Received Date : 21-May-2015
Revised Date : 14-Dec-2015
Accepted Date : 10-Jan-2016
Article type : Standard Papers
Editor: Natalia Norden
Section: Ecosystems Ecology

Ecological equivalence of species within phytoplankton functional groups

Crispin M Mutshinda[1,*], Zoe V Finkel[2], Claire E. Widdicombe[3], Andrew J Irwin[1]

[1]Department of Mathematics and Computer Science, Mount Allison University, Sackville, NB, Canada

[2]Environmental Science Program, Mount Allison University, Sackville, NB, Canada

[3]Plymouth Marine Laboratory, Prospect Place, Plymouth, PL1 3DH, UK

Corresponding author:

Andrew J Irwin

Department of Mathematics and Computer Science,

Mount Allison University,

Sackville, NB, Canada

airwin@mta.ca

Running head: (45 chars)

Ecological equivalence within phytoplankton functional groups

Summary

1. There are tens of thousands of species of phytoplankton found throughout the tree of life. Despite this diversity, phytoplankton are often aggregated into a few functional groups according to metabolic traits or biogeochemical role. We investigate the extent to which phytoplankton species dynamics are neutral within functional groups.
2. Seasonal dynamics in many regions of the ocean are known to affect phytoplankton at the functional group level leading to largely predictable patterns of seasonal

succession. It is much more difficult to make general statements about the dynamics of individual species.

3. We use a 7 year time-series at station L4 in the Western English Channel with 57 diatom and 17 dinoflagellate species enumerated weekly to test if the abundance of diatom and dinoflagellate species vary randomly within their functional group envelope or if each species is driven uniquely by external factors.
4. We show that the total biomass of the diatom and dinoflagellate functional groups is well predicted by irradiance and temperature and quantify trait values governing the growth rate of both functional groups. The biomass dynamics of the functional groups are not neutral and each has their own distinct responses to environmental forcing. Compared to dinoflagellates, diatoms have faster growth rates, and grow faster under lower irradiance, cooler temperatures, and higher nutrient conditions.
5. The biomass of most species vary randomly within their functional group biomass envelope, most of the time. As a consequence, modelers will find it difficult to predict the biomass of most individual species. Our analysis supports the approach of using a single set of traits for a functional group and suggests that it should be possible to determine these traits from natural communities.

Keywords

demographic stochasticity, diatoms, dinoflagellates, English Channel, neutral model, functional types, time series, traits

Introduction

Functional groups are collections of species that share morphological, physiological, and biochemical traits or other defining characteristics (Iglesias-Rodríguez *et al.* 2002; Pena 2003; Le Quéré *et al.* 2005). Species within a functional group perform similar ecosystem services (e.g., fixing nitrogen) or require similar inorganic and organic resources. Grouping similar species into functional groups simplifies analyses and aids in conceptual and quantitative model building. Phytoplankton communities are enormously diverse and the functional group concept allows the aggregation of thousands of species into only a handful of functional groups. Typically these groups are defined based on a combination of higher

phylogenetic grouping and biogeochemical function (e.g., silicifying diatoms, mixotrophic dinoflagellates, nitrogen fixing and non-nitrogen fixing cyanobacteria, and calcifying coccolithophorids) or cell size (Irwin *et al.* 2006; Hood *et al.* 2006; Finkel *et al.* 2010). The physiological traits of these functional groups used in models are usually based on a few model organisms studied in the lab (Moore *et al.* 2002; Litchman & Klausmeier 2008; Barton *et al.* 2013b).

The success of the functional group concept suggests that the species within functional groups may behave similarly enough to be described by a single set of functional traits. A functional trait is defined as a feature of an organism that can be measured and that influences one or more essential functional processes such as reproduction and growth (Weithoff 2003). Functional traits determine an organism's effects on ecosystem processes and its response to environmental forcing, and reflect adaptations to the abiotic and biotic environments as well as tradeoffs among different functions within an organism. It is not known how to identify functional traits for an entire functional group, whether a small number of species can provide trait values representative of the group, or how to identify potentially representative species.

If species within a functional group are very similar, the abundance or biomass dynamics of each species within a functional group may be neutral relative to the overall dynamics of the group. The idea of ecological neutrality or functional equivalence is that the abundance or biomass of each species at one sampling time, relative to the total community abundance or biomass at that time, is only determined by the species' relative abundance or biomass at the previous sampling time. The species' taxonomic identity and environmental conditions provide no information on the relative contribution of individual species to total abundance or biomass (Volkov *et al.* 2003; Hubbell 2005, 2006; Shipley, Vile & Garnier 2006). In a neutral community, all species are identical on a per capita basis in their demographic properties (birth rate, death rate and immigration rate) and have equal competitive abilities. Consequently, the demographic events underpinning fluctuations in relative species abundance (birth, death, migration) are drawn randomly from any one species in proportion to its abundance, causing relative species' abundances to "drift" upward or downward as a random walk called ecological drift (Hubbell 2001). In other words, a neutral community is one where changes in relative species abundances are

essentially due to demographic stochasticity (or demographic noise) irrespective of species' identities or ecological drift. Species neutrality within a functional group is a restricted sense of this idea: the proportion of a species' biomass relative to its functional group's biomass is a random walk not influenced by taxonomic identity or environmental conditions.

Time-series data of phytoplankton biomass and environmental conditions permit a test of the functional group concept in natural phytoplankton communities. Using time-series data we test if species' biomass dynamics are neutral within functional groups. To carry out this analysis, we first develop a Bayesian model of functional group biomass dynamics and extract values of functional traits for two ecologically dominant and diverse phytoplankton functional groups. We assess if these functional traits can adequately predict changes in the biomass of each functional group with changing environmental conditions. We then test for neutrality of species within each functional group by quantifying the extent to which the biomass dynamics of each species within a functional group is consistent with ecological drift. If many species within a functional group are highly non-neutral then a single set of traits for that functional group is likely inadequate to describe its biomass dynamics and it may be necessary to subdivide the functional group.

Materials and Methods

Time-series data.

We analyzed select time-series data from Station L4 (50° 15.00′N, 4° 13.02′W) that forms part of the Western Channel Observatory (WCO) in the Western English Channel (www.westernchannelobservatory.org.uk). Station L4 is in a coastal, temperate environment with strong seasonal cycles and has one of the longest phytoplankton time-series. We used weekly observations of the abundance (cells L^{-1}) of 57 diatom and 17 dinoflagellate species, genera, or morphological classes (Supplementary Table S1, Southward *et al.* 2004; Harris 2010; Widdicombe *et al.* 2010). This relatively high sampling frequency is essential for our time-series modeling. We focused on diatoms and dinoflagellates because they are ecologically important, diverse groups with high quality and high resolution taxonomic data. We chose not to analyse the phytoflagellates because they were not identified to the species level and their identification was primarily made

based on size. Similarly, coccolithophorids were not analyzed because they were dominated by a single species, *Emiliania huxleyi* (Widdicombe *et al.* 2010). We omitted some infrequently observed diatom and dinoflagellate species to ensure that all species had at least 20 observations and were observed in more than 10% of the sampling weeks. For simplicity we refer to each taxonomic unit identified during sampling as a species. Samples were taken at 10 m depth. Species were identified using light microscopy at 200x or 400x and counted from a 200 mL sample following Utermöhl (1958). Full methods are described for phytoplankton counting in Widdicombe *et al.* (2010) and for temperature and nutrients in Smyth *et al.* (2010). Microscopy was used to obtain an estimate of cell volume for each species (Kovala & Larrance 1966). The total biomass (g C m^{-3}) of diatoms and dinoflagellates is computed as the product of abundance (cells m^{-3}) and a fixed carbon quota for each species (g C $cell^{-1}$) derived from cell volume and an allometric relationship (Menden-Deuer & Lessard 2000). Carbon quota varies over a cell's life cycle and with environmental conditions, but this variation is expected to be much smaller than the variation in abundance, so even though we lack time-resolved variation in carbon quota, we expect that the majority of the variation in phytoplankton biomass is captured by these data. Environmental data consist of *in situ* surface water temperature (°C), surface nitrate concentration (μmol L^{-1}) and irradiance measured continuously at the nearby field station (mol m^{-2} d^{-1}). The time series extends over many years with phytoplankton counts beginning in 1992 at station L4. To maximize the period of weekly data and minimize the large gaps in the time series for these observations (phytoplankton counts and environmental data) we restrict our attention to 349 consecutive weeks spanning April 14, 2003 to December 21, 2009. Linear interpolation was used to establish a regular 7-day grid for all data, commonly to adjust observations that are 5-9 days apart and infrequently to fill in missing data from an unsampled week. Interpolating missing values means we will miss some natural variability in the data, but we expect this will have little impact on the results since only about 10% or less of observations were missing.

Model overview.

We developed models of the biomass for each functional group as a function of environmental conditions and of the biomass of individual species within their functional group (diatom or dinoflagellate) biomass envelope assuming species are neutral within

their functional groups. We included an extra layer between these models, which are based on true biomasses, and the observed biomasses to allow for measurement error and fill gaps in the time-series of the biomass of individual species. Since the biomass of each functional group is empirically log normally distributed (Supplementary Fig. S1) we modeled the log biomass for each group. The true log biomass for each week in the time series was modeled as the true log biomass in the previous week, plus the intrinsic growth rate, linear effects due to temperature, irradiance, and nitrate concentration, and losses due to density-dependent grazing. At a species level, ecological drift is modeled by assuming that the proportion of the functional group biomass attributed to each species is on average the same proportion realized in the previous week. Demographic stochasticity is present in both models; in the neutral species model it is a crucial part of the dynamics while in the functional group model demographic stochasticity is frequently insignificant compared to the forcing provided by environmental conditions. A list of symbols used in the models is provided in Table 1.

Functional group biomass model.

We account for sampling error by modeling the observed log biomass of each functional group y_t (time t, $1 \leq t \leq T$) conditionally on the true log biomass g_t. We assumed that y_t is a realization of a normal distribution with mean g_t through the sampling or observation model

$$y_t \mid g_t, \sigma_1^2 \sim \mathrm{N}(g_t, \sigma_1^2), \qquad (1)$$

where σ_1^2 is an unknown functional-group-level sampling variance to be estimated from the data. The normality assumption on the functional group log-biomasses is theoretically justified by the central limit theorem given that we are dealing with species-rich assemblages with many rare species and only a few abundant ones. We validated the log-normality assumption on the functional group biomasses both graphically (Supplementary Fig. S1) and statistically through the 2-sample Kolmogorov-Smirnov goodness-of-fit test (Supplementary Table S1).

Let $E_t = \{Temp_t, PAR_t, NO3_t\}$ denote the set of environmental conditions at time t where $Temp_t$, PAR_t, and $NO3_t$ indicate, respectively, the temperature, irradiance, and nitrate concentration at time t, standardized to have mean zero and variance one. We modeled the actual log biomass, g_t, of each functional group at times $t = 2,3,\ldots,T$ conditionally on its biomass at the previous sampling time t-1 and environmental conditions E_t using a normal distribution with time-dependent mean μ_t as

$$g_t \mid g_{t-1}, E_t \sim N(\mu_t, \sigma_g^2), \qquad (2)$$

where $\mu_t = g_{t-1} + r + \delta g_{t-1} + \beta_1 Temp_t + \beta_2 PAR_t + \beta_3 NO3_t$, $r > 0$ is the intrinsic growth rate (week^{-1}), β_1, β_2 and β_3 are respectively, dimensionless estimates of the temperature, irradiance, and nitrate concentration effects on the functional group log biomass, δ is the density-dependence parameter, and σ_g^2 is the functional-group-level process variance. We use a linear model for μ_t rather than more complex non-linear functions of temperature, irradiance, and nitrate concentrations, since the amount of variation in the environmental data is relatively small and the resulting linear model explained the vast majority of variation in functional group log biomass. From Eq. (2) and the properties of the log-normal distribution, it follows that, conditionally on g_{t-1}, $Temp_t$, PAR_t, $NO3_t$ and δ, the functional group biomass at time t, g_t, is log-normally distributed with mean $\exp(\mu_t + \sigma_g^2/2)$ and variance $\exp(\sigma_g^2 - 1)\exp(2\mu_t + \sigma_g^2)$.

Neutral model for the biomass of each species within each functional group

In parallel with the observation model for functional group biomass we model the observed log biomass of species i, $x_{i,t}$ (times $t = 1,2,\ldots,T$) given its true log-biomass, $s_{i,t}$, as a realization of a normal distribution centered at $s_{i,t}$ through the observation model

$$x_{i,t} \mid s_{i,t} \sim N(s_{i,t}, \sigma_2^2), \qquad (3)$$

where σ_2^2 is the species-level sampling variance to be evaluated from the data. A key advantage of the Bayesian approach is the ease with which we accommodate missing values since Bayesian inference makes no distinction between missing data and

parameters. In a Bayesian framework, all unobserved quantities are assigned priors and estimated from the data. In OpenBUGS one simply needs to extend the model specification with priors on the missing observations which are then automatically imputed with samples from their posterior predictive distributions (Thomas *et al.* 2006; Gelman *et al.* 2013)

Under the neutrality assumption of the biomass of species i drifting randomly within its functional group biomass envelope, the expected biomass $M_{i,t}$ of species i at time t, conditional on the biomass of the functional group, G_t , is

$$M_{i,t} = \gamma_{i,t-1} G_t \quad (4)$$

where $\gamma_{i,t} = S_{i,t} / G_t$ is the proportion of the functional group biomass at time t (t = 1, ..., T) due to species i , $S_{i,t} = \exp(s_{i,t})$ is the true biomass of species i at time t, and G_t is the true functional group biomass at time t, with $g_t = \log(G_t)$. Under this model the true biomass of species i is solely determined by demographic stochasticity (random drift). Hence, conditionally on G_t and $\gamma_{i,t-1}$, we assume the underlying log-biomass of the ith species at time t is drawn from a normal distribution with mean $m_{i,t} = \log(M_{i,t})$ and variance $\sigma^2_{i,t} = v^2_{i,t} / S_{i,t-1}$,

$$s_{i,t} \mid G_t, \gamma_{i,t-1} \sim N(m_{i,t}, \sigma^2_{i,t}) \mathrm{I}(s_{i,t} < g_t). \quad (5)$$

Note that the variance $\sigma^2_{i,t}$ of the focal species' log-biomass at time t is defined as the species' demographic variance, $v^2_{i,t}$, scaled inversely with the species' biomass at the previous sampling time used as a proxy for population size, resulting in a higher variability of low biomasses and *vice-versa*.

Specification of priors

We complete the model specification with explicit statements of fairly uninformative priors on the model parameters and the initial functional group and species biomasses. We placed on the functional-group-level intrinsic growth rate, r , a normal distribution centred at zero with variance 10, truncated at zero to exclude negative values i.e., $r \sim \mathrm{N}(0, 10)\, \mathrm{I}(r > 0)$, where the function $\mathrm{I}(.)$ denotes the indicator function which takes the value 1 when its

argument is true and the value zero otherwise. We assigned centred normal priors with variance 100 on β_1, β_2 and β_3 independently, and a standard normal prior on the density-dependence parameter δ. We imposed $\text{Inverse}-\text{Gamma}(1,1)$ priors independently on each of the variance parameters σ_g^2, σ_1^2, and σ_2^2, and an $\text{Inverse}-\text{Gamma}(\omega_{i,1},\omega_{i,2})$ on $\sigma_{i,t}^2$, where $\omega_{i,1}$ and $\omega_{i,2}$ are species-specific demographic parameters. We assigned $\text{Beta}(\alpha_1,\alpha_2)$ priors on $\gamma_{i,t}$ independently for $t=1,2,...,T$, and independent $\text{Exp}(1)$ distributions on the hyper-parameters α_1, α_2, $\omega_{i,1}$ and $\omega_{i,2}$. Finally, we assumed for initial functional group and species log-biomasses g_1 and $s_{i,1}$ non-informative normal priors centred at zero with variance 1000, i.e., $g_1 \sim N(0,1000)$ and $s_{i,1} \sim N(0, 1000)\ \mathrm{I}(s_1 < g_1)$. The corresponding sampling models are $y_1 \sim N(g_1,\sigma_1^2)$ and $x_{i,1} \sim N(s_{i,1}, \sigma_2^2)\ \mathrm{I}(x_{i,1} < y_1)$ from which $\gamma_{i,1}=\exp(s_{i,1}-g_1)$ and $p_{i,1}=\exp(x_{i,1}-y_1)$ follow.

We use Markov chain Monte Carlo (MCMC) (Gilks, Richardson & Spiegelhalter 1996) to simulate, through OpenBUGS (Thomas *et al.* 2006), the joint posterior distribution of the model parameters. The OpenBUGS code is provided in the Supplementary Material. We ran 20,000 iterations of three parallel Markov chains starting from dispersed initial values, and discarded the first 5,000 samples from each Markov chain as burn-in, thinning the remainder to monitor each 25th sample. We assessed the convergence of the MCMC through visual inspection of traceplots and autocorrelation functions (Supplementary Figs. S2 and S3). The chains reached the target distribution after roughly 1500 iterations. All chains mixed well by jumping freely over the parameter space.

Neutrality index

We evaluated the extent to which individual species biomasses drift randomly within their respective functional groups' biomass envelopes by comparing, for each species *i*, the relative biomass $\gamma_{i,t}=S_{i,t}/G_t$ predicted by the neutral model at time t $(2 \le t \le T)$ with the observed counterpart $p_{i,t}=X_{i,t}/Y_t$. For a species with neutral biomass dynamics, the predicted relative biomass should be evenly spread around the distribution of the observed

relative biomass, so one should be bigger than the other roughly half the time, which we express by testing if $\Pr(\gamma_{i,t} > p_{i,t})$ is close to 0.5. This probability can be straightforwardly evaluated within OpenBUGS defining $\eta_{i,t} = \text{step}(\gamma_{i,t} - p_{i,t})$ where step(u) is 1 if $u > 0$ and 0 otherwise. The posterior mean $\hat{\eta}_{i,t} = \mathrm{E}[\eta_{i,t} \mid data]$ provides an estimate of the required probability $\Pr(\gamma_{i,t} > p_{i,t})$ that takes parameter uncertainty into account. We regard the biomass dynamics of species *i* from time $(t-1)$ to t as consistent with random drift when $0.25 < \hat{\eta}_{i,t} \leq 0.75$. We define a neutrality index ϕ_i as the proportion of $\hat{\eta}_{i,t}$ between 0.25 and 0.75 counting only times Ω_i when species i was observed and not using imputed values of $p_{i,t}$,

$$\phi_i = \frac{1}{\#\Omega_i} \sum_{t \in \Omega_i} I(0.25 < \hat{\eta}_{i,t} \leq 0.75), \qquad (6)$$

where $\#\Omega_i$ indicates the size of the set Ω_i, and use this index to assess the importance of random drift as a driver of individual species' biomass dynamics. We consider $\phi_i \geq 0.75$ and $\phi_i < 0.25$ as supporting respectively the prevalence of random drift and that of non-neutral forces in driving the biomass dynamics of species *i*, whereas $0.25 < \phi_i \leq 0.75$ suggests an interplay of neutral and non-neutral forces in shaping the focal species biomass patterns.

Results

Data overview.

The biomass of diatoms and dinoflagellates at station L4 in the Western English Channel show a strong seasonal cycle, with a great deal of variability superimposed on the annual cycles (Fig. 1). The aggregated biomass (g C m^{-3}) of 57 diatom species and 17 dinoflagellate species varied by a factor of ~1000 and ~10,000 on an annual basis, respectively. The richness of the diatom and dinoflagellate communities is highly variable from week to week with a median richness of 16 and 4 species, respectively. The times of rapid biomass accumulation and maximum biomass density are slightly different between the two functional groups, with diatoms blooming earlier in the season. These strong

seasonal dynamics indicate that it should be possible to predict much of the variation in functional group biomass from environmental conditions.

There are very strong seasonal oscillations in sea surface temperature (°C), irradiance (mol photons m^{-2} d^{-1}), and nitrate concentrations (μmol L^{-1}) (Fig. 1). These three variables strongly affect phytoplankton growth rates, but they are not even approximately independent, so it may not be appropriate to use all three variables in a statistical model. A principal component analysis of these three variables showed that the first two principal components account for 93% of the variation in the three variables. The largest pairwise correlation was between nitrate concentration and irradiance (r = –0.74). The rapid variability and uptake of nitrogen sources means that point estimates of nitrate concentration may not always be a good measure of the nitrate resources available to phytoplankton. After exploring models with all three predictors and omitting either nitrate concentration or irradiance, we decided to omit nitrate concentration (Eq. 2, set $\beta_3 = 0$) from our final model of functional group biomass. Omitting nitrate concentration makes the interpretation of the coefficients for temperature and irradiance more straighforward.

Functional group biomass.

Our models of functional group biomass using temperature and irradiance explained 96% and 98% of the temporal variation in the diatom and dinoflagellate log-biomasses, respectively, with model predictions close to the observed data (Supplementary Fig. S4). The residuals from the model of functional group log-biomass, y_t - g_t, are clustered around zero with no apparent trend and no serial correlation (Supplementary Fig. S5), showing that our model assumptions about the functional group biomass dynamics are sensible and that the seasonal cycles in the functional group biomass data are largely explained by fluctuations in temperature and irradiance.

We obtained a posterior mean for four functional group trait parameters in each functional group model: intrinsic growth rate (r, $week^{-1}$), the effect of temperature on growth (β_1 , dimensionless), the effect of irradiance on growth rate (β_2 , dimensionless), and a coefficient measuring the effect of increased biomass density on growth (δ , dimensionless). The growth parameters in the functional group model (Eq. 2) are well

estimated with narrow posterior distributions and differ between diatoms and dinoflagellates (Fig. 2). Our neutrality model (see below) yielded replicate estimates of each functional group trait parameter (Fig. 2, symbols and 95% credible intervals).

The intrinsic growth rate of diatoms is much higher than the intrinsic growth rate for dinoflagellates (Fig. 2). The net growth rate includes linear loss terms such as respiration but is unaffected by quadratic loss terms such as density-dependent grazing or viral attack which is described by the density-dependent parameter δ. Both functional groups exhibit an increase in their growth rates under elevated temperatures and irradiance but the sensitivity to changes in temperature and irradiance is much greater for dinoflagellate biomass than for diatom biomass. Both functional groups experience strong density dependent limitation on their growth rates as expected from the annual cycle of biomass blooms (Fig. 1). The magnitude of this effect is larger for diatoms, but only differs by about 25% between functional groups which is a smaller relative difference than between any of the other parameters.

Testing the neutrality of species within functional groups.

Our neutrality index ϕ (Eq. 6) measures the proportion of time the week-to-week change in the biomass of a species, relative to the total biomass of its functional group, is well predicted by random drift or neutral model (Fig. 3). Five dinoflagellate and three diatom species are clearly neutral within their functional group biomass envelopes (species with $\phi \geq 0.75$ in Fig. 3). None of the species are clearly driven by non-neutral processes ($\phi \leq 0.25$). The majority of species in both functional groups have a neutrality index between 0.4 and 0.7 indicating that most species exhibit a mixture of neutral and non-neutral dynamics in line with the neutrality-niche continuum hypothesis (Gravel *et al.* 2009, Mutshinda & O'Hara 2011). However, the distribution of this index is skewed to the right for both functional groups implying that most species fall on the neutrality side of the continuum. The cell size and median abundance of these neutral species do not stand out from the rest of the community, but there is an overall trend for species that are observed less frequently to have a larger neutrality index.

Our model linking a species' observed biomass to its true biomass (Eq. 3) allowed us to impute observations of a species when its abundance was below the detection threshold of the sampling protocol, but these data were not used in the computation of the neutrality index of a species as they were not independent of the neutrality hypothesis. Since most species were not observed in the weeks their abundance was low, we do not have a good estimate of the neutrality of species at those times. Thus a more nuanced interpretation of our result is that most species are neutral within functional groups most of the time during the times of the year they are most abundant. There is no significant correlation between a species' neutrality index and its cell size, carbon quota or time-averaged log biomass.

Discussion

The phytoplankton community at station L4 in the Western English Channel is highly dynamic with a strong seasonal oscillation in biomass and community structure (Southward *et al.* 2004). Our analysis clearly establishes that the biomass dynamics of diatoms and dinoflagellates are well characterized by functional-group level traits. These two functional groups respond distinctly to changes in environmental conditions over time due to differences in four eco-physiological traits: intrinsic growth rate (r), temperature (β_1) and irradiance (β_2) effects on growth rate, and a density dependent term (δ) (Fig. 2). Diatoms have relatively large intrinsic growth rates compared to the dinoflagellates, which is consistent with many previous studies (Furnas 1991; Raven, Finkel & Irwin 2005; Irwin, Nelles & Finkel 2012). The intrinsic growth rates represent the net density-independent growth rate of the total biomass of each functional group at the average environmental conditions (temperature and irradiance) and as a result are much smaller than maximum growth rates of individual species in culture conditions. Dinoflagellates were relatively more responsive to warming compared with diatoms. This is consistent with earlier results at this site (Widdicombe *et al.* 2010), across the North Atlantic (Irwin *et al*. 2012), and at a tropical Caribbean site (Mutshinda, Finkel & Irwin 2013a; Mutshinda *et al.* 2013b). This supports the hypotheses of phytoplankton communities restructuring and increasing dominance of dinoflagellates in a warming world (Leterme *et al.* 2005; Finkel *et al.* 2010). Both functional groups at station L4 are strongly affected by density dependent loss rates,

which are likely a combination of grazing pressure, aggregation, sinking, and viral attack, but the pressure on diatoms is about 25% stronger than that on dinoflagellates, consistent with their higher growth rates and similar maximum biomass densities for the two groups. Since temperature and irradiance are correlated with nitrate concentration and in some cases water column stability, sensitivity to temperature and irradiance may reflect functional group responses to these and other correlated variables. In aggregate these traits indicate that diatoms will have higher net biomass accumulation under lower irradiances and lower temperatures and higher nitrate concentrations compared to the dinoflagellates that will tend to have higher increases in biomass in warmer more stratified waters with higher irradiance and lower nitrate concentrations, consistent with previous lab and field work (Smayda & Reynolds 2003; Smayda & Trainer 2010; Irwin *et al* 2012; Barton *et al.* 2013a; Xie *et al.* 2015; Brun *et al.* 2015).

Consistent with previous work (Litchman *et al.* 2007), these results clearly show that phytoplankton communities are structured by traits at the functional group level. Despite the success of predicting phytoplankton blooms at the functional group level and knowledge of traits at the species level (Irwin *et al.* 2012; Edwards *et al.* 2013), it has been very difficult to consistently predict the occurrence of particular species of interest, e.g., the prediction of toxic species blooms (Zingone & Oksfeldt Enevoldsen 2000; Hallegraeff 2010). We hypothesized that the individual species within these functional groups are varying neutrally relative to the biomass envelope of their respective functional group. Our analysis shows that at Station L4, the vast majority of diatom and dinoflagellates species are driven more by neutral dynamics, relative to their total functional group biomass, than by non-neutral factors (Fig. 3). This result provides a resolution to the apparent paradox of the predictability of functional groups and non-predictability of species: stochastic variation dominates the dynamics at the species level within functional groups.

The biomass of diatom and dinoflagellates species relative to the total biomass of their functional groups are affected by a combination of niche-selecting and neutral processes, with the net effect that most species behave neutrally within their functional groups 50 to 75% of the time. Phytoplankton species richness at Station L4 is also determined by a combination of niche and neutral processes (Vergnon, Dulvy & Freckleton 2009). Species will be niche selected within their functional groups when they have unique

traits that affect their biomass dynamics under transient conditions not included in our model, such as susceptibility to viral attack and grazing or the formation of resting stages. The dominance of neutral processes within functional groups is consistent with the observation that many North Atlantic diatom and dinoflagellate species have wide niches relative to environmental variation (Irwin *et al.* 2012). Wide niches and niche overlap among species weakens the effect of niche selection and means that demographic stochasticity is a very important source of biomass variation at the species level. The dominance of neutral processes on the biomass dynamics of phytoplankton species means that it will be challenging to predict the biomass of individual species relative to functional group biomass.

Niche processes control the biomass dynamics of phytoplankton functional groups. By contrast, within their functional groups, phytoplankton species are more often than not ecologically equivalent. As a consequence aggregation of species into functional groups is a sensible approach for modeling how phytoplankton communities respond to environmental forcing. Aggregating species together within functional groups averages out species-level demographic stochasticity. When considering how to identify trait values to define the different phytoplankton functional groups we must take into account that there is some niche differentiation some of the time among species within functional groups. As a result it is risky to use traits from a single species to represent a functional group. In addition, traits determined from lab studies may not adequately describe a functional group in the field due to acclimation to multiple environmental conditions and biotic interactions among species. Rather than use the average of trait values from a few species studied in the lab to represent a functional group, or a broad range of trait values from many species to represent each functional group (Follows et al 2007), potentially the best approach will be to find trait values representative of functional groups as a whole using field data as we have done here or using more complex mechanistic models.

Acknowledgements

Station L4 phytoplankton biomass and environmental data were provided by the Western Channel Observatory which was funded as part of the UK Natural Environmental Research Council's National Capability. We thank A Atkinson and T Smyth for assistance and advice

and D Talmy and anonymous reviewers for comments on an earlier draft. ZVF was supported by the National Science and Engineering Research Council (NSERC) of Canada and the Canada Research Chairs program. AJI was supported by NSERC.

Data accessibility

Source data were obtained from and are available from the Western Channel Observatory www.westernchannelobservatory.org.uk. The L4 environmental data are available at the British Oceanographic Data Centre www.bodc.ac.uk (doi: 10.5285/f1968a39-26bf-55fe-e044-000b5de50f38). OpenBUGS code is available in the supplementary online information.

References

Barton, A.D., Finkel, Z.V., Ward, B.A., Johns, D.G. & Follows, M.J. (2013a) On the roles of cell size and trophic strategy in North Atlantic diatom and dinoflagellate communities. *Limnology and Oceanography*, **58**, 254–266.

Barton, A.D., Pershing, A.J., Litchman, E., Record, N.R., Edwards, K.F., Finkel, Z.V., Kiørboe, T. & Ward, B.A. (2013b) The biogeography of marine plankton traits. *Ecology letters*, **16**, 522–534.

Brun, P., Vogt, M., Payne, M.R., Gruber, N., O'Brien, C.J., Buitenhuis, E.T., Le Quéré, C., Leblanc, K. & Luo, Y.-W. (2015) Ecological niches of open ocean phytoplankton taxa. *Limnology and Oceanography*, 1020-38.

Edwards, K.F., Litchman, E. & Klausmeier, C.A. (2013) Functional traits explain phytoplankton community structure and seasonal dynamics in a marine ecosystem. *Ecology Letters*, **16**, 56–63.

Finkel, Z.V., Beardall, J., Flynn, K.J., Quigg, A., Rees, T.A.V. & Raven, J.A. (2010) Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research*, **32**, 119–137.

Furnas, M.J. (1991) Net in situ growth rates of phytoplankton in an oligotrophic, tropical shelf ecosystem. *Limnology and Oceanography*, **36**, 13–29.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2013) *Bayesian Data Analysis*. CRC press.

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J. (1996) *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC.

Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2009) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters*, **9**, 399–409.

Hallegraeff, G.M. (2010) Ocean climae change, phytoplankton community responses, and harmful algal blooms: A formidable predictive challenge. *Journal of phycology*, **46**, 220–235.

Harris, R. (2010) The L4 time-series: the first 20 years. *Journal of Plankton Research*, **32**, 577–583.

Hood, R.R., Laws, E.A., Armstrong, R.A., Bates, N.R., Brown, C.W., Carlson, C.A., Chai, F., Doney, S.C., Falkowski, P.G., Feely, R.A., Friedrichs, M.A.M., Landry, M.R., Keith Moore, J., Nelson, D.M., Richardson, T.L., Salihoglu, B., Schartau, M., Toole, D.A. & Wiggert, J.D. (2006) Pelagic functional group modeling: Progress, challenges and prospects. *Deep Sea Research Part II: Topical Studies in Oceanography*, **53**, 459–512.

Hubbell, S.P. (2005) Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional ecology*, **19**, 166–172.

Hubbell, S.P. (2006) Neutral theory and the evolution of ecological equivalence. *Ecology*, **87**, 1387–1398.

Iglesias-Rodríguez, M.D., Brown, C.W., Doney, S.C., Kleypas, J., Kolber, D., Kolber, Z., Hayes, P.K. & Falkowski, P.G. (2002) Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids. *Global Biogeochemical Cycles*, **16**, 1100.

Irwin, A.J., Finkel, Z.V., Schofield, O.M. & Falkowski, P.G. (2006) Scaling-up from nutrient physiology to the size-structure of phytoplankton communities. *Journal of Plankton Research*, **28**, 459–471.

Irwin, A.J., Nelles, A.M. & Finkel, Z.V. (2012) Phytoplankton niches estimated from field data. *Limnology and Oceanography*, **57**, 787.

Kovala, P.E. & Larrance, J.D. (1966) *Computation of Phytoplankton Cell Numbers, Cell Volume, Cell Surface and Plasma Volume per Litrer, from Microscopical Counts*. DTIC Document.

Leterme, S. C., Edwards, M., Seuront, L., Attrill, M. J., Reid, P. C., & John, A. W. G. (2005) Decadal basin-scale changes in diatoms, dinoflagellates, and phytoplankton color across the North Atlantic. *Limnology and Oceanography*, **50**, 1244-1253.

Litchman, E. & Klausmeier, C.A. (2008) Trait-based community ecology of phytoplankton. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 615–639.

Litchman, E., Klausmeier, C.A., Schofield, O.M. & Falkowksi, P.G. (2007) The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecology Letters*, **10**, 1–12.

Menden-Deuer, S. & Lessard, E.J. (2000) Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnology and Oceanography*, **45**, 569–579.

Moore, J.K., Doney, S.C., Kleypas, J.A., Glover, D.M. & Fung, I.Y. (2002) An intermediate complexity marine ecosystem model for the global domain. *Deep-Sea Research II*, **49**, 403–462.

Mutshinda, C.M., Finkel, Z.V. & Irwin, A.J. (2013a) Which environmental factors control phytoplankton populations? A Bayesian variable selection approach. *Ecological Modelling*, **269**, 1–8.

Mutshinda, C.M., Troccoli-Ghinaglia, L., Finkel, Z.V., Müller-Karger, F.E. & Irwin, A.J. (2013b) Environmental control of the dominant phytoplankton in the Cariaco basin: a hierarchical Bayesian approach. *Marine Biology Research*, **9**, 247–261.

Mutshinda, C.M. & O'Hara, R.B. (2011) Integrating the niche and neutral perspectives on community structure and dynamics. *Oecologia,* **166**, 241-251.

Pena, M.A. (2003) Plankton size classes, functional groups and ecosystem dynamics: an introduction. *Progress in Oceanography*, **57**, 239–242.

Le Quéré, C., Harrison, S.P., Colin Prentice, I., Buitenhuis, E.T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X. & others. (2005) Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, **11**, 2016–2040.

Raven, J.A., Finkel, Z.V. & Irwin, A.J. (2005) Picophytoplankton: bottom-up and top-down controls on ecology and evolution. *Vie et Milieu*, **55**, 209–215.

Shipley, B., Vile, D. & Garnier, É. (2006) From plant traits to plant communities: A statistical mechanistic approach to biodiversity. *Science*, **314**, 812–814.

Smayda, T.J. & Reynolds, C.S. (2003) Strategies of marine dinoflagellate survival and some rules of assembly. *Journal of Sea Research*, **49**, 95–106.

Smayda, T.J. & Trainer, V.L. (2010) Dinoflagellate blooms in upwelling systems: Seeding, variability, and contrasts with diatom bloom behaviour. *Progress in Oceanography*, **85**, 92–107.

Smyth, T.J., Fishwick, J.R., AL-Moosawi, L., Cummings, D.G., Harris, C., Kitidis, V., Rees, A., Martinez-Vicente, V. & Woodward, E.M.S. (2010) A broad spatio-temporal view of the Western English Channel observatory. *Journal of Plankton Research*, **32**, 585–601.

Southward, A.J., Langmead, O., Hardman-Mountford, N.J., Aiken, J., Boalch, G.T., Dando, P.R., Genner, M.J., Joint, I., Kendall, M.A., Halliday, N.C., Harris, R.P., Leaper, R., Mieszkowska, N., Pingree, R.D., Richardson, A.J., Sims, D.W., Smith, T., Walne, A.W. & Hawkins, S.J. (2004) Long-term oceanographic and ecological research in the Western English Channel. *Advances in Marine Biology* pp. 1–105. Academic Press.

Thomas, A., O'Hara, B., Ligges, U. & Sturtz, S. (2006) Making BUGS open. *R news*, **6**, 12–17.

Utermöhl, H. (1958) Zur vervollkommnung der quantitativen phytoplankton-methodik. *Mitt. int. Ver. theor. angew. Limnol.*, **9**, 1–38.

Vergnon, R., Dulvy, N.K. & Freckleton, R.P. (2009) Niches versus neutrality: uncovering the drivers of diversity in a species-rich community. *Ecology Letters*, **12**, 1079–1090.

Volkov, I., Banavar, J.R., Hubbell, S.P. & Maritan, A. (2003) Neutral theory and relative species abundance in ecology. *Nature*, **424**, 1035–1037.

Weithoff, G. (2003) The concepts of 'plant functional types' and 'functional diversity' in lake phytoplankton – a new understanding of phytoplankton ecology? Freshwater Biology 48: 1669-1675

Widdicombe, C.E., Eloire, D., Harbour, D., Harris, R.P. & Somerfield, P.J. (2010) Long-term phytoplankton community dynamics in the Western English Channel. *Journal of Plankton Research*, **32**, 643–655.

Xie, Y., Tilstone, G.H., Widdicombe, C., Woodward, E.M.S., Harris, C. & Barnes, M.K. (2015) Effect of increases in temperature and nutrients on phytoplankton community structure and photosynthesis in the western English Channel. *Marine Ecology Progress Series*, **519**, 61–73.

Zingone, A. & Oksfeldt Enevoldsen, H. (2000) The diversity of harmful algal blooms: a challenge for science and management. *Ocean & Coastal Management*, **43**, 725–748.

Figure captions

Figure 1.

Key environmental conditions affecting phytoplankton growth rate at Station L4 and biomass of phytoplankton functional groups. (a) temperature (°C), (b) irradiance (mol m^{-2} d^{-1}), (c) nitrate concentration (µmol L^{-1}), (d) aggregated biomass of enumerated diatom and (e) dinoflagellate species at station L4 in the Western English Channel, reported weekly for 349 weeks starting on April 1, 2003 through December 21, 2009. Biomass (log_{10}

g C m^{-3}) was computed by summing over all species the product of abundance (cells m^{-3}) and cell carbon inferred from an allometric size-scaling relationship between cell carbon and cell volume, from Menden-Deuer & Lessard (2000).

Figure 2.
Parameters from the growth models for both functional groups. For each parameter the mean (filled circle) and 95% credible intervals (bars) of the parameter estimates are shown for each of the 57 (diatoms) or 17 (dinoflagellates) models estimated for each species.

Figure 3.
Neutrality index for each species, grouped by functional groups (dinoflagellates on top, diatoms at the bottom) and ordered by the value of the neutrality index within groups. A value larger than 0.75 is strong evidence for neutrality, values below 0.25 (none reported) would be strong evidence against neutrality, and the remainder of the results support a temporal mixture of neutral and non-neutral dynamics. Vertical lines highlight the cutoffs of 0.25 and 0.75 (dotted) and the division between neutral and non-neutral at 0.5 (solid).

Table 1.

List of symbols.

Symbol	Interpretation
t	Time (weeks)
i	Species index
y_t, g_t	Observed, true log biomass of a functional group
Y_t, G_t	Observed, true biomass of a functional group
$x_{i,t}$, $s_{i,t}$, $m_{i,t}$	Observed, true, expected log biomass of species i in week t
$X_{i,t}$, $S_{i,t}$, $M_{i,t}$	Observed, true, expected biomass of species i in week t
σ_1^2, σ_g^2	Functional-group level sampling, process variance
σ_2^2, $\sigma_{i,t}^2$, $v_{i,t}^2$	Species level sampling, process, demographic variance
$Temp_t$, PAR_t, $NO3_t$	Temperature, irradiance, and nitrate concentration in week t
μ_t	Modeled mean log biomass
r	Intrinsic growth rate ($week^{-1}$)
δ	Density dependent loss coefficient
β_1, β_2, β_3	Effect of temperature, irradiance, nitrate concentration on growth rate
$p_{i,t}$, $\gamma_{i,t}$	Observed, expected proportion of functional group biomass due to species i
Ω_i	Set of weeks that species i was observed
ϕ_i	Neutrality index for species i