

Supporting Information

Barton et al. 10.1073/pnas.1519080113

SI Text

Continuous Plankton Recorder Data. The continuous plankton recorder (CPR) survey has sampled the abundances of phytoplankton taxa since 1947 across the subpolar and temperate North Atlantic in shelf and pelagic seas (18). The plankton recorder is towed behind ships of opportunity at ~7 m depth on quasi-regular routes, and phytoplankton are captured on a continuously spooling filter mesh as water passes through the instrument. Phytoplankton are identified by microscope upon return to the laboratory, and environmental abundance along the ship transect is estimated from the number of counted cells on the mesh. The coarse size of the filtering mesh (270 μm on a side) means that small, numerous phytoplankton, including the picoplankton, are not enumerated to the species level and that larger cells or colonies of cells are captured more efficiently than smaller ones. Most of the taxa routinely identified to species or genus level are diatoms and dinoflagellates, and we therefore restrict our conclusions to these taxonomic groups (Table S1), which in this region constitute a significant portion of total phytoplankton biomass (58). We used monthly mean CPR species abundance data for 1947–2006 from 41 standard survey areas spanning the North Atlantic and converted monthly average abundance data to presence-only data. The spatial and temporal averaging of the CPR cruise data allows for more complete data coverage (19, 59) but undoubtedly blurs some finer-grain ecological patterns, such as eddies and fronts (60).

Earth System Model and Data. The Geophysical Fluid Dynamics Laboratory's Earth System Model GOLD (ESM2G) (20, 21) was contributed by GFDL to the Fifth Coupled Model Intercomparison Project (CMIP5) coordinated with the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC-AR5). ESM2G exhibits overall high control simulation fidelity to observed regional and seasonal patterns in SST, SSS, MLD, PAR, and macronutrients (20, 21, 61), yet has some relevant biases relative to observations that warrant discussion here. For example, the regional and seasonal patterns of model MLD are quite similar to observations, as measured by both the World Ocean Atlas and from the Argo float dataset (20, 61), yet model MLD can be too deep in the northeast Atlantic and too shallow in the Labrador Sea compared with observations. A similar pattern is apparent for model surface macronutrients (21, 61), which exhibit general regional and seasonal correspondence with observations but high biases in the northeast Atlantic and low biases in the Labrador Sea.

The overall response of niche model drivers in ESM2G to climate change (Fig. S1) includes several robust features across a broad range physical climate models and ESMs. The global trend, which is also apparent in the subtropical North Atlantic, is toward warmer surface ocean temperatures, increased thermal stratification, and decreased mixing depths and consequently a weaker supply of nutrients to the surface (1–3). In the subpolar North Atlantic in waters south of Greenland, many ESMs including ESM2G indicate that surface freshening may lead to increased stratification (3), reduced surface macronutrients (2), and cooling or stable ocean temperatures surrounded by warm-

ing elsewhere (3). Additionally, there is increasing although still equivocal evidence that the lack of long-term warming in the subpolar North Atlantic (62) may be tied to weakening of the Atlantic Meridional Overturning Circulation (63). Changes in PAR are less documented across models, but dynamic spatial changes that characterize ESM2G arise from more predictable SST and sea ice trends (64). Although fine-scale local details will vary across ESMs, our primary results arise from robust, large-scale North Atlantic patterns of projected environmental change shared across models, including ESM2G. Fifty-year averages of model output were used to minimize the effect of climate variability and focus attention on century-scale climate change trends.

Species Distribution Models and Validation. To be included in our analysis, CPR taxa had to meet three criteria. First, we excluded species with fewer than 15 observations (22) because these species are likely to be under-sampled, have restricted and noisy distributions, and are prone to sampling bias errors (65).

Second, we excluded species with mean area under the receiver operating characteristic curve (AUC; *Methods*) lower than 0.7. We used a minimum AUC threshold to identify species for which MaxEnt has difficulty distinguishing between presence and background locations (17), acknowledging that AUC is an imperfect metric for evaluating model goodness of fit (66, 67). Some species with few observations have high AUC as a result of there being little variation in the data to fit with the model. These species may simply be poorly sampled, whereas others may well have restricted spatiotemporal ranges. Species with the most observations (more than 5,000) tended to have AUC in the range 0.5–0.6, which we interpreted as being a result of several factors. Some CPR taxa are agglomerations of several species (e.g., *Thalassiosira* spp.), resulting in high overall observation rates but also making it difficult to predict them with a single model. Some of the taxa had cosmopolitan distributions, resulting in high rates of observations but low model skill. Finally, frequently observed species tended to have quite a bit of variation in whether or not they were present at a particular location among years, but our climatological environmental data are constant and thus unable to account for this variation.

Third, we visually examined how well the SDM for each species matches the underlying, raw CPR data on a season-by-season basis (Fig. S2). For this inspection, we examined only spring (March–May), summer (June–August), and fall (September–November) seasons because there are relatively few winter observations. After removing species with low AUC and sample numbers, we found that the remaining taxa had plausible modeled distributions. The remaining 87 taxa represent a commonly found and sampled subset of the full diversity of the phytoplankton community.

We experimented with a range of threshold values for the minimum allowable number of observations (10–7) and AUC (0.6–0.8) and discovered that our overall findings (e.g., the large-scale features presented in Figs. 1–4) are not sensitive to the choice of thresholds, even if the number of species meeting the criteria varies.

Barton et al. www.pnas.org/cgi/content/short/1519080113

Table S1. Cont.

Taxon name	No. of training samples	Mean AUC
<i>Phalacroma</i> spp.	18	0.90
<i>Planktoniella sol</i>	24	0.97
<i>Podolampas</i> spp.	160	0.86
<i>Podosira stelliger</i>	53	0.86
<i>Pronoctiluca pelagica</i>	17	0.97
<i>Prorocentrum</i> spp.	1,042	0.81
<i>Ptychodiscus noctiluca</i>	15	0.94
<i>Pyrocystis</i> spp.	17	0.94
<i>Pyrophacus</i> spp.	44	0.90
<i>Rhaphoneis amphiceros</i>	221	0.78
<i>Rhizosolenia pungens</i>	24	0.93
<i>Rhizosolenia acuminata</i>	210	0.86
<i>Rhizosolenia alata curvirostris</i>	90	0.85
<i>Rhizosolenia alata indica</i>	1,643	0.78
<i>Rhizosolenia alata inermis</i>	1,899	0.71
<i>Rhizosolenia bergonii</i>	269	0.87
<i>Rhizosolenia calcar-avis</i>	24	0.95
<i>Rhizosolenia delicatula</i>	673	0.71
<i>Rhizosolenia fragilissima</i>	446	0.74
<i>Rhizosolenia setigera</i>	391	0.71
<i>Rhizosolenia stolterfothii</i>	1,363	0.71
<i>Schroederella delicatula</i>	81	0.84
<i>Scrippsiella</i> spp.	897	0.82
<i>Stephanopyxis</i> spp.	239	0.74
<i>Streptotheca tamesis</i>	19	0.90

The taxon names are consistent with those used by the CPR survey. Also included are the number of training samples, or months and standard survey areas where the taxon was present, and the mean AUC score over 100 bootstrap replicates.

Fig. S1. Modeled annual mean sea surface temperature (SST; °C), sea surface salinity (SSS; psu), mixed layer depth (MLD; m), photosynthetically active radiation (PAR; W·m⁻²), surface nitrate (mmol·m⁻³), phosphate (mmol·m⁻³), and silicate (mmol·m⁻³) in 1951–2000 (A) and 2051–2100 (B) and the change from 2051–2100 to 1951–2000 (C). Surface refers to 0–10 m depth.

[Fig. S1](#)

Fig. S2. Median logistic probability in spring (March–May), summer (June–August), and autumn (September–November) for each taxa included in the analysis and results ($n = 87$). The gray circles indicate the frequency that each taxon was present in the survey at each location, with increasing size implying increasing frequency (1.0 = present 100% of the time, 0.5 = present 50% of the time, and 0.01 = present 1% of the time). The thin gray line indicates the overall coverage of the CPR survey.

[Fig. S2](#)

Fig. S3. The latitudinal change in core biogeographical range from 1951–2000 to 2051–2100 for diatoms (A) and dinoflagellates (B). The diatom and dinoflagellate taxa are arranged in descending order by the median latitudinal change over the year, from largest to smallest. The bar plots at right show the median latitudinal change of the core range for each taxa; red (blue) bars indicate species moving northward (southward).

[Fig. S3](#)

Fig. S4. Annual percent change in species richness, comparing future to historical periods (Future – Historical) for all taxa (A), diatoms (B), and dinoflagellates (C). Richness is defined as the number of species exceeding an annual median probability of 0.2 and is expressed as a percentage of retained taxa from the survey ($n = 87$ for all taxa, $n = 48$ for diatoms, and $n = 39$ for dinoflagellates).

[Fig. S4](#)