Latitudinal phytoplankton distribution and the neutral theory of biodiversity

Guillem Chust[1,*], Xabier Irigoien[2], Jerome Chave[3], Roger P. Harris[4]

[1] Marine Research Division, AZTI-Tecnalia, Pasaia, Spain
[2] Red Sea Research Center, King Abdullah University, Thuwal, Saudi Arabia
[3] Evolution et Diversité Biologique, CNRS/UPS, Toulouse, France
[4] Plymouth Marine Laboratory, Prospect Place, Plymouth, UK

*: Corresponding author : Guillem Chust, email address : gchust@azti.es

Abstract:

Aim : Recent studies have suggested that global diatom distributions are not limited by dispersal, in the case of both extant species and fossil species, but rather that environmental filtering explains their spatial patterns. Hubbell's neutral theory of biodiversity provides a framework in which to test these alternatives. Our aim is to test whether the structure of marine phytoplankton (diatoms, dinoflagellates and coccolithophores) assemblages across the Atlantic agrees with neutral theory predictions. We asked: (1) whether intersite variance in phytoplankton diversity is explained predominantly by dispersal limitation or by environmental conditions; and (2) whether species abundance distributions are consistent with those expected by the neutral model.

Location : Meridional transect of the Atlantic (50° N–50° S).

Methods : We estimated the relative contributions of environmental factors and geographic distance to phytoplankton composition using similarity matrices, Mantel tests and variation partitioning of the species composition based upon canonical ordination methods. We compared the species abundance distribution of phytoplankton with the neutral model using Etienne's maximum-likelihood inference method.

Results : Phytoplankton communities are slightly more determined by niche segregation (24%), than by dispersal limitation and ecological drift (17%). In 60% of communities, the assumption of neutrality in species' abundance distributions could not be rejected. In tropical zones, where oceanic gyres enclose large stable water masses, most communities showed low species immigration rates; in contrast, we infer that communities in temperate areas, out of oligotrophic gyres, have higher rates of species immigration.

Conclusions : Phytoplankton community structure is consistent with partial niche assembly and partial dispersal and drift assembly (neutral processes). The role of dispersal limitation is almost as important as habitat filtering, a fact that has been largely overlooked in previous studies. Furthermore, the polewards increase in immigration rates of species that we have discovered is probably caused by water mixing conditions and productivity.

INTRODUCTION

Unlike sessile species or those dwelling on islands, oceanic planktonic species have no apparent barriers to dispersal (Cermeño & Falkowski, 2009). It also appears that planktonic species are broadly distributed, both in space and in time. Planktonic species also exhibit some of the most striking examples of explosive population growth (blooms) and of fine niche specialization (d'Ovidio *et al.*, 2010). Ecologists have long debated whether the regional distribution of species arises from dispersal limitation (MacArthur & Wilson, 1967) or from niche differentiation (Hutchinson, 1957). The neutral theory of biodiversity (Hubbell, 2001) has generated a great deal of attention because it provides an integrative framework in which to test these alternatives (Duivenvoorden *et al.*, 2002). Initially, tests and applications of the neutral theory of biodiversity and biogeography have been restricted to tropical forests (e.g. Condit *et al.*, 2002; Duivenvoorden *et al.*, 2002; Chave *et al.*, 2006; Chust *et al.*, 2006a), but since then they have also been applied in marine ecology (e.g. Dornelas *et al.*, 2006; Martiny *et al.*, 2011), and more specifically to planktonic species assemblages (Alonso *et al.*, 2006; Pueyo, 2006a,b; Dolan *et al.*, 2007; Vergnon *et al.*, 2009; Irigoien *et al.*, 2011). However, these latter works have only tested the neutral model partially because they did not take into account explicitly the migration rate of species.

The neutral model of biodiversity developed by Hubbell (1997, 2001) was inspired by MacArthur & Wilson's (1967) theory of island biogeography. In Hubbell's model, all individuals are assumed to have the same prospects for reproduction and death (neutrality). The variability in relative abundances across species is solely due to demographic stochasticity or 'ecological drift'. This model further assumes a separation

74   of spatial scales: demographic processes occur at the local scale of an ecological

75   community, where species may go locally extinct through demographic drift. The local

76   diversity is replenished by immigration at rate m of propagules from a regional species

77   pool. In this large regional pool, drift may also cause species to go extinct, and novel

78   species arise through speciation, such that θ new species are produced every generation

79   in this regional pool. If $m = 1$, the local community is a random (Poisson) sample of the

80   regional pool. In contrast, if m is close to zero, the local community is virtually isolated

81   from the regional pool. Hubbell's neutral model thus assumes that limited dispersal,

82   rather than niche specialization, is the main explanation for spatial structure across

83   ecological communities. Under this model, the local species abundance distribution is

84   thus defined by only two model parameters θ, and m. A spatially-explicit version of

85   Hubbell's model has also been developed (Chave & Leigh, 2002), in which dispersal

86   from one locale to another is limited by the geographical distance between these sites.

87   In such a model, taxonomic cross-site similarity (i.e. the opposite of β-diversity)

88   declines logarithmically with increasing geographical distance (Hubbell, 2001; Condit

89   et al., 2002; Chave & Leigh, 2002).

90

91   In contrast, niche theory assumes that differences in species composition among

92   communities is caused by heterogeneity in the environment or limiting resources, and

93   by environmental filtering of species according to their environmental requirements,

94   such as oceanographic conditions, and competition for resources such as nutrient

95   concentrations for marine phytoplankton. In niche-based models, species are able to

96   coexist by avoiding competition through resource and environmental partitioning

97   (Gause, 1934; Chesson, 2000). Testing neutral theory against niche theory has proven

98   challenging, because both environmental variables and species distributions tend to be

4

99 spatially autocorrelated (Legendre *et al.*, 2005). On the one hand, species distributions

100 are most often aggregated spatially because of biotic processes such as reproduction and

101 death. On the other hand, the pelagic environment is primarily structured by ocean

102 currents and oceanographic processes causing spatial gradients. Statistical techniques

103 have been developed to partition variation of diversity due to environmental variability

104 and due to dispersal limitation (Legendre, 1993; Legendre *et al.*, 2005; Chust *et al.*,

105 2006b).

106

107 Recently, Cermeño & Falkowski (2009) have offered a thought-provoking analysis of

108 global patterns of fossil diatom diversity. They suggested that diatom distributions over

109 the oceans show no evidence of dispersal limitation either at present or over long time

110 scales, but rather that environmental filtering explains these spatial distributions. This

111 view is in line with the Baas-Becking hypothesis that 'everything is everywhere – the

112 environment selects'. More evidence in support for this conclusion has been gathered by

113 Cermeño *et al.* (2010). However, this view contradicts findings for lake diatoms where

114 the potential for dispersal-related community structuring has been shown (Verleyen *et*

115 *al.*, 2009). Also, an analysis of the genetic structure of populations of a marine diatom,

116 *Pseudo-nitzschia pungens*, is consistent with a strong isolation by distance pattern,

117 suggesting that dispersal limitation may be an important factor in explaining the spatial

118 structure of extant diatom communities (Casteleyn *et al.*, 2010). These few statistical

119 analyses offer a quantitative glimpse of the relative roles of environment and dispersal

120 for diatom diversity (Verleyen *et al.*, 2009; Cermeño *et al.*, 2010). Further, the

121 implications of these alternative interpretations for species abundance distributions have

122 not yet been examined in light of Hubbell's neutral theory.

123

124   Here we examine the structure of communities of three phytoplankton groups (diatoms,

125   dinoflagellates, and coccolithophores), along a transect across the Atlantic Ocean from

126   nearly 50º North to 50º South, to ascertain the extent to which the structure is consistent

127   with niche assembly or dispersal (neutral) assembly. This latitudinal transect allows for

128   large biological diversity and strong environmental gradients to be covered. All three

129   phytoplankton groups behave as passive organisms and occupy the same trophic level.

130   We seek to understand whether marine phytoplankton comply with neutral theory

131   predictions of the distribution of relative species abundance and of spatial turnover in

132   diversity. The following null hypotheses were formulated to address our main question:

133   1) According to the neutral theory, and when species are dispersal limited, the similarity

134   of phytoplankton species composition should decrease with geographic distance, and

135   the distance decay in similarity is expected to be more important than oceanographic

136   conditions and nutrient concentrations. Here, we assess the relative contribution of

137   dispersal limitation and environmental factors to the explanation of the variance in

138   phytoplankton assemblages. We note that niche assembly mechanisms and neutral

139   processes of drift and dispersal can occur simultaneously, so that results indicating a

140   contribution of dispersal limitation, while supporting the neutral model, do not preclude

141   a role for niche differentiation in phytoplankton assemblages. However, not finding a

142   role of dispersal limitation does not provide any information on the validity, or lack

143   there of, of the neutral model. 2) Assuming neutrality, the phytoplankton species

144   abundance distribution should fit the distribution expected from Hubbell's neutral

145   model. As the neutral theory applies to metacommunities, where local communities

146   interact with each other by an immigration rate, the test has been performed in three

147   regions (see also Cermeño *et al*., 2010). Thus, we test, for the first time, the predictions

148   of neutral theory for the spatial turnover in species composition and for relative species

149   abundance in three of the most important phytoplankton groups.

150

151   **MATERIAL AND METHODS**

152

153   *The AMT surveys and datasets*

154

155   The Atlantic Meridional Transect (AMT) is an ocean observation programme that

156   undertakes biological, chemical and physical oceanographic research over a latitudinal

157   transect of the Atlantic ocean from nearly 50º North to 50º South (Fig. 1), a distance of

158   over 13,500 km (Robinson *et al*., 2006). This transect crosses a range of biome types

159   from sub-polar to tropical and from eutrophic shelf seas and upwelling systems to

160   oligotrophic mid-ocean gyres. We analysed phytoplankton data from the first three

161   AMT surveys, on-board the research ship James Clark Ross: AMT1 (which took place

162   from 21 September to 24 October 1995), AMT2 (between 22 April and 28 May 1996),

163   and AMT3 (between 20 September and 25 October 1996). AMT1 and AMT3 sailed

164   from the UK to Falkland Islands, whereas AMT2 sailed from Falkland Islands to the

165   UK. The AMT surveys included 25 sampling stations, each separated by 4° latitude

166   from the next station.

167

168   Data from AMT surveys are available from the British Oceanographic Data Center

169   (BODC; http://www.amt-uk.org/data.aspx) and is described in Robins *et al*. (1996a,b)

170   and Bale (1996). Specifically, chemical and phytoplankton data were sampled at 7-m

171   depth waters using a rosette (i.e. water sampling device) fitted with 12 10-litre General

172   Oceanics water bottles. Physical and optical data were obtained with a CTD (Neil

173   Brown Mark IIIB, Instrument Systems, Inc.). Environmental data considered in our

174   analysis encompasses physical variables (sea surface temperature, salinity), optical

175   variables (down-welling irradiance at Photosynthetically Active Radiation (PAR)

176   wavelengths, percentage of irradiance at sampling depth, surface solar radiation) and

177   nutrients: nitrate+nitrite (NO_3+NO_2), nitrite (NO_2), phosphate (PO_4), and silicate (SiO_4)

178   concentrations. The percentage of surface irradiance at the sampling depth was inferred

179   from the spectral diffuse attenuation coefficient of light (K) at PAR wavelengths.

180   Geographic data were: latitude and longitude.

181

182   For the collection and identification of phytoplankton, 100 ml samples were taken at

183   each station and preserved in lugol's iodine solution (Robins, 1996b). Examination of

184   the samples was conducted following Uthermol's sedimentation technique under an

185   inverted microscope (Robins, 1996b). The sampling procedure and volume used is the

186   standard one for phytoplankton, considered adequate for repeatable characterizations of

187   oceanic phytoplankton communities (Lund *et al*., 1958). Previous studies using these

188   three AMT datasets (and two other ones, AMT4 and AMT5) showed qualitatively

189   similar productivity-diversity patterns, which indicates that 100 ml sample provides a

190   reasonable representation of the phytoplankton community diversity (e.g. Irigoien *et al*.,

191   2004). Phytoplankton (diatoms, dinoflagellates, and cocolithophorids) were

192   taxonomically classified based on morphological characters at species level, and in

193   some cases at genus level. For the present analysis, the species abundance per 100 ml

194   sample volume was considered in order to work with count data (i.e. number of

195   individuals). Overall, diatoms are the most diverse of the three phytoplankton groups

196   (from 83 to 92 diatom species per survey, 35 to 42 dinoflagellate species, and 34-38

197   coccolithophore species), see Table 1. However, coccolithophores showed the highest

198   average species richness per station (9.8), followed by diatoms (8.3) and dinoflagellates

199   (6.5). Among coccolithophores, the most abundant species was the bloom forming

200   *Emiliania huxleyi* in all three surveys. In contrast, the most abundant diatom and

201   dinoflagellate species varied from one survey to the next. In particular, diatoms varied

202   markedly in abundance and dominance; for instance, the most abundant species on

203   AMT1 was *Thalassiosira gracilis* with 6144.6 individuals per ml, all present on a single

204   station, and absent on both AMT2 and AMT3.

205

206   *Spatial species turnover*

207

208   The relative contribution of environmental factors and geographic distance to

209   phytoplankton composition was estimated using similarity matrices, Mantel tests and

210   variation partitioning of the species composition across sites based upon canonical

211   ordination methods (Legendre & Legendre, 1998). The Jaccard index was used to

212   measure the compositional similarity between pairs of stations. The Jaccard index is the

213   number of species shared between the two plots, divided by the total number of species

214   observed. Distance matrices for environmental variables and geographic distance were

215   measured by the Euclidean distance between values at two stations. We used Mantel

216   tests (Legendre & Legendre, 1998) to determine the correlation between species

217   similarity matrices and environmental and geographic distance. The Mantel test is a

218   nonparametric test based on a boostrap randomization of the matrices, to determine how

219   frequently the observed similarity would arise by chance. This test computes a statistic

220   r_M which measures the correlation between two matrices. The rate of change in species

221   similarity with increasing geographic distance was calculated by fitting a linear model.

222   Also, the latitudinal range of a species was defined as the distance between the observed

223 latitudinal extremes of its occurrence. From the individual species ranges, average

224 latitudinal ranges were then computed for each phytoplankton group. To test the

225 correlation between species similarity and environmental distance, we first selected the

226 best subset of environmental variables, such that the Euclidean distance of scaled

227 environmental variables would have the maximum correlation with community

228 dissimilarities, using the *vegan* package (Oksanen et al. 2011) implemented in the R

229 2.13.1 language (R Development Core Team, 2011). We then compared the $2^p - 1$

230 possible models, where p is the number of environmental variables, for each AMT

231 survey and phytoplankton group. Only environmental variables with values in all

232 stations were considered in the initial model. Subsequently, a partial Mantel test was

233 undertaken to determine the relative contribution of environmental distance (after model

234 selection) and geographic distance in accounting for species variation.

235

236 We partitioned the variance of phytoplankton composition across stations to determine

237 the relative contribution of environmental factors and spatial pattern. Species spatial

238 pattern, as a result of aggregation because of biotic processes, were modelled with third-

239 degree polynomial of geographic coordinates of latitude (X) and longitude (Y): $X, Y,$

240 $X*Y, X^2, Y^2, X^2*Y, Y^2*X, X^3$ and Y^3 (cubic trend surface analysis, Legendre 1993). The

241 total intersite variation in species abundance was decomposed into four components:

242 pure effect of environment, pure effect of geographical distance, combined variation

243 due to the joint effect of environment and geographical distance, and unexplained

244 variation. Since partitioning on distance matrices (Mantel approach) underestimates the

245 amount of variation in community composition (Legendre *et al.*, 2005), we used a

246 canonical (i.e. constrained) ordination analysis (ter Braak & Šmilauer, 1998) to estimate

247 a proportion of the variance of the original phytoplankton table of abundances (sites by

248　species). Canonical ordination analysis is a method to reduce the variation in

249　community composition in which the axes are constrained to be linear combinations of

250　explanatory variables. More specifically, species are assumed to have unimodal

251　response surfaces with respect to explanatory gradients. The variance partitioning

252　analysis, detailed in Legendre *et al*. (2005), proceeds in two steps. First, we selected the

253　best two canonical correspondence models (one for environmental variables, the other

254　for spatial terms) using a stepwise procedure and based upon the Akaike Information

255　Criterion (AIC), with the *vegan* package (Oksanen, 2011) implemented in the R 2.13.1

256　language (R Development Core Team, 2011). Subsequently, a partial canonical analysis

257　(ter Braak & Šmilauer, 1998) was undertaken to determine the relative contribution of

258　environmental factors and spatial terms in accounting for species variation. Specifically,

259　the partial canonical analysis estimates the contribution of environmental factors in

260　accounting for species variation by removing the effect of the spatial term covariable.

261　Because of the presence of environmental missing values (at 29 sites) and low number

262　of stations per AMT survey for this type of analysis, the variation partitioning was

263　undertaken for the overall three AMT surveys (46 sites) restricting the analysis to six

264　environmental variables whose values were available for all sites: sea surface

265　temperature, salinity, percentage of irradiance, NO_2, PO_4, and SiO_4.

266

267　*Neutral theory*

268

269　One radical step toward the construction of a mathematically tractable community

270　model is Hubbell's theory of biodiversity (Hubbell, 2001). This theory is radical in

271　assuming that all individuals have the same prospects of reproduction and death

272　irrespective of their age, size and of the species to which they belong. Hubbell (2001)

11

273   modeled local communities in which each death is replaced, with probability 1-m, by an

274   offspring of a randomly chosen individual in the local community, regardless of species,

275   and with probability m, by an immigrant from the regional species pool. The species of

276   immigrant is determined by the relative abundance of species in the regional pool. In

277   Hubbell's original model, community size remains constant, but in later versions, the

278   size of the local community can vary about a stochastic mean size (Volkov *et al*. 2003).

279   Hence, the species composition fluctuates due to stochastic drift only, but not because

280   of habitat selection or of interspecific competition. The local community is embedded in

281   and connected via migration to the geographic area occupied by the regional species

282   pool, the metacommunity, of size J_M (the number of individuals in the regional pool), so

283   that a fraction m of recruits has immigrated from the regional pool rather than being the

284   offspring of local parents. The local community reaches a dynamic equilibrium between

285   stochastic local species extinction and species replenishment through immigration. At

286   the scale of the regional pool, a similar dynamics occurs; diversity is maintained

287   because extinction is balanced by speciation. Speciation in the regional species pool is

288   modeled simply by assuming that each new recruit has a small probability v of yielding

289   an altogether new species, so that $\theta = v \times J_M$ new species appear in the system on

290   average each generation. Hubbell's (2001) neutral model, thus, has two parameters: the

291   regional diversity parameter θ and the immigration rate m. Etienne (2005) has formally

292   shown that θ can jointly be estimated with m from empirical species abundance data

293   using a maximum likelihood framework.

294

295   Jabot & Chave (2011) have proposed a test of neutrality building upon Etienne's (2005)

296   maximum-likelihood (ML) inference method. Briefly, for any species abundance

297   distribution, a ML estimate of the neutral parameters θ and m may be obtained. Using

298   Hubbell's model as a null model, neutral species abundance distributions are

299   constructed, and only those with the same number of species as in the empirical dataset

300   are retained, until one reaches one thousand simulated communities. These neutral

301   species abundance distributions therefore have the same observed number of species

302   and the same θ and m as do the empirical species abundance distribution. To build a

303   test, Shannon's index is then calculated for both the neutral species abundance

304   distributions and for the empirical one. The rationale for our choice of Shannon's index

305   as a summary statistic is further explained in Jabot and Chave (2011). If the empirical

306   Shannon's index falls outside the distribution of neutral Shannon's indices, then

307   neutrality is rejected. The empirical Shannon index was compared with this null

308   distribution by a t-test. This test of neutrality is based on species abundance

309   distributions only, but it is more robust than previous tests.

310

311   We explored the results of this neutrality test along the latitudinal axis by partitioning

312   the global dataset into three regions: northern temperate zone (>25º), tropical zone

313   (between >-25º and <25º) and southern temperate zone (<-25º), see Fig. 1. The

314   boundary of the northern zone with the tropical coincides with the Westerlies biome and

315   Trade-Winds biome, respectively, defined by the Longhurst Biogeographical Provinces

316   (VLIZ, 2009). The tropical zone so defined had a mean SST above 24.5 ºC (North of

317   the equator) and above ~22 ºC (South of the equator).

318

319   We estimated the neutral model parameters θ and m together with confidence intervals

320   and also performed the above test for the total dataset (including diatoms,

321   coccolithophores and dinoflagellates). This inference was implemented in the Tetame

322   software (Jabot *et al*., 2008). Of the 75 samples, 8 had more than 50,000 individuals,

323 and this resulted in prohibitively long calculations (akin to finding the zeros of a

324 polynomial of degree equal to the number of individuals, see Etienne 2005). For these 8

325 samples, we picked a random sample of 50,000 individuals, and replicated this sampling

326 procedure ten times to ensure its stability. In two cases, the neutral parameters could not

327 be computed due to too small sample sizes. In a majority of tests, neutrality was not

328 rejected; in such cases, assuming neutrality, we explored how the estimated immigration

329 probability (m) varied with latitude throughout the main Atlantic zones.

330

331 **RESULTS**

332

333 *Spatial species turnover*

334

335 Mean similarity among stations was highest for coccolithophores (0.29), followed by

336 dinoflagellates (0.23) and diatoms (0.11), see Table 1. The geographic distance range

337 occupied by a species (on average) is less in diatoms (3352.8 km) than in dinoflagellates

338 (4784.1 km) and coccolithophores (6093.8 km) (Table 1). Similarity of the three

339 phytoplankton groups decreases significantly (p<0.001) in all three groups with

340 geographic distance (Fig. 2; r_M (diatoms) = 0.24-0.28; r_M (dinoflagellates) = 0.20-0.34,

341 r_M (coccolithophores) = 0.29-0.39, and in all three AMT surveys. The Mantel

342 correlation between species similarity and environmental factors (0.37-0.74) was higher

343 than with geographic distance (0.21-0.39), for the three phytoplankton groups and the

344 three surveys (Table 2). The Mantel correlation between species similarity and

345 geographic distance, partialling out environmental factors, was significant (p<0.05) for

346 a majority of cases (in all three groups for AMT1 and AMT2).

347

348   The variation partitioning based upon canonical ordination analysis reveals that

349   environment is the largest main-effect factor contributing to phytoplankton species

350   variation (24%; Fig. 3). However, the spatial component accounted for almost as much

351   variation (17%). However, the interaction of environment and distance explained even

352   more of the variation (26%) than either of the main-effect factors, indicating a role for

353   as yet unexplained covariance between environment and separation distance. In the case

354   of diatoms, environment is clearly higher than the spatial terms (25% vs. 8%,

355   respectively), whereas in dinoflagellates (17% vs. 18%) and coccolithophores (5% vs.

356   6%) the two factors are approximately equivalent.

357

358   *Neutral theory parameters and test*

359

360   The estimates of neutral parameters (θ and m) for each station are shown in Table 3 for

361   the three defined latitudinal regions (see also Appendix S3 for parameters for each

362   station). The test of fit of the phytoplankton species abundance distribution to the

363   neutral communities indicates that the number of communities in which neutrality

364   cannot be rejected is higher (45) than the number in which neutrality can be rejected

365   (28) (Table 3). Communities for which neutrality could not be rejected made up a larger

366   percentage of tropical communities (50 to 100%), than of communities in the northern

367   (40 to 57%) or southern (17 to 71%) zones. Fig. 4 shows six examples of the empirical

368   species abundance distribution compared with that expected by a neutral model given

369   the local community parameters θ and m. These examples are representative of

370   communities in all three latitudinal zones and illustrate variation in the goodness of fit

371   of the neutral expectation. Those communities whose abundance distributions were not

372   fit by the neutral model (e.g., Fig 4b,d,f), generally exhibit too many species in the

373   doubling abundance classes of 3 to 16 individuals per species.

374

375   Because species abundance distribution matches neutral theory a majority of cases

376   (60%), we went on in such cases to plot the immigration probability (m) against latitude

377   (Fig. 5a). This plot revealed that m is consistently lower in tropical zones than in

378   temperate zones. In particular, the probability of immigration is a convex function of

379   latitude ($r^2 = 0.44$, p-value < 0.0001), with a minimum in the tropical zone. We used

380   AIC to select the best-fitting polynomial function (up to 4th order). This result suggests

381   that local plankton communities in the temperate zones receive more immigration from

382   the metacommunity (regional species pool) than do tropical communities.

383

384   **DISCUSSION**

385

386   We tested two predictions of neutral theory against data on the community structure of

387   three marine phytoplankton groups in a latitudinal transect of the Atlantic Ocean. First,

388   the canonical ordination analysis and Mantel tests showed that environment and

389   geographic distance explained variation in diversity for the three phytoplankton taxa

390   (diatoms, dinoflagellates and coccolithophores). These analyses also indicated that

391   environment is slightly more important than geographic distance. Second, the Shannon

392   information test of the fit of neutral theory to observed relative species abundance

393   distributions showed that neutral expectations can not be rejected for 60% of

394   communities. These two findings suggest that phytoplankton communities result from a

395   combination of niche and neutral processes, which is in accordance with the patterns

396   found in an exhaustive phytoplankton time series dataset (Vergnon *et al.*, 2009). Similar

397   conclusions were reached in a study of phytoplankton communities in the Caribbean

398   and Mediterranean seas; Pueyo (2006a) states that both neutral and non-neutral

399   mechanisms co-occur. These recent findings and the results of this paper lead to a new

400   perspective, that niche assembly is not the only, or even always the prevailing, assembly

401   mechanism of plankton communities, in contrast to the views that emerge from

402   previous, global-scale studies of fossil diatom assemblages (Cermeño & Falkowski,

403   2009). To the best of our knowledge, ours is the only approach to combine three

404   important analyses of the same dataset: (i) empirical estimation of dispersal limitation,

405   (2) assessment of the relative contribution of environmental factors and dispersal

406   limitation to community assembly; and (3) estimation of migration rate in the neutral

407   model.

408

409   The estimation of dispersal limitation revealed slight differences between phytoplankton

410   groups. On the one hand, the geographic distance range occupied by one species (on

411   average) is less in diatoms than in dinoflagellates and coccolithophores (Table 1). This

412   suggests that connectivity among population sites is low in diatoms. On the other hand,

413   coccolithophore similarity has a correlation with geographic distance (i.e. distance

414   decay) slightly higher (0.29-0.39) than in diatoms (0.24-0.28), which can be interpreted

415   as high spatial structuring (i.e. patchiness). In a pure neutral metacommunity, high

416   slopes in the distance decay and small ranges of geographic distance occupied by the

417   species, are related and provide a measure of dispersal limitation. In our case, however,

418   diatoms have the lowest latitudinal range and the lowest distance decay slope. This

419   apparent paradox should be due to the fact that diatom occurrences are very low (2 to 3

420   stations on average per AMT survey), with respect to coccolithophores (more than 7).

421   The differential abundance of species, and differing species richness, make it difficult to

422   evaluate the significance of small differences in dispersal in the different groups.

423   Although mobility, sedimentation and growth rates are known to differ among these

424   phytoplankton groups (Broekhuizen, 1999), their functional similarity and co-

425   occurrence in similar environments might result in similar dispersal rates at the

426   community level. This is an aspect that requires further research. A limitation of our

427   dataset is that samples were not repeatedly subsampled, to test for repeatability and the

428   degree to which the species diversity present was accurately represented (Gotelli &

429   Colwell, 2001). The difficulty of detecting the smallest organisms and finding the

430   largest organisms, where are rare in finite volumes, is always problematic (e.g. Vergnon

431   *et al.*, 2009). However, the consistent patterns between AMT surveys in our analysis

432   and previous studies (Irigoien *et al.*, 2004) allow us to conclude that community

433   diversity is well captured and sampling biases are not important.

434

435   The three phytoplankton groups exhibited differences in community metrics, although

436   similar patterns between AMT surveys. Coccolithophores are more diverse in tropical

437   zone, decreasing slightly with latitude (see Appendix S1). Over the entire geographic

438   dataset, they are less diverse than diatoms, although local (per sample) diversity is

439   higher than diatoms. Both abundance and the number of species of coccolithophores are

440   very constant across latitudes, compared with diatoms and dinoflagellates. Concerning

441   the species response strength to the environment, canonical ordination analysis and

442   Mantel tests were consistent in that the environment is slightly more important than

443   geographic distance, although the results of the two statistical analyses differ slightly at

444   the group level. At the current, relatively coarse level of analysis, it is not possible to

445   determine which phytoplankton group responds most strongly to environment. The

446   current wisdom is that diatoms are r-strategists associated with mixed waters and

447   unpredictable conditions (e.g. Margalef, 1978). However, all three taxa exhibit massive

448   blooms, generally taking place in temperate, mixed water zones (Fig. 5b). In each of the

449   three taxa, there is a single species responsible for blooms; among diatoms it is

450   *Thalassiosira gracilis*; among dinoflagellates it is *Gymnodinium galeaeformae*, and

451   among coccolithophores, it is *Emiliania huxleyii*, similar to the findings of Irigoien *et*

452   *al.* (2004). During these massive bloom situations, species richness decreases

453   (Appendix S2), in agreement with previous studies (e.g. Irigoien *et al.*, 2004), which is

454   here interpreted as competitive exclusion (Huisman *et al.*, 1999) because of limiting

455   resources. If this is the case, these exceptional situations escape from the neutral theory

456   assumptions.

457

458   In comparison with other ecosystems, the pelagic environment and remote islands (e.g.

459   islands *sensu stricto*, caves, basins, lakes, estuaries, forest remnants) are the two

460   opposite extremes in terms of population connectivity. Whereas islands could be

461   considered as adimensional points where connectivity is very limited, the pelagic zone

462   could be seen as a three dimensional space with no barriers for marine plankton

463   (Cermeño & Falkowski, 2009), except those imposed by physical heterogeneity (e.g.

464   stratification) and continents. From this point of view, i.e. increasing space dimensions

465   increases potential connectivity, land could act as a two dimensional space for sessile

466   species (e.g. plants), whereas coastlines can limit the dispersal of their inhabitants (e.g.

467   restricted intertidal organisms) in one dimension. For instance, whereas coastal fish

468   species are more likely to remain close to their place of origin, oceanic animal species

469   are highly mobile and live in a continuous habitat with high connectivity (Tittensor *et*

470   *al.*, 2010). Within this general framework, our findings reveal, nevertheless, that overall

471   phytoplankton assemblages are poorly but consistently spatially structured across the

472 Atlantic, indicating that dispersal limitation is playing a non negligible role in global

473 oceanic primary-producer distribution. Our results on dispersal limitation and spatial

474 community structure are intermediate between the strong barriers to dispersal evident in

475 thermophilic Archaea (Whitaker *et al*., 2003), and the other extreme of no limits to

476 dispersal, expressed in the view that below 1 mm body size "everything is everywhere,

477 but the environment selects" (Finlay, 2002). Unlike terrestrial plants, for which

478 ecological drift is potentially a key factor on regional scales, marine phytoplankton

479 species are nearly pan-distributed all over latitudes (at least for species described at the

480 morphological level). Whether the morphologically described species include cryptic

481 species (e.g Kooistra *et al*., 2008), or ecotypes with adaptations at the molecular level

482 (e.g. Johnson *et al*., 2006), and to what extent the consideration of those would improve

483 the percentage of the variance explained by the environment is an aspect that requires

484 further research.

485

486 Another striking finding was that, when fitting the neutral model, immigration rates

487 increase poleward, which is consistent for the three AMT surveys. In tropical zones,

488 where oceanic gyres enclose large stable water masses, communities are relatively

489 constant in species richness and abundance and have low immigration rates. In contrast,

490 communities in temperate areas, out of the oligotrophic gyres, are dominated by

491 blooming spatially-unstructured diatoms and show higher rates of species immigration.

492 Thus, high species immigration probability from the metacommunity seems to be

493 associated with areas of high water mixing and productivity.

494

20

495 **CONCLUSION**

496

497 Phytoplankton communities of diatoms, dinoflagellates and coccolithophores across the

498 Atlantic Ocean are slightly more determined by niche differentiation (24%) than by

499 dispersal limitation (17%). In 60% of communities from tropical to temperate ocean

500 latitudes, neutrality assumption on the species abundance distribution could not be

501 rejected. These two findings suggest that the observed structure of phytoplankton

502 communities is consistent with a mechanism that combines both niche- and neutral-

503 assembly processes. The consistent patterns between AMT surveys allow us to conclude

504 that sampling biases are not important although our dataset was limited by the lack of

505 repeatedly subsamples. We provide the first empirical evidence that the role of dispersal

506 limitation and ecological drift is almost as important in structuring marine

507 phytoplankton communities as niche assembly. Furthermore, we also found that in

508 tropical zones, where oceanic gyres enclose large stable water masses, most

509 communities were characterized as having low species immigration rates when fitting

510 the neutral model. In contrast, communities in temperate areas, out of the oligotrophic

511 gyres, show higher rates of species immigration.

512

526

527 **REFERENCES**

528

529 Alonso, D., Etienne, R. S. & McKane, A. J. (2006) The merits of neutral theory. *Trends*

530 *in Ecology & Evolution,* **21**, 451-457.

531 Bale A.J. (1996) AMT-3 cruise report.

532 Broekhuizen, N. (1999) Simulating motile algae using a mixed Eulerian–Lagrangian

533 approach: does motility promote dinoflagellate persistence or co-existence with

534 diatoms? *J. Plankton Res*., **21**, 1191–1216.

535 Casteleyn, G., Leliaert, F., Backeljau, T., Debeer, AE, Kotaki, Y., Rhodes, L.,

536 Lundholm, N. Sabbe, K., & Vyverman, W. (2010) Limits to gene flow in a

537 cosmopolitan marine planktonic diatom. *Proc Natl Acad Sci USA,* **107**, 12952 – 12957.

538 Cermeño, P., & Falkowski, P. G. (2009) Controls on Diatom Biogeography in the

539 Ocean. *Science*, **325**, 1539-1541.

540 Cermeño, P., C. de Vargas, F. t. Abrantes, and P. G. Falkowski (2010) Phytoplankton

541 Biogeography and Community Stability in the Ocean. *PLoS ONE,* **5**, e10037.

542 Cermeño, P., E. Maranon, D. Harbour, F. G. Figueiras, B. G. Crespo, M. Huete-Ortega,

543 M. Varela, and R. P. Harris. 2008. Resource levels, allometric scaling of population

544 abundance, and marine phytoplankton diversity. *Limnology and Oceanography*, **53**,

545 312-318.

546 Chave J, F. Jabot (2008) TeTame 2.1. Estimation of neutral parameters by maximum

547 likelihood. http://www.edb.ups-tlse.fr/equipe1/tetame.htm.

548 Chave J., D. Alonso, R. S. Etienne (2006) Comparing models of species abundance.

549 Nature 441:E1.

550 Chave, J. & Leigh, E.G (2002) A spatially explicit neutral model of beta-diversity in

551 tropical forests. *Theor. Pop. Biol.,* **62**, 153-168.

552 Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol.*

553 *Syst.*, **31**, 343–366.

554 Chust, G., A. Pérez-Haase, J. Chave, and J. L. Pretus (2006b) Floristic patterns and

555 plant traits of Mediterranean communities in fragmented habitats. *Journal of*

556 *Biogeography*, **33**, 1235-1245.

557 Chust, G., J. Chave, R. Condit, S. Aguilar, S. Lao, and R. Pérez (2006a) Determinants

558 and spatial modeling of the tree B-diversity in a tropical forest landscape in Panama.

559 *Journal of Vegetation Science*, **17**, 83-92.

560 Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B., Nunez, P.,

561 Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H.C., Losos, E. & Hubbell, S.P.

562 (2002) Beta-diversity in tropical forest trees. *Science,* **295**, 666-669.

563 Dolan, J. R., Ritchie, M. R. & Ras, J. (2007) The neutral community structure of

564 planktonic herbivores, tintinnid ciliates of the microzooplankton, across the SE Tropical

565 Pacific Ocean. *Biogeosciences*, **4**, 297–310.

566 Dornelas, M., S. R. Connolly, & T. P. Hughes (2006) Coral reef diversity refutes the

567 neutral theory of biodiversity. *Nature*, **440**, 80-82.

568   Duivenvoorden, J.F., Svenning, J.C. & Wright, S.J. (2002) Beta diversity in tropical

569   forests. *Science*, **295**, 636-637.

570   Etienne, R.S. (2005) A new sampling formula for neutral biodiversity. *Ecology Letters*,

571   **8**, 253-260.

572   Finlay, B. J. (2002) Global dispersal of free-living microbial eukaryote species. *Science,*

573   **296**, 1061-1063.

574   Gause, G.F. (1934) The Struggle for Existence. Williams and Wilkins, Baltimore.

575   Gotelli, N. J. & Colwell, R. K. (2001) Quantifying biodiversity: procedures and pitfalls

576   in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379-391.

577   Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in

578   ecology . *Ecological Modelling*, **135**, 147 – 186.

579   Hubbell, S. P., F. L. He, R. Condit, L. Borda-de-Agua, J. Kellner, and H. ter Steege

580   (2008) How many tree species and how many of them are there in the Amazon will go

581   extinct? *Proceedings of the National Academy of Sciences of the United States of*

582   *America,* **105**, 11498-11504.

583   Hubbell, S.P. (1997) A unified theory of biogeography and relative species abundance

584   and its application to tropical rain forests and coral reefs. *Coral Reefs*, **16**, S9–S21.

585   Hubbell, S.P. (2001) *A unified neutral theory of biodiversity and biogeography*.

586   Princeton University Press, Princeton, NJ.

587   Huisman, J., Jonker, R. R., Zonneveld, C. & Weissing, F. J. (1999) Competition for

588   light between phytoplankton species: experimental tests of mechanistic theory. *Ecology,*

589   80, 211–222.

590   Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on*

591   *Quantitative Biology,* **22**, 415–427.

592 Irigoien, X., G. Chust, J. A. Fernandes, A. Albaina, & L. Zarauz (2011) Factors

593 determining mesozooplankton species distribution and community structure in shelf and

594 coastal waters. *Journal of Plankton Research,* **33**, 1182-1192.

595 Irigoien, X., J. Huisman, et al. (2004) Global biodiversity patterns of marine

596 phytoplankton and zooplankton. *Nature*, **429**, 863-867.

597 Jabot, F., & J. Chave (2009) Inferring the parameters of the neutral theory of

598 biodiversity using phylogenetic information and implications for tropical forests.

599 *Ecology Letters*, **12**, 239-248.

600 Jabot, F., & J. Chave (2011) Analyzing Tropical Forest Tree Species Abundance

601 Distributions Using a Nonneutral Model and through Approximate Bayesian Inference.

602 *American Naturalist,* **178**, E37-E47.

603 Jabot F., Etienne R.S. & Chave J. (2008) Reconciling neutral community models and

604 environmental filtering: theory and an empirical test. *Oikos*, **117**, 1308-1320.

605 Johnson, Z. I., E. R. Zinser, et al. (2006) Niche partitioning among Prochlorococcus

606 ecotypes along ocean-scale environmental gradients. *Science*, **311,** 1737.

607 Kooistra, W. H. C. F., D. Sarno, et al. (2008) Global Diversity and Biogeography of

608 *Skeletonema* Species (Bacillariophyta). *Protist*, **159,** 177-193.

609 Legendre, P., D. Borcard, & P. R. Peres-Neto (2005) Analyzing beta diversity:

610 Partitioning the spatial variation of community composition data. *Ecological*

611 *Monographs*, **75**, 435-450.

612 Legendre, P. & Legendre, L. (1998) *Numerical ecology*. Elsevier, Amsterdam.

613 Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**,

614 1659-1673.

615   Lund, J.W.G., Kipling, C. & Le Cren, E.D. (1958) The inverted microscope method of

616   estimating algal numbers and statistical basis of estimations by counting.

617   *Hydrobiologia*, **11**, 143-170.

618   MacArthur, R.H., Wilson, E.O. (1967) *The Theory of Island Biogeography*. Princeton

619   University Press.

620   Margalef, R. (1978) Life forms of phytoplankton as survival altertives in an unstable

621   environment. *Oceanol. Acta*, **1**, 493–509.

622   Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D., Horner-Devine, M.C. (2011)

623   Drivers of bacterial beta-diversity depend on spatial scale. *Proc Natl Acad Sci USA,*

624   **108**, 7850-7854.

625   Oksanen, J., (2011) Multivariate Analysis of Ecological Communities in R: vegan

626   tutorial. pp. 43.

627   d'Ovidio, F., De Monte, S., Alvain, S., Dandoneau, Y., Lévy, M. (2010) Fluid

628   dynamical niches of phytoplankton types. *Proc Natl Acad Sci USA* 107:18366–18370.

629   Pueyo, S. (2006a) Diversity: between neutrality and structure. *Oikos*, **112**, 392-405.

630   Pueyo, S. (2006b) Self-similarity in species-area relationship and in species abundance

631   distribution. *Oikos*, **115**, 582-582.

632   Robins, D.B. et al. (1996a) AMT-1 cruise report and preliminary results. NASA

633   technical memorandum 104566, vol. 35.

634   Robins, D.B. et al. (1996b) AMT-2 cruise report.

635   Robinson, C., A. J. Poulton, P. M. Holligan, A. R. Baker, G. Forster, N. Gist, T. D.

636   Jickells, G. Malin, R. Upstill-Goddard, R. G. Williams, E. M. S. Woodward, & M. V.

637   Zubkov (2006) The Atlantic Meridional Transect (AMT) Programme: A contextual

638   view 1995-2005. *Deep Sea Research Part II: Topical Studies in Oceanography,* **53**,

639   1485-1515.

640   Schluter, D. (2001) Ecology and the origin of species, *Trends in Ecology & Evolution*,

641   **16**, 372-380.

642   Telford, R.J., Vandvik, V. & Birks, H.J.B. (2006) Dispersal limitations matter for

643   microbial morphospecies. *Science*, **312**, 1015-1015.

644   ter Braak, C.J.F & Šmilauer, P. (1998) CANOCO reference manual and user's guide to

645   Canoco for Windows: Software for canonical community ordination (version 4).

646   Microcomputer Power, Ithaca, NY, US.

647   R Development Core Team (2011) R: A Language and Environment for Statistical

648   Computing, Vienna, Austria, ISBN 3-900051-07-0, http://www.R-project.org.

649   Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Vanden Berghe, E. &

650   Worm, B. (2010) Global patterns and predictors of marine biodiversity across taxa.

651   *Nature*, **466**, 1098-U1107.

652   Vergnon, R., Dulvy, N.K. & Freckleton, R.P. (2009) Niches versus neutrality:

653   uncovering the drivers of diversity in a species-rich community. *Ecology Letters,* **12**,

654   1079-1090.

655   Verleyen, E., W. Vyverman, M. Sterken, D. A. Hodgson, A. De Wever, S. Juggins, B.

656   Van de Vijver, V. J. Jones, P. Vanormelingen, D. Roberts, R. Flower, C. Kilroy, C.

657   Souffreau, & K. Sabbe (2009) The importance of dispersal related and local factors in

658   shaping the taxonomic structure of diatom metacommunities. *Oikos*, **118**, 1239-1249.

659   VLIZ (2009) Longhurst Biogeographical Provinces. Available online at

660   http://www.vliz.be/vmdcdata/vlimar/downloads.php.

661   Volkov, I., Banavar, J.R., Hubbell, S.P. & Maritan, A. (2003) Neutral theory and

662   relative species abundance in ecology. *Nature*, **424**, 1035-1037.

663   Whitaker, R.J., Grogan, D.W. & Taylor, J.W. (2003) Geographic Barriers Isolate

664   Endemic Populations of Hyperthermophilic Archaea. *Science*, **301**, 976-978.

665

666   **BIOSKETCH**

667

668   **Guillem Chust** is a marine ecologist at AZTI Foundation for Marine research (Spain).

669   In 2002, he obtained the PhD from the University of Paul Sabatier (Toulouse, France).

670   His research focuses on the distribution patterns of species and biodiversity, the effects

671   of climate change in marine and coastal ecosystems, and on scale-dependent processes

672   in ecology.

673

674 **Figure legends**

675

676 Fig. 1. Oceanographic sampling stations corresponding to AMT1, AMT2, and AMT3

677 overlain on a satellite image of ocean colour (blue, green, yellow and red represent

678 increasing values of sea surface chlorophyll-a concentration; mean annual of 2010,

679 MODIS sensor). Arrows indicate the main Atlantic oceanographic gyres.

680

681 Fig. 2. Species similarity against the distance between stations for each AMT (AMT-1

682 in (*a*), AMT-2 in (*b*), and AMT-3 in (*c*), and for the three phytoplankton groups

683 (diatoms, dinoflagellates and coccolithophores). Species similarity was averaged at

684 1000 km interval. Error values are the standard deviation divided by two.

685

686 Fig. 3. Variation partitioning (%) of species composition, based on constrained

687 correspondence analysis, according to spatial terms and environmental determinants, for

688 each phytoplankton group.

689

690 Fig. 4. Empirical species abundance distributions and that expected under neutral model

691 of six communities using Preston plots. Grey bars show the binned abundance classes

692 (i.e. 1, 2, 3-4, 5-8, 9-16, …), and black circles represent the expected number of species

693 for each abundance class under neutral model with maximum likelihood estimation of θ

694 and m parameters, and J individuals. *a*) Northern station AMT3.4 ($J = 2294$, $\theta = 3.75$, m

695 $= 0.45$, $p = 0.114$); *b*) Northern station AMT1.4 ($J = 3224$, $\theta = 3.46$, $m = 0.52$, $p = 0.$

696 003); *c*) Tropical station AMT3.9 ($J = 1548$, $\theta = 3.91$, $m = 0.26$, $p = 0.344$); *d*) Tropical

697 station AMT3.12 ($J = 7052$, $\theta = 3.82$, $m = 0.54$, $p = 0.009$); *e*) Southern station AMT2.5

698 ($J = 3436$, $\theta = 7.69$, $m = 0.099$, $p = 0.167$); *f*) Southern station AMT1.20 ($J = 2692$, $\theta =$

699   4.63, $m = 0.44$, $p < 0.001$). Communities in the left side (*a*, *c* and *d*) fitted to neutral

700   model according to the test ($p > 0.05$), and communities in the right side (*b*, *d* and *f*) did

701   not fit to neutral model ($p < 0.05$).

702

703   Fig. 5. (*a*) Immigration rate (*m*) and (*b*) overall abundance across latitude for each AMT

704   survey. Fitted curve is a 4th order polynomial model (for *m*, $r^2=0.44$, p<0.0001; for

705   abundance, $r^2=0.54$, p<0.0001), selected with AIC comparing four polynomial models

706   from first to 4th order.

707

708   Table 1. Statistics of community structure of phytoplankton groups and AMT surveys.

709   Abundance is the total number of individuals (per 100 ml) in all stations and for all

710   species.

711

	Diatoms	Dinoflagellates	Coccolithophores
Mean species richness per station	8.25	6.53	9.77
Species richness (AMT1)	92	35	34
Species richness (AMT2)	83	38	35
Species richness (AMT3)	83	42	38
Abundance (AMT1)	683648	23282	94110
Abundance (AMT2)	1563014	7120	109535
Abundance (AMT3)	568879	5674	104262
Mean similarity (AMT1)	0.095	0.221	0.325
Mean similarity (AMT2)	0.107	0.229	0.241
Mean similarity (AMT3)	0.119	0.231	0.308
Mean similarity (AMT1-3)	0.107	0.227	0.291
Mean number of sites where a species is present (AMT1)	2.46	4.40	7.76
Mean number of sites where a species is present (AMT2)	2.45	3.89	6.31
Mean number of sites where a species is present (AMT3)	2.29	4.66	7.09
Mean number of sites where a species is present (AMT1-3)	2.40	4.32	7.05
Mean range of latitudes occupied (AMT1, in km)	4385.9	5776.0	7285.0
Mean range of latitudes occupied (AMT2, in km)	3078.7	3511.2	4934.7
Mean range of latitudes occupied (AMT3, in km)	2593.7	5065.1	6061.8

712
713
714

715 Table 2. Mantel and partial Mantel tests between species similarity and environmental
716 determinants and geographical distance, for each AMT survey and phytoplankton
717 group. Irrad: Irradiance, Sol: Solar radiance.

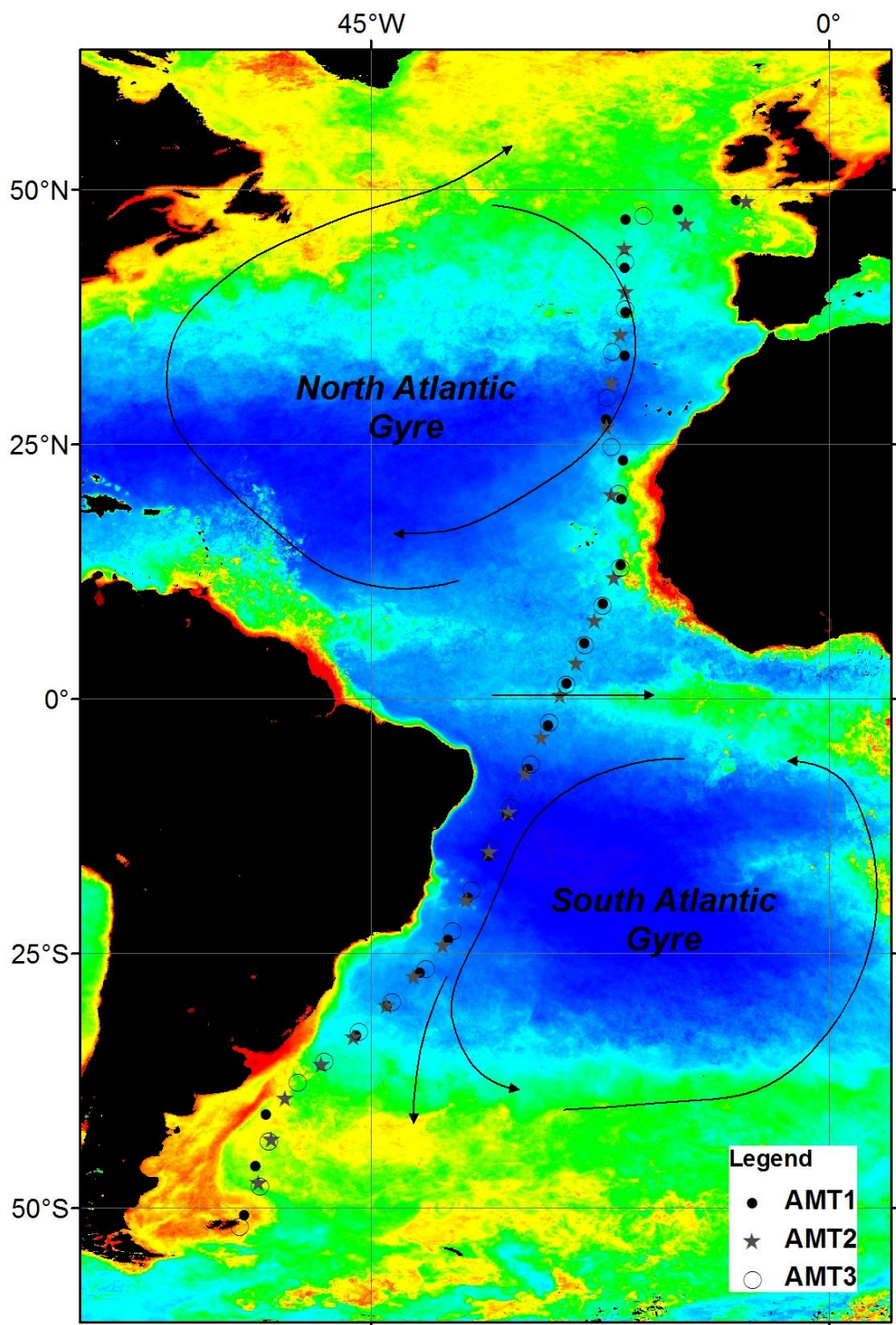|			Mantel r	p-value	Terms selected	Terms entered
AMT1	Diatoms	Jacc × Environ.	0.42	.001	Temperature, Irrad	NO$_3$+NO$_2$, NO$_2$, PO$_4$, Salinity, SiO$_4$, Temperature, Irrad
		Jacc × Distance	0.25	.001		
		Jacc × Environ. (Distance partially out)	0.38	.001		
		Jacc × Distance (Environ. partially out)	0.15	.009		
	Dinoflag.	Jacc × Environ.	0.58	.001	NO$_2$	NO$_3$+NO$_2$, NO$_2$, PO$_4$, Salinity, SiO$_4$, Temperature, Irrad
		Jacc × Distance	0.33	.001		
		Jacc × Environ. (Distance partially out)	0.53	.001		
		Jacc × Distance (Environ. partially out)	0.14	.047		
	Coccolith.	Jacc × Environ.	0.74	.001	NO$_2$, Temperature	NO$_3$+NO$_2$, NO$_2$, PO$_4$, Salinity, SiO$_4$, Temperature, Irrad
		Jacc × Distance	0.39	.001		
		Jacc × Environ. (Distance partially out)	0.68	.001		
		Jacc × Distance (Environ. partially out)	0.15	.030		
AMT2	Diatoms	Jacc × Environ.	0.38	.001	Temperature	NO$_3$+NO$_2$, NO$_2$, PO$_4$, Salinity, SiO$_4$, Temperature, Irrad , Sol
		Jacc × Distance	0.29	.001		
		Jacc × Environ. (Distance partially out)	0.32	.001		
		Jacc × Distance (Environ. partially out)	0.19	.005		
	Dinoflag.	Jacc × Environ.	0.37	.001	NO$_2$, Temperature	NO$_3$+NO$_2$, NO$_2$, PO$_4$, Salinity, SiO$_4$, Temperature, Irrad , Sol
		Jacc × Distance	0.34	.001		
		Jacc × Environ. (Distance partially out)	0.23	.005		
		Jacc × Distance (Environ. partially out)	0.18	.004		
	Coccolith.	Jacc × Environ.	0.60	.001	Temperature	NO$_3$+NO$_2$, NO$_2$, PO$_4$, Salinity, SiO$_4$, Temperature, Irrad , Sol
		Jacc × Distance	0.32	.001		
		Jacc × Environ. (Distance partially out)	0.55	.001		
		Jacc × Distance (Environ. partially out)	0.16	.014		
AMT3	Diatoms	Jacc × Environ.	0.46	.001	Temperature	Salinity, Temperature
		Jacc × Distance	0.24	.004		
		Jacc × Environ. (Distance partially out)	0.41	.001		
		Jacc × Distance (Environ. partially out)	0.07	.199		
	Dinoflag.	Jacc × Environ.	0.47	.001	Temperature	Salinity, Temperature
		Jacc × Distance	0.21	.011		
		Jacc × Environ. (Distance partially out)	0.43	.001		
		Jacc × Distance (Environ. partially out)	0.04	.323		
	Coccolith.	Jacc × Environ.	0.56	.001	Temperature	Salinity, Temperature
		Jacc × Distance	0.29	.001		
		Jacc × Environ. (Distance partially out)	0.51	.001		
		Jacc × Distance (Environ. partially out)	0.10	.091		

718

719 Table 3. Test of fitting phytoplankton Species Abundance Distribution (SAD) to the
720 neutral model for the three AMT surveys and zones. S: species richness; N: total sum of
721 the number of individuals; H: Shannon's index of diversity; θ: the fundamental
722 biodiversity parameter; m: species immigration probability of a local community from
723 the metacommunity. S, H, θ, and m are the mean values for the corresponding zone. See
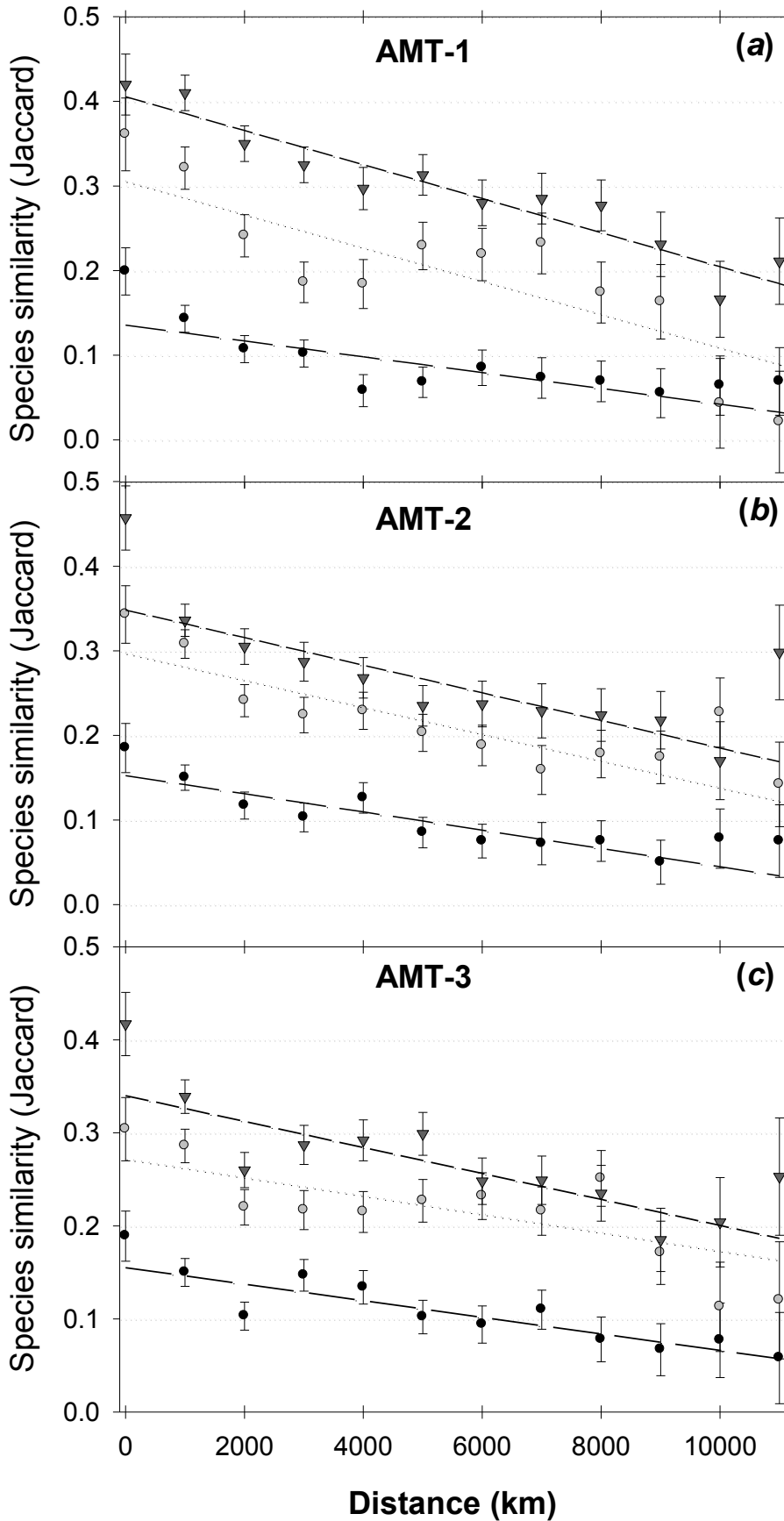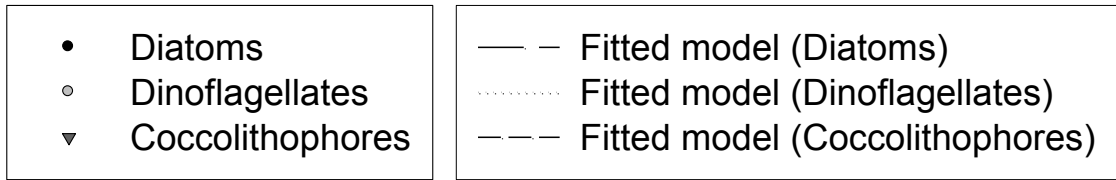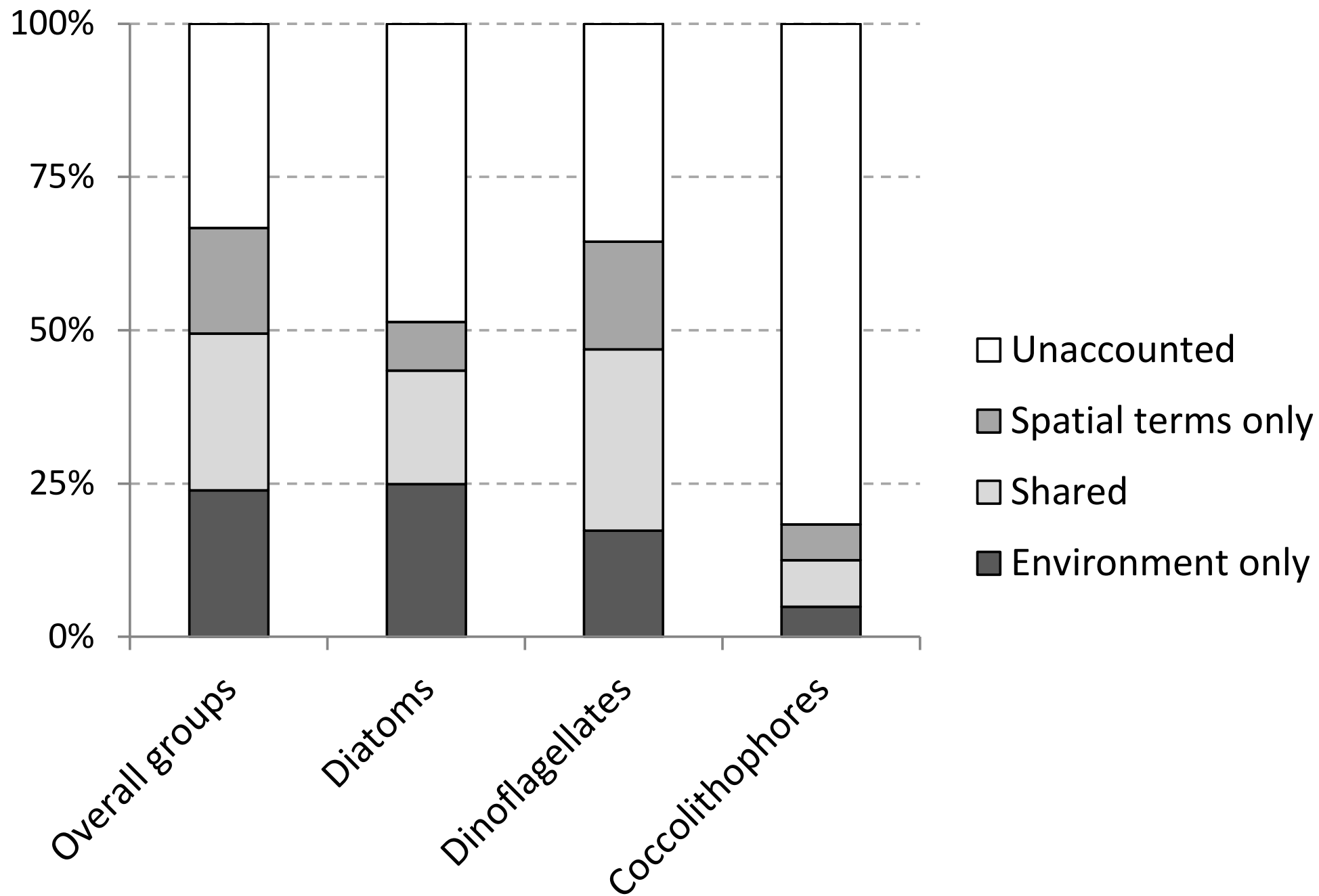724 Appendix S3 for values for each station.

725

	Zone	Number of stations	S	N	H	θ	m	Number of stations with Neutral SAD (p>0.05)
AMT1	**Northern**	7	24.9	2755.1	1.57	4.15	0.45	4
	Tropical	12	20.9	2921.9	1.53	3.77	0.36	6
	Southern	6	35.2	13326.0	1.45	5.52	0.42	1
AMT2	**Northern**	7	22.6	12914.0	1.34	3.22	0.53	4
	Tropical	11	17.6	1137.9	1.92	4.02	0.15	11
	Southern	7	28.7	7161.6	1.77	5.28	0.21	4
AMT3	**Northern**	5	25.0	5776.6	1.54	4.16	0.45	2
	Tropical	10	25.0	5210.6	1.81	4.32	0.23	8
	Southern	7	23.4	10910.8	1.35	3.29	0.51	5
Overall		**73**						**45**

726
727

North Atlantic
Gyre

South Atlantic
Gyre

Legend

● AMT1
★ AMT2
○ AMT3

(a)

m

AMT1
AMT2
AMT3

Latitude (º)

(b)

Abundance (log scale)
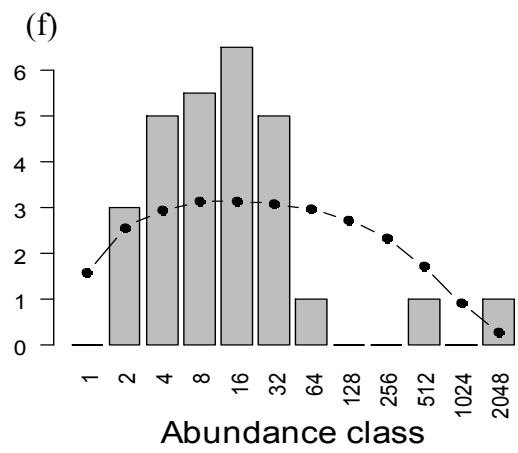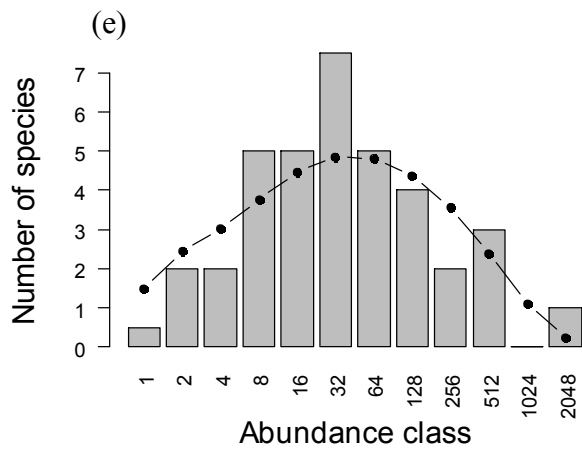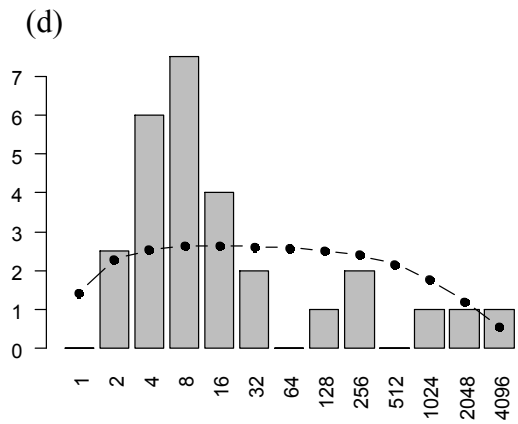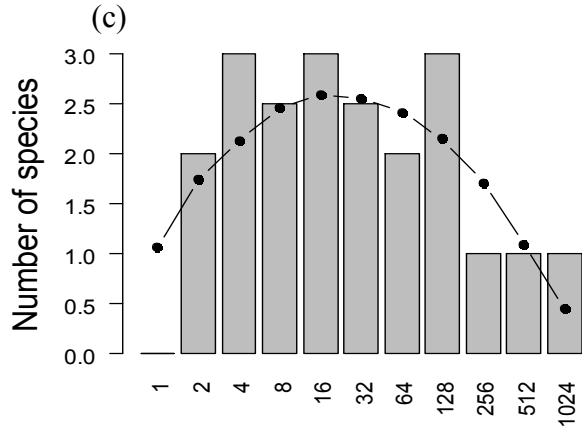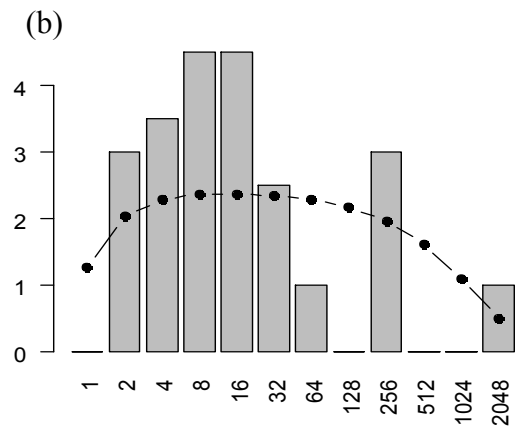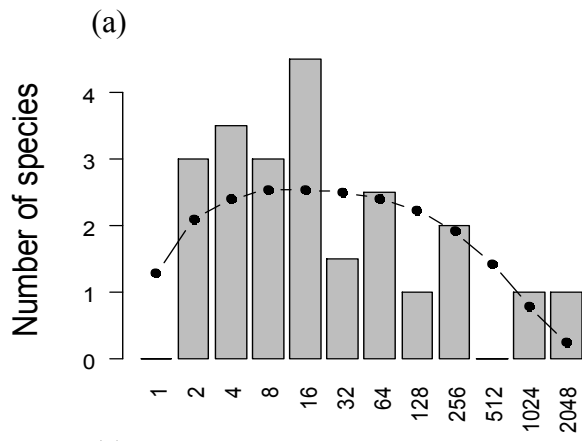
Latitude (º)

Appendix S1. Latitudinal patterns of sea surface temperature, salinity, and species richness of diatoms, dinoflagellates and coccolithophorids.

Appendix S2. Unimodal relation of phytoplankton species richness across biomass ($r^2 = 0.15$, p-value=0.003) and abundance ($r^2 = 0.34$, p-value<0.001).

Appendix S3. Test of fitting phytoplankton species abundance distributions to the neutral model for the each sampling station. *J*: total sum of the number of individuals; S: species richness; H: Shannon's index 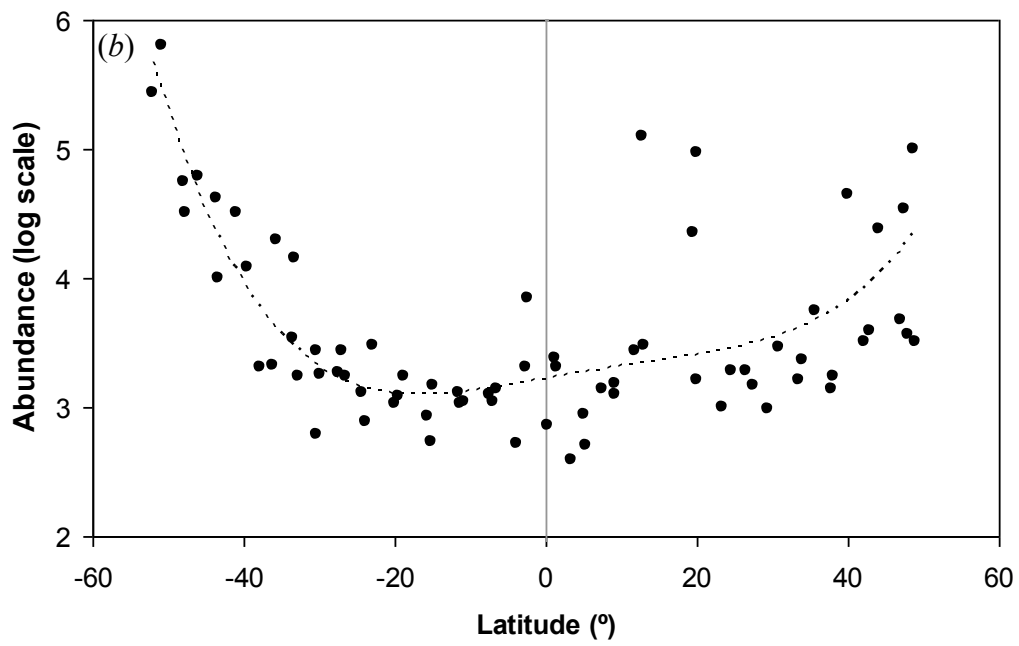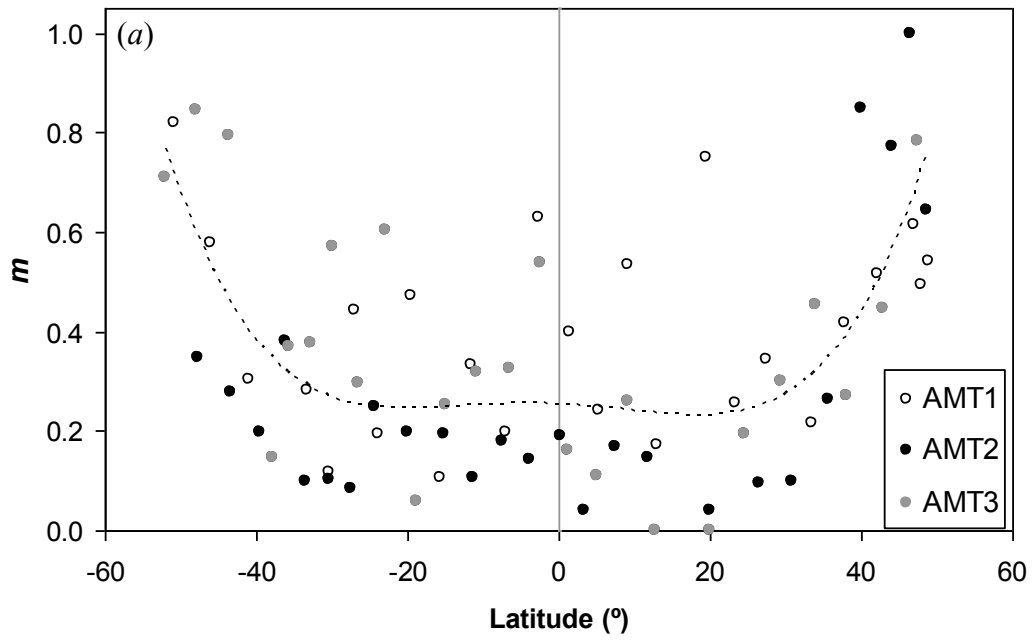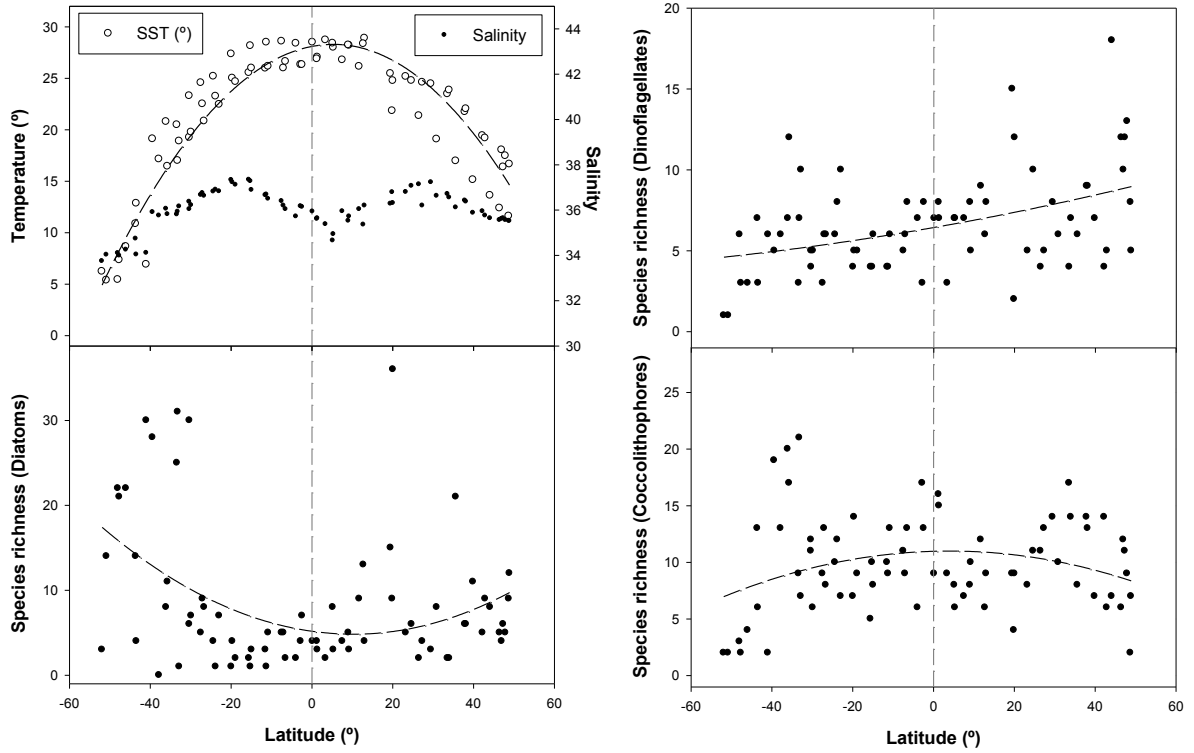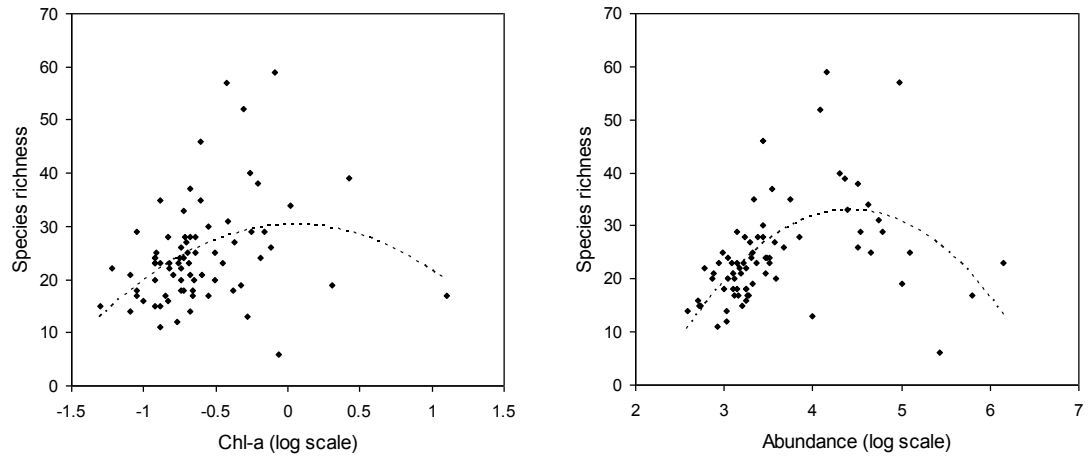of diversity; θ: the fundamental biodiversity parameter; *m*: species immigration probability of a local community from the metacommunity; *p*-value: probability of the neutrality test based upon Shannon's index.

Zone	AMT survey and station	*J*	S	H	θ	*m*	*p*-value
Northern zone	AMT1.1	3196	24	1.155	3.63541	0.54153	0.009
Northern zone	AMT1.2	3647	27	1.583	4.13152	0.49229	0.050
Northern zone	AMT1.3	4718	26	1.723	3.69217	0.61449	0.141
Northern zone	AMT1.4	3224	23	0.939	3.46402	0.51623	0.003
Northern zone	AMT1.5	1391	29	2.065	5.64489	0.41876	0.082
Northern zone	AMT1.6	1641	23	1.836	4.44645	0.21660	0.092
Northern zone	AMT1.7	1469	22	1.665	4.03803	0.34346	0.073
Tropical	AMT1.8	998	18	1.814	3.57938	0.25734	0.256
Tropical	AMT1.9	22635	38	0.850	4.41010	0.74953	0.000
Tropical	AMT1.10	2975	21	1.730	3.58023	0.17204	0.163
Tropical	AMT1.11	1250	18	1.120	3.07614	0.53425	0.013
Tropical	AMT1.12	509	16	1.973	3.74556	0.24146	0.414
Tropical	AMT1.13	2068	25	1.360	4.30817	0.39897	0.004
Tropical	AMT1.14	2040	24	1.387	3.90031	0.62754	0.017
Tropical	AMT1.15	1110	20	1.344	4.16201	0.19894	0.010
Tropical	AMT1.16	1314	17	1.455	3.00126	0.33475	0.118
Tropical	AMT1.17	842	11	1.559	2.26946	0.10769	0.406
Tropical	AMT1.18	1197	23	2.022	4.29514	0.47035	0.260
Tropical	AMT1.19	760	21	1.719	4.97221	0.19558	0.029
Southern zone	AMT1.20	2692	28	0.960	4.63056	0.44443	0.000
Southern zone	AMT1.21	2758	46	2.728	10.51770	0.11720	0.133
Southern zone	AMT1.22	14506	59	2.167	8.70780	0.28042	0.005
Southern zone	AMT1.23	32179	35	1.366	4.59466	0.30258	0.011
Southern zone	AMT1.24	61663	29	1.297	3.12250	0.57787	0.036
Southern zone	AMT1.25	630258	15	0.170	1.54189	0.82094	0.000
Southern zone	AMT2.1	32454	26	0.651	2.96950	0.34781	0.002
Southern zone	AMT2.2	9873	13	1.228	1.54473	0.27934	0.357
Southern zone	AMT2.3	12255	52	2.080	8.03276	0.19643	0.007
Southern zone	AMT2.4	2129	35	1.813	6.54618	0.38104	0.008
Southern zone	AMT2.5	3436	37	2.452	7.69979	0.09899	0.167
Southern zone	AMT2.6	608	22	2.424	6.72432	0.10233	0.561
Southern zone	AMT2.7	1830	17	1.751	3.43551	0.08314	0.240
Tropical	AMT2.8	1304	20	1.944	3.85101	0.24988	0.313
Tropical	AMT2.9	1053	12	1.666	2.17805	0.19631	0.541
Tropical	AMT2.10	532	15	1.519	3.52518	0.19569	0.066
Tropical	AMT2.11	1058	14	1.783	2.95684	0.10465	0.437
Tropical	AMT2.12	1238	21	2.055	4.37661	0.18101	0.323
Tropical	AMT2.13	515	15	1.992	3.81080	0.14274	0.471
Tropical	AMT2.14	729	20	2.032	4.75005	0.18881	0.208
Tropical	AMT2.15	388	14	2.223	5.84715	0.04129	0.871
Tropical	AMT2.16	1390	18	1.844	3.52581	0.16668	0.314
Tropical	AMT2.17	2706	30	2.108	5.86781	0.14536	0.097
Tropical	AMT2.18	1604	15	1.929	3.50292	0.04005	0.548
Northern zone	AMT2.19	1918	17	1.860	3.30704	0.09655	0.379

Northern zone	AMT2.20	2914	24	2.168	4.62893	0.09789	0.386
Northern zone	AMT2.21	5566	35	2.090	5.59426	0.26337	0.108
Northern zone	AMT2.22	44709	24	0.864	2.65332	0.84951	0.013
Northern zone	AMT2.23	24292	33	1.169	3.72220	0.77140	0.008
Northern zone	AMT2.24	1423869	10	0.160	1.05899	1.00000	0.016
Northern zone	AMT2.25	101299	17	1.074	1.56579	0.64474	0.168
Northern zone	AMT3.1	34057	26	0.803	3.31893	0.78361	0.006
Northern zone	AMT3.2	3888	20	1.602	2.87015	0.44806	0.245
Northern zone	AMT3.3	1727	28	1.795	5.45832	0.26896	0.028
Northern zone	AMT3.4	2294	23	1.664	3.74597	0.45542	0.114
Northern zone	AMT3.5	974	25	1.860	5.39180	0.29832	0.032
Tropical	AMT3.6	1930	28	2.149	5.58667	0.19335	0.168
Tropical	AMT3.7	94326	51	0.737	na	0.00031	0.000
Tropical	AMT3.8	125084	18	0.077	na	0.00008	0.000
Tropical	AMT3.9	1548	21	1.990	3.91312	0.25849	0.344
Tropical	AMT3.10	866	23	2.394	6.10525	0.11066	0.565
Tropical	AMT3.11	2373	28	2.240	5.46656	0.16028	0.296
Tropical	AMT3.12	7052	28	1.195	3.81965	0.53635	0.009
Tropical	AMT3.13	1392	23	1.798	4.35725	0.32602	0.110
Tropical	AMT3.14	1114	24	1.973	4.88359	0.31718	0.118
Tropical	AMT3.15	1464	18	1.469	3.27231	0.25196	0.076
Tropical	AMT3.16	1742	16	1.734	3.43278	0.05887	0.257
Tropical	AMT3.17	3046	24	1.134	3.62876	0.60225	0.007
Southern zone	AMT3.18	1761	22	1.656	3.94985	0.29796	0.073
Southern zone	AMT3.19	1789	18	1.642	2.83822	0.57197	0.260
Southern zone	AMT3.20	1740	18	1.456	2.99142	0.37682	0.102
Southern zone	AMT3.21	19936	40	1.752	5.10525	0.37098	0.036
Southern zone	AMT3.22	2060	19	1.734	3.50460	0.14671	0.169
Southern zone	AMT3.23	42038	33	1.187	3.84918	0.79323	0.008
Southern zone	AMT3.24	55292	28	1.073	3.56535	0.84672	0.011
Southern zone	AMT3.25	273563	5	0.339	0.51678	0.70993	0.291