Estimating proportion of clones and genotype richness in aquatic microalgae

Ingrid Sassenhagen[1], Deana Erdner[2], Bryan Lougheed[3], Mindy Richlen[4], and Conny Sjöqvist[5]

[1]Uppsala University Faculty of Science and Technology
[2]The University of Texas at Austin
[3]Uppsala Universitet
[4]Woods Hole Oceanographic Institution
[5] ° Abo Akademi University

March 07, 2024

Abstract

Although the majority of microalgal species reproduce asexually for large parts of the growth season, most population genetic studies have rarely found clones in microalgal blooms. Instead, population genetic studies have identified large intraspecific diversity in most microalgal species. This paradox of frequent asexual reproduction but low number of clones creates challenges when interpreting the proportion of clones and distinct genotypes in natural microalgal populations. To estimate the proportion of clones and genotype richness, we created a computer model that simulates the composition of microalgal populations after a defined period of exponential growth. We simulated the probability of picking clones of the same genotype from this population as a function of initial genotype diversity, intraspecific differences in growth rates and sample size. This model was applied to five microalgal species for which high-resolution population genetic data and population growth rates based on monitoring data were available. The number of distinct genotypes in each population was extrapolated from the model outputs and the observed proportion of clones in the respective population genetic studies. The estimates from our simulation suggested that the genotype richness in most blooms exceeds several thousand distinct genotypes with very high variability among microalgal species. The highest numbers of distinct genotypes (500,000 and 2,000,000 genotypes) were estimated for species with very low numbers of observed clones in population genetic studies ($< 5\%$), but genotype richness was also strongly impacted by intraspecific variability in growth rates. Furthermore, the probability of finding clones and presumably sampling a representative fraction of genotypes increased significantly with higher sample sizes, challenging the detection of differences in genotype diversity between sub-samples with few isolates.

Hosted file

`manuscript8.doc` available at https://authorea.com/users/408556/articles/709168-estimating-proportion-of-clones-and-genotype-richness-in-aquatic-microalgae

Hosted file

`Table1.xlsx` available at https://authorea.com/users/408556/articles/709168-estimating-proportion-of-clones-and-genotype-richness-in-aquatic-microalgae

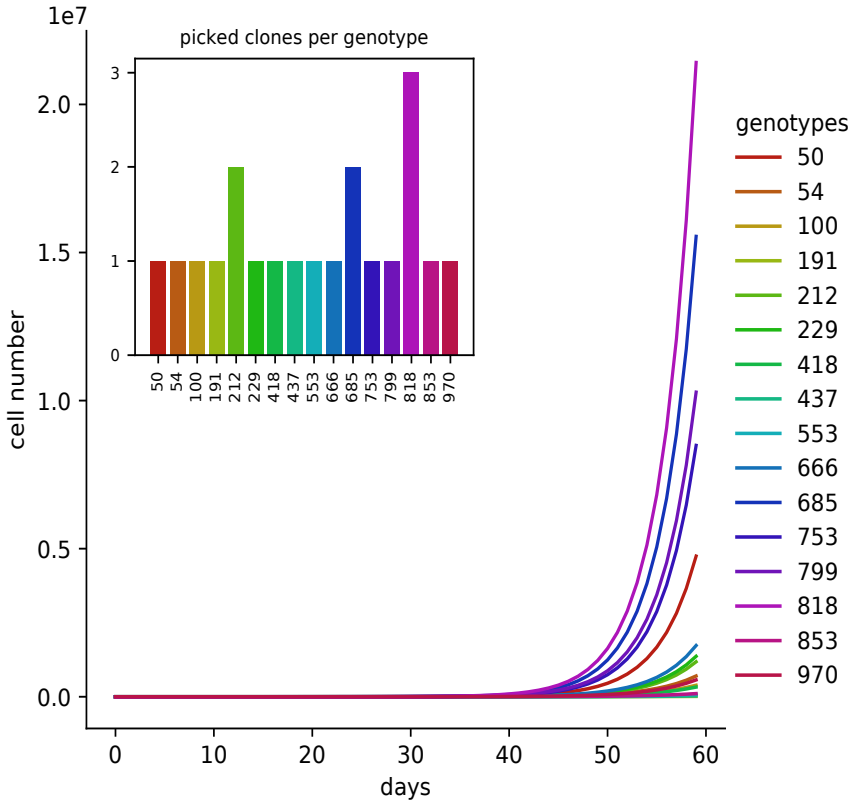

Figure 1: Visualization of the increase in cell numbers of sampled genotypes over 60 days and the resulting number of picked clones. For this model run, 1000 distinct genotypes, a sample size of 20 isolates, and a growth rate distribution with μ=0.2 divisions per day and σ=0.04 were chosen. Every genotype was assigned a number from 0-999.

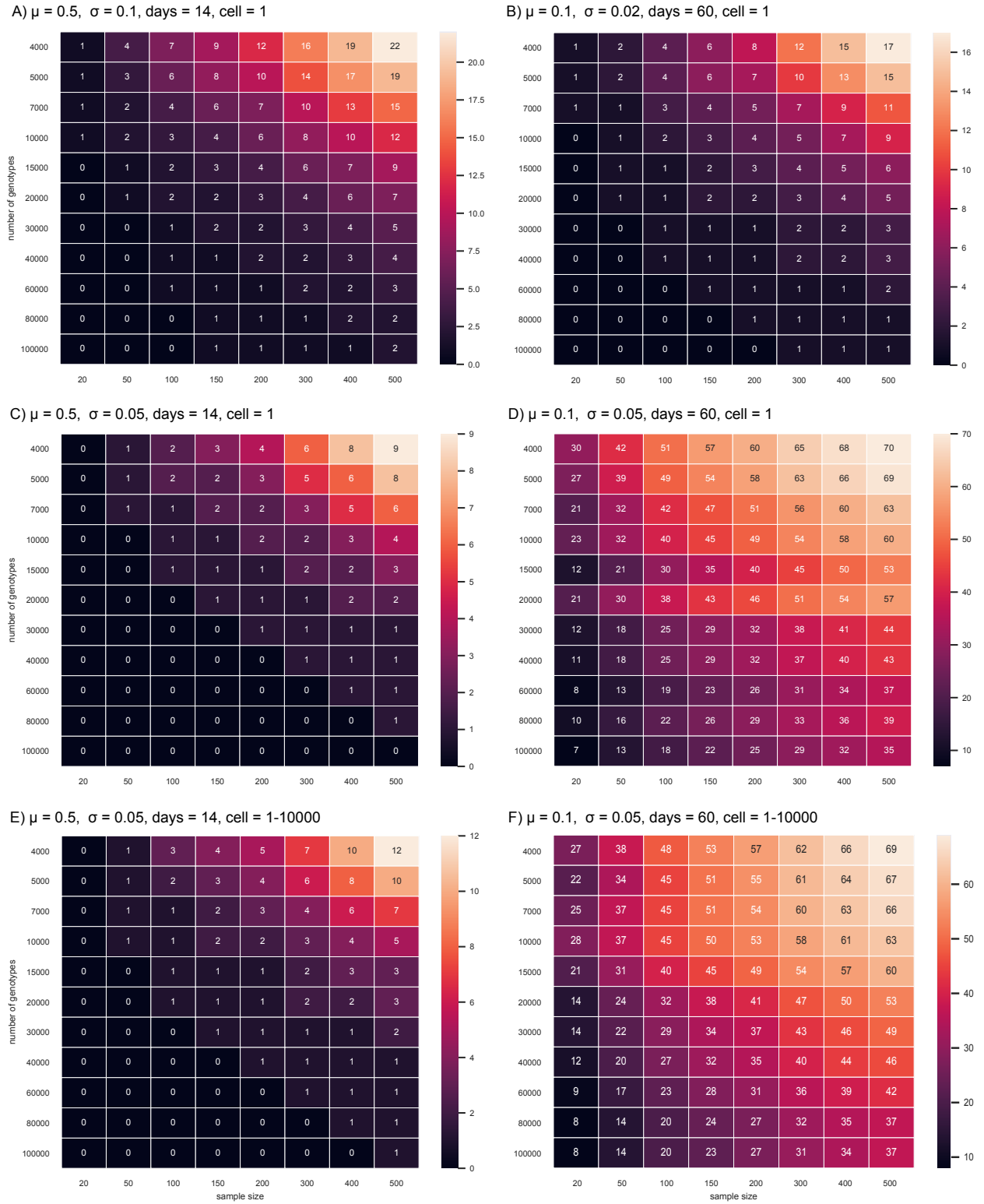

Figure 2: Heatmaps illustrating the probability of picking clones and the impact of intraspecific variability in growth rates (σ) in two ideal model species characterized by high and low average growth rates (μ), and different amount of initial cell concentrations per genotype (cell). A) Fast growing species with intermediate intraspecific variability and initially one cell per genotype. B) Slow growing species with intermediate variability and initially one cell per genotype. C) Fast growing species with low intraspecific variability and initially one cell per genotype . D) Slow growing species with high intraspecific variability and initially one cell per genotype. E) Fast growing species with intermediate intraspecific variability and heterogeneous initial cell number per genotype. F) Slow growing species with intermediate variability and heterogeneous initial cell number per genotype.

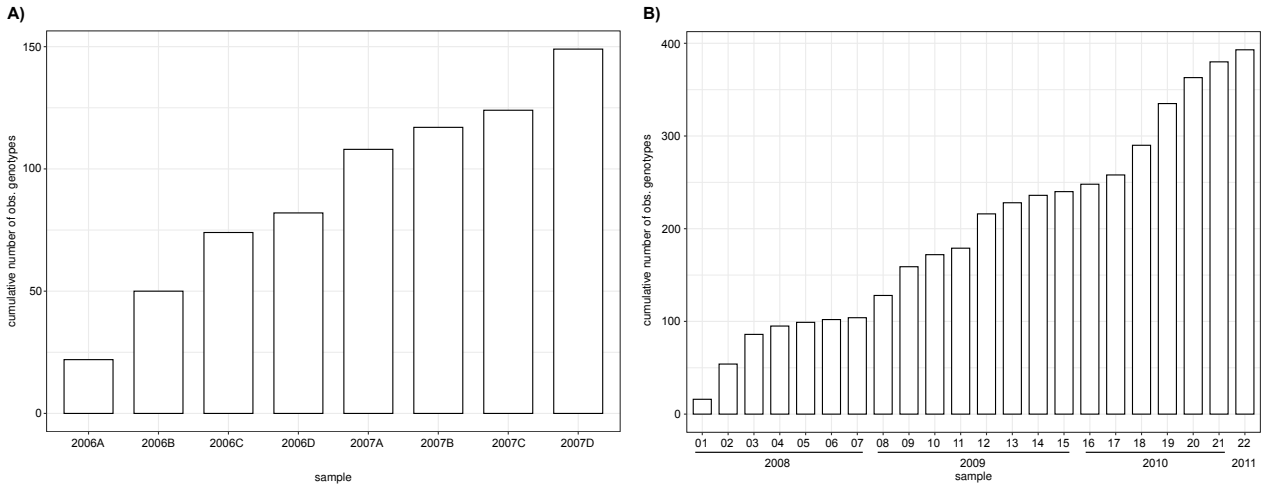

Figure 3: Cumulative number of observed genotypes in *Alexandrium catenella* (A) and *Pseudo-nitzschia multistriata* (B) across samples.

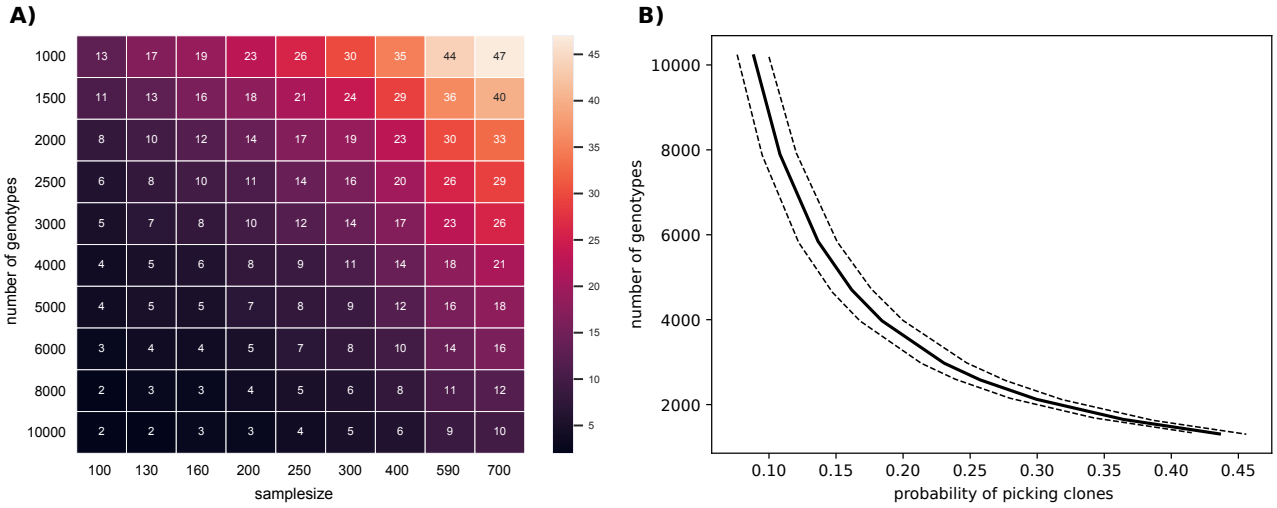

Figure 4: A) Heatmap visualizing the probability of picking clones of *Ditylum brightwellii* depending on number of initial genotypes and sample size. μ = 0.18, σ = 0.05, initial cell number per genotype = 1-100, exponential growth period = 20 days. B) Number of *D. brightwellii* genotypes as a function of the modeled probability to pick clones, assuming a sample size of 590 isolates and an average growth rate of 0.18 with a standard deviation of 0.05. Solid line indicates the mean (R^2=0.99), while dashed lines indicate upper and lower confidence intervals (σ).

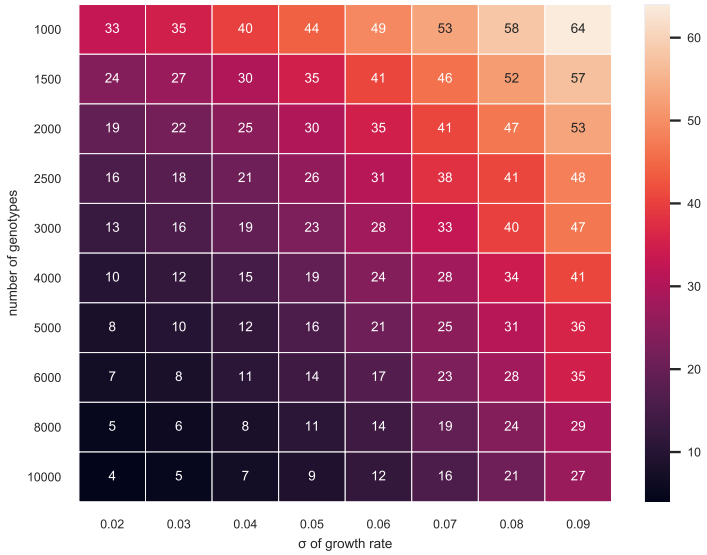

Figure 5: Heatmap visualizing the probability of picking clones of *Ditylum brightwellii* depending on number of initial genotypes and intraspecific variability in growth rates (σ). μ = 0.18, sample size = 590 isolates, exponential growth period = 20 days, initial cell number per genotype = 1-100.

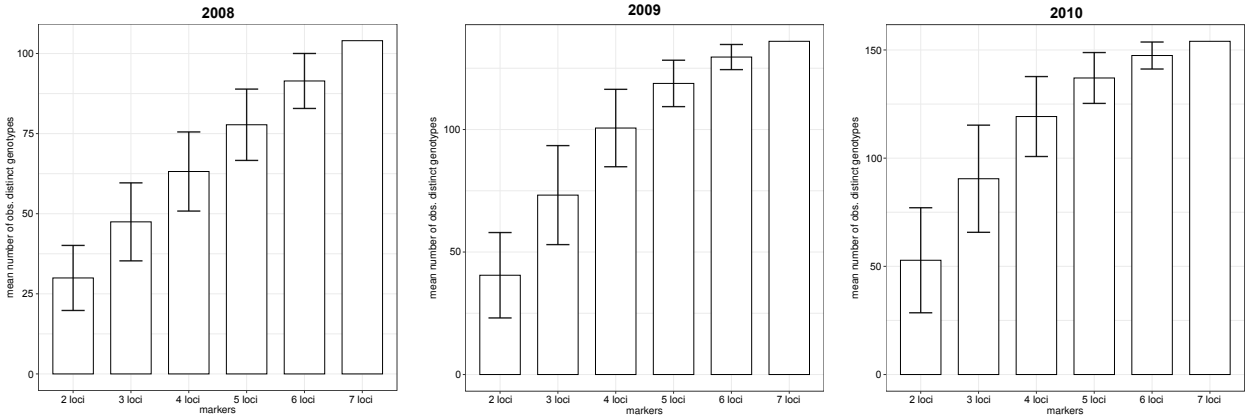

Figure 6: Mean number of observed distinct genotypes of *Pseudo-nitzschia multistriata* (Tesson et al. 2014) depending on the number of applied genetic markers. Error bars indicate standard deviation from the mean calculated from all possible combinations of the number of selected microsatellites.