Musters, C.J.M., O. Ieromina, S.H. Barmentlo, E.R. Hunting, M. Schrama, E. Cieraad, M.G. Vijver, and P.M. van Bodegom. Partitioning the impact of environmental drivers and species interactions in dynamic aquatic communities. Ecosphere.

Appendix S2: Effect of sample size and randomization

The efficiency of a random forest to predict the abundance of a taxon may depend on the ability of the method to distinct random effects from real effects and on the number of samples that it has as a learning set. To study the relationship between the number of samples and the performance of our data analysis, we used a present-absent data set of 206 biotic samples from our research area collected in two other years (2013: 115 sites; 2014: 91 sites), in a limited period of time within the year (viz. April-May). We used the same data set of 206 samples to study the effect of randomization by randomizing the presence of the species over the samples, while keeping the prevalence intact. Our data analysis did not find any taxon to have a R^2 deviating from zero in that case.

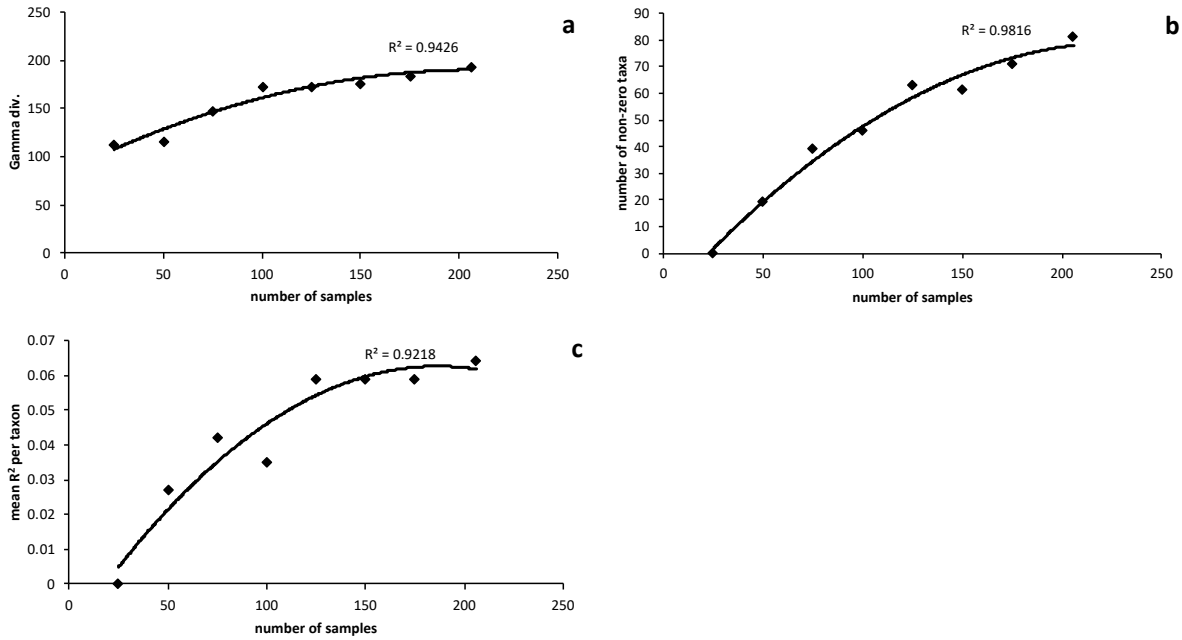

Figure S1 Effect of sample size on a: gamma diversity (total number of taxa), b: number of non-zero, i.e. predicted taxa and c: R^2 averaged over all taxa. Based on presence-absence data of 2013-14.