Environmental heterogeneity determines evolutionary pathways to species coexistence

Supporting information

Zhijie Zhang, Mark van Kleunen, Lutz Becks

December 1, 2025

Table S1: Information of the study species

species name	Abbreviation	ID in the R code	Source*
Chlamydomonas reinhardtii	CR	c	Chlamydomonas Resource Center
Acutodesmus obliquus	AO	13	SAG 276-3a
Chlorella vulgaris	CV	65	Konstanz
Chlorococcus infusionum	CI	G2	SAG 213-2a
Lobochlamys culleus	LC	G6	SAG 17.73
Pediastrum boryanum	PB	17	Konstanz
Scenedesmus sp.	SS	3	Konstanz
Vitreochlamys aulata	VA	G8	SAG 69.72

*SAG: Sammlung von Algenkulturen der Universität Göttingen (with strain number)
Konstanz: collection at the University of Konstanz

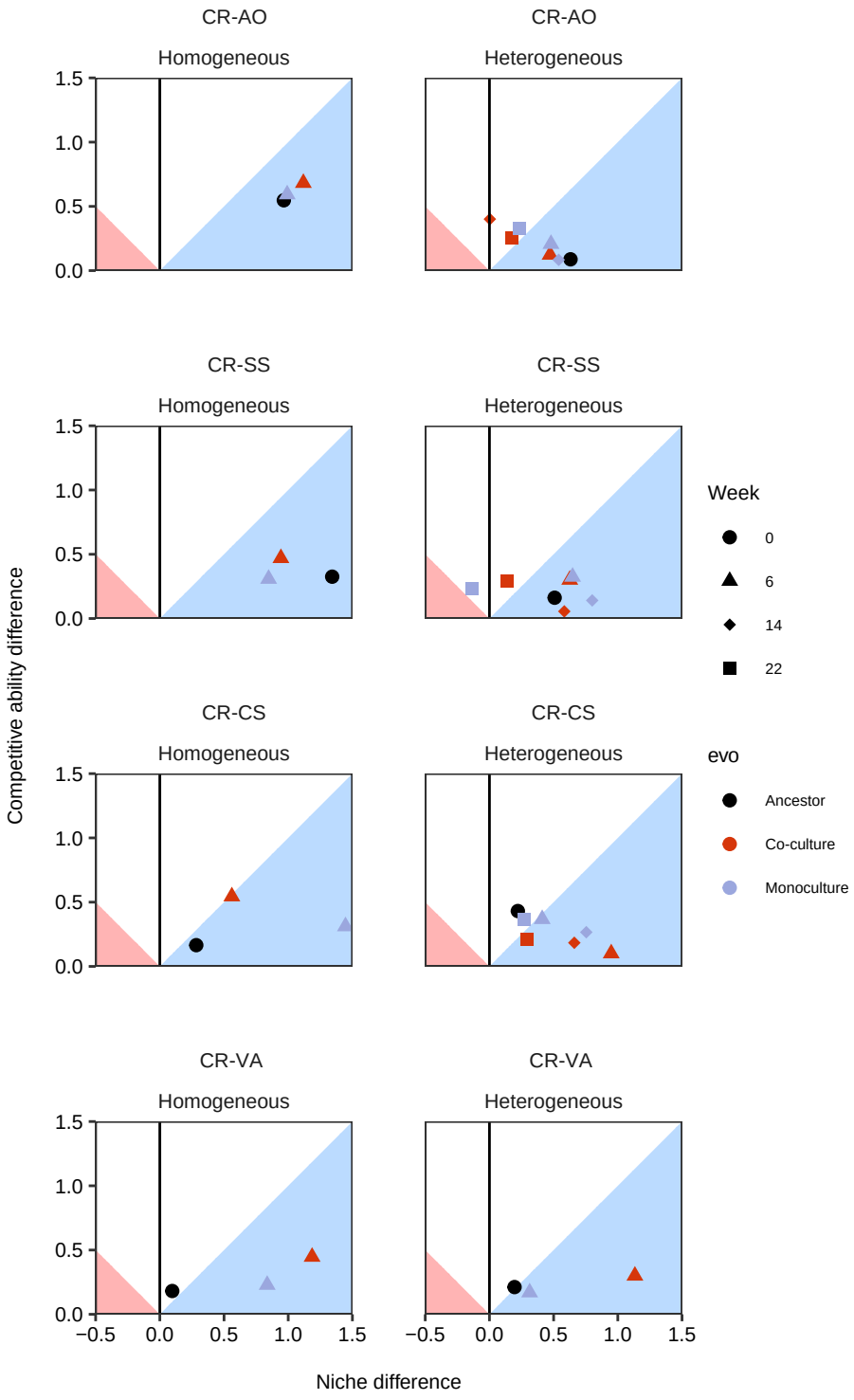

Figure S1: **Evolution of competitive-ability and niche differences of each species pair in the homogeneous and heterogeneous environments**. Abbreviations of species names are shown at the top of each panel. Black color indicates the ancestor. Red color indicates the co-cultures where the species pairs interacted and evolved together. Blue color indicates the monoculture controls where each species did not experience interspecific competition.

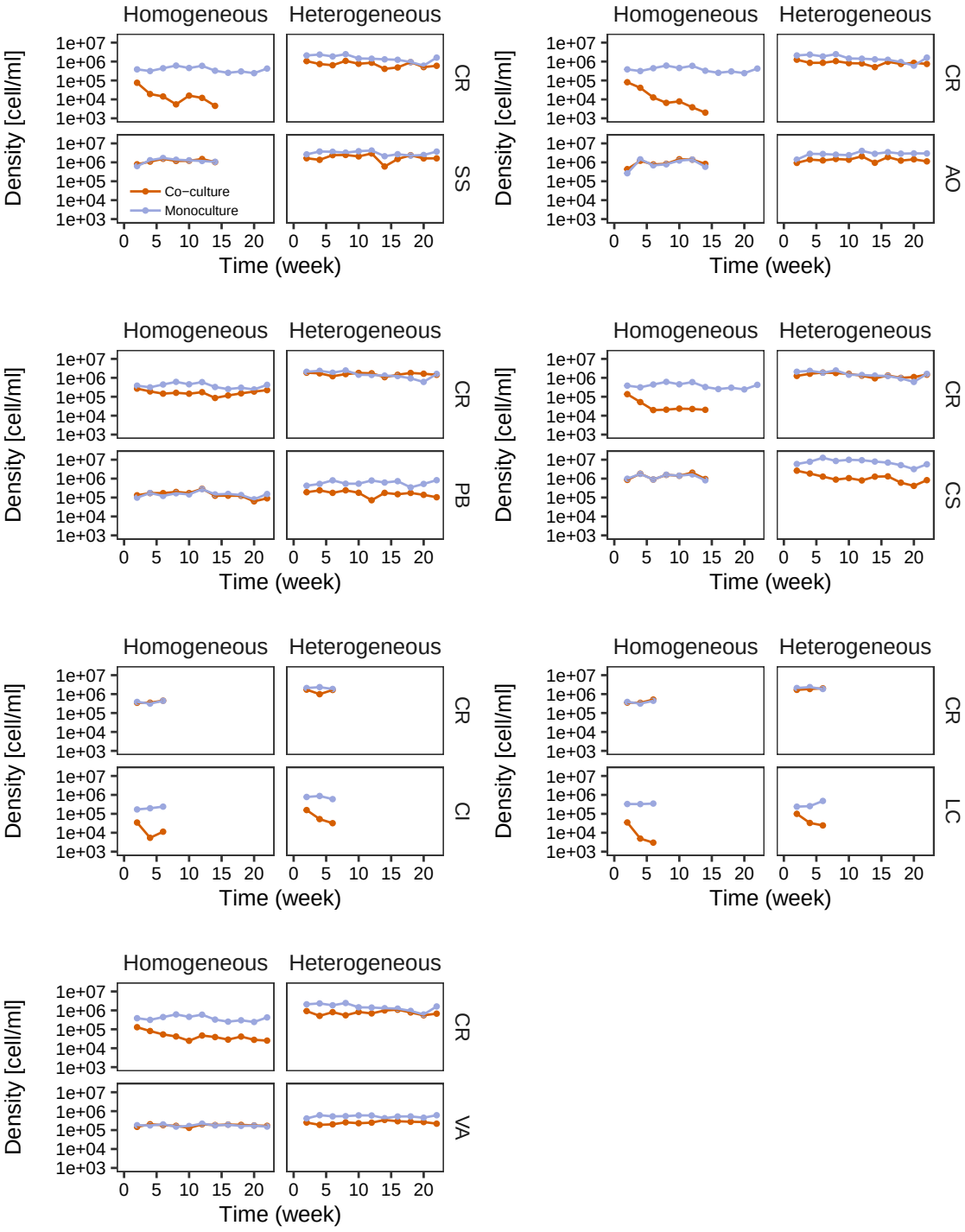

Figure S2: **Population dynamics of the co-cultures (red) and monocultures (blue) over 22 weeks.** In the homogeneous environment, when two species competed, one of them showed much lower density than its monoculture. This difference between co-cultures and monocultures was weaker in the heterogeneous environment. Full species names and their corresponding abbreviations can be found in Table S1.

Supplement Note 1 The relationship between R* and competitive ability

To explore the relationship between R * and competitive ability, we used a Consumer-Resource model where two consumers competed for one resource in chemostat system:

$$\frac{dR}{dt} = D(S - R) - c_1 R N_1 - c_2 R N_2 \tag{S1}$$

$$\frac{dN_1}{dt} = N_1 \left(\frac{w_1 c_1 R}{k_1 + c_1 R} - D \right) \tag{S2}$$

$$\frac{dN_2}{dt} = N_2 \left(\frac{w_2 c_2 R}{k_2 + c_2 R} - D \right) \tag{S3}$$

where S is the supply concentration of the resource and D is the Dilution rate. c_i is the consumption rate of the consumer on resource i. k_i is the half saturation constant at which the consumed resource level the consumer can reach half of its maximum growth rate on the resource i. w_i is the weighing factor that indicates the conversion efficiency of the consumer to the consumed resource i.

Then, we calculated the minimum resource requirement, R_i^*, for each species as

$$\frac{D k_i}{w_i c_i - D c_i} \tag{S4}$$

We calculated the growth rate of each species without any competition (i.e., the resource concentration equals the supply concentration, GR_i) and the invasion growth rate of each species when invading the equilibrium of the other species (IGR_i, i.e., the growth rate of species i at R_j^*). The supply concentration, S, was set to 1 and the dilution rate, D, was set to 0.1. All other parameters were randomly drawn from a uniform distribution ranging from 0 to 1. We repeated this 999 times (i.e., 999 pairs of species).

We calculated the competitive-ability differences, $\mathbf{F}$, following Carroll [1]:

$$\mathbf{F} = |\ln \sqrt{\frac{(GR_i - IGR_i)GR_j}{(GR_j - IGR_j)GR_i}}| \tag{S5}$$

Last, we tested whether competitive-ability differences were correlated with absolute differences in $R*$ between species with the *lm* function in *R*.

We found that competitive-ability differences were positively correlated with $R*$ differences (Fig. S5; Sum of sq = 547.5, $P < 0.001$). This indicates that convergence in resource use will reduce the competitive-ability difference. Note that while we used a simple model (e.g., linear consumption), changing the assumption would not qualitatively affect the results.

Supplement Note 2 Imaging and machine learning

Supplement Note 2.1 Imaging and data acquisition

All samples from the experimental evolution were collected in 96-well plates, fixed with 1% PFA (paraformaldehyde) and 10% GA (glutaraldehyde), and settled overnight for imaging. We applied a 5-fold dilution for the samples from day seven to day 12 in the heterogeneous environment and samples from day 10 to day 12 in the homogeneous environment, as the samples were too dense for imaging.

We imaged each well at 5 sites (covering 18.35% area of the well) using three filter sets and the autofluorescence of the algal cells at 10X magnification using IXM4 (ImageXpress® Micro 4 High-Content Imaging System). The three filters were Cy5 (exposure time: 10 ms), Cy3 (100 ms), and Texas red (100 ms), which excited the cell at three wavelengths and gave us three images per site.

We used custom modules to detect cells in the images. In summary, the cell was detected on the basis of the light intensity difference between the background and the cell in the Cy5 images. Then, we counted for each site the number of cells, and with the three wavelengths, exported 78 cell features that span auto-fluorescence intensity and cell morphology (e.g., cell length and size).

Supplement Note 2.2 Identifying species with machine learning

We built Neural Network Sequential Models to identify different algal species, using the *keras* [2] package in R. We randomly split the data collected from the monoculture experiment into two parts: 80% of the cells were used as the training data and 20% as the test data. Since the predictive accuracy of models typically decreases with the number of classes (e.g., the number of species), we trained a separate model to classify each species pair in the competition experiment rather than training a single model to classify all eight species.

For each model (i.e., species pair), we first filled the missing data with the mean of the corresponding feature. Second, we z-score transformed each of the cell features. Third, we built a Neural Network Sequential Model that contains four layers. The first to third layers contained 156, 60, and 30 units, respectively. The last layer outputted a length n numeric vector (probabilities for each digit) using a softmax activation function, with n equal to the number of species. Fourth, we assessed the accuracy of species identification with the test data. Last, we used the trained model to identify the species in the co-cultures after transforming the cell features in the same way as the training data.

The median accuracy of classifying different species was more than 98%, indicating that our machine learning approach provided reliable data of population dynamics.

Supplement Note 3 Model selection

Supplement Note 3.1 The selection of competition models

We used both the Beverton-Holt model and the Ricker model to parameterize the competition coefficients for each species pair per time point. As described in the main text, the Beverton-Holt model takes the following form:

$$\frac{N_{i,t+1}}{N_{i,t}} = \frac{\lambda_i}{1 + \alpha_{ii}N_{i,t} + \alpha_{ij}N_{j,t}}$$

where N_i is the density of individuals of species *i* at time *t*, λ_i is the intrinsic growth rate (i.e., per capita growth rate in the absence of competitors), and the competition coefficients α_{ii} and α_{ij} are the per capita effects of intraspecific competition of species *i* and interspecific competition of species *j* on species *i*, respectively.

The Ricker model takes the following form:

$$ln\frac{N_{i,t+1}}{N_{i,t}} = r_i - \alpha_{ii}N_{i,t} - \alpha_{ij}N_{j,t}$$

where r_i is the intrinsic growth rate of species *i*.

We found that the Beverton-holt model fitted the data better, with a median R^2 of 62.3%. In comparison, the median R^2 of the Ricker model was 51.2%.

Supplement Note 3.2 The selection of consumer-resource models

We used three sets of Consumer-Resource models to parameterize each species per time point, as described in the equations (1-4) of the main text. The three sets of models assumed that the resource conversion and consumption can be either linear or nonlinear. We did not fit the case where both conversion and consumption are linear, as this is biologically unrealistic. For each species per time point, we chose the best model based on R^2. Most populations ($> 80\%$) showed linear consumption on light but nonlinear consumption on nitrogen and phosphorus. A moderate percentage ($< 40\%$) of the populations showed linear conversion on nitrogen and phosphorus (Fig. S6).

References

1. Carroll, I. T., Cardinale, B. J. & Nisbet, R. M. Niche and fitness differences relate the maintenance of diversity to ecosystem function. *Ecology* **92,** 1157–1165. eprint: https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/10-0302.1. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/10-0302.1 (2011).
2. Arnold, T. B. kerasR: R Interface to the Keras Deep Learning Library. *J. Open Source Softw.* **2,** 296 (2017).

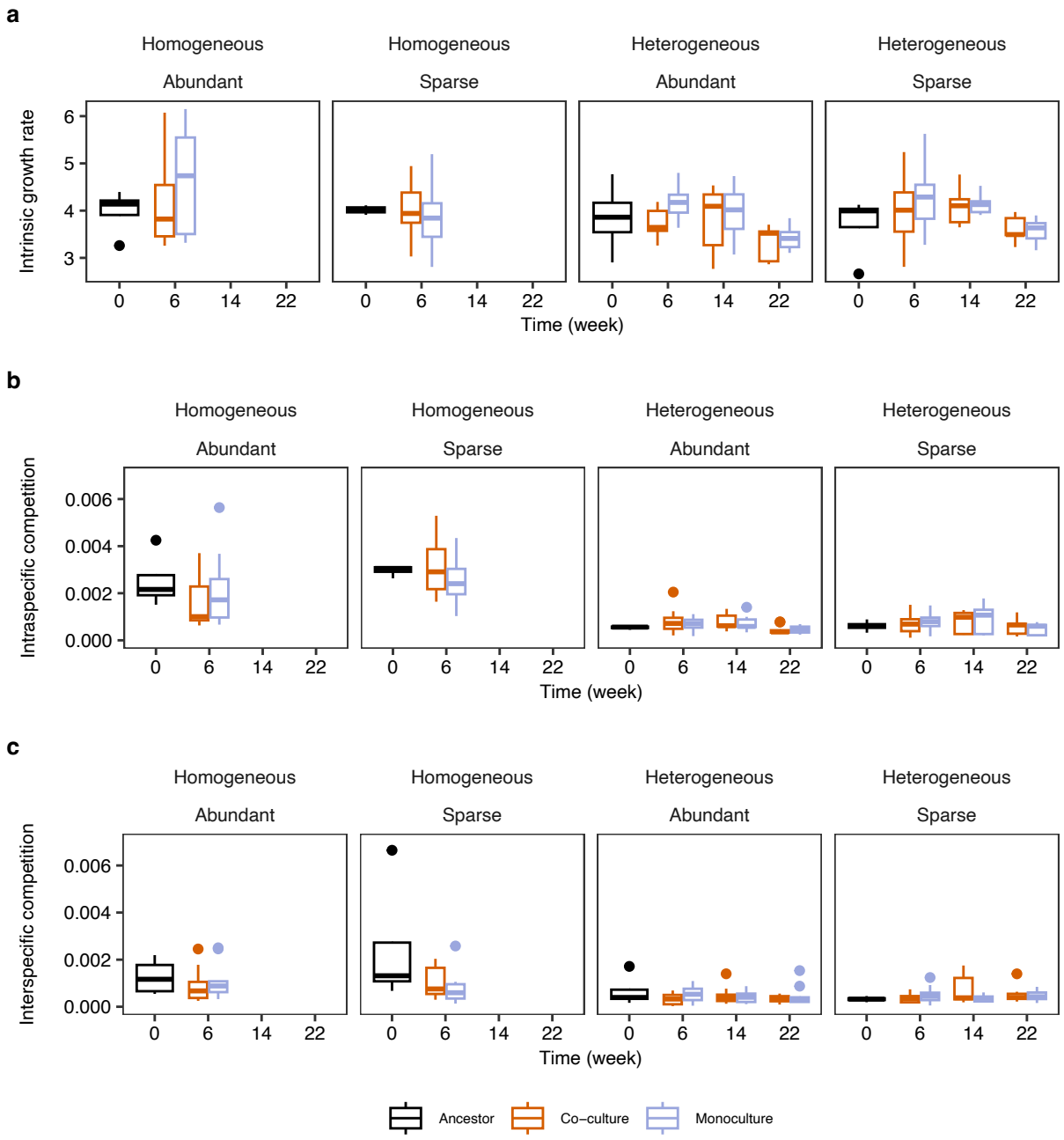

Figure S3: **Evolution of intrinsic growth rate and competition coefficients in the homogeneous and heterogeneous environments**. Black color indicates the ancestor. Red color indicates the co-cultures where the species pairs interacted and evolved together. Blue color indicates the monoculture controls where each species did not experience interspecific competition. We categorized each species per environment as abundant (Frequency $> 50\%$ in the co-culture over the course of experimental evolution) or sparse ($< 50\%$). We found that in the homogeneous environment, the sparse species that evolved in co-cultures were more sensitive to interspecific competition than their monoculture control did (left two panels in c; $\chi^2 = 3.02$, $P = 0.082$). The reverse was found for the abundant species. This indicates that the sparse species may have lost adaptation to tolerate interspecific competition from the abundant species. In the heterogeneous environments, the evolution of sparse and abundant species did not differ strongly with each other, probably because differences in the density between the two types of species were not as strong as the homogeneous environment (Fig. S2).

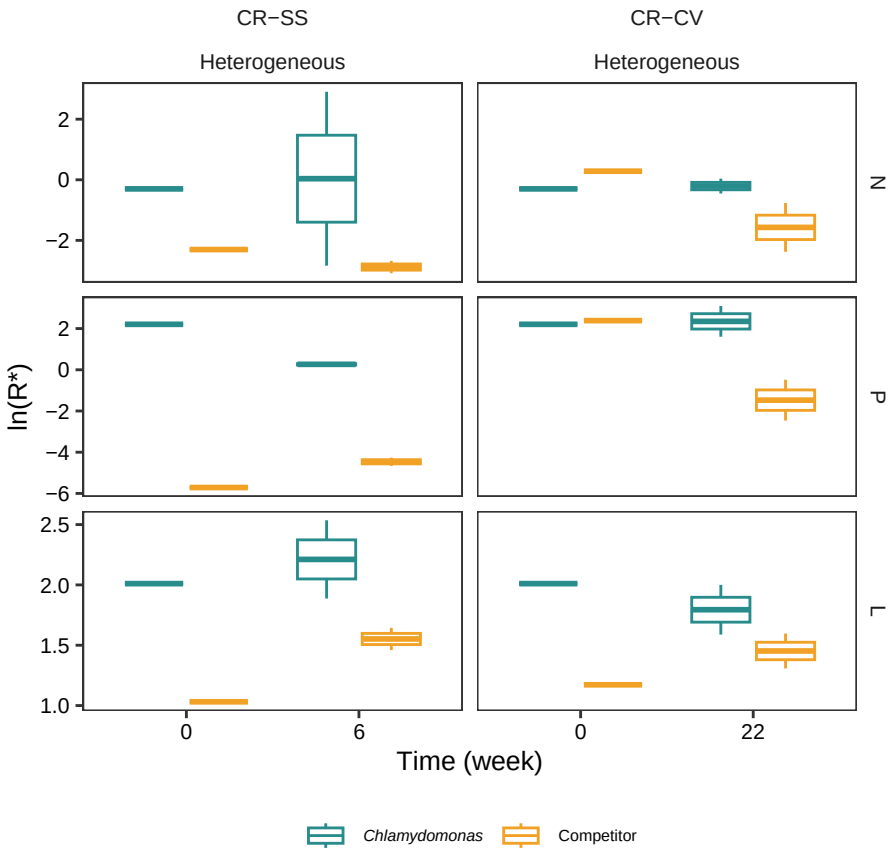

Figure S4: **Evolution of resource requirement for different competing species in the heterogeneous environment**. Left panel shows the pair of *Chlamydomonas reinhardtii* and *Scenedesmus sp.* The two species significantly converged in their minimum requirement of phosphorus (Sum of sq = 3.3904, P = 2.71×10^{-6}). The right panel shows the pair of *C. reinhardtii* and *Chlorella vulgaris*. The two species significantly diverged in their minimum requirement of phosphorus. (Sum of sq = 5.337, P = 0.014)

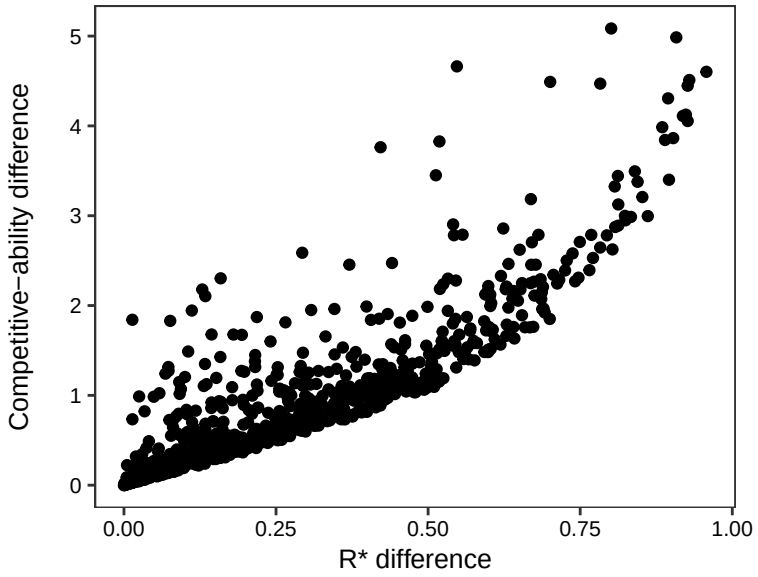

Figure S5: **The relationship between the competitive-ability difference and the difference in minimum resource requirement**. The results are from 999 pairs of species which were randomly generated.

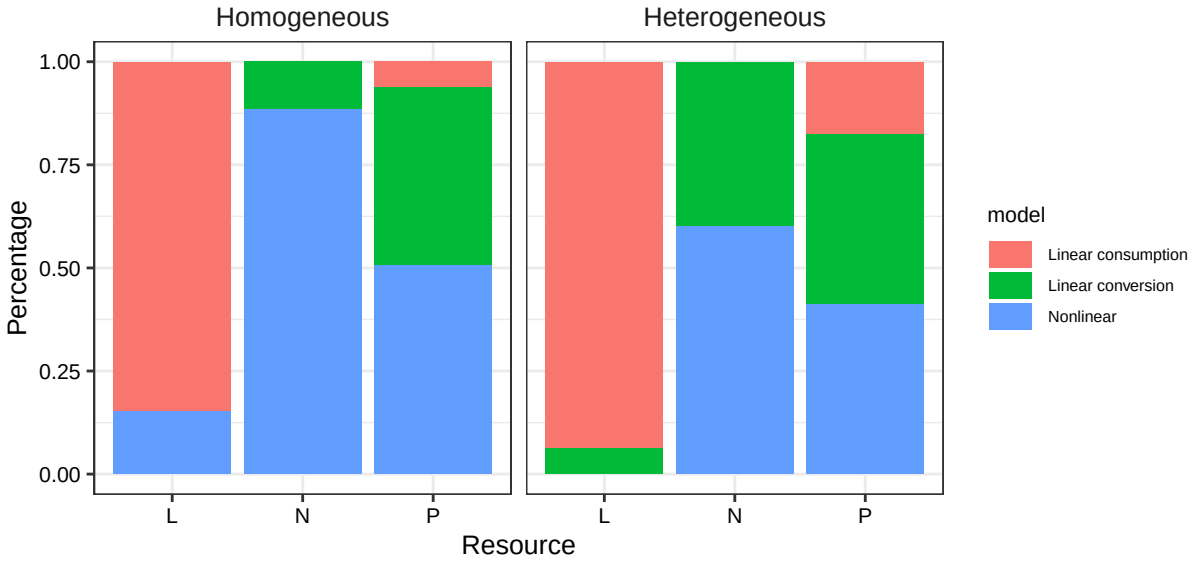

Figure S6: **Best Consumer-Resource models for different populations on light, nitrogen, and phosphorus.** The "linear consumption" model assumed that the consumption rate increased linearly with the resource and the conversion of consumed resource to growth rate increased asymptotically (nonlinear) with consumed resource. The "linear conversion" model assumed that the consumption rate increased asymptotically with resource and the conversion of consumed resource to growth rate was linear. The "nonlinear" model assumed consumption and conversion were nonlinear.