Parasites & Vectors

CrossMark

Environmental and anthropogenic factors affecting the probability of occurrence of *Oncomegas wageneri* (Cestoda: Trypanorhyncha) in the southern Gulf of Mexico

Víctor M. Vidal-Martínez[1*†], Edgar Torres-Irineo[2], David Romero[3], Gerardo Gold-Bouchot[4], Enrique Martínez-Meyer[5], David Valdés-Lozano[6] and M. Leopoldina Aguirre-Macedo[1†]

Abstract

Background: Understanding the environmental and anthropogenic factors influencing the probability of occurrence of the marine parasitic species is fundamental for determining the circumstances under which they can act as bioindicators of environmental impact. The aim of this study was to determine whether physicochemical variables, polyaromatic hydrocarbons or sewage discharge affect the probability of occurrence of the larval cestode *Oncomegas wageneri,* which infects the shoal flounder, *Syacium gunteri*, in the southern Gulf of Mexico.

Methods: The study area included 162 sampling sites in the southern Gulf of Mexico and covered 288,205 km^2, where the benthic sediments, water and the shoal flounder individuals were collected. We used the boosted generalised additive models (boosted GAM) and the MaxEnt to examine the potential statistical relationships between the environmental variables (nutrients, contaminants and physicochemical variables from the water and sediments) and the probability of the occurrence of this parasite. The models were calibrated using all of the sampling sites (full area) with and without parasite occurrences (n = 162) and a polygon area that included sampling sites with a depth of 1500 m or less (n = 134).

Results: *Oncomegas wageneri* occurred at 29/162 sampling sites. The boosted GAM for the full area and the polygon area accurately predicted the probability of the occurrence of *O. wageneri* in the study area. By contrast, poor probabilities of occurrence were obtained with the MaxEnt models for the same areas. The variables with the highest frequencies of appearance in the models (proxies for the explained variability) were the polyaromatic hydrocarbons of high molecular weight (PAHH, 95 %), followed by a combination of nutrients, spatial variables and polyaromatic hydrocarbons of low molecular weight (PAHL, 5 %).

(Continued on next page)

* Correspondence: vvidal@mda.cinvestav.mx
†Equal contributors
1Laboratorio de Parasitología, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Km 6 Carretera Antigua a Progreso, Cordemex, Mérida, Yucatán 97310, México
Full list of author information is available at the end of the article

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 2 of 13

(Continued from previous page)

Conclusions: The contribution of the PAHH to the variability was explained by the fact that these compounds, together with N and P, are carried by rivers that discharge into the ocean, which enhances the growth of hydrocarbonoclastic bacteria and the productivity and number of the intermediate hosts. Our results suggest that sites with PAHL/PAHH ratio values up to 1.89 promote transmission based on the high values of the prevalence of *O. wageneri* in the study area. In contrast, PAHL/PAHH ratio values ≥ 1.90 can be considered harmful for the transmission stages of *O. wageneri* and its hosts (copepods, shrimps and shoal flounders). Overall, the results indicate that the PAHHs affect the probability of occurrence of this helminth parasite in the southern Gulf of Mexico.

Keywords: Cestoda, Trypanorhyncha, Bioindicators, Environmental impact, Flatfish, Contamination, Gulf of Mexico

Background

Understanding the environmental factors influencing the probability of occurrence of parasitic species is fundamental for determining the circumstances under which they can be affected by human activities. This knowledge is necessary because parasites can be useful as bioindicators, i.e., species or communities that are used to assess the quality of the marine environment and how it changes over time [1–4].

Marine biologists traditionally have used free-living, benthic organisms as bioindicators to assess the environmental quality of the oceanic sediments associated with anthropogenic impacts [5–9]. However, there are limitations to the use of benthic, free-living organisms as bioindicators, such as the following: 1) the large number of samples (and funding) required for quantitative sampling; 2) seasonal variation affecting the distribution and abundance of the organisms in addition to the water or the sediment quality; 3) many methods (and indices) are available for the analyses, suggesting a great heterogeneity in the interpretation of the results; and 4) a limited taxonomic knowledge regarding the various groups because of the large number of species in the benthic realm [10–12].

The parasites of aquatic organisms are also affected by anthropogenic and natural environmental influences and have been proposed as alternative bioindicators [2, 4]. There are several advantages to the use of parasites as bioindicators of environmental quality: 1) their communities are far less speciose than those of free-living, benthic organisms; 2) their taxonomy and life cycles are relatively well known; and 3) from a parasite point of view, each individual host is an island (or habitat), and statistically, each host becomes a "sampling unit" with its own set of parasite species [2, 4, 13–18]. At this point, parasites are accepted as an integral part of the environment and suffer the same type of influence from natural and anthropogenic variables on their various life stages (e.g., transmission forms, such as cercariae or coracidia; larval stages, such as metacercariae or cystacanths; and adult stages) [3, 19, 20]. However, even though it has

been possible on a small spatial scale for the freshwater or coastal environments to determine links between the environmental variables affecting the ecological metrics (e.g., species number or relative abundance) of parasites (e.g., [21–23]), the knowledge regarding the environmental factors influencing the probability of occurrence of the parasites at a seascape scale (hectares to thousands of square km^2) is almost lacking. Thus, an important step required to integrate parasites as bioindicators of environmental quality in the oceans at a seascape scale is to determine the environmental or anthropic factors that affect the parasite community or the population metrics.

Ecological niche models (ENMs) are useful for providing insight into the mechanisms by which environmental variables influence the probability of occurrence of both terrestrial and marine organisms [24, 25]. Most of these models are correlative in nature and determine whether there are statistical associations between environmental, biological and/or geographical variables and species abundances or occurrences to establish the sets of conditions under which the species can maintain viable populations [26, 27]. Although ENMs have been profusely used in the terrestrial realm to address a myriad of topics (e.g., [28–30]), these models have a more recent history in the marine environment (e.g., [31, 32]).

During studies in the southern Gulf of Mexico to determine the environmental quality regarding the sediments, water and organisms for the Mexican oil company (PEMEX), we obtained data on the geographical distribution and abundance of the helminth parasites infecting the shoal flounder, *Syacium gunteri*. These studies were developed for PEMEX because inland and offshore oil extraction, together with agriculture, are significant economic activities in the southern Gulf of Mexico, and high concentrations of nutrients from the river runoff, polycyclic aromatic hydrocarbons, and other contaminants, such as pesticides, are released into the environment (e.g., [33, 34]). Thus, to study the potential effects of natural environmental factors (e.g., oxygen concentration and salinity), nutrients and chemical pollutants on the probability of occurrence of parasites, we selected the

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 3 of 13

larval cestode, *Oncomegas wageneri*, because of its high overall prevalence in the study area (84 %). The life cycle of *O. wageneri* is unknown. However, Palm (1995) [35] described that *O. wageneri* occupy a special position within the Eutetrarhynchoidea because its plerocercoid stage lacks a blastocyst. Thus, based on the life cycle of an eutetrarhynchid relative (*Prochristianella hispida*) [36], *O. wageneri* should have copepods and shrimp as first and second intermediate hosts, respectively, while stingrays should be the definitive hosts. In this case, the shoal flounder could act as a paratenic host or as a potential third intermediate host [Dr. Bjoern Schaeffner, Universidade de São Paulo, pers. com.]. Because this parasite has transmission stages and intermediate hosts that are exposed to the environment, its infection parameters should reflect the environmental conditions experienced by all of these organisms. In this sense, the values of the infection parameters, such as the prevalence and mean abundance, can be considered indirect measures of the net colonisation of the fish (sensu [37]). However, even when the fish sampling procedure used was standardised, the number of fish collected was highly uneven among the sampling sites, making it difficult to compare the abundance of *O. wageneri* among the sampling sites. To overcome this problem, we transformed the parasite data from abundance values to presence/absence values. This method was adapted from the field of community ecology, in which estimators of the number of species are frequently biased by the number of units sampled at a specific locality and in which presence/absence descriptors perform well (e.g.,[38, 39]). This method allowed us to use a binomial distribution for all the ENMs to estimate the probability of *O. wageneri* occurrence at the sampling site level based on the potential effect of different environmental and anthropic variables (i.e., pollutants). Thus, the objective of the present paper was to determine whether the probability of the occurrence of *O. wageneri* presented statistical associations with natural physicochemical environmental variables, nutrients and polycyclic aromatic hydrocarbons at the seascape level in the southern Gulf of Mexico.

Methods
Study area and sampling procedures for sediments and water
The study area included 162 sampling stations in the southern Gulf of Mexico (Fig. 1) and covered 288,205 km^2. Bottom sediments were collected between September 6 and October 8, 2005, at water depths between 0.1 and 200 m with a 36 m long shrimp boat using a Smith-McIntyre grab with a weight of 80 kg and a volumetric capacity of 0.018 m^3. For water depths between 201 and 3571 m, the bottom sediments were collected from the oceanographic vessel OV Justo Sierra with a 0.25 m^2

Hessler Sandia MK-III box corer. Water samples were taken at 5 m depth using 1-gallon amber glass bottles that were closed under water to avoid contamination with surface mixtures. We obtained 32 physicochemical parameters from the water and the sediments, including oxygen (mg/L), salinity (UPS), pH, and nitrogen (micromol/g), among others (see Additional file 1: Table S1 in the online supporting material for a complete list). The hydrocarbon sampling procedures have been described elsewhere [34, 40, 41]. The sediment samples were placed in high-density polythene (HDPE) bags and kept at 4 °C for transport to CINVESTAV-IPN, Unidad Mérida. The types of hydrocarbons and metals and the physicochemical characteristics of the sediment were determined in the Geochemistry and the Marine Chemistry laboratories, respectively, using standardised methods [42–44]. Both the environmental and the biological data were obtained from all study sites within a very narrow window of time (within a month). Therefore, the temporal variability in the variables was assumed to be minimal, and most of the variability in the data was considered to be spatial.

Sampling procedures for the flatfishes and helminth parasites
The shoal flounders for the present study were collected between September and October 2005 by professional fishermen based on their commercial fishing permit (issued by the Secretaria de Ganaderia, Desarrollo Rural, Pesca y Alimentación, number: 01067, and available upon request). The collection was conducted using a commercial fishing boat with 4 shrimp trawl nets with a light mesh of 1¾ inches by 1½ inches in the bag and with a turtle excluder device (Super Shooter). The trawls were performed by making circles around the sampling site for 50-60 min at a speed of 0.6-0.7 knots/h and were performed for the shortest period of time possible considering the welfare issues [45, 46]. Because shoal flounders are fragile animals [47], most of them were dead when the nets landed into the boat. The flounders that were still alive were kept in a container with marine water and an oxygen supply and later were euthanised with 100 mg/L of benzocaine until opercular movements ceased; posteriorly, the brain was then severed by spiking [48]. The fishing activities did not involve endangered or protected species according to Mexican regulations (NOM-059-SEMARNAT-2001).

The statement of ethics approval for the present paper was provided by the Institutional Animal Care and Use Committee (IACUC) from the Center of Research and Advanced Studies (Protocol number: 0138-15) and is available upon request.

The fish collected (n = 194) were kept in isolated plastic bags within cold storage coolers (-20 °C) on the vessel until they were transported to CINVESTAV-IPN Unidad
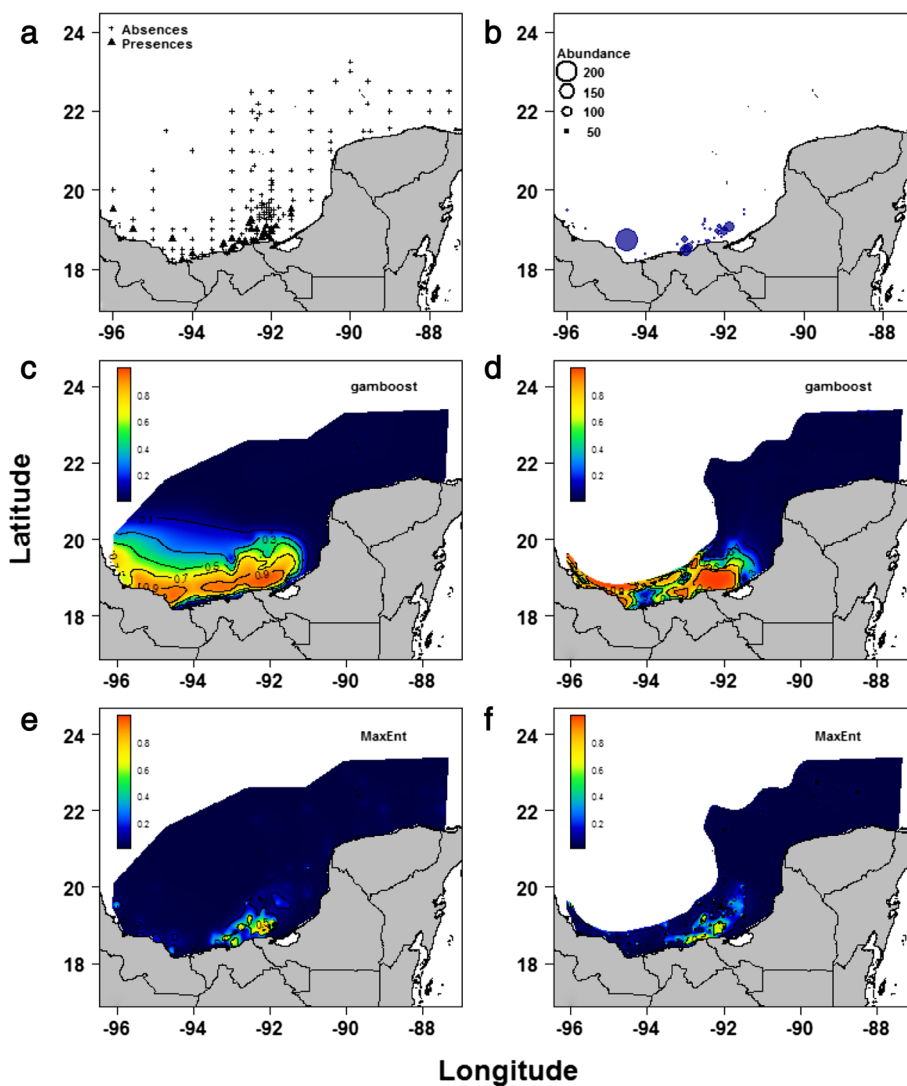
Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 4 of 13

Fig. 1 Geographical distribution of the larval cestode *Oncomegas wageneri* in the southern Gulf of Mexico. **a** Presence (▲) and absence (+) of *O. wageneri* at the 162 sampling sites of the oceanographic expedition Xcambo 2 between September and October 2005. **b** Total number of *O. wageneri* per sampling site. **c** Probability of occurrence of *O. wageneri* using a boosted generalised additive model for the full area (n = 162 sampling sites). **d** Probability of occurrence of *O. wageneri* using a boosted generalised additive model for the polygon area (n = 134 sampling sites at a depth of 1500 m or above). **e** Probability of occurrence of *O. wageneri* using the MaxEnt for the full area. **f** Probability of occurrence of *O. wageneri* using the MaxEnt for the polygon area

Mérida for examination. Once in the laboratory, the total length, the standard length, the maximum height (cm) and weight (g) were recorded for each fish. Subsequently, a helminthological examination was performed to search for *O. wageneri*. The parasitological examinations were conducted using stereomicroscopes, and a 0.7 % saline solution was used as necessary. The abundance of *O. wageneri* was quantified by squashing the mesenteries and muscle fillets between two glass plates. The detected parasites were preliminarily identified in wet mounts prepared with ammonium picrate fixative on semipermanent slides [49, 50]. Additional *O. wageneri* individuals (usually 10-20) were fixed in 4 % formalin in

labelled vials, stained with Mayer's paracarmine, mounted in Canada balsam for subsequent taxonomic identification and identified at the species level using specialised literature (e.g., [51, 52]). In both cases (plerocerci stained with ammonium picrate or Mayer's carmine), the identification of most individuals was based on the scolex measurements, the presence and measurements of the macrohook in the bothrial surface of the basal swelling of the tentacles, the number of hook rows in the basal armature, and the number of hooks in the basal and meta-basal armature of the tentacles. Because the morphological characteristics of all of our plerocerci were within the ranges indicated for *O. wageneri* (see [51]), we assumed that all of them

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 5 of 13

belonged to this species. However because the plerocerci lack adult morphological characteristics, such as the distribution of testes posterior to the ovary, the potential presence of other species cannot be completely ruled out. The host species was identified by ichthyologists at the Necton Laboratory (CINVESTAV-IPN Mérida Unit).

Even when the fish sampling procedure used was standardised, the number of fish collected was highly variable among the sampling sites (Additional file 1: Table S1). For the sampling sites from which a fair number of fish were collected, no more than 10 fish were examined in the search for *O. wageneri*. However, no more than 1-5 individual fish were collected from several of the sampling sites. Clearly, this fact makes it very difficult to compare the abundance of the parasite species among the sampling sites because this parameter depends heavily on the number of fish caught. To overcome this problem, we transformed the parasite data from the mean abundance values to the presence/absence values. We divided the number of *O. wageneri* by the number of fish examined at each sampling site. Then, instead of using the mean abundance value, we used the value of this metric to represent the number of times the parasite was present. For example, if we caught 5 shoal flounders at a specific sampling site and found that they were infected with 5, 2, 3, 0 and 0 *O. wageneri*, this meant that we had 3 presences and 2 real absences at that sampling site. This method of changing abundance data to binary data (presences and absences) allowed us to overcome the problem of the dependence of *O. wageneri* abundance on the number of fish obtained per sampling site and to represent the number of parasites as the number of presences. This method was adapted from the field of community ecology, in which species richness estimators are also frequently biased by the number of sampling units considered at a specific locality, and the presence/absence descriptors perform well (e.g.,[38, 39]). The use of this methodology allowed us to choose a binomial distribution for the dependent variable in the generalised additive model (see below) to estimate the probability of the occurrence of *O. wageneri* in response to the environmental and pollution variables at the sampling site level. In addition to the 32 environmental variables recorded, to factor in the contribution of unknown environmental variables acting at different spatial scales into the analysis, we used a principal coordinates of neighbour matrices (PCNM) analysis to generate a set of spatial variables from the geographical position of each sampling site [53–56]. This set of spatial variables (called PCNM vectors) represents a spectral decomposition of the spatial relationships among the sites that corresponds to all of the spatial scales that can be perceived from the data [56]. The PCNM vectors were grouped into the following three spatial scales: large (58–22 km); medium

(21–15 km) and small (14–2 km). These variables were used as independent variables, and the probability of the occurrence of *O. wageneri* per sampling site was used as the dependent variable in both the generalised additive models (boosted GAMs hereafter) and the maximum entropy algorithm (MaxEnt hereafter).

Data analysis

Calibration of the Generalised Additive Models with Mboost

We fitted the generalised additive models (GAMs) using a boosting algorithm based on component-wise univariate base learners implemented in the R package mboost [57] to examine whether the environmental variables statistically related to the probability of the occurrence of *O. wageneri*. In the mboost package, a GAM has two parts: a distributional part (or assumption in Hofner et al. [57]) and a structural part. The distributional part specifies the conditional distribution of the dependent variable. In this case, we chose the binomial distribution because our data were transformed to parasite presence or absence. The structural part specifies how the predictors (our environmental, pollution and PCNM variables) are related to the dependent variable [57]. The mboost package is essentially a machine-learning algorithm and performs a variable selection during the fitting process [57, 58]. In the boosted GAM, the structure of each predictor is assumed to be additive, so many base-learners (similar to the classical smoothers in the GAMs [59]), such as penalised B-splines or P-splines, can be used. We used base-learners based on B-splines on the covariates (called bbs hereafter). The presence of spatial autocorrelation violates the assumption of observation independence and can bias the results [60, 61]. Because such a situation is common in ecological data [57], we used a spatial base-learner on the geographical coordinates that addresses the spatial autocorrelation in the boosted GAM [61]. Because we used a boosted GAM, the effects in the model were of two types: a global effect (related to the environmental variables) and a local effect (related to the spatial autocorrelation and the nonstationarity assumption), following the suggestions of Kneib et al. [62] and Hothorn et al. [63]. Thus, the boosted GAM was able to perform the following tasks: 1) fit a complex model considering all the available variables; 2) choose the most relevant variables (the ones providing the most information); 3) allocate the information to the local and global components of the model; and 4) divide the input data into a training set and a test set, which is one of the best ways to ensure that the results are not artefacts of overfitting. The boosted GAM implementations [57], as well as the PCNM analyses, are based on the R software for statistical computing [64].

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 6 of 13

Calibration of MaxEnt

The maximum entropy algorithm (MaxEnt) estimates the probability of habitat suitability for the establishment of viable populations by finding a probability distribution that is closest to uniform but restricted by the mean values of the environmental variables at the sampling sites [65]. We concur with Merow et al. [66] in that using only the autofeatures of the MaxEnt could produce misleading results. Thus, we followed their recommendations as much as possible when choosing the MaxEnt settings. The following settings were used for MaxEnt: a cumulative output format; jackknife to measure the variable importance; a random test percentage = 50 %; 0.8 of the regularisation multiplier; a maximum number of background sampling points = 10,000; a replicate number = 100; a replicated run type = cross-validate; and a maximum number of iterations = 5000. Regularisation overcomes the risk of overfitting in MaxEnt. If the selected multiplier is very small (e.g., 0.01), there is great risk of overfitting and an increase in the model complexity, which affects many features in the model (e.g., linear, quadratic, and product, among others [67, 68]). By contrast, a multiplier that is high (e.g., 3) restricts the number of features included in the model too much. Empirically, we found that a multiplier of 0.8 works well with the *O. wageneri* data to neither overfit nor excessively restrict the number of features allowed in the models.

All of the sampling points without flatfish and parasite occurrences were used for the background file in the boosted GAM and MaxEnt. However, the number of sampling sites included in these files is critical for the performance of the model, and there should be good biological reasons to choose the size of this region (called M sensu Soberón and Peterson [67]), as noted elsewhere [69]. In our case, the biological justification to include all the sampling sites in the study area with presences and absences (the full area hereafter) up to a 3571 m depth and a region including only the sampling sites with depths of 1500 m or above (henceforth called the polygon area) was that helminth parasites have been reported up to depths of 5000 m [70, 71]. With respect to *S. gunteri*, the Mexican expeditions in the Gulf of Mexico searching for flatfishes have been unable to explore waters deeper than 200 m [47]; consequently, there is no reason to deny the possibility that *S. gunteri* or other fishes infected with *O. wageneri* may be detected in deeper waters in the near future.

Evaluation of model performance for the boosted GAM and MaxEnt

It is important to bear in mind that in all boosted models (and all models that result from penalised/regularised regression), there is no p-value or significance test. To fairly evaluate the performance of both the boosted GAM and MaxEnt models, we used the following same statistical tests for both methods: Cohen's kappa, AUC (area under the curve, also known as receiver operating characteristic curves or ROC curves) and pROC curves (partial receiver operating characteristic curves). All of these tests are based on a confusion matrix and depend on the departure from ideal scores for true positive and true negative values. Cohen's kappa attempts to correct the degree of agreement by subtracting the portion of the counts that may be attributed to chance [72]. This coefficient ranges from -1 (total disagreement) through 0 (random classification) to 1 (total agreement). The AUC method measures the capacity of a model to determine both when a species is present and when it is absent [73]. The axes of the graph representing the AUC are 1 – specificity or the false positive rate on the X-axis and 1 – commission rate, the sensitivity or true positive rate on the Y-axis [65, 74]. The AUC graph ranges from 0 to 1 on both axes, with a 45° diagonal line between [0,0] and [1,1]. The values below this line indicate a performance no better than random, whereas values above the diagonal line between 0.7 and 0.8 are considered useful, and values >0.8 are considered excellent [75]. The foundation for the proposal of the pROC method is that the AUC method incorrectly assumes that 1 – specificity (the X-axis) spans the entire range [0,1], even when the model predictions may not span that whole range. Thus, Peterson et al. [74] proposed that changes to the AUC curves were required to generate partial ROC curves that span only a subset of the full spectrum of areal predictions. This proposal implied a change in the name of the X-axis to "proportion of area predicted present" and the assumption that the model only functions for part of the areal predictions, accepting at least a 10 % not possible to quantify on that axis [76, 77]. Thus, even when all of our dependent and independent variables were obtained in situ and within a very narrow temporal window, we do recognise the need to consider the potential effect of sampling error on the performance of the ENM models [69, 74, 78, 79]. This sampling error occurs because the spatial predictions of the ENM models present omission errors known as false negatives (omitting known distributional area) and commission errors known as false positives (the inclusion of unsuitable areas for the species distribution in the prediction) [74]. Thus, we considered a commission error (E) of 10 % for both the polygon and full area models for *O. wageneri*. Instead of using the procedure suggested by Peterson [74] and Barve [80] for the calculation of pROC curves, we used the pROC package of Robin [81]. This pROC package can be found at http://cran.r-project.org/web/packages/pROC/pROC.pdf. As mentioned above, we used the R package "dismo" to simultaneously compare the performance of different types of ENMs, including the boosted GAMs and MaxEnt (http://cran.r-project.org/

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 7 of 13

web/packages/dismo/dismo.pdf [82]), while producing its own pROC curves. Finally, we compared the partial ROC curves of each model using a bootstrap test, as suggested by Pepe et al. [83], to evaluate differences in the performance of the models. Regarding the relevance of the independent variables, the mboost package has the capacity to compute bootstrap estimates to undertake a cross-validation to prevent overfitting [57]. This cross-validation involves a variable selection process that provides the frequency at which each variable is selected during the bootstrap process. Then, we used these frequencies as proxies for the importance of each of the variables within the model.

Environmental variable layers

We interpolated each variable to build layers that were used to predict the probability of occurrence of *O. wageneri* from the models fitted (i.e., the boosted GAMs and MaxEnt). The interpolation was performed with ordinary kriging. For this procedure, we built a grid encompassing the full area and another grid for the polygon area for the dependent and independent variables.

Results

A total of 7143 *O. wageneri* were collected from 29 out of the 162 (18 %) sampling sites. *O. wageneri* infected 163 out of 194 shoal flounders, *Syacium gunteri*. The overall prevalence, mean abundance and mean intensity of *O. wageneri* at the 29 sampling sites where the species was present were 84 %, 36 ± 45, and 44 ± 47, respectively. Across all the sampling points, the standard length and weight of the flatfishes were 12 ± 3 cm and 38 ± 45 g, respectively. The prevalence and mean intensity (± standard deviation) of *O. wageneri* per sampling site and the environmental and spatial variables selected by the boosting algorithm for the generalised additive models (boosted GAM) and MaxEnt models are shown in Tables 1 and 2, respectively. The mean abundance values of *O. wageneri* were transformed to presence/absence values as explained in the methodology section.

Oncomegas wageneri boosted GAM and MaxEnt

The number of sampling sites where *O. wageneri* were present and their mean abundance values per sampling site are shown in Fig. 1a and b. Figure 1c shows the probability of *O. wageneri* occurrence for the full area, which had a strong statistical association with high molecular weight polyaromatic hydrocarbons (PAHH) (selected by the bootstrap analysis with a 95 % frequency, Table 2). By contrast, all the remaining spatial and nutrient variables showed a minor contribution (5 %) to the explained variability (Table 2). Thus, for the full area, the final model was: Probability of the occurrence of *O. wageneri* ~ bbs

(PAHH) + all other 16 variables in the full area column in Table 2. The boosted GAM for *O. wageneri* across the full area (Fig. 1c) shows that the probability of the occurrence for this parasite was high in the region between Cayo Arcas and the Coatzacoalcos River mouth (orange and yellow zones), whereas the continental shelf of the Yucatan Peninsula and the oceanic region in the middle of the Gulf of Mexico showed a very low occurrence probability for this parasite. Thus, the probability of *O. wageneri* occurrence based on the boosted GAM for the full area (Fig. 1c) closely resembles the actual spatial distribution of *O. wageneri* in the study area (Fig. 1a and b).

For the polygon area (Fig. 1d), the probability of *O. wageneri* occurrence was similar to the one using the full area (Fig. 1c). However, the number of independent variables differed between the models with 17 for the full area and 20 for the polygon area (Table 2). For the polygon area, the most important component of the model was related to 3 spatial variables (55 % frequency all together), followed by the PAHH (7 % frequency). Thus, for the polygon area, the final model was: Probability of the occurrence of *O. wageneri* ~ bspatial (Lon, Lat) + bbs (PCNM2) + bbs (PCNM58) + bbs (PAHH) + all other 16 variables in the polygon area column in Table 2. The MaxEnt model for *O. wageneri* (Fig. 1e and f) for the full area and for the polygon area poorly predicted the probability of *O. wageneri* occurrence (Fig. 1a). The performance statistics (kappa, AUC and pROC) from the boosted GAM for the full area and the polygon area for *O. wageneri* were all above 0.8 (Table 3). By contrast, all the performance statistics in the MaxEnt models for *O. wageneri* for the full area and the polygon area were below 0.8 (Table 3).

Discussion

The main hypothesis tested in this paper was that because *O. wageneri* has transmission stages and intermediate hosts that are exposed to a polluted environment, the probability of the occurrence of this parasite should reflect the environmental conditions experienced at the seascape level. Our results suggest that this pattern occurred. *O. wageneri* was widely distributed near the coast (Fig. 1c) and was very much influenced, not only by the distribution of the shoal flounder, but also by the PAHH and nutrients present in the shallow waters of the coastal zone (Fig. 1c; Table 2). However, before undertaking a detailed interpretation of the statistical relationships between the dependent and independent variables in the boosted GAMs and the MaxEnt models for this parasite species it is necessary to address two issues: the effect of the size of the background area (full or polygon areas) and the marked differences in the values of the performance statistics between the ENMs for *O. wageneri* (Table 3).

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 8 of 13

Table 1 Geographic position of the sites in the southern Gulf of Mexico, where sediment, water, shoal flounders *Syacium gunteri* and cestodes *Oncomegas wageneri* were collected

Site	Latitude DD	Longitude DD	Number of fish collected	Mean standard length of *S. gunteri* ± SD cm	Infection parameters of *O. wageneri* %	MA ± SD
1	-91.87749	19.063737	6	11.52 ± 0.32	100	97 ± 51
2	-92.924406	18.558687	10	10.41 ± 1.18	100	81 ± 35
3	-92.657425	18.691936	7	10.70 ± 1.42	43	19 ± 28
4	-92.460664	18.711013	4	8.05 ± 1.41	75	10 ± 9
5	-92.50014	19.187617	10	9.01 ± 0.89	80	29 ± 22
6	-91.500028	19.374545	1	23.30	100	13
7	-93.000353	18.753978	11	12.48 ± 1.23	89	73 ± 60
8	-95.501	19.001317	10	10.95 ± 1.17	80	8 ± 6
9	-92.500832	18.999193	9	12.47 ± 2.06	100	20 ± 22
10	-92.006873	18.999577	3	16.87 ± 4.59	67	62 ± 9
11	-91.501037	19.500352	4	23.38 ± 5.54	75	11 ± 13
12	-92.00304	18.937508	6	12.28 ± 0.47	100	94 ± 68
13	-92.375937	18.999565	6	11.18 ± 0.46	100	39 ± 14
14	-95.999833	19.501753	10	12.59 ± 1.31	60	61 ± 108
15	-94.25	18.233333	10	10.97 ± 1.15	80	13 ± 9
16	-92.15496	18.88098	9	11.00 ± 0.71	100	34 ± 15
17	-92.986852	18.483098	4	11.25 ± 0.26	100	110 ± 44
18	-93.00009	18.50345	11	11.35 ± 1.23	82	60 ± 40
19	-93.20285	18.62565	10	17.78 ± 5.62	80	22 ± 23
20	-92.81203	18.691928	11	9.81 ± 0.97	100	33 ± 27
21	-92.563643	19.126108	11	11.32 ± 1.88	90	37 ± 23
22	-92.50009	19.251617	6	13.37 ± 4.15	17	27
23	-92.18765	18.938898	4	10.74 ± 0.82	100	40 ± 46
24	-92.252252	18.772892	3	8.83 ± 2.48	100	25 ± 27
25	-92.12616	19.063458	5	14.16 ± 7.30	80	51 ± 53
26	-94.000493	18.375	6	15.13 ± 0.93	100	26 ± 30
27	-92.390441	18.68429	2	10.45 ± 5.65	100	43 ± 39
28	-92.063815	18.812852	4	10.90 ± 5.53	100	11 ± 7
29	-94.499377	18.753407	1	11.30	100	214

DD = decimal degrees, % = prevalence, MA ± SD = mean abundance ± standard deviation. See Additional file 1: Table S1 for a complete list of physicochemical parameters and contaminants from water and sediments

The background area and the performance of the ENMs

For *O. wageneri*, the size of the background area (or accessible area sensu Barve et al. [69]) considered (full and polygon areas) did not affect the performance of the boosted GAM (Table 3) but did influence the number and identity of the environmental and spatial variables associated with the probability of the occurrence of this parasite (Table 2). By contrast, the background area considered (either full or polygon areas) was not relevant to the poor performance of the MaxEnt models for *O. wageneri*, which was a rather surprising result (Table 3). It is possible that the relatively low number of sampling sites where *O. wageneri* occurred in our study (29 (18 %)

out of 162 sampling sites) affected the performance of the MaxEnt models. This is a very unusual result because the MaxEnt method is normally very reliable even with few occurrences [65]. Even by reducing the size of the background zone from the full area to the polygon area to resemble the size of the actual *O. wageneri* distribution area, there was no improvement in the performance statistics (Table 3). As an extra test, we reduced the size of the background to the size of the *O. wageneri* occurrence region, as suggested by Phillips et al. [84], but that strategy also failed to improve the values of the performance statistics. Thus, it is possible that the boosted GAMs would be a better choice for the analysis

Table 2 Independent variables selected by the boosted general additive model (boosted GAM) for the full area and polygon area models

	Independent variables	Units	Full area Fr (%)	Polygon area Fr(%)
1	bspatial (Lon, Lat)	DD	1.45	18.90
2	bbs (Total depth, S)	m	-	4.80
3	bbs (Temperature, W)	℃	0.09	5.40
4	bbs (Salinity, W)	UPS	0.13	0.60
5	bbs (Oxygen, W)	(mg/L)	0.17	5.00
6	bbs (Alkalinity, W)	meq/l	0.13	2.10
7	bbs (CO2, W)	mmol/l	0.13	2.10
8	bbs (Nitrate, W)	µMolar	0.21	3.00
9	bbs (Phosphate, W)	µMolar	0.09	0.70
10	bbs (Silicate, W)	µMolar	-	1.10
11	bbs(Sigma T, W)	Kg/m3	0.30	-
12	bbs(Sand, S)	%	-	0.50
13	bbs (Silt, S)	%	-	1.10
14	bbs (Clay, S)	%	0.04	-
15	bbs (Phosporus, S)	micromol/g	0.04	0.60
16	bbs (Nitrogen, S)	micromol/g	-	1.30
17	bbs (PAHL, S)	µg/g	0.47	5.70
18	bbs (PAHH, S)	µg/g	94.70	6.90
19	bbs (Aliphatic PAHs, S)	µg/g	-	0.60
20	bbs (PCNM2)	DD	0.30	24.00
21	bbs (PCNM21)	DD	0.47	3.70
22	bbs (PCNM58)	DD	0.68	11.90

Fr (%) is the frequency at which each variable is selected during the bootstrapping, and used as proxy of the importance of each of the variables (expressed as percentage) within the model. bbs = base-learner based on B-splines, bspatial = spatial base-learner on geographical coordinates. W = water, S = sediment; PAHL and PAHH are polyaromatic hydrocarbons of low and high molecular weight respectively; PCNM2, PCNM21 and PCNM58 are the spatial variables of the principal coordinates of neighbour matrices (PCNM) analysis, acting at 2, 21 and 58 km respectively

Table 3 Performance statistics of the boosted general additive model (boosted GAM) and MaxEnt models for *Oncomegas wageneri* for the full area and the polygon area

	Boosted GAM		MaxEnt	
	Full area	Polygon area	Full area	Polygon area
Kappa	0.823	0.823	0.577	0.434
AUC	0.970	0.959	0.679	0.466
pROC	0.837	0.833	0.737	0.649
	(0.722-0.941)[a]	(0.725-0.949)	(0.561-0.912)	(0.474-1.000)

Kappa = Cohen's kappa, AUC = area under the curve and pROC = partial receiver operating characteristic curve. [a] The values in brackets were the ranges obtained by bootstrapping

Considering the polygon area, the performance of the model was also excellent. However, the model was increased to 20 variables; the most important of which were the three spatial ones (frequency = 55 % all together) followed by a combination of environmental and spatial variables with minor frequency values (Table 2). Considering the MaxEnt models for *O. wageneri* sampled in the full area and the polygon area, the values of the Kappa and pROC performance statistics were between 0.4 and 0.7. Because these values indicate a very poor performance [75], no further interpretation of the MaxEnt models or the environmental variables associated with these models was considered necessary. Applying the criterion of parsimony for choosing the best *O. wageneri* model for interpretation, we selected the boosted GAM with a smaller number of variables; that is, we selected the full area model shown in Fig. 1c over that of the polygon area because, despite similarly good values for their performance statistics, the model for the full area had fewer variables (Table 2). Thus, the rest of the discussion concentrates on the interpretation and explanation of the patterns obtained with the boosted GAM for *O. wageneri* sampled in the full area (Fig. 1c).

Values of the infection parameters of *O. wageneri* in the full area

The values of the prevalence and mean abundance of *O. wageneri* infecting the shoal flounder in the full area (Table 1) were high and similar to those obtained for the Mexican flounder *Cyclopsetta chittendeni* [85] (prevalence range 17-100 %; mean abundance range: 8 ± 6-110 ± 44). These results suggest that the transmission of the larval forms of *O. wageneri* in the southern Gulf of Mexico is high. However, even though the larvae of *O. wageneri* have been reported by several authors to infect fish in the Gulf of Mexico [35, 51, 86], none of these authors provided data on the prevalence or mean abundance of this parasite for other zones in the Gulf of Mexico. Thus, it is difficult to know if the high values of the infection parameters of *O. wageneri* in *S. gunteri* that

of problematic data sets of this type. Certainly, it would be desirable (but expensive) to perform more sampling in the region where *O. wageneri* is present for the construction of a comparable background region to obtain a comparable number of sampling sites with absences, as suggested by Phillips et al. [84], to correct for the potential effect of a geographical bias.

Considering the specific case of the boosted GAM for *O. wageneri* sampled in the full area, the performance of the model was excellent (Table 3) because all the values of the Kappa, AUC and pROC performance statistics were between 0.8 and 0.9 [75]. The model included 17 variables; the most influential of which was the PAHH (frequency = 95 %) followed by a minor contribution from a combination of nutrients, the PAHL and the PCNM spatial variables (frequency = 5 %) (Table 2).

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 10 of 13

we found could be similar to those in other regions of the Gulf of Mexico.

Statistical relationships in the boosted GAM for *O. wageneri*

The PAHH variable had the highest frequency (95 %) in the boosted GAM for the full area *O. wageneri* model (Table 2), which suggests that both the shoal flounder and *O. wageneri* have been chronically exposed to these pollutants. The PAHL were present in the *O. wageneri* models but with minor frequency (Table 2). One of the most common indices to determine the main source of PAH is the PAHL/PAHH ratio. If this ratio is <1, the most likely origin is pyrolytic (incomplete combustion of organic matter - combustion of fossil fuel, vehicular engine combustion, smelting, waste incinerators, forest fires and coal combustion), whereas values >1 indicate a petrogenic origin (unburned petroleum and its products - gasoline, kerosene, diesel, lubricating oil and asphalt) [87, 88]. In our data set, only 14 of the 162 sampling sites had PAHL/PAHH ratios >1 (See Additional file 1: Table S1). Thus, most of the PAHs to which the flatfishes and their parasites had been exposed were pyrolytic. Regarding the potential toxicological effects of PAHL and PAHH, none of these compounds exceeded the probable effect level (PEL) established for marine and estuarine sediment quality [89], but several (e.g., benzo[a]pyrene) are considered carcinogenic [90]. Therefore, whether these compounds have a direct effect on *O. wageneri*, the shoal flounder or the definitive hosts (rays) remains an open question.

Another question is why were the PAHH selected preferentially by the *O. wageneri* boosted GAM? It is likely that the presence of these compounds together with the presence of the PAHL, N and P carried from the continent through the river discharge into the marine sediments enhances the growth of hydrocarbonoclastic bacteria; these are common, free-living bacteria in marine environments and include, for example, certain species of the genera *Bacillus*, *Pseudomonas* and *Halomonas* that feed on these compounds [91, 92]. These bacteria feed easily on the PAHL, but it takes them a long time to decompose the PAHH, if they do at all [93], which in turn could be a potential explanation for the persistent presence of the PAHH in the boosted GAM model. In any case, an increase in the number of colonies of these bacteria would in turn enhance primary and secondary productivity in the area with a consequent increase in the number of intermediate hosts, which has been suggested for regions affected by oil spills, such as the Prestige spill [16, 17]. The above-mentioned provision of nutrients and the PAHH from the continent into the coastal zones (up to 200 m depth) has been widely documented [33] for the marine region from Terminos Lagoon to the

Coatzacoalcos River zone. However, based on the available data, it is difficult to infer what deleterious effects are induced by the PAHH (and other pollutants) on the transmission stages (coracidia, plerocercoids), on the first and second intermediate hosts (copepods and shrimps, respectively), on the shoal flounder (acting as paratenic or third intermediate host) or even on the definitive host (stingrays). Thus, experimental exposure of the components of the life cycle of *O. wageneri* (hosts and transmission stages) will be necessary to clarify the relative effects of these persistent organic pollutants (POPs) because previous work has stressed their potential deleterious effects on the life cycles of marine parasites [94, 95].

The frequency of the spatial component in the boosted GAM of *O. wageneri* (summing the bspatial and the PCNMs variability) was very small (1.45 %) (Table 2). This result suggests that the spatial autocorrelation had a small effect on the models for *O. wageneri*, which most likely occurred because the shoal flounder also likely had a fairly high site fidelity, meaning that these individuals do not move large distances as the pelagic fishes do. These flounders (e.g., *Syacium gunteri*) do not enter the coastal lagoons but complete their life cycle in a relatively narrow region with small migrations from coastal to deeper oceanic waters (~100 m depth) for feeding and reproduction [47, 96]. For these flatfishes and their parasites, the environmental variables that occur at the local level (in the very same place where they live) were far more important than the environmental variables acting at larger spatial scales. This interpretation agrees with that of Hothorn et al. [63], who suggested that the mobility of the species under study has a large influence on the relevance of the spatial component in the boosted GAMs. These researchers found that in highly vagile birds, the spatial component was very important; however, for dragonflies with a low spatial range, this component was largely irrelevant.

Conclusions

Our results have led us to two conclusions. One involves a methodological issue, and the other concerns a biological issue. The conclusion regarding methodology was that the boosted GAMs applied to the full area were excellent for describing the probability of the occurrence of *O. wageneri* in the southern Gulf of Mexico, as suggested by the values of the Kappa, AUC and pROC performance statistics (Table 3) and based on the scale proposed by Hosmer and Lemeshow [75]. By contrast, the models produced by the MaxEnt performed very poorly for *O. wageneri*. Furthermore, decreasing the background zone to the polygon area so that it was the same size as the parasite occurrence area, which included only the sample sites shallower than 1500 m or even smaller, (as suggested by Phillips et al.

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 11 of 13

[84]), did not improve performance. Thus, the conclusion at this time is that the MaxEnt is not a good tool to describe the probability of the occurrence of *O. wageneri* infecting the shoal flounder in the southern Gulf of Mexico.

With respect to our biology-related conclusion, 10 out of the 14 sampling sites with PAHL/PAHH ratios values >1 had significantly lower values for the prevalence of *O. wageneri* than those sites with values < 1 (Fisher's exact test; *p* <0.0001). In the case of the mean abundance, it was not possible to calculate whether there were differences between sites with PAHL/PAHH ratio values >1 or < 1 because there were only four sites with PAHL/PAHH ratio values >1 and few individual fish for comparison (Additional file 1: Table S1). Altogether, our results suggest relatively high prevalence values of *O. wageneri* for sites with PAHL/PAHH ratios values between 0 and 1.89 (Additional file 1: Table S1). The origin of these PAHH is most likely petrogenic, but the problem is apparently still not extreme, judging by the high prevalence values of *O. wageneri* in the present study (Table 1 and Additional file 1: Table S1). The 10 sites with PAHL/PAHH ratio values > 1 (range 1.90 to 3.13), had neither fish nor parasites. We concluded that sites with a PAHL/PAHH ratio value up to 1.89 had an enhanced transmission based on the high values of the prevalence of *O. wageneri* (Additional file 1: Table S1), while the sites with PAHL/PAHH ratio values ≥ 1.90 were harmful for both the fish and the parasites because, apparently, they are not able to persist in those sites (Additional file 1: Table S1), which in turn negatively affects their probability of occurrence.

Additional file

Additional file 1: Table S1. Geographic positions, the number of shoal flounders (*Syacium gunteri*) collected and the environmental variables recorded for each sampling site. The variables from Depth to Sigma-T were obtained from water, and those from Carbon % to Total hydro from sediments. PAHL and PAHH are polyaromatic hydrocarbons of low and high molecular weight respectively; Oxygen sat is oxygen saturation; pHW and pHS are pH values for water and sediment respectively; UCM is unresolved complex mixture and Total hydro is the sum of Total PAHs, Aliphatics and UCM. (XLSX 75 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

VMVM conceived of the study, performed part of the statistical analysis and drafted the manuscript. ETI carried out part of the statistical analysis, produced the figures and helped to draft the manuscript. DR participated in the design of the study, produced the figures and helped to draft the manuscript. GGB provided the data on hydrocarbons and heavy metals, and helped to draft the manuscript. EMM helped to draft the manuscript. DVL provided the physicochemical data, and helped to draft the manuscript. MLAM participated in the design of the study and helped to draft the manuscript. All authors read and approved the final manuscript.

Author details

[1]Laboratorio de Parasitología, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Km 6 Carretera Antigua a Progreso, Cordemex, Mérida, Yucatán 97310, México. [2]Laboratorio de Tecnologías Geoespaciales, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Km 6 Carretera Antigua a Progreso, Cordemex, Mérida, Yucatán 97310, México. [3]Posgrado de Geografía. Facultad de Filosofía y Letras, Universidad Nacional Autónoma de México, Circuito Interior, Ciudad Universitaria 04510, México DF, México. [4]Oceanography Department and GERG, Texas A&M University, College Station, TX, USA. [5]Laboratorio de Análisis Espaciales, Dpto. Zoología, Instituto de Biología, Universidad Nacional Autónoma de México, Apdo. Postal 70-153, 04510 México, DF, México. [6]Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Mérida, Km 6 Carretera Antigua a Progreso, Cordemex, Mérida, Yucatán 97310, México.

References

1. Burge CA, Mark EC, Friedman CS, Froelich B, Hershberger PK, Hofmann EE, et al. Climate change influences on marine infectious diseases: implications for management and society. Annu Rev Mar Sci. 2014;6:249–77.
2. Lafferty KD. Environmental parasitology: what can parasites tell us about human impacts on the environment? Parasitol Today. 1997;13:251–5.
3. Hudson PJ, Lafferty KD, Dobson AP. Parasites and ecological systems: Is a healthy ecosystem an infected one? Trends Ecol Evol. 2006;21:381–5.
4. Vidal-Martínez VM, Pech D, Sures B, Purucker ST, Poulin R. Can parasites really reveal environmental impact? Trends Parasitol. 2010;26:44–51.
5. Holt EA, Miller SW. Bioindicators: using organisms to measure environmental impacts. Nat Educ. 2011;3:8.
6. Borja Á, Muxika I. Guidelines for the use of AMBI (AZTI's Marine Biotic Index) in the assessment of the benthic ecological quality. Mar Pollut Bull. 2005;50:787–9.
7. Diaz RJ, Solan M, Valente RM. A review of approaches for classifying benthic habitats and evaluating habitat quality. J Environ Manage. 2004;73:165–81.
8. Ranasinghe JA, Weisberg SB, Smith RW, Montagne DE, Thompson B, Oakden JM, et al. Calibration and evaluation of five indicators of benthic community condition in two California bay and estuary habitats. Mar Pollut Bull. 2009;59:5–13.
9. Teixeira H, Weisberg SB, Borja Á, Ranasinghe JA, Cadien DB, Velarde RG, et al. Calibration and validation of the AZTI's Marine Biotic Index (AMBI) for Southern California marine bays. Ecol Indic. 2012;12:84e95.
10. Warwick RM. Environmental impact studies on marine communities: pragmatical considerations. Austral J Ecol. 1993;18:63–80.
11. Wright J. Biomonitoring with aquatic benthic macroinvertebrates in southern Costa Rica in support of community based watershed monitoring. Ontario, Canada: MSc Thesis, York University; 2010.

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 12 of 13

12. Tweedley JR, Warwick RM, Clarke KR, Potter IC. Family-level AMBI is valid for use in the north-eastern Atlantic but not for assessing the health of microtidal Australian estuaries. Estuar Coastal Shelf S. 2014;141:85–96.

13. Kuris A. Hosts as Islands. Am Nat. 1980;116:570–86.

14. Hechinger RF, Lafferty KD, Huspeni TC, Brooks AJ, Kuris AM. Can parasites be indicators of free-living diversity? Relationships between species richness and the abundance of larval trematodes and of local benthos and fishes. Oecologia. 2007;151:82–92.

15. MacKenzie K, Williams HH, Williams B, McVicar AH, Siddall R. Parasites as indicators of water quality and the potential use of helminth transmission in marine pollution studies. Adv Parasit. 1995;35:85–144.

16. Perez-del Olmo A, Raga JA, Kostadinova A, Fernandez M. Parasite communities in *Boops boops* (L.) (Sparidae) after the Prestige oil-spill: detectable alterations. Mar Poll Bull. 2007;54:266–76.

17. Pérez-del-Olmo A, Fernández M, Raga JA, Kostadinova A, Morand S. Not everything is everywhere: the distance decay of similarity in a marine host- parasite system. J Biogeogr. 2009;36:200–9.

18. Marcogliese DJ. Parasites of the superorganism: Are they indicators of ecosystem health? Int J Parasit. 2005;35:705–16.

19. Anderson TK, Sukhdeo MVK. The relationship between community species richness and the parasite community in *Fundulus heteroclitus*. J Parasitol. 2013;99:391–6.

20. Huspeni TC, Lafferty KD. Using larval trematodes that parasitize snails to evaluate a salt-marsh restoration project. Ecol Appl. 2004;14:795–804.

21. Hechinger RF, Lafferty KD, Kuris AM. Trematodes indicate animal biodiversity in the Chilean intertidal and Lake Tanganyika. J Parasitol. 2008;94:966–8.

22. Pech D, Vidal-Martínez VM, Aguirre-Macedo ML, Gold-Bouchot G, Herrera-Silveira JA, Zapata-Pérez O, et al. The checkered puffer (*Spheroides testudineus*) and its helminths as bioindicators of chemical pollution in Yucatan coastal lagoons. Sci Total Environ. 2009;407:2315–24.

23. Altman I, Byers JE. Large-scale spatial variation in parasite communities influenced by anthropogenic factors. Ecology. 2014;95:1876–87.

24. Elith J, Leathwick J. Conservation prioritization using species distribution models. In: Moilanen A, Wilson KA, Possingham H, editors. Spatial conservation prioritization: quantitative methods and computational tools. Oxford: Oxford University Press; 2009. p. 70–93.

25. Kempf A, Stelzenmüller V, Akimova A, Floeter J. Spatial assessment of predator-prey relationships in the North Sea: the influence of abiotic habitat properties on the spatial overlap between 0-group cod and grey gurnard. Fish Oceanogr. 2013;22:174–92.

26. Thuiller W, Münkemüller T. Habitat suitability modelling. In: Moller AP, Fiedler W, Berthold P, editors. Effects of Climate Change on Birds. Oxford: Oxford University Press; 2005. p. 77–85.

27. Araujo MB, Towsend-Peterson A. Uses and misuses of bioclimatic envelope modeling. Ecology. 2012;93:1527–39.

28. Peterson AT, Ortega-Huerta MA, Bartley J, Sánchez-Cordero V, Soberón J, Buddemeier RH, et al. Future projections for Mexican faunas under global climate change scenarios. Nature. 2002;416:626–9.

29. Zambrano L, Martínez-Meyer E, Menezes N, Peterson AT. Invasive potential of common carp (*Cyprinus carpio*) and Nile tilapia (*Oreochromis niloticus*) in American freshwater systems. Can J Fish Aquat Sci. 2006;63:1903–10.

30. Trejo I, Martínez-Meyer E, Calixto-Pérez E, Sánchez-Colón S. Analysis of the effects of climate change on plant communities and mammals in México. Atmósfera. 2011;24:1–14.

31. Elith J, Leathwick J. Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol S. 2009;40:677–97.

32. Rombouts I, Beaugranda G, Dauvinb JC. Potential changes in benthic macrofaunal distributions from the English Channel simulated under climate change scenarios. Estuar Coastal Shelf S. 2012;99:153–61.

33. García-Cuellar JA, Arreguín-Sánchez F, Hernández-Vázquez S, Lluch-Cota DB. Impacto ecológico de la industria petrolera en la sonda de Campeche, México, tras tres décadas de actividad: una revisión. Interciencia. 2004;29:311–9.

34. Vidal-Martínez VM, Aguirre-Macedo ML, Del Rio-Rodríguez R, Gold-Bouchot G, Rendón-Von Osten J, Miranda-Rosas G. The pink shrimp *Farfantepenaeus duorarum*, it symbions and helmints as bioindicators of chemical pollution in Campeche Sound, Mexico. J Helminthol. 2006;80:159–74.

35. Palm H. Untersuchungen zur Systematik von Russelbandwurmern (Cestoda: Trypanorhyncha) aus atlantischen Fischen. Doktortitel. Institute fur Meereskunde an der Christian-Albrechts-Universitat Kiel. 1995.

36. Mattis TE. Development of two tetrarhynchidean cestodes from the northern Gulf of Mexico. Dissertation. University of Southern Mississippi. 1986.

37. Vidal-Martínez VM, Kennedy CR, Aguirre-Macedo ML. The structuring process of the macroparasite community of an experimental population of *Cichlasoma urophthalmus* through time. J Helminthol. 1998;72:199–208.

38. Hortal J, Borges PAV, Gaspar C. Evaluating the performance of species richness estimators: sensitivity to sample grain size. J Anim Ecol. 2006;75:274–87.

39. Torres-Irineo E, Justin-Amande J, Gaertner D, Delgado De Molina A, Murua H, Chavance P, et al. Bycatch species composition over time by tuna purse seine fishery in the eastern tropical Atlantic Ocean. Biodivers Conserv. 2014;23:1157–73.

40. Gold-Bouchot G, Sima-Alvarez R, Zapata-Pérez O, Güemez-Ricalde J. Histopathological Effects of Petroleum Hydrocarbons and Heavy Metals on the American Oyster (*Crassostrea virginica*) from Tabasco, Mexico. Mar Poll Bull. 1995;31:439–45.

41. Gold-Bouchot G, Zavala-Coral M, Zapata-Pérez O, Ceja-Moreno V. Hydrocarbon concentrations in oysters (*Crassostrea virginica*) and recent sediments from three coastal lagoons in Tabasco, Mexico. B Environ Contam Tox. 1997;59:430–7.

42. Strickland JDH, Parsons TR. A practical Handbook of Seawater Analysis, vol. 167. Otawa: Fisheries Research Board of Canada Bulletin; 1972.

43. Comisión Oceanográfica Intergubernamental. Determinacion de los hidrocarburos del petróleo en los sedimentos. Manuales y Guías 11. Paris; 1982.

44. UNEP/FAO/IAEA/IOC. Sampling of selected marine organisms and sample preparation for the analysis of chlorinated hydrocarbons. Reference Methods for Marine Pollution Studies No. 12 Rev. 2. Paris; 1991.

45. FSBI. Fish Welfare. Briefing Paper 2, Fisheries Society of the British Isles, Granta Information Systems, Cambridge: UK; 2002.

46. Suuronen P. Mortality of fish escaping trawls gears. Rome: FAO Fisheries Technical Paper 478; 2005. FAO.

47. Sánchez-Gil P, Arreguín-Sanchez F, García-Abad MC. Ecological Strategies and Recruitment of *Syacium gunteri* (Pisces: Bothidae) in the southern Gulf of Mexico. Neth J Sea Res. 1994;32:433–9.

48. Mood A. Worse thing happen at sea: the welfare of wild-caught fish. In: Reef Environmental Education Foundation (REEF). 2010. http://fishcount.org. Acceded 29 December, 2014.

49. Scholz T, Aguirre-Macedo ML. Metacercariae of trematodes parasitizing freshwater fish in Mexico: a reappraisal. In: Salgado-Maldonado G, García-Aldrete AN, Vidal-Martínez VM, editors. Metazoan parasites in the Neotropics: ecological, systematic and evolutionary perspective. Commemorative Volume of the 70th Aniversary of the Instituto de Biología, Universidad Nacional Autónoma de México. México DF: Universidad Nacional Autónoma de México; 2000. p. 85–99.

50. Vidal-Martínez VM, Aguirre-Macedo ML, McLaughlin JP, Hechinger RF, Jaramillo AG, Shaw JC, et al. Digenean metacercariae of fishes from the lagoon flats of Palmyra Atoll, Eastern Indo-Pacific. J Helminthol. 2012;86:493–509.

51. Toth LM, Campbell RA, Schmidt GD. A revision of Oncomegas Dollfus, 1929 (Cestoda: Trypanorhyncha; Eutetrarhynchidae), the description of two new species and comments on its classification. Syst Parasitol. 1992;22:167–87.

52. Schaeffner BC, Beveridge I. Description of a new trypanorhynch species (Cestoda) from Indonesian Borneo, with the suppression of *Oncomegoides* and the erection of a new genus *Hispidorhynchus*. J Parasitol. 2012;98:408–14.

53. Borcard D, Legendre P, Drapeau P. Partialling out the spatial component of ecological variation. Ecology. 1992;73:1045–55.

54. Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H. Dissecting the spatial structure of ecological data at multiple scales. Ecology. 2004;85:1826–32.

55. Borcard D, Legendre P. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecol Model. 2002;153:51–68.

56. Santana-Piñeiros AM, Pech D, Vidal-Martínez VM. Spatial structure of the helminth parasite communities of the tonguefish, *Symphurus plagiusa*, from the Campeche coast, southern Mexico. Int J Parasitol. 2012;42:911–20.

57. Hofner B, Mayr A, Robinzonov N, Schmid M. Model-based Boosting in R – A hands-on tutorial using the R Package mboost. Comput Stat. 2012;29:3–35.

58. Bühlmann P. Boosting for high dimensional linear models. Ann Stat. 2006;34:559–83.

59. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed Effects Models and Extensions in Ecology with R. New York: Springer Science + Business Media; 2009.

60. Maloney KO, Schmid M, Weller DE. Applying additive modelling and gradient boosting to assess the effects of watershed and reach characteristics on riverine assemblages. Methods Ecol Evol. 2012;3:116–28.

Vidal-Martínez *et al. Parasites & Vectors* (2015) 8:609

Page 13 of 13

61. Legendre P. Spatial autocorrelation: trouble or new paradigm? Ecology. 1993;74:1659–73.

62. Kneib T, Hothorn T, Tutz G. Variable selection and model choice in geoadditive regression models. Biometrics. 2009;65:626–34.

63. Hothorn T, Müller J, Schroëder B, Kneib T, Brandl R. Decomposing environmental, spatial, and spatiotemporal components of species distributions. Ecol Monogr. 2011;81:329–47.

64. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. http://www.R-project.org. Accessed 21 October 2014.

65. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190:231–59.

66. Merow C, Smith MJ, Silander JA. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography. 2013;36:1058–69.

67. Soberón J, Peterson AT. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers Inform. 2005;2:1–10.

68. Lobo JM, Jiménez-Valverde A, Real R. AUC: a misleading measure of the performance of predictive distribution models. Global Ecol Biogeogr. 2008; 17:145–51.

69. Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, et al. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol Model. 2011;222:1810–9.

70. Campbell RA, Haedrich RL, Munroe TA. Parasitism and ecological relationships among deep-sea benthic fishes. Mar Biol. 1980;57:301–13.

71. Bray RA. The bathymetric distribution of the digenean parasites of deep-sea fishes. Folia Parasit. 2004;51:268–74.

72. Ben-David A. Comparison of classification accuracy using Cohen's Weighted Kappa. Expert Sys Appl. 2008;34:825–32.

73. Elith J, Graham HC, Anderson PR, Dudík M, Ferrier S, Guisan A. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006;29:129–51.

74. Peterson AT, Papeş, Soberón J. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecol Model. 2008, 213: 63-72.

75. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley; 2000.

76. Walter SD. The partial area under the summary ROC curve. Stat Med. 2005; 24:2025–40.

77. Liu C, White M, Newell G. Measuring and comparing the accuracy of species distribution models with presence-absence data. Ecography. 2011; 34:232–43.

78. Peterson AT. Niche modeling: model evaluation. Biodivers Inform. 2012;8:41.

79. Escobar LE, Peterson AT, Favi M, Yung V, Pons DJ, Medina-Vogel G. Ecology and geography of transmission of two bat-borne rabies lineages in Chile. PLoS Negle Trop Diss. 2013;7:e2577.

80. Barve N. Tool for Partial-ROC. Version 1. 2008. http://kuscholarworks.ku.edu/dspace/handle/1808/ 10059. Accesed 11 July 2014.

81. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.

82. Hijmans RJ, Phillips S, Leathwick J, Elith J. Package 'dismo'. 2011. http://cran.r-project.org/web/packages/dismo/index.html. Accesed: 31 July 2014.

83. Pepe M, Longton G, Janes H. Estimation and comparison of receiver operating characteristic curves. Stata J. 2009;9:1–16.

84. Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, et al. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl. 2009;19:181–97.

85. Centeno-Chalé OA, Aguirre-Macedo ML, Gold-Bouchot G, Vidal-Martínez VM. Effects of oil spill related chemical pollution on helminth parasites in Mexican flounder *Cyclopsetta chittendeni* from the Campeche Sound, Gulf of Mexico. Ecotox Environ Safe. 2015;119:162–9.

86. Thatcher VE. Studies on the Cestoda of elasmobranch fishes of the northern Gulf of Mexico. Part I. Proc La Acad Sci. 1961;23:65–74.

87. Hasanati M, Savari A, Nikpour Y, Ghanemi K. Assessment of the sources of polycyclic aromatic hydrocarbons in Mousa Inlet by molecular ratios. J Environm Studies. 2011;37:1–3.

88. Ţigănuş D, Coatu V, Lazăr L, Oros A, Spînu AD. Identification of the Sources of Polycyclic Aromatic Hydrocarbons in Sediments from the Romanian Black Sea Sector. Cercetări Marine. 2013;43:187–96.

89. Canadian Council of Ministers of the Environment. Canadian sediment quality guidelines for the protection of aquatic life: Polycyclic aromatic hydrocarbons (PAHs). Winnipeg: Canadian environmental quality guidelines; 2004.

90. Juhasz AL, Naidu R. Bioremediation of high molecular weight polycyclic aromatic hydrocarbons: a review of the microbial degradation of benzo[a]pyrene. Int Biodeter Biodegr. 2000;45:57–88.

91. Revelo-Romo D, Gómez-Perdomo MI, Concha-Obando M, Bravo D, Fernández-Izquierdo P. Characterization of hydrocarbonoclastic marine bacteria using the 16 s rrna gene: a microcosm case study. Dyna. 2013;80:122–9.

92. Singh AK, Sherry A, Gray ND, Jones DM, Bowler BFJ, Head IM. Kinetic parameters for nutrient enhanced crude oil biodegradation in intertidal marine sediments. Front Microbiol. 2014;5:1–13.

93. Atlas R, Bragg J. Bioremediation of marine oil spills: when and when not-the Exxon Valdez experience. Microbiol Technol. 2009;2:213–21.

94. Pietrock M, Marcogliese DJ. Free-living endohelminth stages: at the mercy of environmental conditions. TREPAR. 2003;19:293–9.

95. Le Yen TT, Rijsdijk L, Sures B, Hendriks AJ. Accumulation of persistent organic pollutants in parasites. Chemosphere. 2014;108:145–51.

96. García-Abad MC, Yáñez-Arancibia A, Sánchez-Gil P, Tapia-Garcia M. Distribución, reproducción y alimentación de *Syacium gunteri* Gingsburg (Pisces: Bothidae), en el Golfo de México. Rev Biol Trop. 1992;39:27–34.