Supporting Information for *A metacommunity model for phytoplankton biodiversity maintenance with a seed bank and facilitative interactions* by C. Picoche & F. Barraquand.

S1 Intrinsic growth rate modeling

The intrinsic growth rate can be decomposed in 2 elements:

$$r_i(T) = E(T)f_i(T) \tag{1}$$

with $E(T)$ the response to temperature common to all species and $f_i(T)$ the niche part of the model. This section provides detailed information regarding the growth function $E(T)$, the taxon-specific functions $f_i(T)$ that characterize species' differential responses to temperature variations, and species interaction estimates that affect the realized growth rate of phytoplanktonic organisms in a community.

Common response to temperature

Phytoplanktonic growth rates cover a broad range of values: between 0.2 and 1.78 day^{-1} for diatoms in Reynolds (2006), even reaching 3 day^{-1} in the meta-analysis of 308 experiments by Edwards *et al.* (2015). These values are often computed from measurements on isolated species or on small communities in laboratory conditions, in a constant environment. A broader perspective is therefore necessary to understand general responses to changes in the environment (Bissinger *et al.*, 2008; Edwards *et al.*, 2016), especially temperature.

We used the equation by Scranton & Vasseur (2016) as a starting point, but it was not able to reproduce the observed values found in Edwards *et al.* (2015) (see low values of the growth rate in Fig. 1 a). In this context, we decided to use the formula by Bissinger *et al.* (2008) to compute the maximum growth rate response to the temperature. There are two reasons for this choice. First, their model is a general function that can be applied to all species. Second, Bissinger *et al.* (2008) is an update of the seminal work of Eppley (1972) that was used in Scranton & Vasseur (2016).

The relationship between temperature and growth rate is then $E(T) = 0.81\mathrm{e}^{0.0631T_{\circ C}}$, with $T_{\circ C}$ in Celsius degrees. In this case, growth rates vary between 0.81 and 3.9 day^{-1} for temperatures between 0 and 25°C, in line with previous observations. However, these daily growth rates need to be proportional to the daylength as no growth occurs at night: we therefore halve the Bissinger value in our models.

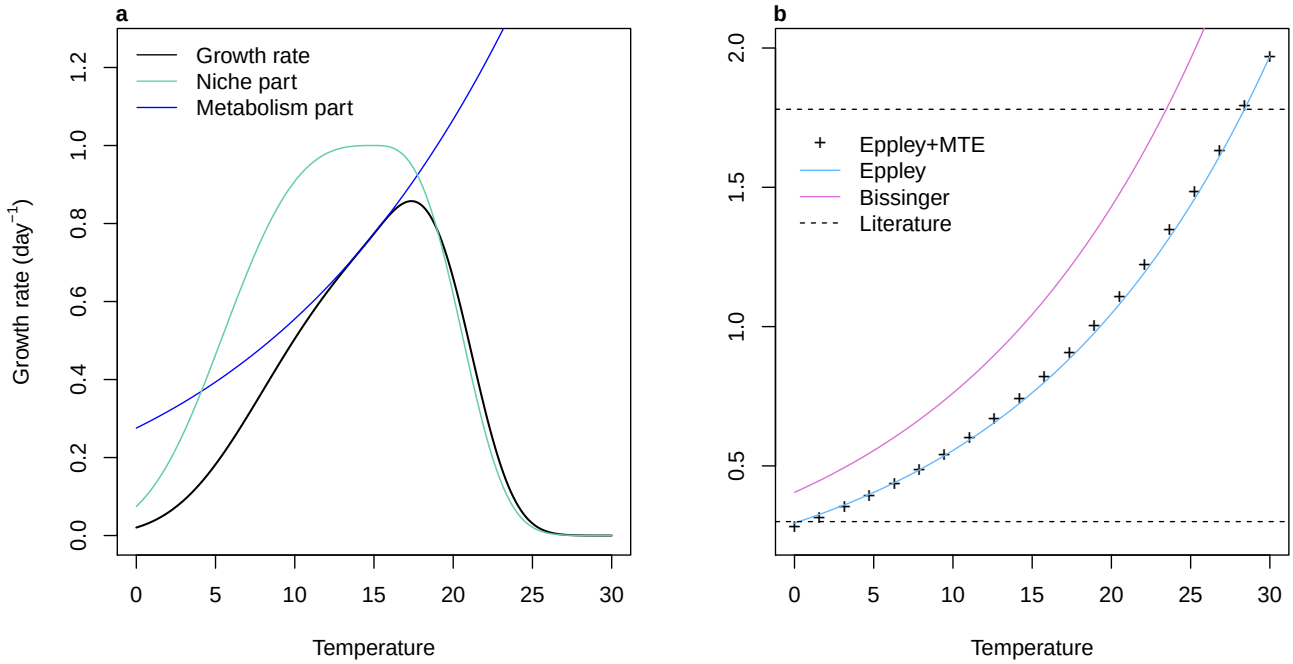

Figure 1: Decomposition of the Scranton & Vasseur (2016) growth rate formula (a). The black line indicates the final growth rate with their model. The blue line corresponds to the taxon-specific response to temperature for a thermal optimum of 15°C and the green line is the maximum achievable growth rate, a composite of the metabolic theory of ecology and the formula by Eppley (1972). This formula is shown by black crosses in (b) and compared to the Eppley (1972) curve in blue and Bissinger *et al.* (2008) formula in purple. Horizontal lines show limits found the literature (Reynolds, 2006).

Taxon-specific response to temperature

The niche part of the growth rate $f_i(T)$ is mainly defined by two parameters: the thermal optimum $T_{K,i}^{opt}$ in Kelvin and a proxy of the niche width b_i, which drives the phenology of the taxon (eq. 2).

$$f_i(T) = \begin{cases} \exp(-|T_K - T_{K,i}^{opt}|^3/b_i), & T_K \leq T_{K,i}^{opt} \\ \exp(-5|T_K - T_{K,i}^{opt}|^3/b_i), & T > T_{K,i}^{opt} \end{cases} \quad (2)$$

The annual dynamics of phytoplanktonic organisms is usually characterized by a blooming period and a lower concentration during the rest of the year. The bloom can be triggered by a combination of nutrient and light input, as well as a sufficient temperature. All parameters being more or less dependent on seasonality, it is reasonable to restrain this study to the effect of temperature.

We base our estimates of $T_{K,i}^{opt}$ and b_i on field observations. For each taxon and each year, the beginning of the bloom of a given taxon is defined by the date at which its abundance exceeds its median abundance over the year. The duration of the bloom is the number of days between the beginning and the date where abundance falls below the median value. Taxa are then separated into two groups. In the field, generalists are characterized by one long bloom in the year or several blooms during which the abundances oscillate around their median. Specialists tend to appear only once or twice in the year for shorter amounts of time. A genus is therefore defined as a generalist if its cumulated blooms over a year last more than the average duration of all blooms (137 days) for at least 15 years over the 20 years of the time series, and as a specialist if they fall below this threshold.

In the models, we assume that generalists have a niche width between 15 and 30°C and (coldwater) specialists, between 5 and 10°C. We assume that temperatures outside of this range lead to a growth rate at least 10 times

inferior to the growth rate obtained at their thermal optimum ($\exp(-|7.5|^3/b_i = 0.1$ for a niche width of 15°C). This leads to values of b_i between 180 and 1500 for generalists, and 7 and 55 for specialists. A set of b values is drawn from a uniform distribution within these boundaries. Meanwhile, taxa are ordered as a function of the mean cumulated bloom length and larger niche values are attributed to longer mean bloom length, i.e. $\sum \overline{L}_{i,b} > \sum L_{j,b} \Rightarrow b_i > b_j$ where $\overline{L}$ is the mean over 20 years of the annual cumulated lengths of the bloom

The thermal optimum T_i^{opt} was first defined as the mean minimum temperature of the bloom throughout the whole time series. However, this value led to blooms occuring mainly in the winter and needed to be increased by 5°C in order to simulate realistic phytoplankton cycles.

It should be noted that a variation in niche width also affects the final shape of thermal preferences. Indeed, when b_i increases, the niche term f_i has smaller variation in values around the thermal optimum. In this case, the final value of the growth rate is driven by the metabolism part of the equation (Fig. 2).

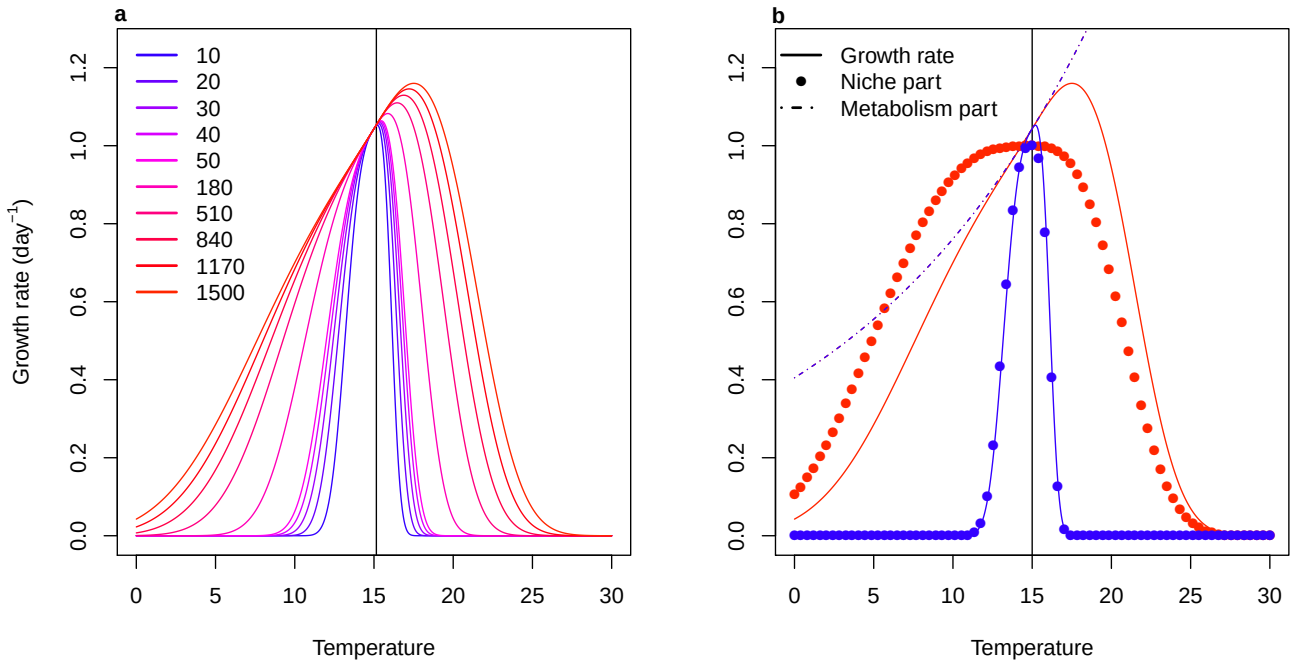

Figure 2: Relationship between daily growth rates and temperature with different values of niche width b (which values are indicated in the legend, corresponding to specialist and generalist taxa) and the same thermal optimum, 15°C, indicated by the solid black line (right). On the left panel, only the two extreme values of b (10 and 1500) are shown in blue and red respectively. Coloured lines then correspond to the final growth rate, points correspond to $f_i(T)$ values (see eq. 1) and dotted lines correspond to $E(T)$ values.

The shift of the maximum growth rate may lead to a deformation of the expected thermal niche of each taxon. We provide actual shapes of $f_i(T)$ for modeled taxa below (Fig 3).

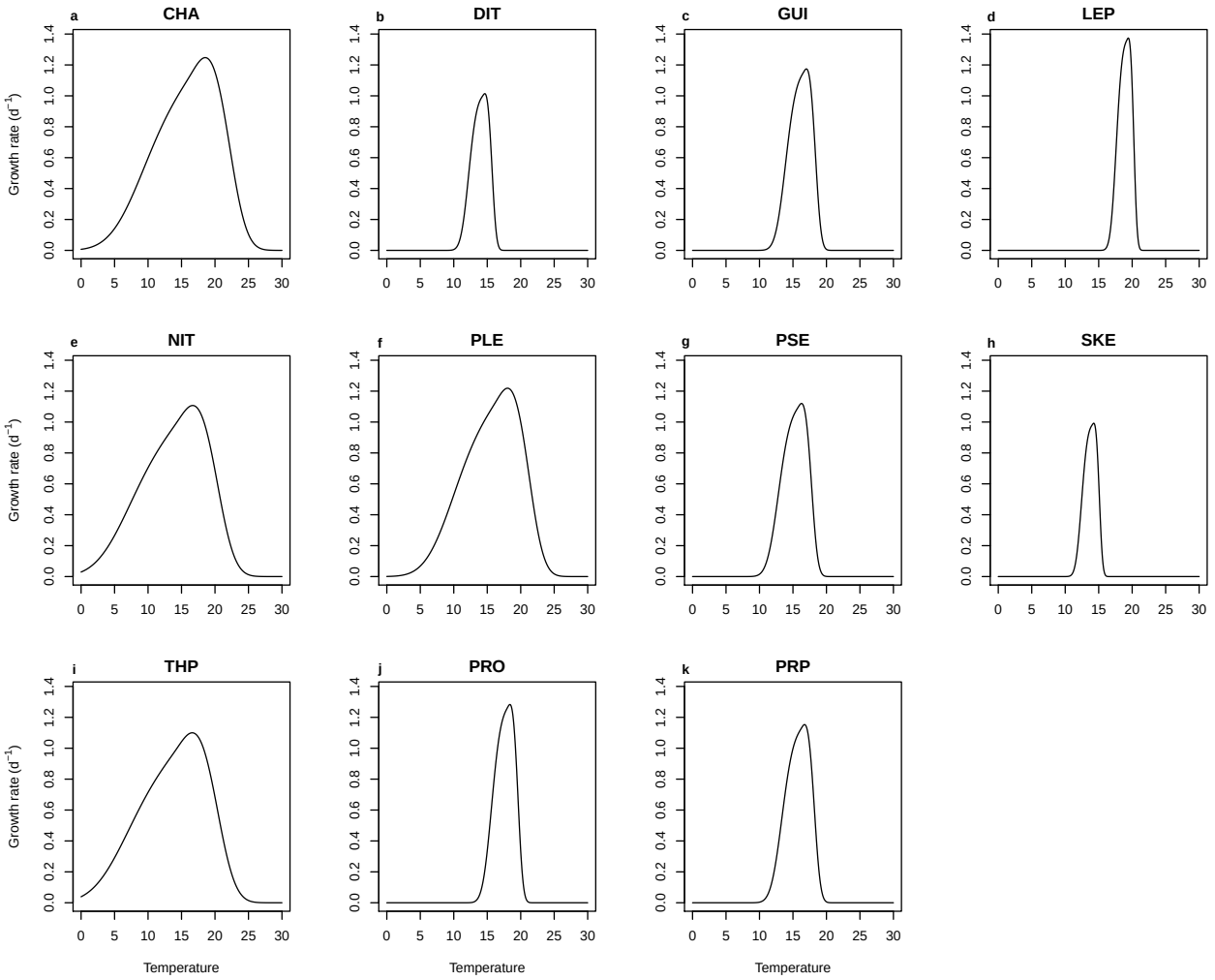

Figure 3: Growth rate as a function of temperature for the taxa used in the model

S2 Initial estimation of interaction values

Model I: Lotka-Volterra interactions

Interactions between taxa have previously been computed with a Multivariate AutoRegressive model (eq. 3, Picoche & Barraquand, 2020).

$$\mathbf{n}_{t+1} = \mathbf{B}\mathbf{n}_t + \mathbf{C}\mathbf{u}_{t+1} + \mathbf{e}_t, \mathbf{e}_t \sim \mathcal{N}_{\mathcal{S}}(0, \mathbf{Q}) \tag{3}$$

where $\mathbf{n}_t$ is the $1 \times S$ vector of log-abundance of phytoplankton taxa, $\mathbf{B}$ is the $S \times S$ interaction matrix with elements b_{ij}, $\mathbf{C}$ is the $S \times V$ environment matrix describing the effects of variables $\mathbf{u}_{t+1}$ on growth rates and the noise $\mathbf{e}_t$ is a $1 \times S$ noise vector following a multivariate normal distribution with a variance-covariance matrix $\mathbf{Q}$. The interaction model we use in the present paper is a Beverton-Holt multispecies model (eq. 1), also called at times Leslie-Gower. In Picoche & Barraquand (2020)'s Supporting information, we showed that MAR and BH interaction coefficients, respectively b_{ij} and α_{ij}, could map once abundances at equilibrium N_i^* are defined.

$$\begin{cases} b_{ii} - 1 = & \frac{-\alpha_{ii} N_i^*}{1+\sum_l \alpha_{il} N_l^*} \\ b_{ij, i \neq j} = & \frac{-\alpha_{ij} N_j^*}{1+\sum_l \alpha_{il} N_l^*} \end{cases}$$

Let's define $\tilde{b_{ij}}$ with $\tilde{b_{ii}} = b_{ii} - 1$, and $f_A(i) = \sum_l \alpha_{ij} N_l^*$.

$$\tilde{b_{ij}}(1 + f_A(i)) = -\alpha_{ij} N_j^*$$

We then sum on columns (on j).

$$\sum_j [\tilde{b_{ij}}(1 + f_A(i))] = -f_A(i)$$

$$\Leftrightarrow -f_A(i)(1 + \sum_j \tilde{b_{ij}}) = \sum_j b_{ij}$$

$$\Leftrightarrow f_A(i) = -\frac{\sum_j \tilde{b_{ij}}}{(1 + \sum_j \tilde{b_{ij}})}$$

$$\Leftrightarrow \alpha_{ij} = -\frac{1}{N_j^*} \tilde{b_{ij}} (1 - \frac{\sum_j \tilde{b_{ij}}}{1 + \sum_j \tilde{b_{ij}}})$$

$$\Leftrightarrow \alpha_{ij} = -\frac{1}{N_j^*} \frac{\tilde{b_{ij}}}{1 + \sum_j \tilde{b_{ij}}}$$

This gives an exact correspondance between α_{ij} and b_{ij}. In the multispecies BH model, the presence of mutualistic interactions can lead to an orgy of mutual benefaction (May, 1981). We impose a minimum value of 1 to the denominator of the BH formulation, meaning that the growth rate cannot be higher than the maximum growth rate calculated $r_i(T)$. We acknowledge the possibility of overyielding, i.e. an increase in growth due to the presence of other species, but the phenomenon seems quite rare for phytoplanktonic communities (see Schmidtke et al. 2010 for underyielding and see Shurin et al. 2014 for an observed but not frequent, and small, overyielding).

Model II: saturating interactions

We now move to a model with saturating interactions between taxa.

$$N_{t+1,i} = \frac{e^{r_i(T)} N_{t,i}}{1 + \sum_{j/a \in \mathbb{C}} \frac{a_C N_{t,j}}{H_{ij} + N_{t,j}} + \sum_{j/a \in \mathbb{F}} \frac{a_F N_{t,j}}{H_{ij} + N_{t,j}}} \tag{4}$$

where coefficients a_C and a_F are the maximum interaction strength for competition and facilitation respectively, H_{ij} coefficients are the abundance of species j to reach half of the maximum effect in the interaction of j on i, and $\mathbb{C}$ and $\mathbb{F}$ are the sets of competitive and facilitative interactions. This formula can be linked with the Unique Interaction Model by Qian & Akçay, 2020, i.e. a model where each taxon provides a unique type of benefit or disadvantage to the focus species.

There is no unique solution for making a correspondence between the **B** matrix of the MAR model and the new formulation including H_{ij}, a_C and a_F. We approximate the maximum interaction strength a_C as the average sum of all species effects $\alpha_{ij} N_i$ exerted on a given species if all interactions were competitive (eq. 5). To compute a_F, we can make two assumptions: on average, a) there is 70% facilitation in our dataset and b) the growth rate $r_i(T)$

should not exceed, as in model I. We consider that the relationships that apply to individual interactions α_{ij} should also apply to the saturation point (eq. 6).

$$a_C = \frac{1}{S}\sum_i \left(\sum_j \alpha_{ij} N_{j,max} \right) \quad (5)$$

$$(1-0.7)a_C + 0.7a_F = 0 \quad (6)$$

At low abundances, we can consider that interactions are far from saturation. Taking the tangent of the function at this point, H_{ij} can be approximated by $f\frac{a_{C/F}}{\alpha_{ij}}$, where $f = 2$ is a correction factor that takes into account the fact that the slope at origin for the type II response is likely higher than the slope for a linear effect of density.

S3 Choice of parameters derived from literature

This section contains additional information on "fixed" parameters definitions and chosen values. Note that these values are then subjected to sensitivity analysis.

Loss rate The loss rate corresponds to multiple mortality processes. The Lotka-Volterra model of Scranton & Vasseur (2016) considered a rate around 0.04 day^{-1}. In Jewson *et al.* (1981), washout (0.5%), parasitism (4% of cells are infested and die) and grazing still remained low (about 0.05%) when compared to growth rates. Li *et al.* (2000) found values between 0.02 and 0.1 day^{-1} for natural mortality only, while a review by Sarthou *et al.* (2005) indicated a loss of daily primary productivity between 0.45 and 1.1 due to grazing only, 0.13 being potentially due to cell autolysis (in the absence of nutrients, or because of viral charge). Trying to make a compromise emerge from the literature above, we set as the reference value to 0.2 day^{-1}.

Sinking rate Among the hydrodynamics processes that drive sinking rate, turbulence and eddies – themselves driven by tidal currents, the shape of the coast or wind conditions – are influential in keeping the cells at the top of the water column. For that reason, laboratory experiments on sinking rates are not sufficient to calibrate a field-based model. We therefore chose sinking rate values from field studies. In the Gotland Basin (central Baltic Sea), Passow (1991) measured a large variability in sinking rates, even within the same genus (e.g., between 1 and 30% for *Chaetoceros* spp.). However, a pattern could be highlighted, with a small number of genera that sank more than the rest of the community. The mean sinking rate for *Chaetoceros* and *Thalassiosira* was around 10% while it was around 1% for the other species. Sinking rate values around 10% are consistent with the loss rates in Kowe *et al.* (1998) in a river and Wiedmann *et al.* (2016) in an estuary (mouth of Adventfjorden). When estimating changes of the sinking rate over time, values between 4 and 50% were obtained (Jewson *et al.*, 1981). We therefore chose to represent the sinking values with a Beta-distribution (Fig. S4) which accounts for observed maximum and mean values, while still allowing a highly skewed distribution of sinking rates between species. High sinking rates are attributed to the morphotypes corresponding to *Chaetoceros* (CHA) and *Thalassiosira* (THP).

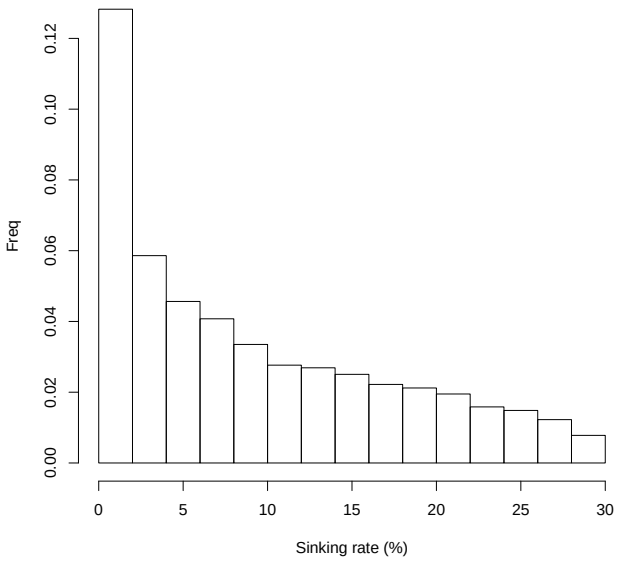

Figure 4: Possible beta-distribution of sinking rates

Cyst mortality and burial McQuoid *et al.* (2002) present maximum and mean depth of sediment at which germination of diatoms and dinoflagellates could still occur when incubated. The authors also present sediment datation according to depth. Depth can therefore be related to maximum and mean age of phytoplankton cysts before death.

Assuming m is the probability of mortality and the survival follows a geometric law, the life expectancy of a cyst is $\frac{1}{m} \Leftrightarrow m = \frac{1}{L_{mean}}$where L_{mean} is the average duration of cyst viability.

Another way to look at the process is that life expectancy L follows the distribution $p(L > l) = e^{-ml}$.With maximum values, we can arbitrarily choose that for these values $p(L > l_{max}) = 0.05$. In this, $m = -\frac{ln(p(L>l_{max}))}{l_{max}}$.In both cases, $m \propto 10^{-4}\mathrm{d}^{-1}$.

As we highlight in the main text, burial is a very important process that controls the availability of cysts, conditional to their survival in the sediment. However, burial rate will be almost entirely context-dependent on the local sedimentation and no generally applicable literature could be found. We varied it between 0.001 and 0.1 per day.

Resuspension As mentioned in the main text, resuspension values are mostly taken from models or data for inorganic particles. Rates vary greatly from one publication to another: in Fransz & Verhagen (1985), in a coastal area, the resuspension rate of sediments is evaluated around 5×10^{-5} day^{-1} in winter and decreases in summer, with a link between resuspension and the light extinction coefficient. In Kowe *et al.* (1998), the resuspension rate of diatoms is evaluated around 1.9×10^{-5} day^{-1}. In Le Pape *et al.* (1999), resuspension rate of sediments and dead diatoms is 0.002 day^{-1}. In this paper, we explore values between 10^{-5} (stratified water column) to 0.1 (highly mixed environment).

Finally, it should be noted that cyst burial, sinking rate and resuspension are all highly contingent upon the local hydrodynamics and therefore represent intermingled processes.

S4 Phytoplankton time series with model II

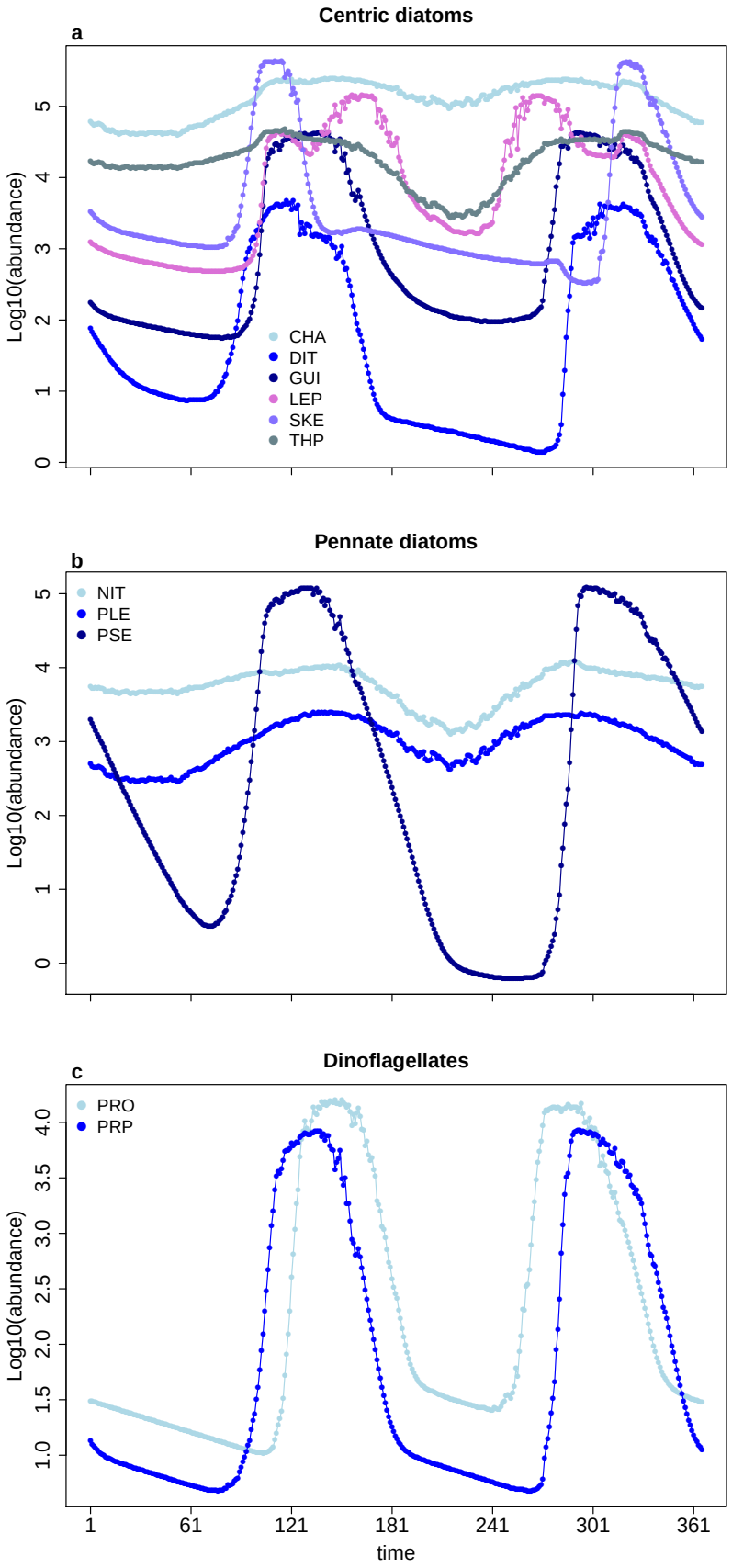

Figure 5: Simulated phytoplankton time series for a year in model II (with saturating interactions). Each panel corresponds to a cluster of interacting species.

S5 Growth rates and survival without a cyst bank

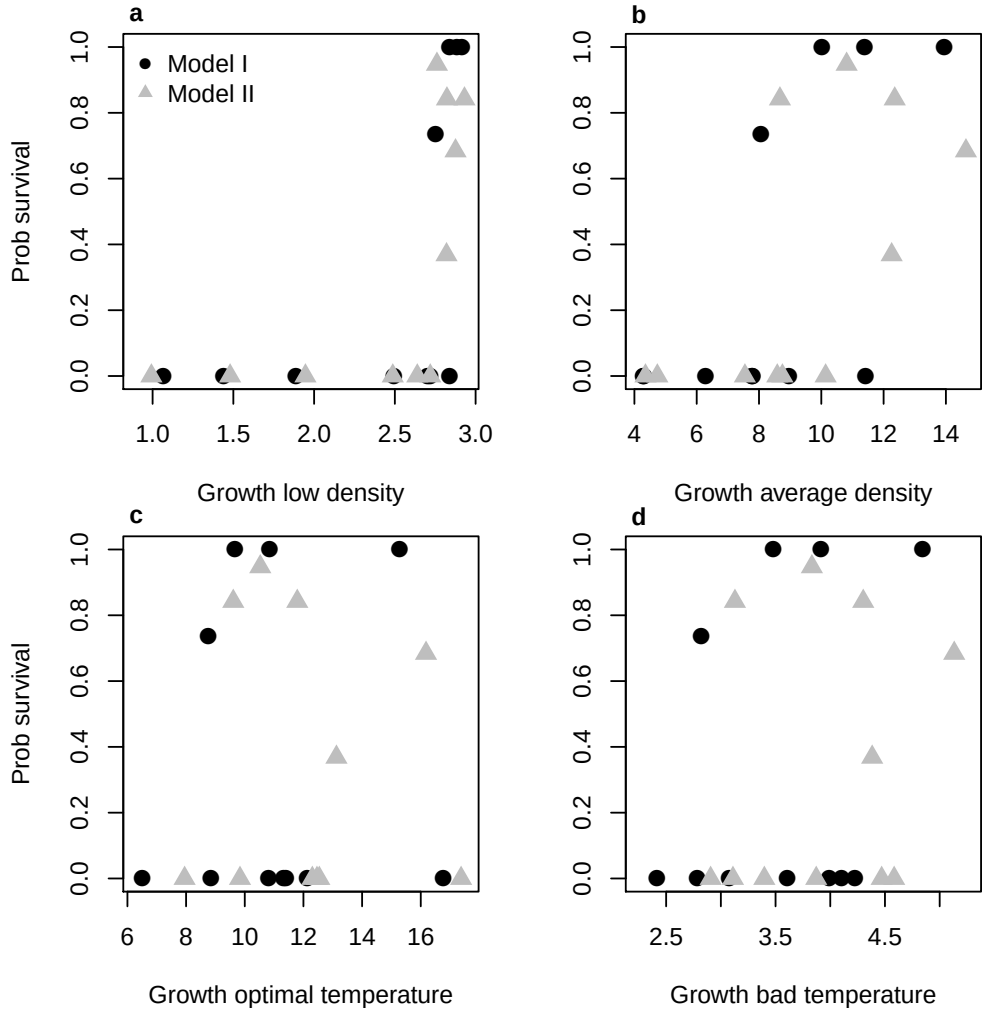

Figure 6: Probability of survival in the absence of a cyst bank as a function of the growth rate in different conditions (mean temperature and low densities, i.e., all species density are 1 C/L, or average densities, i.e., all species are at their simulated average density values; or average densities and optimal temperature for each species, or suboptimal temperatures - here, 30°C). It seems that opportunistic species are the ones with the highest probability of survival.

References

Bissinger, J., Montagnes, D., Harples, J. & Atkinson, D. (2008). Predicting marine phytoplankton maximum growth rates from temperature: Improving on the Eppley curve using quantile regression. *Limnology and Oceanography*, 53, 487–493.

Edwards, K., Thomas, M., Klausmeier, C. & Litchman, E. (2015). Light and growth in marine phytoplankton: allometric, taxonomic, and environmental variation. *Limnology and Oceanography*, 60, 540–552.

Edwards, K., Thomas, M., Klausmeier, C. & Litchman, E. (2016). Phytoplankton growth and the interaction of light and temperature: A synthesis at the species and community level. *Limnology and Oceanography*, 61, 1232–1244.

Eppley, R. (1972). Temperature and phytoplankton growth in the sea. *Fishery Bulletin*, 70, 1063–1085.

Fransz, H. & Verhagen, J. (1985). Modelling research on the production cycle of phytoplankton in the Southern Bight of the North Sea in relation to riverborne nutrient loads. *Netherlands Journal of Sea Research*, 19, 241–250.

Jewson, D.H., Rippey, B.H. & Gilmore, W.K. (1981). Loss rates from sedimentation, parasitism, and grazing during the growth, nutrient limitation, and dormancy of a diatom crop. *Limnology and Oceanography*, 26, 1045–1056.

Kowe, R., Skidmore, R., Whitton, B. & Pinder, A. (1998). Modelling phytoplankton dynamics in the River Swale, an upland river in NE England. *Science of The Total Environment*, 210, 535–546.

Le Pape, O., Jean, F. & Ménesguen, A. (1999). Pelagic and benthic trophic chain coupling in a semi-enclosed coastal system, the Bay of Brest (France): a modelling approach. *Marine Ecology Progress Series*, 189, 135–147.

Li, M., Gargett, A. & Denman, K. (2000). What determines seasonal and interannual variability of phytoplankton and zooplankton in strongly estuarine systems? *Estuarine, Coastal and Shelf Science*, 50, 467–488.

May, R.M. (1981). *Theoretical Ecology: Principles and Applications*. Oxford University Press UK.

McQuoid, M.R., Godhe, A. & Nordberg, K. (2002). Viability of phytoplankton resting stages in the sediments of a coastal Swedish fjord. *European Journal Phycology*, 37, 191–201.

Passow, U. (1991). Species-specific sedimentation and sinking velocities of diatoms. *Marine Biology*, 108, 449–455.

Picoche, C. & Barraquand, F. (2020). Strong self-regulation and widespread facilitative interactions between genera of phytoplankton. *Journal of Ecology*.

Qian, J. & Akçay, E. (2020). The balance of interaction types determines the assembly and stability of ecological communities. *Nature Ecology & Evolution*, 4, 356–365.

Reynolds, C.S. (2006). *The ecology of phytoplankton*. Cambridge University Press.

Sarthou, G., Timmermans, K.R., Blain, S. & Tréguer, P. (2005). Growth physiology and fate of diatoms in the ocean: a review. *Journal of Sea Research*, 53, 25–42.

Scranton, K. & Vasseur, D.A. (2016). Coexistence and emergent neutrality generate synchrony among competitors in fluctuating environments. *Theoretical Ecology*, 9, 353–363.

Wiedmann, I., Reigstad, M., Marquardt, M., Vader, A. & Gabrielsen, T. (2016). Seasonality of vertical flux and sinking particle characteristics in an ice-free high arctic fjord-Different from subarctic fjords? *Journal of Marine Systems*, 154, 192–205.