, J.J., Huttner, J., and Bovee,
722 E.C. (Eds). 2002. An illustrated guide to the Protozoa, 2nd ed. *Society of Protozoologists* and Allen
723 Press, Lawrence, Kansas. Pp. 1086-1097.

724 Guinotte, J.M., and Davies, A.J. 2012. Predicted deep-sea coral habitat suitability for the U.S. West
725 Coast. *Report to NOAA-NMFS*. 85 pp.

726 Guisan, A., Graham, C.H., Elith, J., and Huettmann, F. 2007. Sensitivity of predictive species
727 distribution models to change in grain size. *Diversity and Distributions* 13: 332–340.

728 Haeckel, E. 1889. Report on the deep-sea keratosa. Report on the scientific results of the voyage of
729 H. M. S. Challenger during the years 1873–76. *Zoology* 32**:** 1–92.

730 Hori, S., Tsuchiya, M., Nishi, S., Arai, W., Yoshida, T., and Takami, H. 2013. Active bacterial flora
731 surrounding Foraminifera (Xenophyophorea) living on the deep-sea floor. *Bioscience, Biotechnology*
732 *and Biochemistry* 77: 381-384.

733 Hughes, J.A., and Gooday, A.J. 2004. Associations between living benthic foraminifera and dead
734 tests of *Syringammina fragilissima* (Xenophyophorea) in the Darwin Mounds region (NE Atlantic).
735 *Deep-Sea Research I* 51: 1741-1758.

736 Jenness, J. 2012. DEM Surface Tools. Jenness Enterprises.

737 Jones, K.H. 1998. A comparison of algorithms used to compute hill slope as a property of the DEM.
738 *Computers and Geosciences* 24:315-323.

739 Kamenskaya, O.E., Melnik, V.F., and Gooday, A.J. 2013. Giant protists (xenophyophores and
740 komokiaceans) from the Clarion-Clipperton ferromanganese nodule field (Eastern Pacific). Biology
741 Bulletin Reviews 3: 388-398.

742 Koslow, J.A., Boehlert, G.W., Gordon, J.D.M., Haedrich, R.L., Lorance, P., and Parin, N. 2000.
743 Continental slope and deep-sea fisheries: implications for a fragile ecosystem. *Journal of Marine*
744 *Science* 57: 548-557.

745 Laureillard, J., Mejanelle, L., and Sibuet, M. 2004. Use of lipids to study the trophic ecology of
746 deep-sea xenophyophores. *Marine Ecology Progress Series* 270: 129-140.

747 Lecroq, B., Gooday, A.J., Tsuchiya, M., and Pawlowski, J. 2009. A new genus of xenophyophores
748 (Foraminifera) from Japan Trench: morphological description, molecular phylogeny and elemental
749 analysis. *Zoological Journal of the Linnaean Society* 156: 455-464.

750 Lemche, H., Hansen, B., Madsen, F.J., Tendal, O.S., and Wolff, T. 1976. Hadal life as analysed
751 from photographs. *Vidensk Meddr Dansk Naturh Foren* 139: 263-336.

752 Levin, L.A. 1991. Interactions between metazoans and large, agglutinating protozoans: implications
753 for the community structure of deep-sea benthos. *American Zoologist* 31: 886-900.

754 Levin, L.A. 1994. Paleoecology and ecology of xenophyophores. *Palaios* 9: 32-41.

755 Levin, L.A., and Gooday, A.J. 1992. Possible roles for xenophyophores in deep-sea carbon cycling.
756 In: Rowe, G.T., and Pariente, V. (Eds.). 1992. Deep-sea food chains and the global carbon cycle.
757 Springer Netherlands. pp. 93-104.

758 Levin, L.A., and Thomas, C.L. 1988. The ecology of xenophyophores (Protista) on eastern Pacific
759 seamounts. *Deep Sea Research* 35: 2003-2027.

760 Levin, L.A., DeMaster, D.J., McCann, L.D., and Thomas, C.L. 1986. Effects of giant protozoans
761 (class: Xenophyophorea) on deep-seamount benthos. *Marine Ecology – Progress Series* 29: 99-104.

762 Mohn, C., Rengstirf, A., White, M., Duineveld, G., Mienis, F., Soetaert, K., and Grehan, A. 2014.
763 Linking benthic hydrodynamics and cold-water coral occurrences: high-resolution model study at
764 three cold-water coral provinces in the NE Atlantic. *Progress in Oceanography* 122: 92-104

765 Oey, L.-Y. 1997. Eddy energetic in the Faroe-Shetland channel: a model resolution study.
766 *Continental Shelf Research* 17: 1929-1944.

767 Ortega-Huerta, M.A., and Peterson, A.T. 2008. Modelling ecological niches and predicting
768 geographic distributions: a test of six presence-only methods. *Revista Mexicana de Biodiversidad*
769 79: 205-216.

770 Orton, J.H. 1920. Sea-temperature, breeding and distribution in marine animals. *Journal of the*
771 *Marine Biological Association of the United Kingdom* 12: 339-366.

772 Pawlowski, J., Holzmann, M., Fahrni, J., and Richardson, S.L. 2003. Small subunit ribosomal DNA
773 suggests that the xenophyophorean *Syringammina corbicula* is a Foraminiferan. *Journal of*
774 *Eukaryote Microbiology* 50: 483-487.

775 Pfannkuche, O. 1993. Benthic response to the sedimentation of particulate organic matter at the
776 BIOTRANS station, 47°N, 20°W. *Deep-Sea Research II* 40: 135-149.

777 Phillips, S.J., Anderson, R.P., and Schapire, R.E. 2006. Maximum entropy modelling of species
778 geographic distributions. *Ecological Modelling* 190: 231-259.

779 Phillips, S.J., and Dudik, M. 2008. Modelling of species distributions with Maxent: new extensions
780 and a comprehensive evaluation. *Ecography* 31: 161-175.

781 Rengstorf, A.M., Grehan, A., Yesson, C., and Brown, C. 2012. Towards high-resolution habitat
782 suitability modelling of vulnerable marine ecosystems in the deep-sea: resolving terrain attribute
783 dependencies. *Marine Geodesy* 35: 343-361.

784 Rengstorf, A.M., Yesson, C., Brown, C., and Grehan, A.J. 2013. High-resolution habitat suitability
785 modelling can improve conservation of vulnerable marine ecosystems in the deep sea. *Journal of*
786 *Biogeography* 40: 1702-1714.

787 Richardson, S.L. 2001. *Syringammina corbicula* sp. nov. (Xenophyophorea) from the Cape Verde
788 Plateau, E. Atlantic. *Journal of Foraminiferal Research* 31: 201-209.

789 Riemann, F., Tendal, O.S., and Gingele, F.X. 1993. *Reticulammina Antarctica* nov. spec.
790 (Xenophyophora, Protista) from the Weddell Sea, and aspects of the nutrition of xenophyophores.
791 *Polar Biology* 13: 543-547.

792 Roden, G.I. 1987. Effects of seamounts and seamount chains on ocean circulation and thermohaline
793 structure. In: Keating, B. Fryer, P., Batizar, R., and Boehlert, G. (Eds.). 1987. Seamounts, islands,
794 and atolls. Geophysical Monograph No. 43, *American Geophysical Union*. pp. 335-354.

795 Ross, R.E., and Howell, K.L. 2012. Use of predictive habitat modelling to assess the distribution and
796 extent of the current protection of 'listed' deep-sea habitats. *Diversity and Distributions* 19: 433-
797 445.

798 Schulze, F.E. 1904. Über den Bau und die Entwicklung gewisser Tiefsee-
799 organismen. *Sitzungsberichte der Königlich-preussischen Akademie der Wissenschaften zu*
800 *Berlin* 53: 1387.

801 Schulze, F.E. 1907. Die Xenophyophoren: Eine besondere Gruppe der Rhizopoden. Wissenschaftliche
802 Ergebnisse der Deutschen Tiefsee-Expedition auf dem Dampfer "Valdivia" 1898-1899. G. Fischer.

803 Shumchenia, E.J., and King, J.W. 2010. Comparison of methods for integrating biological and
804 physical data for marine habitat mapping and classification. *Continental Shelf Research* 30: 1717-
805 1729.

806 Smith, C.R., and Demopoulos, A.W.J. 2003. The deep Pacific ocean floor. In: Tyler, P.A. (Ed.) 2003.
807 Ecosystems of the deep oceans. *Ecosystems of the world* 28: 179-218. Elsevier: Amsterdam. ISBN.

808 Tendal, O.S. 1972. A monograph of the Xenophyophoria (Rhizopodea, Protozoa). *Doctoral*
809 *dissertation, Danish Science Press.*

810 Tendal, O.S. 1979. Aspects of the biology of Komokiacea and Xenophyophoria. *Sarsia* 64: 13-17.

811 Tendal, O.S. 1996. Synoptic checklist and bibliography of the Xenophyophorea (Protista), with a
812 zoogeographical survey of the group. *Galathea Report* 17 (1995-1996): 79-101.

813 Tendal, O.S., and Gooday, A.J. 1981. Xenophyophoria (Rhizopoda, Protozoa) in bottom photographs
814 from the bathyal and abyssal NE Atlantic. *Oceanologica Acta* 4: 415-422.

815 Tendal, O.S., and Lewis, K.B. 1978. New Zealand xenophyophores: Upper bathyal distribution,
816 photographs of growth position, and a new species. *New Zealand Journal of Marine and Freshwater*
817 *Research* 12: 197-203.

818 Thrush, S.F., and Dayton, P.K. 2002. Disturbance to marine benthic habitats by trawling and
819 dredging: implications for marine biodiversity. *Annual Review of Ecology and Systematics* 33: 449-
820 473.

821 Tittensor, D.P., Baco, A.R., Hall-Spencer, J.M., Orr, J.C., and Rogers, A.D. 2010. Seamounts as
822 refugia from ocean acidification for cold-water stony corals. *Marine Ecology* 31 (Suppl. 1): 212-225.

823 Warren, D.L., and Seifert, S.N. 2011. Ecological niche modelling in Maxent: the importance of
824 model complexity and the performance of model selection criteria. *Ecological Applications* 21: 335-
825 342.

826 Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., and Grehan, A.J. 2007. Multiscale terrain
827 analysis of multibeam bathymetry data for habitat mapping on the Continental Slope. *Marine*
828 *Geodesy* 30:3-35.

829 Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., and NCEAS Predicting
830 Species Distributions Working Group. 2008. Effects of sample size on the performance of species
831 distribution models. *Diversity and Distributions* 14: 763-773.

832 Yesson, C., Taylor, M.L., Tittensor, D.P., Davies, A.J., Guinotte, J., Baco, A., Black, J., Hall-
833 Spencer, J.M., and Rogers, A.D. 2012. Global habitat suitability of cold-water octocorals. *Journal of*
834 *Biogeography* 39: 1278-1292.

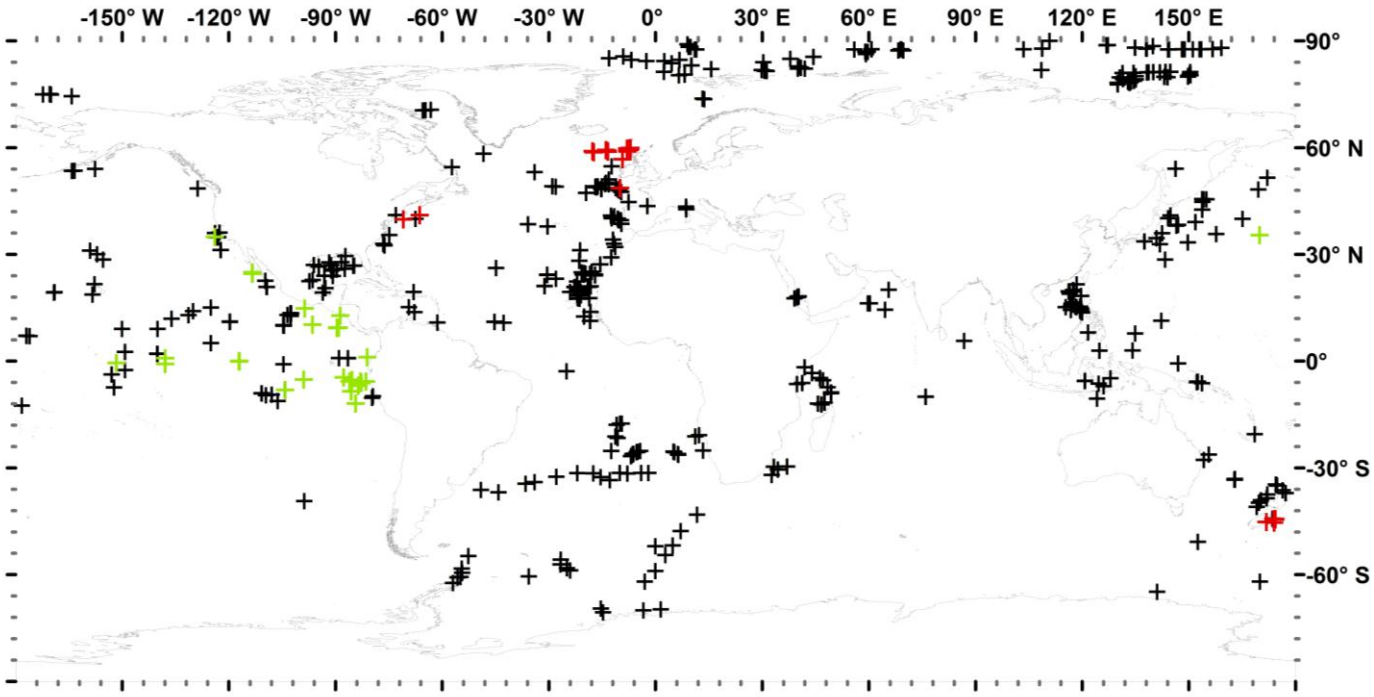

Figure 1: Global sampling locations for xenophyophores. Taxa are colour-coded: *Syringammina fragilissima* – red; *Stannophyllum zonarium* – green; remaining Xenophyophorea – black.

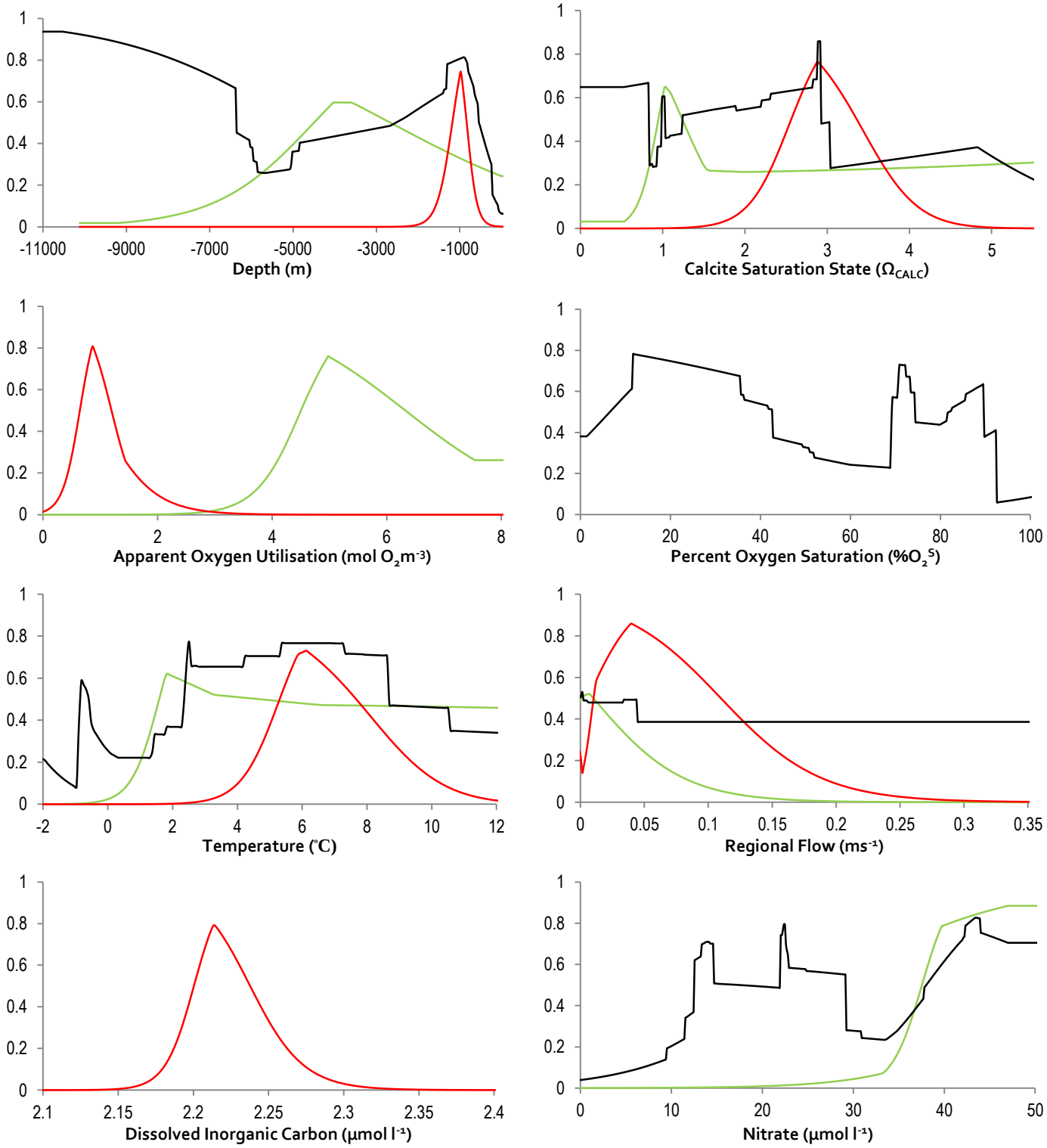

Figure 2: Variable response curves for global Maxent habitat suitability models of xenophyophore taxa. Note Y axis is habitat suitability – 0 (min) to 1 (max) in all cases. Taxa are colour-coded as in Figure 1: Xenophyophorea – black; *Syringammina fragilissima* – red; *Stannophyllum zonarium* – green.

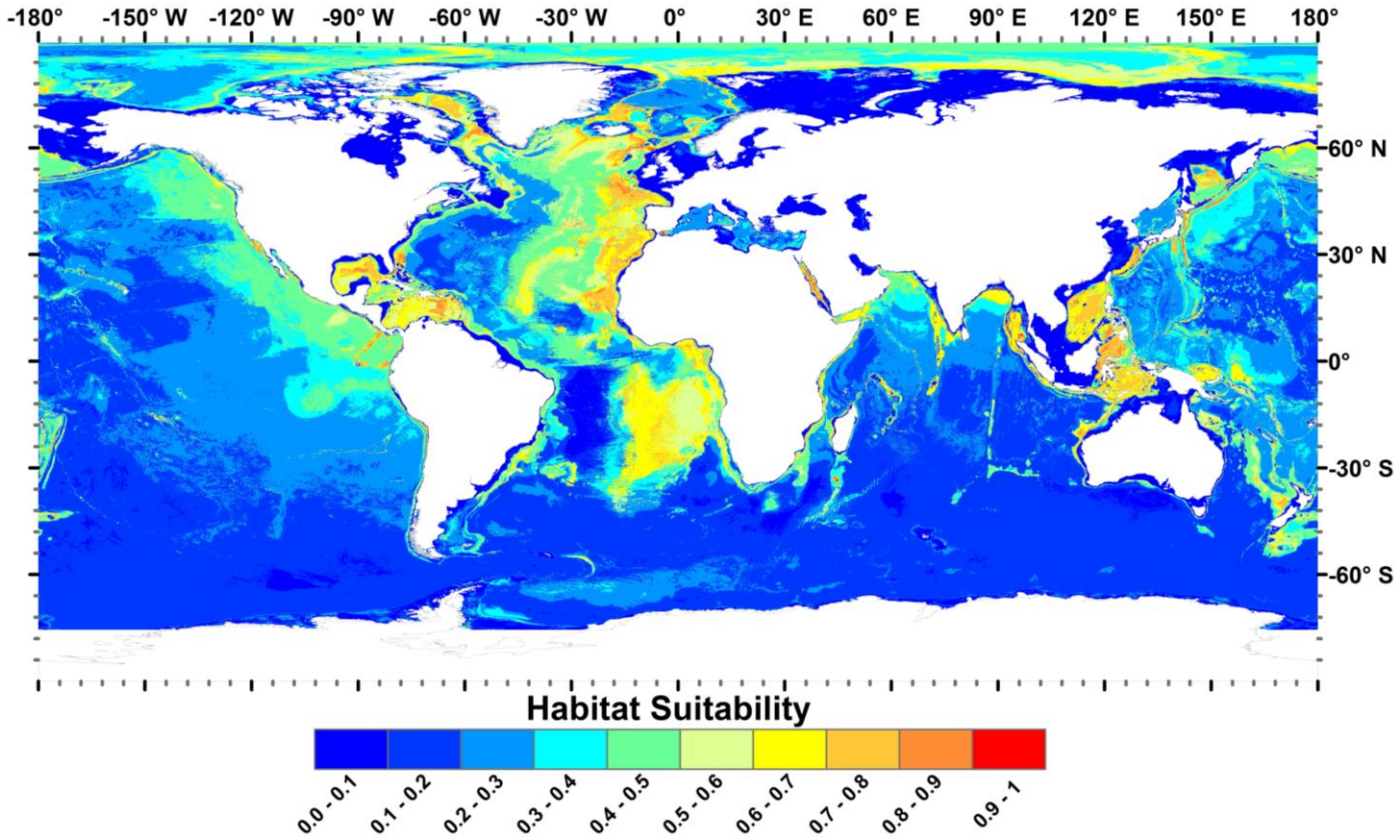

Figure 3: Global habitat suitability for Xenophyophorea at 30 arc-second resolution. Based on Maxent output (logistic). Habitat suitability values of 0 illustrate minimally suitable environmental conditions in an area. Habitat suitability values of 1 illustrate maximally suitable environmental conditions in an area.

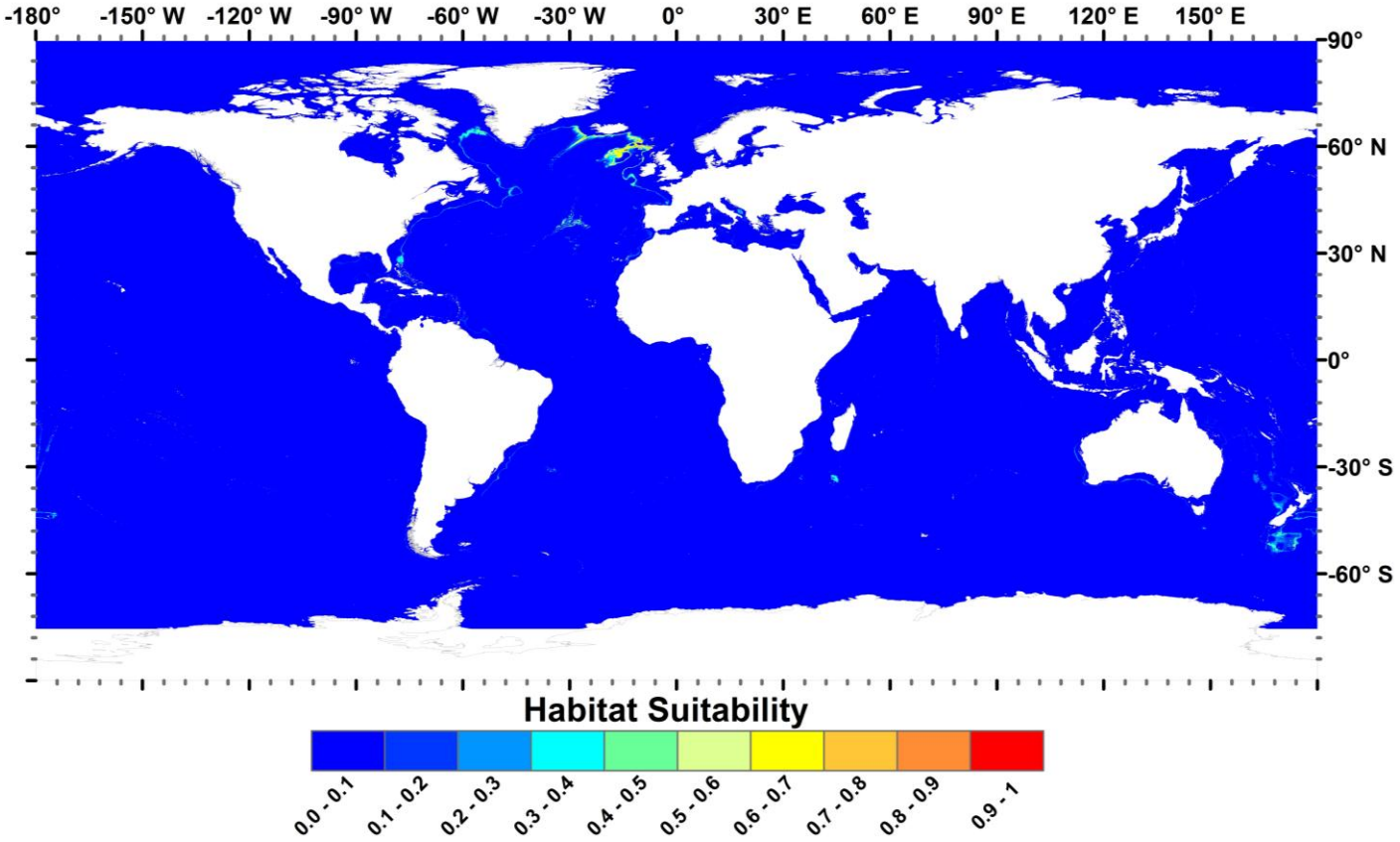

Figure 4: Global habitat suitability for the xenophyophore species *Syringammina fragilissima* at 30 arc-second resolution. Based on Maxent output (logistic). Habitat suitability values of 0 illustrate minimally suitable environmental conditions in an area. Habitat suitability values of 1 illustrate maximally suitable environmental conditions in an area.

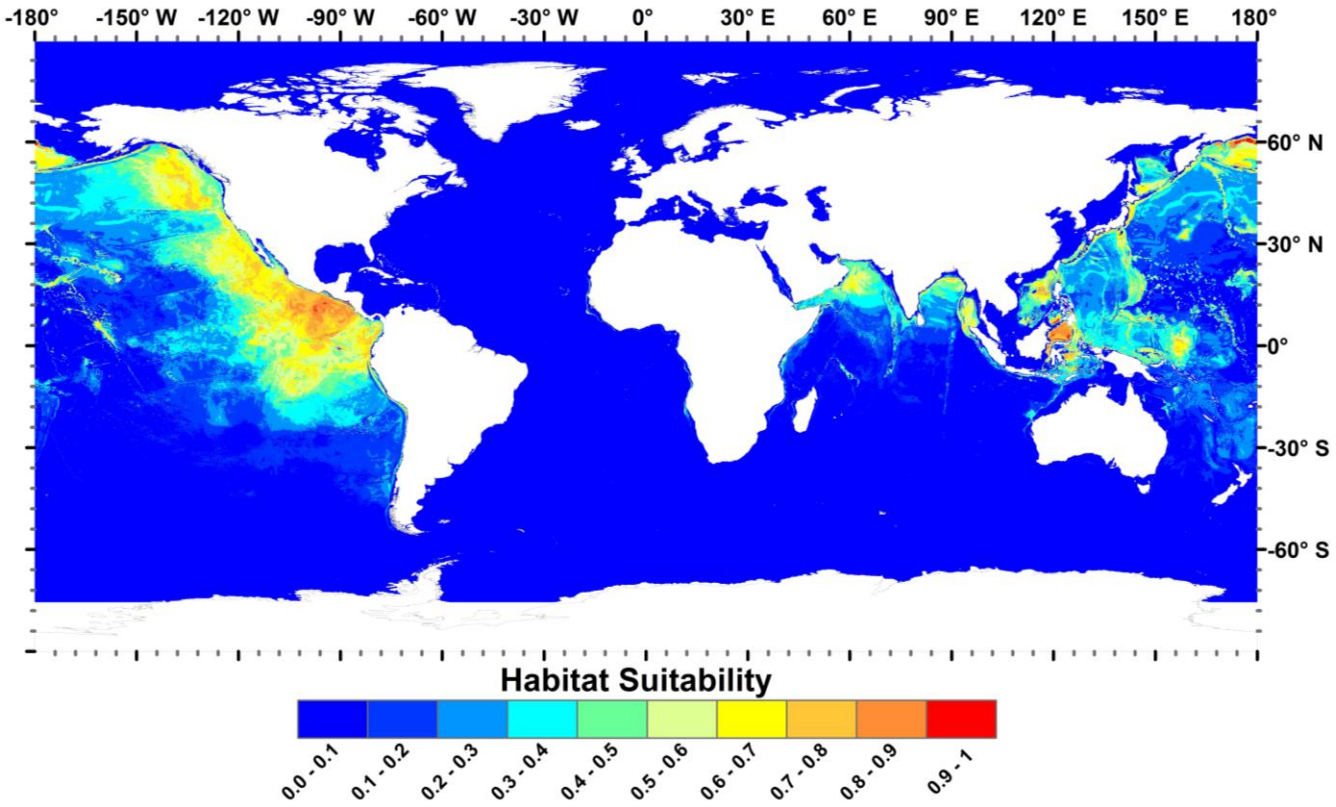

Figure 5: Global habitat suitability for the xenophyophore species *Stannophyllum zonarium* at 30 arc-second resolution. Based on Maxent output (logistic). Habitat suitability values of 0 illustrate minimally suitable environmental conditions in an area. Habitat suitability values of 1 illustrate maximally suitable environmental conditions in an area.

Table 1: Number of geo-referenced records available per xenophyophore taxon. Only one record per 30 arc-second cell was retained for Maxent analyses. Number of species for which geo-referenced records were available = 68.

Genus	Species	No. records	No. records retained in analysis
Aschemonella	*carpathica*	5	
	catenata	4	
	composita	12	
	grandis	9	
	ramuliformis	85	
	scabra	80	
	Unknown	58	
Cerelasma	*gyrosphaera*	2	
	lamellosa	1	
	massa	7	
	Unknown	1	
Cerelpemma	*radiolarium*	4	
Galatheammina	*calcarea*	9	
	discoveryi	4	
	erecta	13	
	lamina	1	
	microconcha	3	
	tetraedra	4	
	Unknown	7	
Holopsamma	*argillaceum*	1	
	cretaceum	1	
Homogammina	*lamina*	13	
	maculosa	20	
	Unknown	3	
Maudammina	*arenaria*	1	
Nazareammina	*tenera*	2	
Occultammina	*profunda*	1	
	Unknown	1	
Psammetta	*arenocentrum*	1	
	erythrocytomorpha	3	
	globosa	6	
	Unknown	4	
Psammina	*delicate*	6	
	fusca	1	
	globigerina	3	
	nummulina	4	
	plakina	1	
	sabulosa	3	
	zonaria	1	
	Unknown	12	
Psammopemma	*calcareum*	1	
Reticulammina	*antarctica*	1	
	cerebreformis	10	
	cretacea	1	
	labyrinthica	34	
	lamellata	3	
	maini	1	
	novazealandica	3	
	plicata	1	
	Unknown	26	
Semipsammina	Unknown	2	
Shinkaiya	*lindsayi*	1	
Spiculammina	*delicata*	1	
Stannarium	*concretum*	1	
Stannoma	*alatum*	1	
	coralloides	5	
	dendroides	11	
	Unknown	1	
Stannophyllum	*alatum*	4	
	annectens	1	
	concretum	1	
	flustraceum	2	
	fragilis	1	
	globigerinum	19	
	granularium	11	
	indistinctum	3	
	mollum	9	
	pertusum	1	
	radiolarium	3	

	reticulatum	2	
	setosum	1	
	venosum	1	
	zonarium	31	31
	Unknown	3	
Syringammina	*corbicula*	4	
	fragilissima	49	40
	minuta	1	
	reticulata	2	
	tasmanensis	7	
	Unknown	10	
Unknown		170	
Total		837	569

Table 2: Summary of geophysical and environmental variables used in this study. All variables are stored in an ArcGIS file geo-database. Superscript notes indicate particular analysis or treatment of data.

Variable group	Variable	Units	Reference
Bathymetric Variables[1]			
	Aspect	Degrees	Jenness (2012)
	Aspect - Eastness [2,3]	Degrees	Wilson *et al.* (2007)
	Aspect - Northness [2,4]	Degrees	Wilson *et al.* (2007)
	Curvature - Plan [5,7]		Jenness (2012)
	Curvature - Profile [5,6]		Jenness (2012)
	Curvature - Tangential [5,8]		Jenness (2012)
	Depth	m	Becker *et al.* (2009)
	Roughness [9]		Wilson *et al.* (2007)
	Rugosity [5]		Jenness (2012)
	Slope [5]	Degrees	Jenness (2012)
	Terrain Ruggedness Index [9]		Wilson *et al.* (2007)
	Topographic Position Index [9]		Wilson *et al.* (2007)
Carbonate chemistry variables			
	Calcite saturation state [10,11]	Ω_{CALC}	Steinacher *et al.* (2009)
Chemical variables			
	Alkalinity [10]	µmol l^{-1}	Steinacher *et al.* (2009)
	Dissolved inorganic carbon [10]	µmol l^{-1}	Steinacher *et al.* (2009)
	Nitrate [10]	µmol l^{-1}	Garcia *et al.* (2006b)
	Phosphate [10]	µmol l^{-1}	Garcia et al. (2006b)
	Salinity [10]	pss	Boyer *et al.* (2005)
	Silicate [10]	µmol l^{-1}	Garcia *et al.* (2006b)
Hydrodynamic variables			
	Regional flow [12]	m s^{-1}	Carton *et al.* (2005)
	Vertical flow [12]	m s^{-1}	Carton *et al.* (2005)
Oxygen variables			
	Apparent oxygen utilisation [10]	mol m^{-3}	Garcia *et al.* (2006a)
	Dissolved oxygen concentration [10]	ml l^{-1}	Garcia *et al.* (2006a)
	Percent oxygen saturation [10]	% O_2^S	Garcia *et al.* (2006a)
Temperature variables			
	Temperature [10]	°C	Boyer *et al.* (2005)

[1] Derived from SRTM30 bathymetry.

[2] Calculated in ArcGIS 10.

[3] Modified calculation from Wilson *et al.* (2007) using Sin((Aspect * π) /180, to produce 1 = east and -1 = west orientation.

[4] Modified calculation from Wilson *et al.* (2007) using Cos((Aspect * π) /180, to produce 1 = north and -1 = south orientation.

[5] Calculated using the 4 cell method in Jenness (2012).

[6] Longitudinal curvature in Jenness (2012) and defined as "Longitudinal curvatures are set to positive when the curvature is concave (i.e. when water would decelerate as it flows over this point). Negative values indicate convex curvature where stream flow would accelerate." Zero indicates an undefined value.

[7] Defined in Jenness (2012) as "Plan curvatures are set to positive when the curvature is convex (i.e. when water would diverge as it flows over this point). Negative values indicate concave curvature where stream flow would converge." Zero indicates an undefined value.

[8] Defined in Jenness (2012) as "Tangential curvatures are set to positive when the curvature is convex (i.e. when water would diverge as it flows over this point). Negative values indicate concave curvature where stream flow would converge." Zero indicates an undefined value.

[9] Calculated using GDAL DEM Tool. Values at zero indicate flat areas, higher values indicate rough and variable terrain.

[10] Variable creation process followed the Davies and Guinotte (2011) upscaling approach.

[11] Created using SRES1B scenario data from the years 2000-2010.

[12] SODA data extracted from version 2.0.4, monthly means for the years 1990-2007.

Table 3: Test AUC values for global Maxent habitat suitability models of xenophyophore taxa based on single variables. The highest AUC scores in each variable group are highlighted in bold and underlined for each taxon.

Variable group	Variable	Xenophyophorea	*Syringammina fragilissima*	*Stannophyllum zonarium*
Bathymetric Variables	Aspect	0.535	0.672	0.615
	Eastness of aspect	0.514	0.509	0.507
	Northness of aspect	0.518	0.614	0.529
	Plan curvature	0.562	0.500	0.500
	Profile curvature	0.577	0.500	0.500
	Tangential curvature	0.560	0.500	0.500
	Depth	**0.686**	**0.987**	**0.696**
	Roughness	0.571	0.548	0.666
	Rugosity	0.544	0.583	0.688
	Terrain ruggedness index	0.561	0.574	0.659
	Topographic position index	0.598	0.393	0.408
	Slope	0.554	0.577	0.673
Carbonate chemistry variables	Calcite saturation state	**0.654**	**0.957**	**0.784**
Chemical variables	Alkalinity	0.669	0.961	0.808
	Dissolved inorganic carbon	0.689	**0.977**	0.840
	Nitrate	**0.728**	0.903	**0.913**
	Phosphate	0.703	0.935	0.891
	Salinity	0.714	0.834	0.318
	Silicate	0.717	0.931	0.908
Hydrodynamic variables	Regional flow	**0.500**	**0.810**	**0.633**
	Vertical flow	0.491	0.500	0.500
Oxygen variables	Apparent oxygen utilisation	0.733	**0.910**	**0.907**
	Dissolved oxygen concentration	0.725	0.873	0.895
	Percent oxygen saturation	**0.747**	0.891	0.897
Temperature variables	Temperature	**0.720**	**0.986**	**0.776**

Table 4: Model evaluation statistics for global Maxent habitat suitability models of xenophyophore taxa based on multiple variables. The three most important variables for each taxon (jack-knife of regularised training gain) are highlighted in bold and underlined. ***** indicates the variable that reduced the training gain most when omitted and therefore contained the most information that was not present in other variables. † indicates the variable with the highest training gain when used in isolation and which thus had the most useful information by itself. Thresholds are based on the maximum sensitivity plus specificity of the test dataset.

Statistic	Xenophyophorea	*Syringammina fragilissima*	*Stannophyllum zonarium*
Model evaluation			
Test AUC	0.836	0.997	0.941
Test gain	0.841	4.580	1.768
Entropy	8.632	4.512	7.551
Threshold			
Logistic value	0.379	0.100	0.276
Test omission (%)	0.200	0.000	0.000
Fractional predicted area	0.244	0.010	0.124
Probability	4.83×10^{-64}	7.59×10^{-25}	6.70×10^{-9}
Regularised training gain in isolation			
Depth	0.228	**3.153**	0.109
Calcite saturation state	0.114	**2.115**	0.172
Dissolved inorganic carbon	-	1.848	
Nitrate	**0.331**	-	**1.254**
Regional flow	0.004	0.608	0.016
Apparent oxygen utilisation	-	1.618	**1.450†***
Percent oxygen saturation	**0.286**	-	
Temperature	**0.339†***	**3.241†***	**0.441**