Supplement 2: Data consistency analysis

We are aware that long-term community datasets can potentially be biased by data inconsistency, which may originate from (1) changes in the methodological approach applied; (2) change in the person conducting the analysis (different experience/taxonomical knowledge); or (3) from long-term learning in taxonomy (Straile, Jochimsen & Kümmerlin, 2015).

For 1-2), we would like to emphasize that the Danube phytoplankton dataset has been produced using the same methodology (Utermöhl, 1958), sampled with similar sampling frequency (~once a week); and its counting counted by one person during the entire 34-year period (Keve T. Kiss; last author). For 3), we would like to emphasize that in the Danube phytoplankton dataset, we observed decrease in taxonomic richness. This is particularly important since long-term learning cannot lead to observation of long-term decrease in taxonomic richness.

We verified data consistency by performing asymptotic species richness estimates using the iNEXT R package for (A) each individual year based on presence-absence data: iNEXT(x, q=0, datatype="incidence_raw", endpoint=500); and (B) for individual samples based on abundance data: iNEXT(x, q=0, datatype="abundance", endpoint=NULL) (Chao *et al.*, 2014; Hsieh, Ma & Chao, 2016).

In A), estimated and observed richness of years were highly correlated (Fig 1A; Pearson cor., r=0.896, n=34, p<0.001); and time trend analysis indicated no significant change in the Estimated/Observed species richness ratio over time (Fig 1B; Mann-Kendall test, tau: 0.05, p=0.72). This may indicate that the sampling strategy and the counting approach were adequate to follow long-term changes in taxonomy consistently.

In B) the abundance-based estimated and observed species richness did not show difference (Fig 1C, Pearson cor., r=1, n=34, p<0.001). This may indicate that the estimated abundance based on counted individuals was highly adequate to cover the taxonomic richness of phytoplankton at sample level. On the other hand, the estimated bootstrap SE was very slightly higher between 2000 and 2010 (Fig 1D; GAM), which might be related to recent low abundances of phytoplankton in the middle Danube.
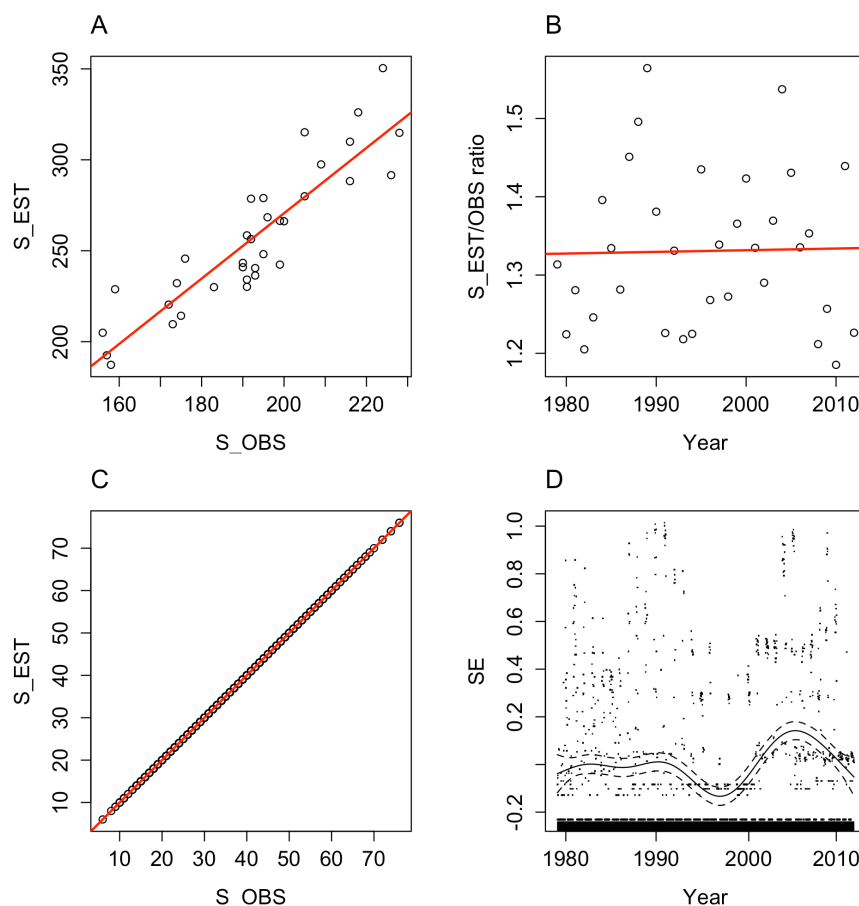
FIG 1 A) Correlation between estimated (in iNEXT using presence/absence data) and observed species richness in each year (Pearson, r=0.896, n=34, p<0.001); **B)** Time-trend in the yearly Estimated/Observed species richness ratio along individual years (Mann-Kendall trend analysis, tau=0.05, p=0.72); **C)** Correlation between estimated (in iNEXT using abundance data) and observed species richness of individual samples (Pearson, r=1, n=34, p<0.001); **D)** Time trend based on Generalized Additive Modeling (GAM) in bootstrap SE of asymptotic species richness estimates at the level of individual samples (GAM, p<0.001).

We also tested how consistency in taxonomic identification might affect the values of functional diversity components. The Danube dataset was organized in 2 main steps: 1) correction of taxonomic names (i.e. spelling mistakes), and then 2) harmonization among synonyms. These steps allowed us to compare the values of FD components using a larger (721 taxa following the 1st step) and a smaller (605 taxa following the 2nd step) taxonomic list; therefore with different levels in taxonomic harmonization. The least affected FD component was functional dispersion, while the FG approach

seemed to be more robust than the FT approach (Fig 2). Overall, however, values of FD components were robust based on both functional approaches.
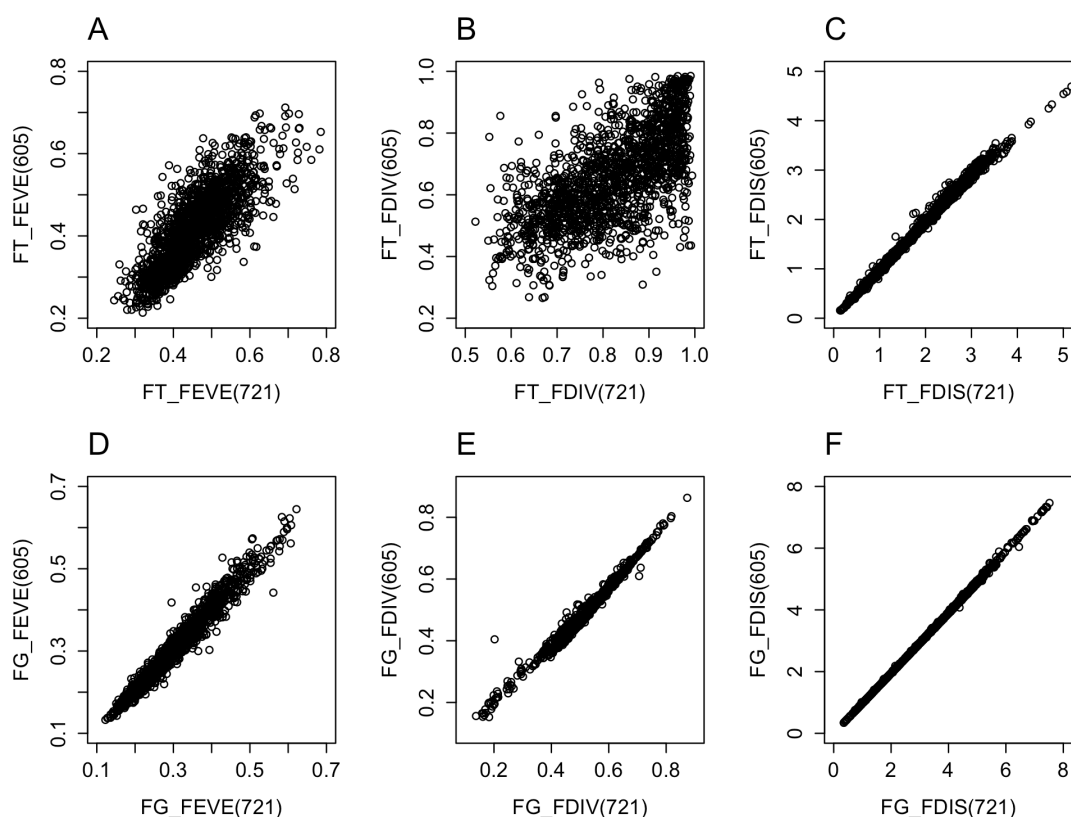

FIG 2 Correlation for each functional diversity component in the Danube phytoplankton dataset based on data harmonization steps 1) 721 taxa following spelling corrections and 2) 605 taxa following synonym corrections. **A)** FT_FEVE (Pearson corr., r=0.804, p<0.001); **B)** FT_FDIV (Pearson corr., 0.660, p<0.001); **C)** FT_FDIS (Pearson corr., 0.998, p<0.001); **D)** FG_FEVE (Pearson corr., 0.982, p<0.001); **E)** FG_FDIV (Pearson corr., 0.988, p<0.001); **F)** FG_FDIS (Pearson corr., 0.999, p<0.001); n=1608 in all cases.

We would like to emphasize that time trend analysis of each functional trait (FT) and functional group (FG) were performed individually (see Fig 6 and 7 of the manuscript). Therefore, independently of taxonomic names, abundance was pulled based on functional characteristics. This analysis therefore even less sensitive to taxonomy than the calculation of FD components.

We also included time-trend analysis for genus richness (G) as merged taxonomic units in each separate season (see Fig 5, Table 3 in the manuscript). However, genus level data cannot be used for functional diversity analysis. This is because functional characteristics of taxa can differ largely within genus (i.e. *Fragilaria construens* is a benthic, but *Fragilaria crotonensis* is a planktonic taxon; therefore they belong to different FGs (coda TB & P) and have different FT characteristics).

Given that the Danube phytoplankton monitoring has been focused on taxonomy, we consider the overall decrease in taxonomic richness, as well as its seasonal differences robust in our dataset.

References

Chao A., Gotelli N.J., Hsieh T.C., Sander E.L., Ma K.H., Colwell R.K. & Ellison A.M. (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs,* **84,** 45-67.

Hsieh T.C., Ma K.H. & Chao A. (2016) iNEXT: iNterpolation and EXTrapolation for species diversity. R package version 2.0.14 URL: http://chao.stat.nthu.edu.tw/blog/software-download/.

Straile D., Jochimsen M.C. & Kümmerlin R. (2015) Taxonomic aggregation does not alleviate the lack of consistency in analysing diversity in long-term phytoplankton monitoring data: a rejoinder to Pomati et al. (2015). *Freshwater Biology,* **60,** 1060-1067.

Utermöhl H. (1958) Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *Mitteilungen der Internationalen Vereinigung für theoretische und angewandte Limnologie,* **9,** 1-38.